Uncertainty, Imperfect Information, and Expectation Formation over the Firm's Life Cycle*

Cheng Chen Tatsuro Senga Chang Sun Hongyong Zhang†

July 2020

Abstract

Using a long-panel dataset of Japanese firms that contains firm-level sales forecasts, we provide evidence on firm-level uncertainty and imperfect information over their life cycle. We find that firms make non-negligible and positively correlated forecast errors. However, they make more precise forecasts and less correlated forecast errors when they become more experienced. We then build a model of heterogeneous firms with endogenous entry and exit where firms gradually learn about their demand by using a noisy signal. In our model, informational imperfections lead firms to enter the market without being fully informed. Moreover, young firms tend to wait long before entering or exiting the market faced with high uncertainty about their demand. The former learning effect, combined with the latter real-options effect, adversely affect firms' entry decisions and thus resource allocation. Our quantitative exercise substantiates the importance of accumulation of experience for firms' post-entry dynamics and aggregate productivity.

Keywords. firm expectations, forecast errors, uncertainty, learning, productivity

JEL Classification. D83; D84; E22; E23; F23; L2

*This research was conducted as a part of a research project funded by the Research Institute of Economy, Trade and Industry (RIETI). We thank Costas Arkolakis, Nick Bloom, Vasco Carvalho, Gene Grossman, Kyle Handley, Christian Hellwig, Aubhik Khan, Caliendo Lorenzo, Yulei Luo, Eduardo Morales, Steve Redding, Jane Ryngaert, Edouard Schaal, Peter Schott, Wing Suen, Stephen Terry, Mirko Wiederholt, and participants at various seminars and conferences for helpful comments. Financial support from HKGRF (project codes: 17500618, 17507916 and 27502318), JSPS KAKENHI (grant numbers: 17H02531, 17H02554), RIETI and Princeton University is greatly appreciated.

†Chen: Clemson University, cchen9@clemson.edu. Senga: Queen Mary University of London, ESCoE, and RIETI, t.senga@qmul.ac.uk. Sun: University of Hong Kong, sunc@hku.hk. Zhang: Research Institute of Economy, Trade and Industry, zhang-hong-yong@rieti.go.jp.

1 Introduction

A growing literature has highlighted the importance of uncertainty and imperfect information in driving firm dynamics and aggregate productivity.[1] In fact, firms face uncertainty and imperfect information when making almost all decisions in a dynamic environment, including investment, hiring, and market entry.[2] A key part of these decisions is to form expectations of future outcomes, such as sales and profits. However, as we seldom observe firms' expectations over their life cycles directly, how firms respond to and resolve uncertainty and imperfect information over time remains unknown. This makes it difficult to quantify the degree of informational imperfections and to evaluate how much they matter for aggregate outcomes such as aggregate productivity.

In this paper, we make empirical progress by using panel data with quantitative measures of sales expectations at the firm level. Unlike other datasets featuring old and large firms, ours is designed to include young firms since their inception, allowing us to track how firms resolve uncertainty and form expectations over their life cycle.[3] Our main finding is that firms become better informed of future sales over their life cycle. Specifically, we present evidence that the precision of forecasts increases with firms' experience. Moreover, although each firm's forecast errors are positively autocorrelated, which implies imperfect information, the autocorrelation of forecast errors declines with firms' experience, a new fact overlooked by existing studies. This is where our paper makes a distinct point in that we provide direct evidence suggesting that firms accumulate experience and make more informed decisions over time, the way of thinking in the theoretical literature as in Jovanovic (1982).

To investigate its aggregate implications, we then build a model of heterogeneous firms with endogenous entry, and embed Jovanovic-type learning into this framework. Firms' entry/exit decisions may be distorted because firms are not fully informed in the presence of uncertainty and imperfect information, which leads to productivity losses. We use our dataset to calibrate the model and show that these productivity losses are substantial. To our knowledge, our analysis is the first to robustly isolate and succeed in reproducing the age-declining characteristics of firm-level forecast errors, shedding light on the quantitative importance of uncertainty and imperfect information as drivers of firm dynamics and, in

[1] See, for example, Bloom (2009) and Bloom et al. (2018), for seminal works.

[2] It is commonly understood that uncertainty matters for individual-level decision making, such as investment (Guiso and Parigi, 1999), hiring (Bertola and Caballero, 1994), market entry (Dixit, 1989), and trade (Handley, 2014).

[3] This is true for the dataset used in Coibion et al. (2020). Although datasets used in Bloom et al. (2017) and Coibion et al. (2018) cover small firms, these datasets do not cover firms over a long time. Finally, the Ifo business climate survey used in Bachmann et al. (2013), Bachmann and Elstner (2015), and Enders et al. (2019b) covers a wide range of firms over a long time, it contains only qualitative firm-level expectations.

turn, aggregate productivity.

The dataset we use is a parent firm-affiliated firm-matched 20-year panel dataset on Japanese multinational firms—taken from annual business surveys conducted by the Japanese government. Our dataset has three distinctive features: (1) it contains both quantitative expectations and realized outcomes for each firm, which allows us to calculate forecast errors at the firm level; (2) its panel structure and the inclusion of many young firms enable the analysis of within-firm variation of forecast errors over the firm's life cycle; and (3) it is a confidential, mandatory survey enforced by the government, which leads to high response rates (70% on average) and high quality.

Exploring our dataset, we show the following features of forecast errors made by an individual firm regarding its sales. First, firms make small forecast errors on average, and firm-level components explain most of the variations in forecast errors (compared with aggregate components, e.g., country-year and industry-year fixed effects). Second, firms make more precise forecasts as they become more experienced in the market, either through multinational production or exporting before it. Moreover, this result survives even when we control for firm size and measures of market/product diversification at the firm level, as older firms tend to be larger than young firms and have more diversified businesses and products than those firms because of better management. Third, past forecast errors are positively correlated with current and future forecast errors. This is similar to Coibion and Gorodnichenko (2012)'s finding regarding households' and professional forecasters' forecast errors; however, we also show that this positive autocorrelation declines over the life cycle of firms. In summary, we find that the longer firms operate, the more informed they become. This is captured by the declining variance and serial correlation of forecast errors as firms become more experienced.

In light of this evidence, we build a model that integrates Jovanovic (1982) type of learning into a standard monopolistic competition model. Firms face a downward sloping demand curve in a setting where the demand is heterogeneous across firms and the firm-specific demand is the sum of a permanent component and a transitory component as in Arkolakis et al. (2018). We extend this model with two key ingredients.

First, we introduce a noisy signal, thereby allowing the firm to *gradually* learn its permanent demand component. In particular, we follow Kydland and Prescott (1982) to assume that the firm cannot differentiate the information noise from its permanent demand component, but it can differentiate the transitory and permanent demand components. This additional information structure relative to Jovanovic (1982) allows us to reproduce the age-declining serial correlation of forecast errors in our model, allowing us to achieve clean mapping from data to the model parameters.[4]

Second, we introduce endogenous entry and exit. This element, coupled with imperfect information, is essential to reproducing the above-mentioned life cycle of firms consistent with our dataset, thereby allowing us to investigate the effects of accumulation of experience on firm dynamics and aggregate productivity. To fully utilize the international context of our dataset, we focus on firms' decisions on entering multinational production by switching from exporting exporting and exiting from exporting. Exporting requires that firms pay iceberg trade costs and multinational production requires sunk entry costs but no such iceberg trade costs.[5] This implies that a firm's optimal behavior for entering multinational production is described by Ss rules.[6]

In our model, the Ss rules are belief-driven.[7] Firms start multinational production when their expected permanent demand draws are above a certain threshold. Different from a perfect information model where firms sort by the permanent demand draws perfectly, such *sorting* is imperfect in the learning model. As a result, it creates productivity losses, and this mechanism is called the learning effect. Moreover, uncertainty implies a negative impact of real options on multinational production. That is, young exporting firms have high uncertainty due to the lack of experience, which is captured by a large posterior variance of their belief about permanent demand. For those firms, the option value of waiting is high, and accordingly, they adopt a "wait-and-see" rule (i.e., inaction in the sense of continuing to export) for multinational production and for exiting. This inaction also creates productivity losses, holding some similarity with Bloom (2009) and Bloom et al. (2018).[8]

In the final part of the paper, we quantify these productivity losses using our model and dataset. A defining feature of our parameterization is that we use cross-sectional moments of forecast errors to pin down the key parameters that determine the degree of informational imperfections in the model. Moreover, we infer the degree of informational imperfections on a region by region basis by using the cross-country aspect of our dataset. We first group countries into eight regions based on geographic and economic similarities. For each region, we then back out the parameters by matching the declining pattern of the absolute value and the autocorrelation of forecast errors over the firm's life cycle. Next, we calculate

[4]The age-declining variance of forecast errors can be reproduced in a model of perfect information by making volatility of productivity/demand shifter age-dependent as in Atkeson and Kehoe (2005). However, the age-declining serial correlation of forecast errors requires imperfect information because these models exhibit zero serial correlation of forecast errors.

[5]See Helpman et al. (2004) for this modeling approach.

[6]See, for example, Dixit and Pindyck (1994) and Abel and Eberly (1996).

[7]See Baley and Blanco (2019) for such pricing rules by firms under uncertainty.

[8]The difference here is that they study variations of uncertainty over time, and we focus on how uncertainty evolves over the firm's life cycle.

the signal-to-noise ratio (variance of the demand draw divided by variance of the signal noise), an inverse measure for the degree of informational imperfections, for the eight regions respectively. This reveals substantial differences in the degree of informational imperfections across regions. For instance, the signal-to-noise ratio is the highest in the United States (U.S.), followed by Western Europe. The lowest and the second-lowest signal-to-noise ratios are in the Middle East and (Sub-Saharan) Africa. Taken together, the signal-to-noise ratio in the United States is four times as high as its value in Africa and 86 times as high as its value in the Middle East.

To put this in context, we exploit the variation of the signal-to-noise ratio in our model and show that it leads to substantial differences in aggregate productivity. Aggregate productivity in the model increases by 64% when the signal-to-noise ratio increases from its lower bound (the Middle East) to its upper bound (the U.S.) implied by our data. The corresponding number is 7.3% when the signal-to-noise ratio increases from its second-lowest value (Africa) to its upper bound. In addition, when we move from the perfect information benchmark to our imperfect information model with the signal-to-noise ratio set at its upper bound, aggregate productivity decreases by 3.7%. In the same exercise, with the signal-to-noise ratio set at the level of Africa, the corresponding productivity loss is 11.2%. In summary, efficiency losses in production due to imperfect information and uncertainty are substantial.

Our results closely echo the literature of information and technology adoption.[9] We study a choice between multinational production and exporting as the multicountry aspect of our dataset guides us; though, this is indeed a problem of technology adoption, as starting multinational production is thought of as equivalent to adopting a new technology relative to exporting that is more costly in terms of the variable cost. One related strand of the literature has emphasized that similarities between technologies and other factors can facilitate firms to adopt new technologies, because using existing technologies yields information value that can be used to exploit new technologies.[10]

Moreover, there are two aspects that distinguish our paper in the literature. First, we utilize direct expectations data. While economists have long speculated on how agents form expectations, it is the lack of direct expectations data that has made the treatment of agents' expectations an assumption-based approach. A growing literature breaks with this tradition by collecting and analyzing direct expectations data as in Bloom et al. (2017) and Altig et al.

[9] Foster and Rosenzweig (2010) include an extensive review of the literature that substantiated the importance of learning in determining the adoption of new technologies.

[10] See Irwin and Klenow (1994), Parente and Prescott (1994), and Jovanovic and Nyarko (1996) for the analysis of learning and information spillovers between technologies in the context of firms.

(2019), among others.[11] Notably, the seminal works by Coibion and Gorodnichenko (2012) and Coibion and Gorodnichenko (2015) have demonstrated how to best model and calibrate a theoretical framework and thus highlighted the usefulness of such a direct-measure-oriented approach.[12] Our paper differs from these studies because we study firms expectations of their future circumstances and do not rely on macroeconomic expectations. Moreover, we focus on the life-cycle properties of firms expectations and provide insights into young firms post-entry dynamics as emphasized by, for example, Sedláček and Sterk (2017) and Foster et al. (2016).

Second, we identify and quantify misallocation due to imperfect information along the life cycle of firms. Our focus on the life-cycle of firms is reminiscent of Hsieh and Klenow (2014), and our quantitative exercise are related to David et al. (2016) who also substantiated the role of imperfect information in determining allocative efficiency. What distinguishes our paper from theirs is that we show that productivity losses through extensive margin dynamics—firms' entries and exits—are substantial. Regarding the importance of the extensive margin, our paper complements the results by Midrigan and Xu (2014) and Buera et al. (2011), among other papers on misallocation, although they have focused on financial frictions.[13]

These two innovations can be extended in future research because firm-level expectations data concerning firm-specific variables are increasingly available and misallocation has been shown to matter for macro development, which crucially depends on technology adoptions.[14]

2 Empirical Facts

In this section, we document a set of stylized facts about firms' expectations over their life cycles. We construct our panel of Japanese firms operating in foreign markets to study the properties of the forecast errors and their relationship with firms' experience. First, the forecast errors made by firms become smaller as they become more experienced. Second, the forecast errors are positively autocorrelated, but the serial correlation declines as they become more experienced. Overall, the results presented in this section indicate that firms become better informed as they accumulate more experience.

[11] Other papers that study micro-level expectations include Gennaioli et al. (2016) for U.S. firms, Bachmann and Elstner (2015), Bachmann et al. (2017), and Enders et al. (2019a) for German firms, Boneva et al. (2020) for firms in the United Kingdom, Coibion et al. (2020) and Ma et al. (2019) for Italian firms, and Tanaka et al. (2019) for Japanese firms.

[12] Recent papers that have studied how agents form expectations and respond to shocks include Baker et al. (2018), Bordalo et al. (2018) and Coibion et al. (2018).

[13] Related literature includes Khan and Thomas (2013) and Buera and Moll (2015) who studied the role of financial frictions in generating capital misallocation and its aggregate implications.

[14] See Hsieh and Klenow (2009) and Hsieh and Klenow (2014) for evidence.

2.1 Data and the Reliability of Sales Forecasts

This subsection describes our panel data of Japanese firms operating in foreign markets. Our main data source is the Basic Survey on Overseas Business Activities ("foreign activities survey" hereafter) conducted by the Ministry of Economy, Trade and Industry (METI). The survey contains information on overseas affiliated firms of Japanese parent companies, including the affiliated firms' location, industry, sales, and employment. The survey covers two types of overseas businesses: (1) direct (first-tier) affiliated firms with more than 10% of the equity share capital owned by Japanese parent companies, and (2) second-tier affiliated firms with more than 50% of the equity share capital owned by Japanese parent companies. Our baseline sample contains on average 1784 parent companies and 6950 affiliated firms in a typical year from 1995 to 2014. The unit of analysis in our empirical investigation is the affiliated firm by year. We slightly change the terminology: We refer to the affiliated firms as "firms" and to all the affiliated firms belonging to the same parent company as a "business group".

For some of our analysis, we combine the foreign activities survey with the Basic Survey of Japanese Business Structure and Activities ("domestic activities survey" hereafter), another survey conducted by METI. The latter survey can be matched to the foreign activities survey and provides additional information on the parent companies; it contains data on each parent companys exports to seven regions—North America, Latin America, Asia, Europe, the Middle East, Oceania, and Africa—which allows us to measure each firms export experience (inherited from the parent companies) even before its entrance into the foreign market for multinational production.

The unique feature of the foreign activities survey is that each firm reports its sales forecast for the next year when it fills out the survey of the current year. Because such information is rarely available in firm-level datasets, we show that the sales forecasts are reliable and contain useful information that affect actual firm decisions.

First, we show that firms do not use naive rules to make their sales forecasts. In Table 1, we present the expected growth rates, calculated as the ratio of the firm's forecast for year $t + 1$ to its realized sales in year t minus one. If firms simply use their realized sales in year t to predict their sales next year, the expected growth rate will be zero. In Table 1, only 3.34% of the observations in our sample have a zero expected growth rate. The shares of the other frequent cases are all extremely low. For the firms reporting zero expected growth rates, it is difficult to tell whether they are making a naive forecast or making a serious forecast with the expectation that their sales growth will be close to zero. We therefore conduct robustness checks of our main regressions in Online Appendix Tables OA.5, OA.14 and OA.18 by dropping all observations with zero expected growth rates. Our main empirical results remain largely unchanged.

Table 1: The Most Frequent Values of Expected Growth Rates

Top 1-5		Top 6-10	
$E_t(R_{t+1})/R_t - 1$	Freq. (%)	$E_t(R_{t+1})/R_t - 1$	Freq. (%)
0.0000	3.34	0.0714	0.11
0.1111	0.22	0.0417	0.11
0.2500	0.20	0.3333	0.11
0.0526	0.17	0.0870	0.11
0.2000	0.14	1.0000	0.10

Notes: The most frequent values of expected growth rates among all firm-year observations. The total number of observations is 166,200.

Second, we show that the sales forecasts have statistically significant and economically strong impacts on future firm outcomes. Specifically, we regress the realized sales in year $t+1$ on the sales forecast made in year t and a set of fixed effects, and the results are reported in Table 2. The first three columns of Table 2 show that the sales forecasts in year t positively and significantly predict the realized sales in year $t+1$. Importantly, the effect of the sales forecast does not disappear when we include the realized sales in year t as a control variable in Column 2. Its coefficient is actually much larger than the coefficients of realized sales in the previous year. Further including the realized sales in year $t-1$ does not change this pattern (Column 3). Columns 4-6 show that the sales forecasts also have strong predicative power for future employment, even if we control for current and past employment. These findings easily reject the hypothesis that firms fill out this survey question with random guesses. By contrast, firms take these forecasts seriously, and the forecasts contain more information on the firms' future conditions than realized outcomes in the past.

Finally, the foreign activities survey is mandated by METI under the Statistics Law; thus, the information in the survey cannot be applied to purposes beyond the scope of the survey, such as tax collection. Firms have no incentive to misreport because of tax purposes. Moreover, unlike earnings forecasts announced by public firms, the sales forecasts reported to METI are confidential; thus, firms have no incentive to misreport strategically and manage the expectations of the stock market. In total, the aforementioned empirical patterns assure us that the sales forecasts contained in the foreign activities survey are reliable and suitable for our empirical analysis.

Table 2: Sales Forecasts Predict Firms' Future Outcomes

Dep. Var.	log total sales $\log(R_{i,t+1})$			log employment $\log(L_{i,t+1})$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\log E_t(R_{i,t+1})$	0.675^a	0.551^a	0.583^a	0.306^a	0.132^a	0.132^a
	(0.0111)	(0.0123)	(0.0147)	(0.0137)	(0.00648)	(0.00687)
$\log R_{it}$		0.140^a	0.0819^a			
		(0.00848)	(0.0157)			
$\log R_{i,t-1}$			0.0641^a			
			(0.00725)			
$\log L_{it}$					0.519^a	0.508^a
					(0.0117)	(0.0146)
$\log L_{i,t-1}$						0.0596^a
						(0.00732)
Country-Year FE	Y	Y	Y	Y	Y	Y
Industry-Year FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
N	129202	127532	107017	127751	126803	107060
# of B-groups (cluster)	4960	4940	3868	4946	4932	3848
Within R-squared	0.481	0.488	0.495	0.166	0.393	0.398
R-squared	0.961	0.964	0.967	0.957	0.969	0.972

Notes: The dependent variable is firm i's log total sales or total employment in year $t+1$. We use R to denote sales and L to denote employment. $E_t(R_{i,t+1})$ refers to the firm's expectation in year t for its sales in year $t+1$. Standard errors are clustered at the business group level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

2.2 Forecast Errors

We now describe how firms' forecast errors evolve over their life cycles. Our main measure of forecast errors is the log point deviation of the realized sales from the sales forecast as

$$FE_{t,t+1}^{\log} \equiv \log\left(R_{t+1}/E_t\left(R_{t+1}\right)\right),$$

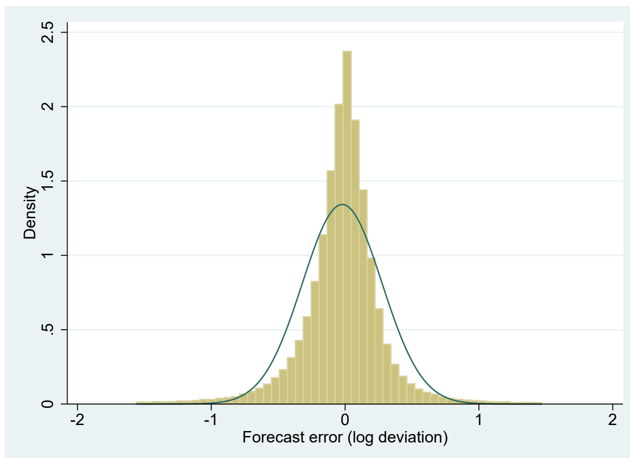
where R_{t+1} is the realized sales in period $t+1$ and $E_t\left(R_{t+1}\right)$ denotes a firm's time t forecast of its sales next period. A positive (negative) forecast error means that the firm under-predicts (over-predicts) its sales. In the online appendix, we show that our key empirical results are robust to two alternative definitions of forecast errors: the percentage deviation, and the residual of raw forecast errors after removing aggregate components such as industry and country-year fixed effects.[15] We also trim the top and bottom one percent of observations of the forecast errors, to exclude outliers.

In Figure 1, we plot the distribution of our leading measure of forecast errors, $FE_{t,t+1}^{\log}$,

[15]The aggregate components explain approximately 11% of the variation in forecast errors; however, recent work has substantiated that firms may have heterogeneous exposure to aggregate shocks, which implies that the simple residual forecast errors we construct may still be affected by the aggregate economic conditions. Therefore, we construct alternative residual forecast errors by explicitly considering firms heterogeneous exposure to aggregate shocks. For these alternative residual forecast errors, aggregate components explain approximately 23% of the variation in forecast errors, but our main empirical findings are robust to these alternative measures. Detailed discussions are in Section 1.2.6 of the online appendix.

across all firms in all years. The forecast errors are centered around zero, and the distribution appears to be symmetric. The shape of the density is similar to a normal distribution, although the center and the tails have more mass than the fitted normal distribution (solid line in the graph). The average forecast error across all firm-year observations is -0.024, with a median of -0.005 and a standard deviation of 0.298. The absolute value of $FE^{\log}_{t,t+1}$ is 0.2, which implies that firms on average overforecast or underforecast their sales by 20%.

Figure 1: Distribution of the forecast errors



Notes: Histogram of $FE^{\log}_{t,t+1}$ with the fitted normal density (solid line).

Fact 1: Precision of Forecasts Increases as Firms Become More Experienced

Figure 2 presents the average absolute value of forecast errors by age cohorts, where age is top-coded at ten. The precision of sales forecasts increases as the firm ages. Specifically, as firms age from one to ten years, the absolute forecast errors decline from 36% to 18% on average. Moreover, the decline occurs mainly in the first five years after entry. Since we examine the impact of parent companies' exporting experience later, we also present these statistics for a subsample in the manufacturing sector.[16] The patterns are similar.

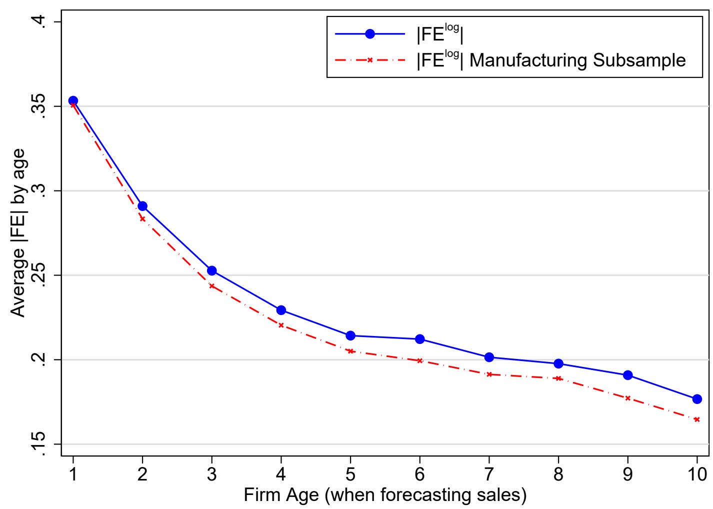
We further confirm these patterns formally by an OLS regression of firm i's absolute forecast error in year t:

$$|FE^{\log}_{it,t+1}| = \delta_n + \beta X_{it} + \delta_{ct} + \delta_s + \varepsilon_{it}, \tag{1}$$

[16] Following Conconi et al. (2016), our manufacturing subsample includes firms in manufacturing or wholesale or retail industries whose parent companies are in manufacturing. For these firms, their parent companies exporting experience helps them gather information on the demand in the destination market and produce forecasts with improved precision. We refer to this sample as the "manufacturing subsample" in our analysis.

where δ_n is a vector of age dummies, δ_{ct} represents the country-year fixed effects, and δ_s represents the industry fixed effects. Time-varying controls such as firm size are denoted by X_{it}. We use age one as the base category; therefore, the age fixed effects represent the difference in the absolute forecast errors between age n and age one. To further control for heterogeneity across firms, we also run regressions with firm fixed effects δ_i instead of the industry fixed effects δ_s.

Figure 2: $|FE^{\log}|$ declines with firm age



Note: Average absolute value of $FE^{\log}$ by age cohorts. The manufacturing subsample is defined in footnote 16.

Column 1 in Table 3 shows the baseline specification with industry and country-year fixed effects. As firms become older, the absolute forecast errors decline. On average, firms that are at least ten years old have absolute forecast errors 17 log points lower. In columns 2 and 3, we control for the size of the firms and their parent companies in Japan (measured by log employment). Although larger firms tend to have smaller absolute forecast errors, the age effects survive.[17]

One possible explanation of the age effects is that firms gradually improve their management quality or capability such that they can obtain a more diversified portfolio of destination markets and products. We show that the age effects are not simply driven by such diversification. To examine the importance of market diversification, we control for the concentration of firm sales across markets, measured by the Herfindahl-Hirschman Index (HHI) of the shares of firm sales in six markets: the host country (local market), Japan, Asia,

[17]Tanaka et al. (2019) report that older firms make more precise forecasts than younger firms do, on the basis of cross-section results. By contrast, our finding is based on within-firm variation with the firm fixed effects, thereby pointing to the life-cycle pattern of forecast errors.

North America, Europe and the rest of the world. Column 4 shows that a higher value of market-level concentration increases the absolute forecast errors, but controlling for it has a limited impact on the age effects. Unfortunately, our main data source does not provide a breakdown of firm sales by products. To examine the importance of product diversification, we focus on the subset of firms in China which we managed to match with the Chinese customs data between 2000 and 2009. For the matched observations, we calculate a product sales HHI of using the customs' records on firm exports at the HS 6-digit product level. Column 5 shows that a higher value of product concentration also has a positive impact on absolute forecast errors, but it is not significant because of the small sample size. The age effects, however, are robust to controlling for the product concentration.[18]

To evaluate the robustness of our results, we restrict our sample to (1) surviving entrants and (2) firms in our manufacturing subsample. Column 6 reports the result for a subsample of firms that have survived and continuously appeared in the data from age one to seven, which shows that our results are not driven by endogenous exits and non-reporting. Column 7 focuses on the manufacturing subsample, and the results are similar. In Online Appendix 1.2, we further show that the age effects (1) are robust to various alternative measures of forecast errors, including those that explicitly take firms' heterogeneous exposure to aggregate shocks into account and (2) are not driven by firms entering in different months of a year so that age-one firms actually have fewer than 12 months of experience (the so-called "partial-year effects"), and (3) are not due to age-dependent biases in the level of forecast errors.[19]

An alternative means to accumulate experience in the foreign market is exporting. Other studies have found that multinational firms use "exporting" to test the market and learn about consumer demand before setting up a firm (Conconi et al., 2016). We now examine how previous export experience of the parent companies affects the precision of forecast errors in our manufacturing subsample. Because export data at the firm-destination country level are not available, we obtain information on parent companies' export experience at the regional level by using the domestic activities survey data. In our data, approximately 73% of the firms inherited export experience from their parents before entry. We refer to these firms as "experienced firms." The other firms are unexperienced firms, although unexperienced

[18] We provide details on the construction of these measures, the matching between China customs data and our main data source, and additional robustness checks regarding product and market diversification in Online Appendix 1.2.4.

[19] Age-dependent biases in the level of forecast errors occur, for example, when young firms over-predict future sales as they are too optimistic. This can cause biases of the age effects on the absolute forecast errors. We design a two-step procedure to manage such potential biases. In particular, we run a first-stage regression on the level of forecast errors, and project the squared residuals from the first stage on the same set of independent variables in the second stage. These regressions show that the conditional variance of forecast errors declines significantly with firm age. Interested readers can refer to Section 1.2.2 of the online appendix.

Table 3: Age effects on the absolute forecast errors

Sample:	All Firms				China	Survivors	Manufacturing
Dep.Var: $\lvert FE^{\log}_{t,t+1}\rvert$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$1(\text{Age}_t = 2)$	-0.066^a	-0.059^a	-0.064^a	-0.049^a	-0.047	-0.069^a	-0.066^a
	(0.007)	(0.007)	(0.007)	(0.009)	(0.030)	(0.010)	(0.009)
$1(\text{Age}_t = 3)$	-0.102^a	-0.089^a	-0.089^a	-0.066^a	-0.074^b	-0.088^a	-0.095^a
	(0.007)	(0.007)	(0.008)	(0.009)	(0.029)	(0.011)	(0.009)
$1(\text{Age}_t = 4)$	-0.127^a	-0.113^a	-0.111^a	-0.085^a	-0.088^a	-0.100^a	-0.115^a
	(0.007)	(0.007)	(0.008)	(0.009)	(0.029)	(0.012)	(0.010)
$1(\text{Age}_t = 5)$	-0.140^a	-0.123^a	-0.117^a	-0.093^a	-0.088^a	-0.112^a	-0.119^a
	(0.007)	(0.007)	(0.008)	(0.009)	(0.029)	(0.012)	(0.010)
$1(\text{Age}_t = 6)$	-0.141^a	-0.123^a	-0.115^a	-0.091^a	-0.084^a	-0.111^a	-0.121^a
	(0.007)	(0.007)	(0.008)	(0.009)	(0.031)	(0.013)	(0.010)
$1(\text{Age}_t = 7)$	-0.152^a	-0.132^a	-0.122^a	-0.100^a	-0.087^a	-0.120^a	-0.128^a
	(0.007)	(0.007)	(0.008)	(0.009)	(0.032)	(0.013)	(0.010)
$1(\text{Age}_t = 8)$	-0.155^a	-0.132^a	-0.121^a	-0.099^a	-0.090^a	-0.110^a	-0.125^a
	(0.007)	(0.007)	(0.009)	(0.009)	(0.033)	(0.015)	(0.011)
$1(\text{Age}_t = 9)$	-0.159^a	-0.135^a	-0.124^a	-0.103^a	-0.098^a	-0.106^a	-0.129^a
	(0.007)	(0.007)	(0.009)	(0.010)	(0.034)	(0.016)	(0.011)
$1(\text{Age}_t \geq 10)$	-0.171^a	-0.136^a	-0.122^a	-0.100^a	-0.093^b	-0.106^a	-0.123^a
	(0.007)	(0.006)	(0.009)	(0.010)	(0.036)	(0.018)	(0.011)
$\log(\text{Emp})_t$		-0.021^a	-0.024^a	-0.023^a	-0.036^a	-0.036^a	-0.023^a
		(0.001)	(0.002)	(0.002)	(0.010)	(0.006)	(0.002)
$\log(\text{Parent Emp})_t$		0.001	0.001	-0.001	-0.000	0.012	0.001
		(0.001)	(0.003)	(0.003)	(0.009)	(0.008)	(0.003)
HHI Market Sales at t				0.016^a			
				(0.005)			
HHI HS6 Product Exports at t					0.006		
					(0.014)		
Industry FE	Y	Y					
Country-year FE	Y	Y	Y	Y	Y	Y	Y
Firm FE			Y	Y	Y	Y	Y
N	131271	128444	123224	104690	8475	19314	87057
R^2	0.097	0.116	0.360	0.368	0.381	0.336	0.359

Notes: Standard errors are clustered at the business group level, c 0.10 b 0.05 a 0.01. The dependent variable is the absolute value of forecast errors in all regressions. Age refers to the age of the firm when making the forecasts. Regressions in columns 1-4 include all firms. Column 5 only includes firms in China that can be matched to the Chinese customs data. Survivors (Column 6) refer to firms that have continuously appeared in the sample from age one to seven, and the manufacturing subsample (Column 7) is defined in footnote 16.

firms can accumulate experience by directly producing in the foreign market after entry.

In Table 4, we provide evidence that previous export experience reduces the forecast errors made by firms that enter a destination market for the first time (among all firms in the same business group). In columns 1 and 2, we use dummy variables that equal one if and only if the firm's parent company exported to the same region in the year (or in one of the two years) before entry. In Column 3, we use a more sophisticated definition of export experience to capture the lumpiness of exporting,[20] and the dummy variable equals one if and only if export experience is positive. These regression results show that having previous export experience reduces absolute forecast errors by 13-16 log points. In Column 4, we use a continuous measure of export experience instead of indicator variables. On average, one additional year of export experience reduces absolute forecast errors by 1.4 log points.

Table 4: Forecast error and previous exporting

Dep.Var: $\lvert FE_{1,2}^{\log}\rvert$	(1)	(2)	(3)	(4)
$Exp_0 > 0$	-0.161^a			
	(0.055)			
$Exp_0 > 0$ or $Exp_{-1} > 0$		-0.150^a		
		(0.054)		
$Exp \quad Expe. > 0$			-0.133^b	
			(0.059)	
$Exp \quad Expe.$				-0.014^a
				(0.005)
Industry FE	Y	Y	Y	Y
Country-year FE	Y	Y	Y	Y
N	438	444	517	517
R^2	0.310	0.319	0.290	0.292

Notes: Standard errors are clustered at the business group level, c 0.10 b 0.05 a 0.01. The dependent variable is a firm's initial forecast error, which is calculated as the absolute log deviation of the realized sales at age = 2 from the sales forecast made at age = 1. We only include firms that are first-time entrants into a particular host country.

In Online Appendix 1.3, we provide a battery of robustness checks of the aforementioned regressions. We show that the results are robust to (1) using the sample of first-time entrants into each region, (2) controlling for the firm and its parent companys size, (3) using the subsample of firms with most of the sales in the local market and excluding intrafirm exports when measuring export experience, (4) using alternative measures of forecast errors, and (5) using the aforementioned two-step procedure. Based on the evidence in this section, our research shows that forecasts become more precise as firms become more experienced.

[20] Specifically, we define export entry as when the parent company does not export to the region for two consecutive years and starts exporting afterward. The variable of export experience is positive if the parent company exported in the past and has not stopped exporting for more than two years.

Fact 2: Forecast Errors are Positively Autocorrelated but Less So as Firms Become More Experienced

A growing literature has highlighted the serial correlation of forecast errors in various contexts. For example, Ryngaert (2017) demonstrated that professional forecasters' forecast errors of future inflation rates are autocorrelated, indicating their imperfect information on macroeconomic conditions.[21] Instead of using expectations data on macroeconomic outcomes, we utilize data on the sales expectations of individual firms and show that their forecast errors are positively autocorrelated over time. Importantly, we document that the serial correlation of forecast errors declines with the firm's age.

Table 5: Correlation of $FE^{\log}_{t,t+1}$ and $FE^{\log}_{t-1,t}$, overall and by age group

Sample	All ages	Age 2-4	Age 5-7	Age ≥ 8
All industries	0.136	0.171	0.150	0.119
	[96489]	[10704]	[14031]	[71754]
Manufacturing	0.136	0.177	0.130	0.121
	[68215]	[7066]	[9632]	[51517]

Notes: $FE^{\log}_{t,t+1}$ is the log deviation of the realized sales in year $t+1$ from the sales forecast made in year t. Age is measured at the end of year t. The manufacturing subsample is defined in footnote 16. Number of observations used for each correlation is shown in the brackets below. All correlation coefficients are significant at the 1% level.

Table 5 presents the serial correlation of forecast errors, for the entire sample and different age groups. Among all firm-year observations, we find that the correlation coefficient between $FE^{\log}_{t,t+1}$ and $FE^{\log}_{t-1,t}$ is 0.136. This result suggests that firms tend to make systematic errors in forecasting their sales. The remaining three columns present the serial correlation for different age groups. When firms become more experienced, the positive correlation is reduced, indicating that firms become more informed and make smaller systematic errors when forecasting. Such patterns are robust when focusing on the manufacturing subsample. We find similar patterns when using alternative definitions of forecast errors (see Online Appendix Table OA.15).

We next confirm this pattern by running the AR(1) type of regressions at the firm level. This allows us to control for the time-varying firm characteristics and various sets of fixed effects to rule out confounding factors. In particular, we run the following regression

$$FE^{\log}_{i,t+1,t+2} = \beta_1 FE^{\log}_{i,t,t+1} + \beta_2 FE^{\log}_{i,t,t+1} \times Age_{it} + \beta_3 X_{it} + \delta_n + \delta_s + \delta_{ct} + \delta_g + u_{it}, \quad (2)$$

[21] See also Coibion and Gorodnichenko (2012) and Andrade and Le Bihan (2013).

where Age_{it} denotes the firm's age at time t and X_{it} denotes the firm's other time-varying characteristics such as employment at time t. In all regressions, we control for the firm's age, industry, country-year, and business group fixed effects, denoted by $\delta_n, \delta_s, \delta_{ct}$ and δ_g, respectively.

Table 6: AR(1) Regressions and the Effect of Age

Sample:	All Firms				Manufacturing			
Dep.Var: $FE^{\log}_{t+1,t+2}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$FE^{\log}_{t,t+1}$	0.102^a	0.103^a	0.098^a	0.098^a	0.098^a	0.101^a	0.094^a	0.095^a
	(0.014)	(0.017)	(0.014)	(0.017)	(0.018)	(0.023)	(0.018)	(0.023)
$\times$ max{Age$_t$, 10}	-0.006^a		-0.005^a		-0.006^b		-0.005^b	
	(0.002)		(0.002)		(0.002)		(0.002)	
$\times$ log(Age$_t$)		-0.020^a		-0.017^b		-0.020^b		-0.018^c
		(0.007)		(0.007)		(0.010)		(0.010)
log(Emp)$_t$			0.002^c	0.002^c			0.002	0.002
			(0.001)	(0.001)			(0.001)	(0.001)
log(Parent Emp)$_t$			-0.008^c	-0.008^c			-0.008	-0.008
			(0.004)	(0.004)			(0.006)	(0.006)
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Country-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Business Group FE	Y	Y	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y	Y	Y
N	95274	95274	93539	93539	67491	67491	66787	66787
R^2	0.190	0.190	0.193	0.193	0.207	0.207	0.209	0.209

Notes: Standard errors are clustered at the business group level, c 0.10 b 0.05 a 0.01. The manufacturing subsample is defined in footnote 16.

Table 6 shows the regression results. To capture the non-linear effect of the firm's age, we use either age top-coded at ten or the log of age. The specifications also differ in their sample coverage (all sectors vs. the manufacturing subsample) and whether they control the firms and their parents' employment. In all specifications, the interaction term between the firm's age and past forecast error is significantly negative, confirming our finding that the autocorrelation in forecast errors declines as the firms become more experienced. For example, the estimates in Column 1 suggest that the AR(1) coefficient starts at 0.096 at age 1 and declines by 0.006 for firms' each additional year of experience.

In Online Appendix 1.4.3, we explore how export experience affects the AR(1) coefficient and perform other regression specifications. Exporting experience has a large negative impact on the AR(1) coefficient, but the estimates are insignificant due to the small sample. Overall, the evidence in this section suggests that firms tend to make persistent forecast errors, and such persistence declines as firms become more experienced.

3 Model

We develop a dynamic industry equilibrium model with Jovanovic (1982) type learning embedded as in Arkolakis et al. (2018) but extend it to allow for endogenous choices of production modes between exporting and multinational production (Helpman et al., 2004). Importantly, we depart from the original Jovanovic-learning model by introducing information shocks alongside the i.i.d. transitory real shock as in Kydland and Prescott (1982). This setup helps us fully match the aforementioned stylized facts and provides a framework to quantify informational imperfections using our data.

3.1 Setup: Demand and Supply

In our model, there are two countries: Japan and a foreign country. Each Japanese firm produces a differentiated variety and decide whether to serve the foreign market, and if so, whether through exporting or multinational production. We focus on the foreign production and sales of these Japanese firms but do not explicitly model the domestic market.

In the foreign country, the representative consumer has the following nested-CES preferences, where the first nest is among the composite goods produced by firms from different countries, indexed by i,

$$U_t = \left(\sum_i \chi_i^{\frac{1}{\delta}} Q_{it}^{\frac{\delta-1}{\delta}}\right)^{\frac{\delta}{\delta-1}},$$

and the second nest is among the varieties $\omega \in \Omega_{it}$ produced by firms from each country i,

$$Q_{it} = \left(\int_{\omega\in\Omega_{it}} e^{\frac{a_t(\omega)}{\sigma}} q_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}. \tag{3}$$

In the first nest, χ_i is the demand shifter for the goods of country i, and δ is the Armington elasticity between goods produced by firms from different countries. In the second nest, σ is the elasticity between different varieties, and $a_t(\omega)$ is the demand shifter for variety ω. We assume that firms differ in their demand shifter, $a_t(\omega)$. After denoting foreign consumers' total expenditure as $\tilde{Y}_t$, we can express the demand for a particular Japanese variety, ω, as:

$$q_t(\omega) = \tilde{Y}_t \tilde{P}_t^{\delta-1} \chi_{jp} P_{jp,t}^{\sigma-\delta} e^{a_t(\omega)} p_t(\omega)^{-\sigma}, \tag{4}$$

where $\tilde{P}_t$ is the aggregate price index for all goods, and $P_{jp,t}$ is the ideal price index for Japanese goods.[22] In our model, we assume that the Japanese varieties are a small fraction

[22]When the Armington elasticity δ equals one, the first nest is Cobb-Douglas, and the expenditure on Japanese goods no longer depend on $P_{jp,t}$. When $\sigma = \delta$, the elasticities in the two nests are the same, as in

of foreign consumers' consumption and treat $\tilde{Y}_t$ and $\tilde{P}_t$ as exogenous.[23] As a result, we can combine the exogenous terms in expression (4), $\tilde{Y}_t\tilde{P}_t^{\delta-1}\chi_{jp}$ into one variable, Y_t, and refer to it as the aggregate demand shifter. In addition, as we focus only on Japanese firms, we suppress the subscript jp in the following analysis and derive the ideal price index of Japanese goods as

$$P_t \equiv \left(\int_{\omega\in\Omega_t} e^{a_t(\omega)} p_t(\omega)^{1-\sigma} d\omega\right)^{1/(1-\sigma)}. \tag{5}$$

We assume that the firm-specific demand shifter, $a_t(\omega)$, is the sum of a time-invariant permanent demand draw $\theta(\omega)$ and a transitory demand shock $\xi_t(\omega)$:

$$a_t(\omega) = \theta(\omega) + \xi_t(\omega). \tag{6}$$

Firms understand that $\theta(\omega)$ is drawn from a normal distribution $N\left(\bar{\theta}, \sigma_\theta^2\right)$, and the i.i.d. transitory demand shock, $\xi_t(\omega)$, is drawn from another normal distribution $N\left(0, \sigma_\xi^2\right)$. We assume that the transitory demand shock is realized after a firm makes the output decision and the firm can observe it. However, following Kydland and Prescott (1982) and the noisy information and rational inattention literature, we assume that the firm only receives a *noisy* signal concerning its permanent demand draw θ at the end of each period:

$$s_t(\omega) = \theta(\omega) + \varepsilon_t(\omega), \tag{7}$$

where i.i.d. noise term $\varepsilon_t(\omega)$ is drawn from a normal distribution $N(0, \sigma_\varepsilon^2)$.[24] The noise term reflects that firms have a limited capacity to digest and analyze freely available information. Therefore, the firm *cannot* precisely back out the implied demand draw $\theta(\omega)$, even if all related information (e.g., sales, profits) are freely available. Instead, the firm receives a noisy signal for the demand draw every period and needs to learn about it over the life cycle.

As discussed later, our model can reproduce the aforementioned age-declining serially correlated forecast errors (Fact 2).[25] This is because $\varepsilon_t(\omega)$, which makes learning gradual, are purely informational and do not directly affect firms' per-period profits. It is exactly this payoff irrelevance nature that leads to positively and serially correlated forecast errors; and, as in Online Appendix 2.3.2, we prove that forecast errors are serially uncorrelated if $\varepsilon_t(\omega)$

Melitz (2003). In our calibration, δ is set to be between one and σ.

[23] In 2009, the total value of Japan's exports and multinational sales is only equivalent to 2.4%, 2.5%, and 2.7% of the total gross output of China, the United States and all 36 countries in the World KLEMS dataset.

[24] See, among others, Sims (2003) and Mackowiak and Wiederholt (2009).

[25] There are models that generate serially correlated forecast errors as in Ma et al. (2019) and Bordalo et al. (2018), but the age-declining patterns are absent in these models.

are real shocks and payoff relevant as in Jovanovic (1982).[26] Nonetheless, our model predicts gradual resolution of uncertainty and can be used to match the decline in absolute forecast errors (Fact 1). This life-cycle property distinguishes our model from existing models in the noisy information and rational inattention literature. Finally, we introduce a real temporary demand shock $\xi_t(\omega)$ that does not affect learning and ensures that firms still have forecast errors when they are old enough, as in our data.

The industry structure features monopolistic competition. There is an exogenous mass of potential entrants J that decide whether to enter the foreign market each period. Each entrant makes a permanent demand draw θ from a normal distribution, $N\left(\bar{\theta}, \sigma_\theta^2\right)$, and a sunk multinational production entry cost f_m^e from a log-normal distribution $\log N(\mu_{f_m^e}, \sigma_{f_m^e}^2)$.[27] The entrant knows f_m^e but does not know θ. If the firm enters the foreign market, it also has to decide how to serve the market by choosing between exporting, which involves a sunk cost of f_x^e, and setting up an firm with the entry cost of f_m^e. Both sunk costs are paid in units of domestic labor. If neither mode is profitable, the potential entrant does not enter and obtains zero payoff.

In each period, the incumbents first receive an exogenous death shock with probability η. For surviving firms, they decide whether to change their mode of service. They can keep their service mode unchanged, or switch to another mode (e.g., from exporting to multinational production). They can also choose to exit permanently. We assume that incumbent multinational firms can switch to exporting without paying the sunk entry cost of exporting, as they have already established their appearance in the foreign market. Each period, firms also must pay a fixed cost, f_x, to export or, f_m, to conduct multinational production.

Output is linear in labor and firms hire workers in a perfectly competitive labor market. Exporters employ domestic workers to produce at a constant wage w, and multinational firms employ foreign workers at a constant wage w^*.[28] For firms that serve the foreign market, they decide how much to produce in period t before the overall demand shifter, a_t,

[26] In Online Appendix 2.3.1, we also show that the perfect information model cannot be used to rationalize the existence of serially correlated forecast errors, even when firms endogenously exit the market. In Online Appendix 2.3.3, we also show that the serial correlation of forecast errors is still zero, even if the permanent demand, θ, is time-variant and follows an AR(1) process in a Jovanovic-type learning model.

[27] We assume ex ante heterogeneity in entry costs by following Das et al. (2007), which helps us reproduce the co-existence of firms with and without exporting experience, but we abstract from ex ante firm heterogeneity in labor productivity. As a robustness check, we also introduce ex ante firm heterogeneity in labor productivity into the model in Online Appendix 3.3. We find that the quantitative implications of our model barely change.

[28] We assume both wages are exogenous since the value of Japanese goods sold abroad is small relative to the total output of the rest of the world (see footnote 23) and the export-to-domestic (gross) output ratio is also small for Japan, ranging from 6% to 8% from 2005 to 2009.

is realized. After the demand shifter in period t is realized, they choose the price p_t to sell all the products produced, because we assume there is no storage technology and firms cannot accumulate inventories.

3.2 Belief Updating

In this subsection, we discuss how a firm forms the ex post belief for its permanent demand. At the beginning of a period, a firm that is $n+1$ ($n \geq 1$) years old has observed noisy signals of the permanent demand draw in the past n periods: $s_1, s_2, \ldots, s_n$. Since both the prior and the noisy signals are normally distributed, Bayes' rule implies that the posterior belief about θ is normally distributed with mean μ_n and variance σ_n^2

$$\mu_n = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}\bar{\theta} + \frac{n\sigma_\theta^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}\bar{s}_n, \quad \sigma_n^2 = \frac{\sigma_\varepsilon^2 \sigma_\theta^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}, \tag{8}$$

where the history of signals $(s_1, s_{2,} \ldots, s_n)$ is summarized by age n and the average signal of the permanent demand draw:

$$\bar{s}_n \equiv \frac{1}{n}\sum_{i=1}^{n} s_i \text{ for } n \geq 1; \quad \bar{s}_0 \equiv \bar{\theta}.$$

For age-one firms (i.e., entrants), their belief for the mean and variance of θ is the same as the prior belief:

$$\mu_0 = \bar{\theta}, \;\; \sigma_0^2 = \sigma_\theta^2.$$

3.3 Static Optimization of Per-Period Profit

In this subsection, we study the firm's static optimization problem. As we focus on firm's behavior in the steady state (i.e., the stationary equilibrium) in what follows, we omit the subscript t whenever possible. In each period, conditional on the mode of service, the firm's output decision is a static problem. Given the belief about a_t, an age-n ($n \geq 1$) firm hires labor and produces q_t quantity of output to maximize its expected per-period profit in period t, $E_{a_t|\bar{s}_{n-1},n}(\pi_{o,t})$. As a result, the realized per-period profit for the firm ($o = m$) or an exporter ($o = x$) is

$$\pi_{o,t} = p_t(a_t)q_t - MC_o \times q_t - wf_o,$$

where the marginal cost of production, MC_o, depends on the mode of service.[29] Firms set the price after observing the realized demand a_t to sell all the output. Maximizing

[29] $MC_x = \tau w$ and $MC_m = w^*$ where w and w^* denote the domestic and foreign wages, respectively.

$E_{a_t|\bar{s}_{n-1},n}(\pi_{o,t})$, the optimal output choice is

$$q_{o,t} = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{b\left(\bar{s}_{n-1},n-1\right)}{MC_o}\right)^{\sigma}\frac{Y}{P^{\delta-\sigma}}, \tag{9}$$

where

$$b\left(\bar{s}_{n-1},n-1\right) = E_{a_t|\bar{s}_{n-1},n-1}\left(e^{a_t/\sigma}\right) = \exp\left\{\frac{\mu_{n-1}}{\sigma}+\frac{1}{2}\left(\frac{\sigma_{n-1}^2+\sigma_{\xi}^2}{\sigma^2}\right)\right\}, \tag{10}$$

and n is the firm's effective market experience. The resulting price and per-period profit function are

$$p_{o,t}\left(a_t\right) = \frac{\sigma}{\sigma-1}e^{a_t/\sigma}\frac{MC_o}{b\left(\bar{s}_{n-1},n-1\right)}; \tag{11}$$

$$E\pi_{o,t} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}\frac{b\left(\bar{s}_{n-1},n-1\right)^{\sigma}}{MC_o^{\ \sigma-1}}\frac{Y}{P^{\delta-\sigma}} - wf_o. \tag{12}$$

3.4 Dynamic Optimization and Equilibrium Definition

In each period, an entrant or incumbent firm chooses among three service modes: exiting, exporting and multinational production. To become an exporter or a multinational firm, a firm must pay a sunk cost. A firm's state variables include the service mode last period, o; the entry cost of multinational production, f_m^e; its current effective market experience, n; and the history of the noisy signals, $\bar{s}_{n-1}$. As firms make optimal decisions based on their belief about θ rather than the true value of θ, these variables are sufficient to characterize the firm's value function and policy function.

We first define the choice-specific value function for the incumbent firm ($n \geq 2$) as

$$v(o', o, f_m^e, n, \bar{s}_{n-1}) = \begin{cases} E_{n-1}\pi_x + \beta(1-\eta)E_{n-1}V\left(x, f_m^e, n+1, \bar{s}_n\right) & \text{if } o' = x, \\ E_{n-1}\pi_m - wf_m^e\mathbf{1}(o = x) + \beta(1-\eta)E_{n-1}V\left(m, f_m^e, n+1, \bar{s}_n\right) & \text{if } o' = m, \\ 0 & \text{if } o' = exit, \end{cases} \tag{13}$$

where o is the mode of service in the previous period, and o' is the choice of mode this period. The expectation is based on information available at the end of age $n-1$, and $V(\cdot)$ is the value function. Note that previous service mode cannot be *entry* for firms older than two. For an entrant ($n = 1$), its choice-specific value function is

$$v(o', ent, f_m^e, 1, \bar{s}_0) = \begin{cases} E_0\pi_x - wf_x^e + \beta(1-\eta)E_0\ V\left(x, f_m^e, 2, \bar{s}_1\right) & \text{if } o' = x, \\ E_0\pi_m - wf_m^e + \beta(1-\eta)E_0\ V\left(m, f_m^e, 2, \bar{s}_1\right) & \text{if } o' = m, \\ 0 & \text{if } o' = exit. \end{cases} \tag{14}$$

With these choice-specific value functions in hand, the value functions must satisfy

$$V(o, f^e_m, n, \bar{s}_{n-1}) = \max_{o' \in \{x,m,exit\}} \{v(o', o, f^e_m, n, \bar{s}_{n-1})\}, \tag{15}$$

and we denote the corresponding policy function as $o'(o, f^e_m, n, \bar{s}_{n-1})$. The definition of equilibrium is contained in the Appendix.

3.5 Properties of Forecasts and Forecast Errors

In this subsection, we show how our model matches the stylized facts presented in Section 2. We illustrate the intuition by using a special case in which there is no endogenous switching of production modes.

Proposition 1 *When there is no endogenous switching of production modes, the forecasts and forecast errors of exporters and multinational firms' sales have the following properties*

1. *The variance of forecast errors declines with years of experience.*

2. *Forecast errors made in two consecutive periods by the same firm are positively correlated. The positive covariance of forecast errors made in two consecutive periods declines with years of experience. Moreover, the autocorrelation coefficient also declines with years of experience, if the variance of the transitory demand shock is large relative to that of the information shock.*

Proof. See Online Appendix 2.2. ■

Life-cycle learning contributes to the age-declining variance of forecast errors. Thanks to learning, firms accumulate more experience and thus have clearer information on their permanent demand when operating in the market for a longer period of time. Because exporting helps firms accumulate sales experience as well, it is a natural result the learning mechanism also rationalizes that experienced multinational firms start with a smaller variance of forecast errors than unexperienced multinational firms do.

The above proposition also rationalizes the finding of the serially correlated forecast errors presented in Section 2.2, as firms adjust their posterior beliefs *gradually* over their life cycles. In other words, firms incorporate their new signals partially into their posterior beliefs. As a result, the firm is more likely to under-predict (or over-predict) its next year's sales at the end of this year if it has underpredicted (or overpredicted) its current year's sales at the end of last year. This leads to the positive autocorrelations of forecast errors.[30] Moreover, as a more

[30] However, this *does not* mean that firms make non-zero forecast errors on average, as positive and negative errors are canceled out across firms.

experienced firm makes smaller forecast errors, the autocovariance of forecast errors declines with years of experience. Finally, if the auto-covariance of forecast errors declines faster than the variance of forecast errors (which is the case when the variance of the transitory demand shock is large relative to that of the information shock), the autocorrelation coefficient of forecast errors also decreases with years of experience.

4 Quantitative Analysis

In this section, we quantitatively assess the aggregate implications of imperfect information and uncertainty. To this end, we first describe the procedures used for calibrating our model, to show the mapping from the model elements to the empirical facts presented in Section 2. We show that the calibrated model can capture the dynamics of firms' forecast errors as observed in the data, as well as other features of the data.[31] The calibrated model is then used to examine the quantitative implications of theoretical channels at work in the model.

4.1 Calibration

We first normalize a set of parameters not separately identified from others. Specifically, aggregate demand shifter, Y; the wage rate in Japan, w; and the wage rate in the foreign country, w^*; are normalized to one. The mean of the logarithm of the permanent demand, $\bar{\theta}$, is normalized to zero. We also normalize the entry costs of exporting, f_x^e, to zero, as we abstract from modeling Japanese domestic firms in the paper.[32]

Next, we calibrate a set of parameters without solving our model (Table 7). We set the elasticity of substitution between the varieties, σ, to four, a common value in the literature (see Bernard et al., 2003). The Armington elasticity among goods from different countries, δ, is set to two, a value in line with the median estimate across sectors in Feenstra et al. (2017).[33] We set the discount factor, β, to 0.96, which implies a real interest rate of 4%.[34]

The exogenous death rate η and the per-period fixed costs of multinational production

[31] In Online Appendix 3.1, we show that the calibrated model is able to capture more of the features found in the data, such as exporters' sales growth and declining exit rates over their life cycles, which we do not directly target in the calibration.

[32] Specifically, moments that can be used to pin down f_x^e, such as the share of exporters relative to domestic firms, are not available. We interpret the entry costs of multinational production, f_m^e, as the entry costs of multinational production *relative to exporting*.

[33] Feenstra et al. (2017) estimated the Armington elasticities for eight sectors by using two-stage least square and two-step GMM approaches. The median estimate across sectors is 2.06 when using the former method and 1.65 when using the latter method.

[34] Data from the World Bank show that the real interest rate in Japan during 1995-2014 is between 2% and 4%.

f_m are crucial for the exit rates of multinational firms. As there is strong selection in the model, firms' exit rates decline over their life cycles if the per-period fixed costs are positive. However, in Online Appendix Table OA.22, we do not find a significant decline in firms' exit rates over their life cycle, even for firms without export experience. Therefore, we postulate that $f_m = 0$ and set η to 0.03 so that the model can match the average exit rate (3%). As a result, multinational firms exit only because of exogenous death shocks.

Table 7: Parameters calibrated without solving the model

Parameters	Description	Value	Source
σ	Elasticity of substitution between different varieties of Japanese goods	4	Bernard et al. (2003)
δ	Armington elasticity between goods from different countries	2	Median estimate in Feenstra et al. (2017)
β	Discount factor	0.96	4% real interest rate
η	Exogenous death rate	0.03	Average exit rates of the affiliated firms
f_m	Per-period fixed costs of multinational production	0	Flat profile of the affiliated firms' exit rate over their life cycles

Three parameters that govern informational imperfection and learning can also be backed out without calibrating the model. Since we have shut down the endogenous exits for multinational firms, there is no selection on the permanent demand draw among multinational firms after entry. As all entrants have the same prior belief for their permanent demand, they choose to enter multinational production based on the entry costs, which are uncorrelated with the permanent demand draws. We derive closed-form expressions for the variance and auto-covariance of forecast errors by market experience while assuming no selection (see Online Appendix 2.1), and these formulas can be directly applied to firms without exporting experience (unexperienced firms). In particular, we target the standard deviation of forecast errors of age-one unexperienced firms and that of unexperienced firms older than ten, as these two moments are the most informative about σ_θ and σ_ξ, respectively. We also target the autocovariance of forecast errors, because the noise ε prevents firms from learning the true value of θ immediately.

Table 8: Parameters related to the forecast errors and moments

Parameters	Value	Description	Moments	Data	Model
σ_θ	2.03	Std of time-invariant shock	Std of FE at age 1	0.51	0.51
σ_ξ	1.13	Std of transitory shock	Std of FE at age 10	0.31	0.31
σ_ε	1.49	Std of noise	Auto-cov. of FE at all ages	0.012	0.012

Notes: All data moments are calculated by using the sample of first-time entrants into countries whose parent companies do not have previous export experience.

The remaining four parameters are jointly calibrated by solving the equilibrium and matching the four moments. The parameters are as follows: per-period fixed cost of exporting, f_x; the mean and standard deviation of the log entry cost of multinational production, $\mu_{f_m^e}$ and $\sigma_{f_m^e}$; and the iceberg trade costs, τ. The four targeted moments are the average exit rate of exporters, the fraction of exporters among active firms, the fraction of experienced firms at age one and the share of exports in Japanese firms' total foreign sales.

In Table 9, we list the moments in an order such that loosely, the moment provides the most information on the parameter in the same row. A higher export per-period fixed cost raises the exporter exit rate, and the iceberg trade costs have a larger impact on the exporters sales share (intensive margin) than on the exit rates (extensive margin). A higher average entry cost of multinational production increases the fraction of exporters among all firms selling in the foreign market. Finally, in our model, only firms with sufficiently small entry costs of multinational production enter without exporting experience. Thus, a higher $\sigma_{f_m^e}$ raises the share of unexperienced firms.

Table 9: Parameters calibrated by solving the model and matching moments

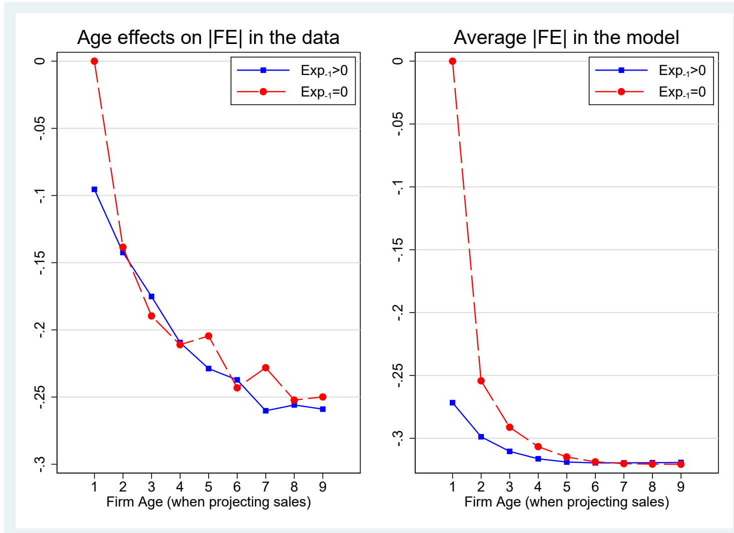
Parameters	Value	Description	Moments	Data	Model
f_x	0.0049	export fixed cost	average exit rate of exporters	0.07	0.07
τ	1.18	iceberg trade cost	Exporter sales share	0.28	0.29
$\mu_{f_m^e}$	0.10	mean of log FDI entry cost	fraction of exporters among active firms	0.71	0.69
$\sigma_{f_m^e}$	1.97	Std of log FDI entry cost	fraction of experienced MNEs at age 1	0.72	0.73

4.2 Dynamics of Forecast Errors

In Figure 3, we examine one set of untargeted moments in our model: the changes in the absolute forecast errors over firms' life cycles . We first estimate the age effects on $\left|FE_{t,t+1}^{\log}\right|$ for firms that enter a foreign market for the first time. In the regression, we interact the age fixed effects with a dummy variable that indicates whether the firm has export experience prior to entry (i.e., being an experienced firm). We plot the estimated fixed effects for experienced and unexperienced firms in the left panel, using the age-one unexperienced firms as the base group. In the right panel, we plot the average $\left|FE_{t,t+1}^{\log}\right|$ by the firm's age in the calibrated model, normalizing the average $\left|FE^{\log}\right|$ to zero for unexperienced firms at age one.

Though not directly targeted in the calibration, our model predicts that the average $|FE^{\log}|$ declines over the firms' life cycle and that the initial $\left|FE^{\log}\right|$ is lower for firms with export experience. However, the model predicts a smaller decline in $\left|FE^{\log}\right|$ for experienced firms and a larger difference between experienced and unexperienced firms at age one. The reason for this is that firms in the model learn fast — there is almost no uncertainty about θ after four periods. This implies experienced firms start with very precise forecasts about

Figure 3: Forecast error — age profile: data vs. model



Notes: The left panel shows the estimated age effects on average $|FE^{\log}_{t,t+1}|$ for the firms in the data; the right panel shows the average $|FE^{\log}_{t,t+1}|$ by firms' ages in the model. To calculate the average $|FE^{\log}|$ at age t in the model, we adjust the partial-year effects by averaging the forecast errors of the firms at age $t-1$ and age t, and most firms enter into the destination market in the middle of each fiscal year. The solid line shows the estimated age effects for the firms that inherited export experience from their parent companies (i.e., experienced firms, with $Exp_{-1} > 0$); the dashed line shows the estimated age effects for unexperienced firms ($Exp_{-1} = 0$). The age effect of age-one unexperienced firms is normalized to zero.

θ and that the main source of forecast errors is the transitory shock. A possible improvement for matching the dynamics of the forecast errors is to introduce learning about both demand and supply and assume that exporting helps firms learn about demand—but not supply—conditions in the destination market, at the expense of complicating the model and quantitative analysis.

In Online Appendix Section 3.1, we compare other untargeted moments in the model and in the data. We show that our model predicts growth in exporters' sales, decline in exporters' exit rates and decline in the volatility of sales growth rates over firms' life cycles, which are consistent with the data. We further simulate a panel of firms and run firm-level regressions. We find the regression coefficients, such as the AR(1) coefficient of forecast errors and the coefficient of the interaction between the firms' experience and previous forecast errors, have the correct signs and magnitudes compared with those estimated from the data.

4.3 Implications for Allocation and Productivity

In our model, if the signals are noisier (with higher σ_ε), learning becomes slower with greater uncertainty. This distorts the selection of firms between exporting and multinational production (and exit), which in turns lowers aggregate productivity in the model.[35] Below we evaluate this quantitatively by varying σ_ε in our calibrated model and highlight two main effects at work in our model.

To gauge the possible range of σ_ε, we calibrate the three variance parameters for eight regions of the world. As in the calibration for all regions in Table 8, we use moments of forecast errors of unexperienced multinational firms and calculate the moments without solving the full model. However, the number of observations greatly decreases when we calculate the standard deviation of forecast errors of age-one firms for eight countries/regions respectively. To improve the precision of the moments, we instead target the standard deviation of forecast errors of both age-one and age-two firms.[36]

Table 10: Learning parameters by country/region

Region	Data			Calibrated			
	$SD(FE_n), n \leq 2$	$SD(FE_{10})$	$Cov(FE_{n-1}, FE_n), \forall n$	λ	σ_ε	σ_θ	σ_ξ
Middle East	0.39	0.36	0.046	0.04	5.69	1.17	1.03
Africa	0.43	0.29	0.017	0.96	1.87	1.83	0.99
Eastern Europe	0.54	0.25	0.022	2.09	2.04	2.95	0.75
China	0.46	0.25	0.014	2.23	1.60	2.38	0.86
Rest Asia-Pacific	0.46	0.27	0.011	2.73	1.44	2.38	0.98
Latin America	0.44	0.30	0.009	2.58	1.28	2.05	1.14
Western Europe	0.45	0.25	0.009	3.89	1.25	2.46	0.93
United States	0.45	0.26	0.008	4.64	1.15	2.49	0.95

Notes: For each country/region, we solve for the parameters $\sigma_\xi, \sigma_\varepsilon$ and σ_θ by matching three moments: standard deviation of forecast errors of age-one and age-two firms, standard deviation of forecast errors of firms above age ten, and covariance of forecast errors of all firms. We only include unexperienced firms in our sample, as in the calibration in Table 8. "Rest Asia-Pacific" excludes China.

In Table 10, we present the calibrated parameters for each region. Firms in less-developed economies such as the Middle East and Africa have the lowest signal-to-noise ratio, λ, and firms in advanced economies such as Western Europe and the United States tend to have higher signal-to-noise ratios and lower informational noisiness. Note that the calibrated λ depends on both σ_ε and σ_θ. However, the variation in λ largely aligns with the variation in σ_ε (Table 10). Therefore, we perform quantitative analysis using different values of σ_ε implied by various values of λ (across regions) while keeping all other parameters (including σ_θ) the same as in the baseline calibration.[37]

[35] It is the aggregate output divided by their total labor input, including domestic and foreign labor used for production, paying fixed costs and entry costs. The details are in the Appendix.

[36] In Online Appendix Table OA.24, we perform the calibration by targeting the standard deviation of forecast errors of firms aged from one to three instead of age-one and age-two firms. The parameter values barely change.

[37] σ_θ equals 2.03 in the baseline calibration; thus, the implied values of σ_ε are 10.17, 2.08, 1.36, and 0.94 for the Middle East ($\lambda = 0.04$), Africa ($\lambda = 0.96$), China ($\lambda = 2.23$), and the United States ($\lambda = 4.64$), respectively.

Table 11: Aggregate outcomes when σ_ε varies

σ_ε	Aggregate productivity	Probability of Entrants Becoming Unexperienced MNEs	Average Allocation Efficiency Index MP v.s. Export	Average Allocation Efficiency Index Active v.s. Exiter
(1)	(2)	(3)	(4)	(5)
10.17 (Middle East)	3.609	0.109	0.721	0.602
2.08 (Africa)	5.513	0.063	0.891	0.762
1.36 (China)	5.785	0.051	0.910	0.757
0.94 (USA)	5.913	0.042	0.927	0.767
0 (Perfect Informaiton)	6.141	1	1	1

Notes: The values of σ_ε correspond to the levels of information frictions in the United States, China, Africa, and the Middle East and are implied by the signal-to-noise ratios calibrated in Table 10. We keep the other model parameters the same as our baseline calibration. The last row ($\sigma_\varepsilon = 0$) presents the outcomes of the perfect information model, in which firms know θ upon entry and they sort into MP, export, and exit according to θ. The details of the perfect information benchmark model can be found in Online Appendix 2.4.

In Table 11, we compare equilibrium outcomes under different values of σ_ε as guided by the data. When we reduce σ_ε from 10.17 (the Middle East) to 0.94 (the U.S.), aggregate productivity increases by 64% (3.609 $\rightarrow$ 5.913) in Column 2. When we compare the United States with Africa, the region with the second highest value in σ_ε, information noisiness lowers aggregate productivity by 6.8% (5.913 $\rightarrow$ 5.513). Compared with the perfect information benchmark ($\sigma_\varepsilon = 0$), aggregate productivity is lower by 3.7% (6.141 $\rightarrow$ 5.913) in the United States, which is in line with other studies such as David et al. (2016). In Africa, the corresponding number is 11.2% (6.141 $\rightarrow$ 5.513) – a significant productivity loss due to informational imperfections and uncertainty. What causes these significant productivity losses in our model? We argue that the effect of a noisier signal σ_ε is twofold, and next we highlight two channels that are distinct in our model, creating misallocation along the extensive margins: (1) the learning effect and (2) the real options effect.

To highlight the learning effect, we benchmark our model against the perfect information model with the same set of parameters, except that σ_ε is zero.[38] In the perfect information model, firms sort into different production modes based on their true demand draws by following the pecking order. Firms with the best demand draws conduct multinational production, and those with the worst demand draw exit; firms with mediocre demand draw export. This pecking order result does not hold in the imperfect information model. Specifically, firms with good permanent demand draws might exit because of a series of bad information shocks at young ages, while firms with bad permanent demand draws might conduct multinational production because of a series of good information shocks at young ages. This leads to the first source of misallocation at the extensive margin and creates productivity losses. We call this effect the learning effect.

[38] See Online Appendix 2.4 for details.

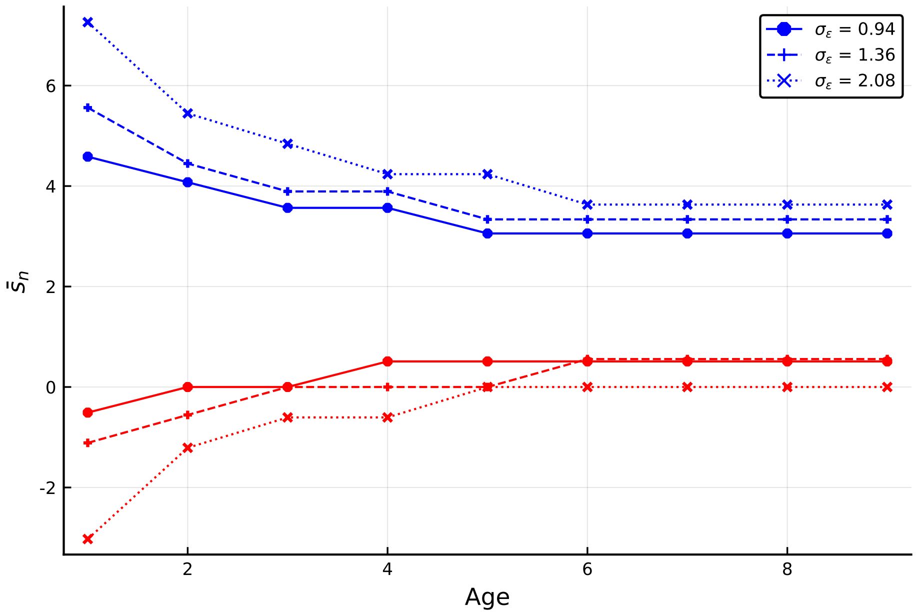
Second, firms choose their optimal production modes *immediately* after entry in the perfect information model. In the imperfect information model, even if firms with good permanent demand draws eventually become multinational firms, they might export for a long time because of the high uncertainty concerning their θ's at young ages. Similarly, even if firms with bad permanent demand draw exit eventually, they might stay in the market for a long time due to the same reason. This creates an inaction region (for entry into multinational production and exiting from the market) over exporting firms' prior beliefs about θ, which shrinks as the firm ages. This second effect is the real-options effect, which leads to the second source of misallocation at the extensive margin and leads to productivity losses.

We use the other statistics presented in Table 11 to illustrate the learning effect. In Column 3, we compute an "allocation efficiency index" to measure the misallocation at the margin of entering multinational production, inspired by the rank test in Bowen et al. (1987). This index equals the probability that a randomly selected exporter has a lower permanent demand draw than a randomly selected multinational firm in the steady-state.[39] In a perfect-information benchmark model in which firms know θ before choosing their production modes, this index equals its maximum of one because only firms with the highest θs enter multinational production. The allocation efficiency index improves as σ_ε declines, because firms make more informed decisions under lower values of σ_ε. Column (4) presents the allocation efficiency index summarizing the ordering of active firms' demand draws (both multinational firms and exporters) relative to exiting firms' demand draws. Naturally, it also improves as σ_ε declines.

In Figure 4, we further illustrate the real-options effect. Specifically, we show how an exporting firm's extensive margin choices are affected by σ_ε along the age dimension. The blue and red curves correspond to the upper and lower bounds of the inaction regions with respect to the cumulative historical signal, $\bar{s}_n$, which determines the posterior beliefs of exporters when they choose production modes. When exporting firms are young, there is a wide range of posterior beliefs within which they continue to export. Crucially, this inaction region shrinks when firms become older and/or when σ_ε declines, which implies a lower value of wait-and-see. By contrast, firms immediately figure out their optimal modes of service after entry in a perfect information model. Therefore, a smaller inaction region in our imperfect information model implies less misallocation at the extensive margins because of faster learning. In total, the allocation of firms into different production modes becomes less distorted when the variance of the information shocks decreases.

[39] In Online Appendix 2.5, we detail the construction of this index and show some of its desired properties.

Figure 4: Exporters' inaction region expands with σ_ε



Note: The blue and red curves indicate upper and lower bounds of the inaction region (keep exporting) under different values of σ_ε. As in Table 11, the values of σ_ε correspond to Africa, China and the United States (highest to lowest), implied by the signal-to-noise ratio calibrated in Table 10. We keep the other model parameters the same as our baseline calibration.

5 Conclusion

In this paper, we use a unique dataset of Japanese multinational firms, which contains information on sales forecasts, to detect imperfect information and learning over a firms life cycle. We document several new, important stylized facts concerning firms forecasts and forecast errors. We view these stylized facts as direct evidence for the existence of age-dependent firm-level uncertainty, imperfect information, and learning. We then build a dynamic industry equilibrium model of trade and multinational production, featuring noisy information and learning, to explain the documented facts. We find that more volatile informational noises reduce aggregate productivity in the model through the learning channel and the real-options channel along the age dimension. We view understanding the cross-country and cross-region variations in the level of firm-level informational frictions as a fruitful venue for further research.

6 Appendix

6.1 Definition of equilibrium

Definition 1 *A steady-state equilibrium of the model is defined as follows:*

1. *value functions* $V(o, f_m^e, n, \bar{s}_{n-1})$*, choice-specific value functions* $v(o', o, f_m^e, n, \bar{s}_{n-1})$ *and policy functions* $o'(o, f_m^e, n, \bar{s}_{n-1})$ *that satisfy equations (13), (14) and (15);*

2. *policy functions of optimal output* $q_o, o \in \{m, x\}$ *that satisfy equation (9);*

3. *prices in the current period* $p_o, o \in \{m, x\}$ *that satisfy equation (11);*

4. *a measure function of firms* $\lambda(o, f_m^e, n, \bar{s}_{n-1}, \theta)$*,* $o \in \{x, m, ent\}$ *that is consistent with the aggregate law of motion. This measure function of firms is defined at the beginning of each period (i.e., after the exogenous exit takes place but before the endogenous mode switching happens). In particular, in each period, an exogenous mass* J *of entrants draw* θ *and* $\log(f_m^e)$ *from normal distributions. Therefore, the measure of entrants with state variables (*f_m^e*,* θ*) is*

$$\lambda(ent, df_m^e, 1, \bar{s}_0, d\theta) = (1-\eta) J g_\theta(\theta)\, d\theta \times g_{f_m^e}(f_m^e)\, df_m^e,$$

   where $g_\theta(\cdot)$ *and* $g_{f_m^e}(\cdot)$ *are the density functions of the distributions for* θ *and* f_m^e*, respectively. The measure function for exporters and multinational firms should be a fixed point of the aggregate law of motion, i.e., given any Borel set of* $\bar{s}_n$*,* Δ_s*, measures of firms with* $n \geq 2$ *satisfy*

$$\lambda(o', f_m^e, n+1, \Delta_s, \theta) = \sum_{o \in \{x,m\}} \int_{\bar{s}_{n-1}, a} \begin{array}{c} \mathbf{1}(\bar{s}_n \in \Delta_s, o'(o, f_m^e, n, \bar{s}_{n-1}) = o') \times \\ (1-\eta) \Pr(\bar{s}_n | \bar{s}_{n-1}, \theta)\, \lambda(o, f_m^e, n, d\bar{s}_{n-1}, \theta) \end{array}.$$

   Firms with $n = 1$ *were entrants in the previous period:*

$$\lambda(o', f_m^e, 2, \Delta_s, \theta) = \int_a \begin{array}{c} \mathbf{1}(\bar{s}_1 \in \Delta_s, o'(ent, f_m^e, 1, \bar{s}_0) = o') \times \\ (1-\eta) \Pr(\bar{s}_1 | \bar{s}_0, \theta)\, \lambda(ent, f_m^e, 1, \bar{s}_0, \theta) \end{array}.$$

5. *the price index* P *is constant over time and must be consistent with consumer optimization (5):*

$$P^{1-\sigma} = \sum_{n \geq 1} \sum_{\substack{o \in \{x,m,ent\} \\ o' \in \{x,m\}}} \int_{f_m^e, \bar{s}_{n-1}, \theta, a} \begin{array}{c} \mathbf{1}(o'(o, f_m^e, n, \bar{s}_{n-1}) = o')\, e^a p_{o'}(a, q_{o'}(b(\bar{s}_{n-1}, n-1)))^{1-\sigma} \\ \times (1-\eta) \lambda(o, df_m^e, n, d\bar{s}_{n-1}, d\theta)\, d\Pr(a|\theta) \end{array}.$$

where o' is the policy function $o'(o, f_m^e, n, \bar{s}_{n-1})$.

6.2 Aggregate Labor Productivity

We define aggregate labor productivity as the aggregate output by Japanese firms divided by their total labor input, including domestic and foreign labor used for production, paying fixed costs and entry costs.

The aggregate output follows our definition of the CES composite of different varieties of Japanese goods in equation (3) in the paper. In the steady state, we can express the CES composite integrating over the mass of firms with different state variables:

$$Q = \left(\sum_{n \geq 1} \sum_{\substack{o \in \{x,m,ent\} \\ o' \in \{x,m\}}} \int_{f_m^e, \bar{s}_{n-1}, \theta, a} \begin{array}{c} \mathbf{1}\left(o'\left(o, f_m^e, n, \bar{s}_{n-1}\right) = o'\right) e^{a/\sigma} q_{o'}\left(b\left(\bar{s}_{n-1}, n-1\right)\right)^{\frac{\sigma-1}{\sigma}} \\ \times (1-\eta)\lambda\left(o, df_m^e, n, d\bar{s}_{n-1}, d\theta\right) d\Pr(a|\theta) \end{array} \right)^{\frac{\sigma}{\sigma-1}},$$

where o' is the policy function $o'(o, f_m^e, n, \bar{s}_{n-1})$. Because of the demand shifter is heterogeneous across firms, aggregate labor productivity is affected by the joint distribution of state variables despite our assumption that all firms have the same labor productivity.

Domestic labor L is used for (1) production by exporters (L_x^{prod}) (2) paying entry costs by firms that switched to MP ($L_{o \to m}^{entry}, o \in \{ent, x\}$) and (3) paying fixed costs by exporters (L_x^{fixed}). We can express each component as:

$$L_x^{prod} = \tau \sum_{n \geq 1} \sum_{o \in \{x,m,ent\}} \int_{f_m^e, \bar{s}_{n-1}, \theta} \begin{array}{c} \mathbf{1}\left(o'\left(o, f_m^e, n, \bar{s}_{n-1}\right) = x\right) q_x\left(b\left(\bar{s}_{n-1}, n-1\right)\right) \\ \times (1-\eta)\lambda\left(o, df_m^e, n, d\bar{s}_{n-1}, d\theta\right) \end{array}, \tag{16}$$

$$L_{x \to m}^{entry} = \sum_{n \geq 2} \int_{f_m^e, \bar{s}_{n-1}, \theta} \mathbf{1}\left(o'\left(x, f_m^e, n, \bar{s}_{n-1}\right) = m\right) f_m^e (1-\eta)\lambda\left(x, df_m^e, n, d\bar{s}_{n-1}, d\theta\right), \tag{17}$$

$$L_{ent \to m}^{entry} = \int_{f_m^e, \theta} \mathbf{1}\left(o'\left(ent, df_m^e, 1, \bar{s}_0\right) = m\right) f_m^e (1-\eta)\lambda\left(ent, df_m^e, 1, \bar{s}_0, d\theta\right), \tag{18}$$

$$L_x^{fixed} = \sum_{n \geq 1} \sum_{o \in \{x,m,ent\}} \int_{f_m^e, \bar{s}_{n-1}, \theta} \begin{array}{c} \mathbf{1}\left(o'\left(o, f_m^e, n, \bar{s}_{n-1}\right) = x\right) f_x \\ \times (1-\eta)\lambda\left(o, df_m^e, n, d\bar{s}_{n-1}, d\theta\right) \end{array}. \tag{19}$$

Foreign labor, L^*, is used for multinational production only. Therefore,

$$L^* = L_m^{prod} = \sum_{n \geq 1} \sum_{o \in \{x,m,ent\}} \int_{f_m^e, \bar{s}_{n-1}, \theta} \begin{array}{c} \mathbf{1}\left(o'\left(o, f_m^e, n, \bar{s}_{n-1}\right) = m\right) q_m\left(b\left(\bar{s}_{n-1}, n-1\right)\right) \\ \times (1-\eta)\lambda\left(o, df_m^e, n, d\bar{s}_{n-1}, d\theta\right) \end{array}.$$

Finally, aggregate labor productivity is defined as

$$\frac{Q}{L+L^*} = \frac{Q}{\sum_{o=m,x} L_o^{prod} + \sum_{o=ent,x} L_{o\to m}^{entry} + L_x^{fixed}}.$$

Note that we do not differentiate between domestic and foreign labor. However, since we have normalized both domestic and foreign wages to one, the current concept of total labor is equivalent to total efficiency units of labor when adjusted using wages.

References

Abel, A. B. and Eberly, J. C. (1996). Optimal Investment with Costly Reversibility. *The Review of Economic Studies*, 63(4):581–593.

Altig, D., Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B. H., and Parker, N. (2019). Surveying business uncertainty. Technical report, National Bureau of Economic Research.

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, 60(8):967–982.

Arkolakis, C., Papageorgiou, T., and Timoshenko, O. A. (2018). Firm learning and growth. *Review of Economic Dynamics*, 27:146–168.

Atkeson, A. and Kehoe, P. j. (2005). Modeling and Measuring Organization Capital. *Journal of Political Economy*, 113(5):1026–1053.

Bachmann, R. and Elstner, S. (2015). Firm Optimism and Pessimism. *European Economic Review*, 79:297–325.

Bachmann, R., Elstner, S., and Hristov, A. (2017). Surprise, Surprise – Measuring Firm-Level Investment Innovations. *Journal of Economic Dynamics and Control*, 83:107–148.

Bachmann, R., Elstner, S., and Sims, E. R. (2013). Uncertainty and Economic Activity: Evidence from Business Survey Data. *American Economic Journal: Macroeconomics*, 5(2):217–249.

Baker, S. R., McElroy, T. S., and Sheng, X. S. (2018). Expectation formation following large unexpected shocks. *Review of Economics and Statistics*, pages 1–45.

Baley, I. and Blanco, A. (2019). Firm Uncertainty Cycles and the Propagation of Nominal Shocks. *American Economic Journal: Macroeconomics*, 11(1):276–337.

Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and Productivity in International Trade. *American Economic Review*, 93(4):1268–1290.

Bertola, G. and Caballero, R. J. (1994). Irreversibility and Aggregate Investment. *The Review of Economic Studies*, 61(2):223–246.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685.

Bloom, N., Davis, S. J., Foster, L., Lucking, B., Ohlmacher, S., and Saporta-Eksten, I. (2017). Business-Level Expectations and Uncertainty. SSRN Scholarly Paper 3085377, Social Science Research Network, Rochester, NY.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really Uncertain Business Cycles. *Econometrica*, 86(3):1031–1065.

Boneva, L., Cloyne, J., Weale, M., and Wieladek, T. (2020). Firms' Price, Cost and Activity Expectations: Evidence from Micro Data. *The Economic Journal*, 130(627):555–586.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2018). Over-reaction in Macroeconomic Expectations. NBER Working Paper 24932.

Bowen, H. P., Leamer, E. E., and Sveikauskas, L. (1987). Multicountry, Multifactor Tests of the Factor Abundance Theory. *The American Economic Review*, 77(5):791–809.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review*, 101(5):1964–2002.

Buera, F. J. and Moll, B. (2015). Aggregate implications of a credit crunch: The importance of heterogeneity. *American Economic Journal: Macroeconomics*, 7(3):1–42.

Coibion, O. and Gorodnichenko, Y. (2012). What Can Survey Forecasts Tell Us About Information Rigidities? *Journal of Political Economy*, 120(1):116–159.

Coibion, O. and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review*, 105(8):2644–78.

Coibion, O., Gorodnichenko, Y., and Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, 108(9):2671–2713.

Coibion, O., Gorodnichenko, Y., and Ropele, T. (2020). Inflation Expectations and Firm Decisions: New Causal Evidence. *The Quarterly Journal of Economics*, 135(1):165–219.

Conconi, P., Sapir, A., and Zanardi, M. (2016). The internationalization process of firms: From exports to FDI. *Journal of International Economics*, 99:16–30.

Das, S., Roberts, M. J., and Tybout, J. R. (2007). Market Entry Costs, Producer Heterogeneity, and Export Dynamics. *Econometrica*, 75(3):837–873.

David, J. M., Hopenhayn, H. A., and Venkateswaran, V. (2016). Information, Misallocation, and Aggregate Productivity. *The Quarterly Journal of Economics*, 131(2):943–1005.

Dixit, A. (1989). Entry and exit decisions under uncertainty. *Journal of political Economy*, 97(3):620–638.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment Under Uncertainty*. Princeton university press.

Enders, Z., Hünnekes, F., and Müller, G. J. (2019a). Firm expectations and economic activity. CESifo Working Paper. tex.publisher: CESifo Working Paper.

Enders, Z., Hünnekes, F., and Müller, G. J. (2019b). Monetary policy announcements and expectations: Evidence from german firms. *Journal of Monetary Economics*.

Feenstra, R. C., Luck, P., Obstfeld, M., and Russ, K. N. (2017). In Search of the Armington Elasticity. *The Review of Economics and Statistics*, 100(1):135–150.

Foster, A. D. and Rosenzweig, M. R. (2010). Microeconomics of technology adoption. *Annu. Rev. Econ.*, 2(1):395–424.

Foster, L., Haltiwanger, J., and Syverson, C. (2016). The slow growth of new plants: learning about demand? *Economica*, 83(329):91–129. Publisher: Wiley Online Library.

Gennaioli, N., Ma, Y., and Shleifer, A. (2016). Expectations and investment. *NBER Macroeconomics Annual*, 30(1):379–431.

Guiso, L. and Parigi, G. (1999). Investment and Demand Uncertainty. *The Quarterly Journal of Economics*, 114(1):185–227.

Handley, K. (2014). Exporting under trade policy uncertainty: Theory and evidence. *Journal of international Economics*, 94(1):50–66.

Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus FDI with Heterogeneous Firms. *The American Economic Review*, 94(1):300–316.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2014). The life cycle of plants in India and Mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084.

Irwin, D. A. and Klenow, P. J. (1994). Learning-by-Doing Spillovers in the Semiconductor Industry. *Journal of Political Economy*, 102(6):1200–1227.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica*, 50(3):649–670.

Jovanovic, B. and Nyarko, Y. (1996). Learning by doing and the choice of technology. *Econometrica*, 64(6):1299.

Khan, A. and Thomas, J. K. (2013). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 121(6):1055–1107.

Kydland, F. E. and Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica*, 50(6):1345–1370.

Ma, Y., Ropele, T., Sraer, D., and Thesmar, D. (2019). A quantitative analysis of distortions in managerial forecasts. Working Paper.

Mackowiak, B. and Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American economic review*, 104(2):422–58.

Parente, S. L. and Prescott, E. C. (1994). Barriers to Technology Adoption and Development. *Journal of Political Economy*, 102(2):298–321.

Ryngaert, J. (2017). What Do (and Don't) Forecasters Know about U.S. Inflation? Working Paper.

Sedláček, P. and Sterk, V. (2017). The growth potential of startups over the business cycle. *American Economic Review*, 107(10):3182–3210.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Tanaka, M., Bloom, N., David, J. M., and Koga, M. (2019). Firm performance and macro forecast accuracy. *Journal of Monetary Economics*.