Journal Pre-proof

Uncertainty, imperfect information, and expectation formation over the firm's life cycle

Cheng Chen, Tatsuro Senga, Chang Sun, Hongyong Zhang

PII: S0304-3932(23)00087-9
DOI: https://doi.org/10.1016/j.jmoneco.2023.08.001
Reference: MONEC 3508

To appear in: *Journal of Monetary Economics*

Received date : 13 July 2022
Revised date : 6 August 2023
Accepted date : 7 August 2023

Please cite this article as: C. Chen, T. Senga, C. Sun et al., Uncertainty, imperfect information, and expectation formation over the firm's life cycle. *Journal of Monetary Economics* (2023), doi: https://doi.org/10.1016/j.jmoneco.2023.08.001.

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

Cheng Chen Tatsuro Senga Chang Sun Hongyong Zhang†

August 2023

Abstract

Using a long panel data set on Japanese firms, we find that firms make more precise forecasts and less autocorrelated forecast errors as they gain more experience. Then, we build a firm dynamics model where firms gradually learn about their demand by using a noisy signal. Using expectations data over time, we cleanly isolate the learning mechanism from other mechanisms and find that it accounts for 20%–40% of the overall decline in forecast errors over the life cycle. Productivity gains from removing information frictions range from 3% to 12%, with firm entry and exit playing prominent roles.

Keywords. Firm expectations; Forecast errors; Learning, Life cycles, Productivity

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

*This research was conducted as a part of a research project "Studies on Firm Management and Internationalization under the Growing Fluidity of the Japanese Economy" funded by the Research Institute of Economy, Trade, and Industry (RIETI). We thank the editor, the referee, Costas Arkolakis, Nick Bloom, Vasco Carvalho, Gene Grossman, Kyle Handley, Christian Hellwig, Aubhik Khan, Nicholas Kozeniauskas, 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), ESRC (grant number: ES/V012266/1), RIETI, and Princeton University is greatly appreciated.

†Chen: Clemson University, cchen9@clemson.edu. Senga: Queen Mary University of London and Keio University, 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

This paper proposes a new approach to quantifying the impact of imperfect information on aggregate productivity in a firm dynamics model with learning. The exercise involves measuring the gap in aggregate productivity between the status quo and the perfect information scenario as in existing research. For this exercise to be valid, one needs to identify the extent of imperfect information as the status quo based on a mapping from observed data on the endogenous variables of the model to parameters that govern the degree of informational imperfection. Although existing studies construct such a mapping based on moments such as input and output choices, this approach has limitations in terms of identifying certain types of information frictions and it requires simplifying assumptions on the information environment.[1] The contribution of this paper is the use of direct expectations data to establish such a clean mapping, and thereby to identify the extent of the dynamic (i.e., life-cycle) imperfect information in an enriched model environment and to demonstrate its detrimental impact on aggregate productivity.

We consider that each individual firm faces uncertainty about idiosyncratic demand and productivity. Demand is time-invariant and learned gradually, whereas productivity follows a first-order autoregressive (AR(1)) process with a variance that declines with firm age (the volatility effect) and is perfectly revealed at the end of each period. The consequences are twofold. First, the learning generates the autocovariances of forecast errors that decline with firm age because firms correct past forecast errors about demand partially and gradually over time. The volatility effect has no impact on the autocovariance because firms correct forecast errors about productivity perfectly and instantly in each period. Second, both the learning and volatility effects yield a variance of forecast errors that declines with firm age. Thus, we can use our model as an accounting device to decompose the contributions of learning and volatility to firms' uncertainty: the contribution of learning is entirely responsible for

[1] For instance, the seminal work of David, Hopenhayn, and Venkateswaran (2016) and David and Venkateswaran (2018) focuses on *static* (or within-period) information frictions, as information is perfectly revealed at the end of each period.

the autocovariance and partially responsible for the variance, with the remaining variance attributed to the volatility effect.

Our paper proceeds in three steps. First, we use a 20-year panel data set on Japanese multinational firms that matches parent firms with affiliated firms. The data set is based on annual business surveys conducted by the Japanese government. Exploring our data set, we document that the variance of forecast errors declines with firms' experience of operations, and that this is robust to controlling for firm size and measures of market/product diversification. Crucially, we document that although each firm's forecast errors are positively autocorrelated, the autocorrelation of forecast errors declines with firms' experience, a new fact that has not been uncovered by existing studies. In addition, we find that firms in countries with better management and/or smaller time differences from Japan make less serially correlated forecast errors. These stylized facts suggest that firms learn about their demand and become better informed as they accumulate more experience. In addition, low-quality management and barriers to within-firm communication are likely to be drivers of information frictions. We believe that the stylized facts about firm-level forecasts are useful for disciplining dynamic firm life-cycle models even if future researchers decide to adopt different setups.

In the second part of the paper, we integrate Jovanovic (1982)-style learning into an otherwise standard industry equilibrium model of heterogeneous firms. Firms face a downward-sloping demand curve in a setting where the firm-specific time-invariant demand shifter is heterogeneous across firms and unknown to them (i.e., never observed by the firm). We depart from Jovanovic (1982) in two crucial ways. First, we assume that firms face information constraints and thus learn about their demand from a noisy signal, which is purely informational and *does not* affect firms' per-period profits (i.e., it is payoff irrelevant).[2] Second, we

[2]We show that this information structure allows us to reproduce the age-declining variance and autocorrelation of forecast errors. The age-declining autocorrelation of forecast errors implies a deviation from full-information rational expectations, but it can reflect either a deviation from full information or departures from rational expectations. We account for the age-declining autocorrelation of forecast errors by a model of information constraints under rational expectations in the spirit of (Coibion and Gorodnichenko, 2012; Coibion, Gorodnichenko, and Kumar, 2015).

introduce an idiosyncratic shock to firm-level productivity in every period, which is revealed to the firm at the end of each period, and we assume that its volatility decreases as firms become older, following Atkeson and Kehoe (2005). Therefore, not only learning but also the age-declining volatility contribute to the age-declining variance of forecast errors in the model, reflecting factors other than learning in much the same way as reality, where there are many other factors that explain the age-declining variance of forecast errors.[3]

In the final part of the paper, we use our model to implement three empirical/quantitative exercises. First, our decomposition exercise shows that the contribution of learning to the *change* in the variance of forecast errors over the firm's life cycle ranges between 20% to 40%. To the best of our knowledge, our decomposition exercise is the first to isolate the evolution of firms' beliefs over their life cycle directly from panel data on expectations and to succeed in isolating the learning channel from other factors that contribute to the age-declining volatility of the firm.

Second, we demonstrate our approach incorporating both the learning and other channels by calibrating our model to infer the learning parameters and other key parameters governing firm dynamics. Our counterfactual experiment of eliminating imperfect information reveals not only a substantial gain in overall productivity, but also the role of endogenous selection in driving it. In our calibrated model, firms incur a positive fixed cost of operation every period. Therefore, only firms with high productivity and/or high perceived demand enter and stay active, whereas other firms may not enter, or may exit soon after entry. This endogenous selection is shut down in a version of our model with zero fixed operation costs, such that all firms enter and only exit due to a random shock. With selection at work in the model, providing better information leads not only to more informed static decisions such as employment, but also to more informed dynamic decisions on entry and exit, implying larger gains in overall productivity. For instance, the productivity gain is 3.49% when we assume

[3] This paper focuses on learning, not modeling all other mechanisms. Thus, we use the (exogenously) age-declining volatility to capture all other mechanisms that affect firms' forecast errors without taking a stand on particular sources.

away the fixed cost, and it becomes 6.35% in our baseline model with the fixed cost. Another important implication is that gains (and the improved selection effect) from removing the information friction are disproportionally larger among younger firms, as life-cycle learning plays a more prominent role for young firms.

Finally, we implement a cross-regional analysis, where we calibrate our model to match data moments for eight regions/countries in the world. We show that the degree of imperfect information and the associated aggregate implications vary across regions/countries. Eliminating the life-cycle information friction leads to average productivity gains ranging from 3% to 12% for different regions. For example, the productivity loss due to information frictions in Africa is almost twice as large as that in Western Europe. Our results are broadly consistent with the view that low-quality management and inefficient within-firm communications can lead to more severe information frictions and therefore larger productivity losses.[4]

Related literature: Our work contributes to the literature on misallocation due to imperfect information, particularly in the life cycle of firms (Hsieh and Klenow (2014)). Our approach and focus differ from previous work, such as David et al. (2016) and David and Venkateswaran (2018), as we directly quantify the productivity loss caused by the life-cycle information friction. This allows us to separately measure the degree of volatility and information frictions over the firm's life cycle, which can have different policy implications.[5] We show that productivity losses through extensive margin dynamics, such as firms' entries and exits, are substantial and highlight the detrimental effect of informational imperfection on young firms.[6] Finally, existing research uses data on public firms to quantify the gain from eliminating information frictions (e.g., David et al. (2016) and Ma, Ropele, Sraer, and Thes-

[4]The caveat here is that our data only contain Japanese firms in various regions/countries, and Japanese firms in one region/country may not be representative of all firms in that region/country. Therefore, the calibrated parameters and the implied productivity gains across regions should be taken with caution.

[5]Life-cycle learning is more relevant for young firms. Therefore, policies that specifically help young firms (e.g., subsidizing their training programs for managers/workers) can lead to productivity gains by alleviating the information problem. For the static information friction studied in David et al. (2016), reducing the labor/capital adjustment costs can reduce the productivity loss due to imperfect information.

[6]Our paper complements the results of other studies on misallocation, including those on financial frictions, such as Buera, Kaboski, and Shin (2011) and Midrigan and Xu (2014).

mar (2019)). As we show that the severity of informational imperfection is higher among younger firms, the productivity gain from eliminating information frictions for the entire economy might be understated by the previous papers.

Although economists have long speculated on how agents form expectations, recent studies, such as Bloom, Davis, Foster, Lucking, Ohlmacher, and Saporta-Eksten (2020) and Altig, Barrero, Bloom, Davis, Meyer, and Parker (2020), have begun to collect and analyze direct expectations data. This approach is useful in modeling and calibrating theoretical frameworks, as shown by the seminal works of Coibion and Gorodnichenko (2012) and Coibion et al. (2015).[7] Our paper contributes to this literature by examining firms' expectations of idiosyncratic objects, such as their own sales, following the approach of Enders et al. (2022), Born, Enders, Müller, and Niemann (2022), and Bachmann, Carstensen, Lautenbacher, and Schneider (2021). We use the expectations data to understand the dynamics of young firms.[8]

Our paper contributes to the literature on firm-level uncertainty by examining how it evolves over the life cycle. Our work aligns with studies by Baley and Blanco (2019), Baley, Figueiredo, and Ulbricht (2022), and Ilut, Valchev, and Vincent (2020), who examine uncertainty fluctuations within firms, although, as noted previously, our paper is unique in separating learning from volatility. In a business-cycle context, the role of information accumulation at the firm level has been studied by Ilut and Saijo (2021), who also use forecast data to validate the structural model.

2 Empirical Facts

In this section, we construct our panel of Japanese firms operating in foreign markets to document the properties of the forecast errors and their relationship with firms' experience. The facts that we will present indicate that firms become better informed as they accumulate

[7]Recent papers that have studied how agents form expectations and respond to shocks include Coibion, Gorodnichenko, and Kumar (2018), Baker, McElroy, and Sheng (2020), and Enders, Hünnekes, and Müller (2022).

[8]We focus on the roles of learning and volatility effects in driving young firm dynamics. For other drivers of young firm dynamics, see Sedláček and Sterk (2017) and Foster, Haltiwanger, and Syverson (2016).

more experience, and that management and within-firm communication could be one driver of information frictions.

2.1 Data and the Reliability of Sales Forecasts

Our main data source is the Basic Survey on Overseas Business Activities (the "foreign activities survey" hereinafter) 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 equity share capital owned by Japanese parent companies, and (2) second-tier affiliated firms with more than 50% of equity share capital owned by Japanese parent companies. The survey is designed to include all Japanese overseas affiliates that satisfy either of the above criteria. Although some firms do not respond to the survey, the response rate is high (71.3% in 2013). The survey is conducted in July and August each year to collect firm-level data on the previous fiscal year (April of the previous year to March of the current year) and their expectations for the current fiscal year.[9] We discuss the expectations data in detail later.

After dropping tax haven countries documented in Gravelle (2009), our baseline regression sample contains, on average, 1,781 parent companies and 6,922 affiliated firms in a typical year during the period from 1995 to 2013. Our sample covers Japanese firms operating in 96 countries and 29 industries, including both manufacturing and services. In Appendix Section A.1, we report descriptive statistics regarding subsamples in different time periods and the distribution of firms across regions and industries in a typical year. The unit of analysis in our empirical investigation is the affiliated firm by year. We slightly change the terminology,

[9]We provide a detailed timeline of the survey in Appendix Section C.3. In addition, we discuss the implications of firms obtaining three months of experience before they make their current forecast. We perform a robustness check by making firms effectively "older by three months" in our sample; after re-calibrating the model, we find slightly larger effects of information friction and gains from moving to perfect information than we find for our benchmark calibration.

referring to the affiliated firms as "firms" and to all the affiliated firms belonging to the same parent company as a "business group."

The unique feature of the foreign activities survey is that each firm reports its sales forecast for the current fiscal year when it fills out the survey. Because such information is rarely available in firm-level data sets, we show that the sales forecasts are reliable and contain useful information that affects actual firm decisions. First, we find that firms do not use naive rules to make their sales forecasts. In Appendix Table A-5, only 3.35% of the observations use their sales in year t as a forecast of sales in $t+1$. Our main regression results are almost unchanged after dropping these observations. Second, we show in Appendix Table A-6 that sales forecasts strongly predict future firm outcomes, even after we control for 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 for purposes beyond the scope of the survey, such as tax collection. Firms have no incentive to misreport because of tax purposes or because they want to manage the expectations of stock market investors. We provide more details about these validation exercises in Appendix Section A.2.

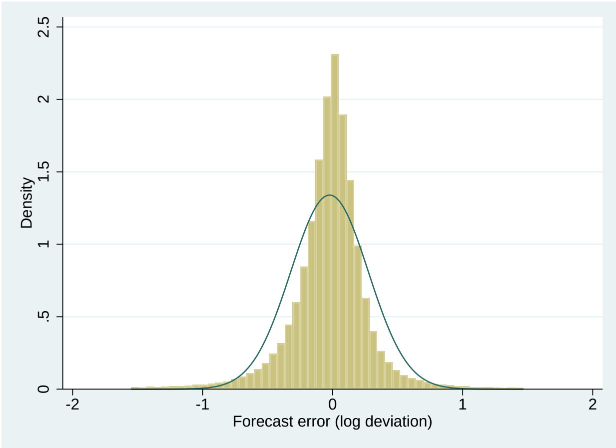
2.2 Forecast Errors

Now, we 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, expressed as:

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

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 in the next period. A positive (negative) forecast error means that the firm under-predicts (over-predicts) its sales. In Appendix Tables A-9, A-10, A-18, and A-19, 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.[10] We trim the top and bottom 1% of observations of the forecast errors to exclude outliers.

Figure 1: Distribution of the Forecast Errors



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

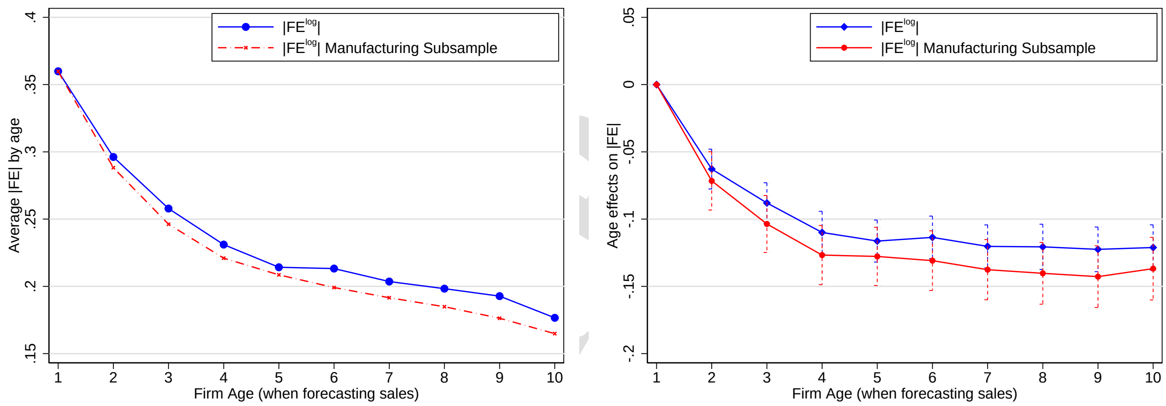
In Figure 1, we plot the distribution of our baseline measure of forecast errors, $FE^{\log}_{t,t+1}$, across all firms in all years. The forecast errors are centered around zero, and the distribution appears to be symmetric. The shape of the density is similar to a normal distribution, although the 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 over- or under-forecast their sales by 20%.

[10]The aggregate components explain approximately 11% of the variation in forecast errors. Recent work has substantiated that firms may have heterogeneous exposure to aggregate shocks, which implies that the "simple" residual forecast errors that we construct may 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 provided in Appendix Section A.4.3.

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

Panel (a) of Figure 2 presents the average absolute value of forecast errors by age cohorts, where age is top-coded at 10.[11] The precision of sales forecasts increases as the firm ages. Specifically, as the firm's age increases from one to 10 years, the absolute forecast errors decline from 36% to 18% on average. Moreover, the decline occurs mainly in the first five years after entry. For concreteness, we also present these statistics for a subsample in the manufacturing sector. The patterns are similar.

Figure 2: $|FE^{\log}|$ Declines with Firm Age



(a) Average $|FE^{\log}|$ by Age (b) Age Dummy Coefficients of $|FE^{\log}|$

Note: Panel (a) plots the average absolute value of $FE^{\log}$ by age cohorts for all firms and for the manufacturing subsample. Panel (b) plots the coefficients of age dummies in the regression specified in equation (A-1). Other than age dummies, we control for firm employment, parent firm employment, industry-year fixed effects, country-year fixed effects, and firm fixed effects. Age-one firms are used as the base category and the coefficient of $\mathbb{1}(\text{Age}_t = 1)$ is normalized to zero. The capped spikes indicate the 95% confidence intervals of the estimates. The two lines correspond to the results in Columns 3 and 5 in Appendix Table A-8.

We further confirm these patterns by using an ordinary least squares regression of firm i's absolute forecast error in year t:

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

[11] We top code the age at 10, as the average absolute forecast error does not decline much after this age, especially in regressions where we control for a set of fixed effects and firm size (see Panel b of Figure 2).

where δ_n is a vector of age dummies, and δ_{ct}, δ_{st}, and δ_i represent the country-year, industry-year, and firm fixed effects, respectively. Time-varying controls, such as firm and parent firm employment, 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. We plot the coefficients of age dummies and their confidence intervals in Panel (b) of Figure 2. It is clear that the absolute forecast errors decline significantly with firm age. We report the detailed regression results in Appendix Table A-8.

We interpret the decline in absolute forecast errors as an improvement in firms' information about their own capability and perform a battery of robustness checks to rule out alternative explanations. As shown in Column 4 of Appendix Table A-8, we find similar results for firms that have survived and continuously appeared in the data from age one to seven, suggesting that our results are not driven by endogenous exits or nonreporting. We further show that our results are (1) robust to alternative measures of forecast errors, including those that explicitly take into account firms' heterogeneous exposure to aggregate shocks (Appendix A.4.2 and A.4.3); (2) robust to controlling for product and market diversification (Appendix A.4.4); (3) not due to age-dependent biases in the level of forecast errors (Appendix A.4.5); and (4) not driven by a "partial-year effect," that is, firms entering in different months of a fiscal year (Appendix A.4.6).

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, Coibion and Gorodnichenko (2012) and Ryngaert (2017) demonstrated that professional forecasters' forecast errors of future inflation rates are autocorrelated, indicating the existence of information frictions related to macroeconomic conditions. 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.

In Appendix Table A-17, we present 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 0.137. This result suggests that firms tend to make systematic errors in forecasting their sales. The remaining three columns show that such serial correlation becomes weaker when firms gain more experience, indicating that firms become more informed and make smaller systematic errors when forecasting. Such patterns are robust to using alternative definitions of forecast errors and to using the manufacturing subsample.[12]

Next, we confirm this pattern by running the AR(1) 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:

$$FE^{\log}_{i,t+1,t+2} = \beta_1 FE^{\log}_{i,t,t+1} + \beta_2 FE^{\log}_{i,t,t+1} \times Age_{it} + \beta_3 X_{it} + \delta_{st} + \delta_{ct} + \delta_g + u_{it}, \tag{2}$$

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 industry-year, country-year, and business group fixed effects, denoted by δ_{st}, δ_{ct}, and δ_g, respectively. In some regressions, we replace the business group fixed effects with business group–firm age fixed effects.

Table 1 shows the regression results. To capture the nonlinear effect of the firm's age, we use either age top-coded at 10 or the log of age. According to the estimates in Column 1, the AR(1) coefficient starts at 0.098 at age one and each additional year of experience reduces it by 0.006. When controlling for business group–firm age fixed effects instead of business group fixed effects, the AR(1) coefficients as well as the impact of firm age are higher. The results are similar when we focus on firms in the manufacturing sample (Columns 5–8).

[12] Importantly, our results are robust when using percentage forecast errors, $\frac{R_{t+1}-E_t(R_{t+1})}{E_t(R_{t+1})}$, and are not an artifact of the log transformation.

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

Dep. Var: $FE^{\log}_{t+1,t+2}$	All firms				Manufacturing			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$FE^{\log}_{t,t+1}$	0.104^a	0.100^a	0.131^a	0.123^a	0.114^a	0.112^a	0.137^a	0.134^a
	(0.014)	(0.013)	(0.018)	(0.017)	(0.019)	(0.019)	(0.025)	(0.025)
$\times \max\{\text{Age}_t, 10\}$	-0.006^a		-0.008^a		-0.008^a		-0.009^a	
	(0.002)		(0.002)		(0.002)		(0.003)	
$\times \log(\text{Age}_t)$		-0.018^a		-0.023^a		-0.027^a		-0.031^a
		(0.006)		(0.007)		(0.009)		(0.011)
$\log(\text{Emp})_t$	0.003^a	0.003^a	0.002^b	0.002^b	0.002^c	0.002^c	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(\text{Parent Emp})_t$	-0.010^b	-0.010^b	-0.010^b	-0.010^b	-0.010^c	-0.010^c	-0.014^b	-0.014^b
	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
Industry-year 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		
Busi.Group-Age FE			Y	Y			Y	Y
N	93478	93478	84839	84839	58630	58630	52510	52510
R^2	0.205	0.205	0.274	0.274	0.229	0.229	0.300	0.300

Notes: Standard errors are clustered at the business group level. Significance levels: c: 0.10, b: 0.05, a: 0.01.

Fact 3: Potential Drivers of Information Frictions

Our data cover a wide range of countries where Japanese firms operate. This subsection explores how serial correlation of forecast errors are correlated with various characteristics of each country, using similar specifications to those in Table 1 to shed light on the potential drivers of underlying differences in informational imperfection across countries.

We focus on three country characteristics: (1) management, (2) time zone differences, and (3) real gross domestic product (GDP) per capita. As suggested by Bloom, Kawakubo, Meng, Mizen, Riley, Senga, and Van Reenen (2021), better-managed firms have superior monitoring practices and can make more accurate forecasts about their own sales growth than can poorly managed firms. Therefore, we use country-level average management scores from Bloom, Lemos, Sadun, Scur, and Van Reenen (2014) as a measure of the management quality in each country. Second, the literature has identified time zone differences as barriers to communication within (multinational) firms (Gumpert, 2018; Bahar, 2020) that possibly lead to more information frictions. Finally, we examine real GDP per capita at the beginning of our sample (1995), which is a proxy for the overall development level of the countries.[13] We

[13]Low GDP per capita may capture a shortage of good managers, as discussed in Hjort, Malmberg, and

interact the country characteristics with the (one-period) lagged forecast error to observe how they affect the AR(1) coefficient. These results are by no means causal and the list of drivers that we study here is not exhaustive. Nevertheless, they help elucidate why information frictions at the firm level differ.

Table 2: AR(1) Coefficient and Country Characteristics

	Dep.Var: $FE^{\log}_{t+1,t+2}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$FE^{\log}_{t,t+1}$	0.1264^a	0.1121^a	0.1077^a	0.0701^a	0.0643^a	0.0606^a	0.0837^a	0.0705^a	0.0670^a
	(0.0080)	(0.0068)	(0.0070)	(0.0094)	(0.0075)	(0.0078)	(0.0101)	(0.0083)	(0.0085)
× Management Score (WMS 2015)	-0.0131^c			-0.0087			-0.0229^a		
	(0.0070)			(0.0071)			(0.0081)		
× Time Diff from Japan		0.0098			0.0142^b			0.0116	
		(0.0066)			(0.0066)			(0.0073)	
× log GDP p.c. 1995			-0.0112^c			-0.0077			-0.0178^a
			(0.0058)			(0.0058)			(0.0066)
Industry-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Country-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Business Group FE				Y	Y	Y			
Busi.Group-Age FE							Y	Y	Y
N	62005	96100	96100	61200	95152	95152	53433	86271	86271
R^2	0.130	0.135	0.135	0.207	0.201	0.201	0.283	0.270	0.270

Notes: Standard errors are clustered at the business group level. Significance levels: c: 0.1, b: 0.05, a: 0.01. The management score is from the World Management Survey up to 2015 . The management score, time zone differences, and log GDP per capita are all standardized to facilitate interpretation of the coefficients.

Table 2 reports the regression results. Country characteristics are all standardized to facilitate interpretation. In Columns 1–3, we control for industry-year and country-year fixed effects, and in the other columns, we further control for business group or business group-firm age fixed effects. In general, we find that the management score and GDP per capita are negatively associated with the AR(1) coefficient of forecast errors, whereas time zone differences affect the coefficient positively.[14] If we view the AR(1) coefficient as a measure for information frictions, these results are consistent with our hypotheses that better management, more similar time zones, and higher income levels are negatively associated with the severity of firm-level information frictions.

Schoellman (2021), but it can also reflect other barriers to collecting information about firm-level fundamentals.

[14]In Appendix A.6, we perform extra robustness checks. We show that our results are robust to controlling for interaction terms between previous forecast errors and firm age. In addition, we run horse race regressions between the time zone difference and GDP per capita, and find results that are similar to those obtained when we include them in the regressions separately.

3 Model

In this section, we develop a dynamic industry equilibrium model with Jovanovic (1982)-type learning embedded as in Arkolakis, Papageorgiou, and Timoshenko (2018). We use this model to rationalize the aforementioned stylized facts and to quantify the role of imperfect information in determining productivity losses in the aggregate using our firm-level data.

3.1 Setup

In our model, time is discrete with periods t = 1,2,..., and the representative consumer spends income Y_t on goods produced by monopolistically competitive firms. Consumer utility from consuming $q_t(\omega)$ units of different products ω can be expressed using the quantity of the following constant elasticity of substitution (CES) aggregate:

$$Q_t = \left(\int_{\omega\in\Omega_t} e^{\frac{\theta(\omega)}{\sigma}} q_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \tag{3}$$

where σ is the elasticity of substitution between different varieties, $\theta(\omega)$ is the demand shifter for variety ω, and Ω_t denotes the set of varieties available at time t. We can express the demand for a particular variety, ω, as:

$$q_t(\omega) = Y_t P_t^{\sigma-1} e^{\theta(\omega)} p_t(\omega)^{-\sigma}, \tag{4}$$

where $P_t \equiv \left(\int_{\omega\in\Omega_t} e^{\theta(\omega)} p_t(\omega)^{1-\sigma} d\omega\right)^{1/(1-\sigma)}$ is the price index of the industry.

The firm-specific demand, $\theta(\omega)$, is unknown to the firm but the firm understands that $\theta(\omega)$ is drawn from a normal distribution $N\left(\bar{\theta}, \sigma_\theta^2\right)$. We assume that the firm cannot fully uncover its permanent demand draw $\theta(\omega)$ from its realized sales, given that it is faced with constraints in collecting and processing information. Instead, the firm receives a *noisy* signal about the permanent demand draw $\theta(\omega)$ and needs to learn about it over the life cycle:

$$s_t(\omega) = \theta(\omega) + \varepsilon_t(\omega), \tag{5}$$

where $\varepsilon_t(\omega)$ is an independent and identically distributed (i.i.d.) noise term drawn from a normal distribution $N\left(0, \sigma_\varepsilon^2\right)$. The noise term can reflect errors in managing and sharing financial data inside the firm, and thus managers are unable to *precisely* determine the implied demand draw $\theta\left(\omega\right)$ from available information, such as realized sales.

At the beginning of each 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: $s_1, s_2, \ldots, s_n$. Because 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:

$$\mu_n = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}\bar{\theta} + \frac{n\sigma_\theta^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}\bar{s}_n, \quad \sigma_n^2 = \frac{\sigma_\varepsilon^2 \sigma_\theta^2}{\sigma_\varepsilon^2 + n\sigma_\theta^2}, \tag{6}$$

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: $\bar{s}_n \equiv \frac{1}{n}\sum_{i=1}^{n} s_i$ for $n \geq 1$ and $\bar{s}_0 \equiv \bar{\theta}$. For age-one firms (i.e., entrants), their belief about the mean and variance of θ is the same as the prior belief, i.e., $\mu_0 = \bar{\theta}, \;\; \sigma_0^2 = \sigma_\theta^2$.

As we will show below, our chosen information structure generates the aforementioned age-declining serially correlated forecast errors about sales (Fact 2).[15] The key is that $\varepsilon_t(\omega)$ is payoff irrelevant, being purely informational and orthogonal to firms' per-period profits. If $\varepsilon_t(\omega)$ is payoff relevant (as in Jovanovic (1982) and Arkolakis et al. (2018)), we show that sales forecast errors are serially *uncorrelated* in Appendix B.6.[16]

Output is linear in labor with $q_t = \varphi_t l_t$ and firms hire workers at the wage rate of w. Firms' labor productivities follow an AR(1) process, where the variance of the shock is age-

[15] Not only a constraint in collecting and processing information but also a lack of knowledge about underlying model structures can lead to serially correlated forecast errors. To make our quantitative decomposition of forecast errors transparent, we incorporate only the former but not the latter. See Ryngaert (2017) for the quantitative importance of each channel for inflation forecasts.

[16] If the firm-specific demand shifter has both permanent and transitory components, then the forecast error in period $t+1$ is independent of its lagged values, as all past forecast errors are linear functions of realized and forecasted demand shifters up to period t, which are in the firm's information set by the end of period t. However, in our model, the realized demand shifter, θ, is never observed, so the forecast error in period $t+1$ is only orthogonal to the forecast made in period $t-1$, but not to the forecast error in period t. Further details are available in Appendix B.6. Alternative shock processes such as learning about a time-varying firm demand θ_t that follows an AR(1) process also imply zero forecast errors (see Appendix B.7).

dependent. Replacing the time subscript with firm age n, we can write the productivity process as

$$\log \varphi_n = (1-\rho)\mu_\varphi + \rho \log \varphi_{n-1} + \nu_n, \quad \nu_n \sim N(0, \sigma^2_{\nu_n}).$$

Following Atkeson and Kehoe (2005), we assume age-dependent volatility. Specifically, we model the decline of σ_{ν_n} using a quadratic function up to an age cutoff. This captures the decline in the variance of forecast errors over the firm's life cycle owing to mechanisms other than learning (e.g., customer accumulation and product diversification). We include this term in our model to tease out the contribution of life-cycle learning to the (total) decline in the variance of forecast errors. Information on autocovariance and variance of forecast errors helps us separately identify the learning mechanism and the age-dependent volatility, as age-dependent productivity shocks do not generate autocovariance.

At t, incumbents receive an exit shock (η) randomly, and surviving firms decide whether to stay in the market by paying a fixed cost (f). Then, firms decide on the number of workers (l_t) before labor productivity (φ_t) is realized. The price (p_t) is set at the end of the period, assuming no storage technology. Next, firms observe new signals (s_t) and update their beliefs.

In each period, a unit mass of potential entrants decides whether to enter the market. They draw a permanent demand shifter (θ) and initial labor productivity (φ_0) from the normal and log-normal distributions, respectively. Potential entrants know the distribution of θ and have perfect information about φ_0. Entrants with a sufficiently high φ_0 choose to enter and produce in the market.

3.2 Static and Dynamic Optimization

In this subsection, we study the firm's static optimization problem. As we focus on firms' behavior in the steady state (i.e., the stationary equilibrium) in what follows, we omit the subscript t whenever possible, and use the age subscript n when necessary. In each period, the firm's output decision is a static choice. Given the belief about θ and φ_n, an age-n

firm hires l_n workers to maximize its expected per-period profit, $E(\pi_n|\varphi_{n-1}, \bar{s}_{n-1}, n)$. The realized per-period profit is $\pi_n = p_n q_n - wl_n - wf$, where $q_n = \varphi_n l_n$ and firms set price p_n to clear the market according to equation (4). Maximizing $E(\pi_t|\varphi_{t-1}, \bar{s}_{n-1}, n)$, the optimal employment is:

$$l_n = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma} \left(\frac{b(\varphi_{n-1}, \bar{s}_{n-1}, n-1)}{w}\right)^{\sigma} YP^{\sigma-1}, \tag{7}$$

where:

$$\begin{aligned} b(\varphi_{n-1}, \bar{s}_{n-1}, n-1) &\equiv E\left(e^{\frac{\theta}{\sigma}} \varphi_n^{\frac{\sigma-1}{\sigma}} |\varphi_{n-1}, \bar{s}_{n-1}, n\right) \\ &= \exp\left\{\frac{\mu_{n-1}}{\sigma} + \frac{\sigma_{n-1}^2}{2\sigma^2} + \frac{\sigma-1}{\sigma}\left((1-\rho)\mu_{\varphi} + \rho\log\varphi_{n-1}\right) + \frac{(\sigma-1)^2\sigma_{\nu_n}^2}{2\sigma^2}\right\}, \end{aligned} \tag{8}$$

and n is the firm's age. We write the resulting price function in Appendix Section B.2, and the expected per-period profit function is:

$$E\pi_n = (\sigma-1)^{\sigma-1}\sigma^{-\sigma} YP^{\sigma-1} \frac{b(\varphi_{n-1}, \bar{s}_{n-1}, n-1)^{\sigma}}{w^{\sigma-1}} - wf. \tag{9}$$

In each period, the potential entrant chooses whether to enter the market and the incumbent firm chooses whether to stay in the market. For an incumbent firm that is $n+1$ years old, its state variables include the labor productivity φ_n, the history of demand signals summarized by $\bar{s}_n$, and its age n in the last period.[17] The incumbent firm's value function (after the random death shock is realized) satisfies:

$$V(\varphi_n, \bar{s}_n, n) = \max\{0, E_n\pi_{n+1} + \beta(1-\eta)E_n V(\varphi_{n+1}, \bar{s}_{n+1}, n+1)\}, \quad n \geq 1. \tag{10}$$

If the firm chooses to exit permanently, it receives a value of zero.

For an entrant that survives the exogenous death shock, its value function has the same format as equation (10) as long as we set $n = 0$. We denote the corresponding policy function as $o(\varphi_n, \bar{s}_n, n)$, which applies to staying or exiting. The definition of equilibrium is contained

[17] As all the distributions are normal (e.g., the demand shifter and the noisy signal), the firm only needs to forecast the mean and variance of the demand shifter. As a result, the average signal of the demand shifter and firm age (which pins down the subjective variance of the demand shifter) are the two state variables that are sufficient to formulate the learning problem faced by the firm.

in Appendix B.2.

4 Decomposing Forecast Errors

In this section, we show how our model matches Facts 1 and 2 presented in Section 2. As will become clear below, learning contributes to both (1) the age-declining variance of forecast errors, and (2) the age-declining autocovariance of forecast errors, while the age-dependent volatility only generates the former. This insight from our model allows us to decompose the variance of forecast errors into learning and age-dependent volatility components. We illustrate the intuitions by using a special case in which the per-period fixed cost is set to zero. In this case, the value of being active in a market is positive for all potential entrants and incumbents. Therefore, all potential entrants enter, and firms do not exit unless they are hit by the exogenous exit shock. We sometimes refer to this case as "no (endogenous) selection."

Proposition 1 *When the per-period fixed cost, f, is set to zero, the forecasts and forecast errors of firm sales have the following properties in the steady state:*

1. *The variance of forecast errors declines with age.*
2. *Forecast errors made in two consecutive periods by the same firm are positively correlated. The positive covariance declines with age.*
3. *The difference between the variance of forecast errors (made at age n) and the autocovariance of forecast errors (made at age $n-1$ and n) has a one-to-one relationship with the (age-dependent) volatility of productivity shocks.*

Proof. See Appendix B.1. ■

Both life-cycle learning and age-dependent volatility contribute to the age-declining variance of forecast errors. First, firms accumulate more experience and thus have clearer information on their permanent demand when they become older, which makes the variance of forecast errors smaller. Second, as we assume that the variance of productivity shocks, σ_{ν_n}, declines with firm age, the variance of forecast errors declines exogenously when the firm becomes older.

The above proposition also rationalizes the finding of the serially correlated forecast errors presented in Section 2.2, as firms adjust their posterior beliefs on the demand shifter *gradually*. In other words, firms incorporate new demand signals partially into their posterior beliefs. As a result, a firm is more likely to under-predict (or over-predict) its next-year sales if it has under-predicted (or over-predicted) its current-year sales. This leads to the positive autocorrelation of forecast errors.[18] Moreover, as a more experienced firm makes smaller forecast errors, the autocovariance of forecast errors declines with years of experience.

In more detail, the reason why *only* the forecast error of the demand shifter is serially correlated is related to the firm's information set. In our model, the realized demand shifter (θ) is never observed by the firm, and only past signals and forecasts are in the firm's information set. Thus, the forecast error of the demand shifter in period $t+1$ (for the forecast made in period t) is orthogonal only to the forecast made in period $t-1$, not orthogonal to *the forecast error in period* t (which equals θ minus the forecast made in period $t-1$). For the forecast error of productivity, both the realized productivity and past forecasts are in the firm's information set. Thus, the forecast error of productivity in period $t+1$ is orthogonal to both the productivity in period t and the forecast made in period $t-1$. As the forecast error in period t is the difference between these two, it is orthogonal to the forecast error in period $t+1$.

Finally, we emphasize that the full-information rational expectation (FIRE) models cannot rationalize the serially and positively correlated forecast errors. In Appendix B.4, we show that FIRE models *without endogenous selection* imply zero autocorrelation in forecast errors. Moreover, in Appendix B.5, we show that FIRE models *with endogenous selection* generate negatively correlated forecasting errors under perfect information with AR(1) type

[18] However, this *does not* mean that firms make biased forecast errors on average. Specifically, positive and negative forecast errors are canceled out over time when we take the average forecast error.

productivity/demand shocks.[19]

4.1 Nonparametric Decomposition

The above proposition illustrates how we can back out the learning parameters (σ_θ and σ_ε) and age-dependent volatility separately by using the panel data of forecast errors. To make the intuitions salient, we assume zero per-period fixed costs (no selection).[20] Under this assumption, the forecast errors of sales at age n are:

$$FE_{n,n+1} \equiv \log \frac{R_{n+1}}{E_n R_{n+1}} = \underbrace{\frac{\theta}{\sigma} - \log E_n(e^{\frac{\theta}{\sigma}})}_{FE^{\theta}_{n,n+1}} + \underbrace{\frac{\sigma-1}{\sigma}\log \varphi_{n+1} - \log E_n(\varphi_{n+1}^{\frac{\sigma-1}{\sigma}})}_{FE^{\varphi}_{n,n+1}}, \tag{11}$$

where the first two terms, denoted by $FE^{\theta}_{n,n+1}$, represent the forecast errors that arise because of the firm's imperfect information about θ. The third and fourth terms, denoted by $FE^{\varphi}_{n,n+1}$, represent the forecast errors that arise from the unpredictable innovation in the firm's AR(1) productivity process. As shown in Appendix B.1, the term $FE^{\varphi}_{n,n+1}$ is linear in the innovation term ν_{n+1}, which is uncorrelated with $FE^{\varphi}_{n-1,n}$ (linear in ν_n). By contrast, the term $FE^{\theta}_{n,n+1}$ is serially correlated because firms never observe θ and gradually update their beliefs about θ with noisy signals. The calculation shows that the covariance and variance of $FE_{n,n+1}$ are:

$$Cov(FE_{n-1,n}, FE_{n,n+1}) = \frac{\sigma_n^2}{\sigma^2}; \quad Var(FE_{n,n+1}) = \frac{\sigma_n^2}{\sigma^2} + \frac{(\sigma-1)^2\sigma^2_{\nu_n}}{\sigma^2}, \tag{12}$$

where σ_n^2 is the perceived variance of the demand shifter of age-$n+1$ firms and σ is the elasticity of substitution.

We can perform a nonparametric decomposition of $Var(FE_{n,n+1})$ into the learning component and the age-dependent volatility component using the two formulas together. Specif-

[19] In this case, we compute autocorrelations of forecast errors for firms that have survived for at least two consecutive periods. If a firm receives a more positive productivity shock in period t, it can afford a more negative productivity in period $t+1$ and remain in the market. Selection leads to negatively correlated productivity shocks in periods t and $t+1$ conditional on firm survival. As forecast errors come from the unpredictable productivity shocks, the forecast errors in two consecutive periods are also negatively correlated.

[20] In Section 5, we show that these moments continue to tightly pin down the parameters related to learning and volatility effects when per-period fixed costs are strictly positive.

ically, the covariance of forecast errors is only related to learning, as age-dependent volatility does not enter into the expressions. When we take the difference between the variance and the autocovariance of forecast errors, the only term that is left is the (age-dependent) variance of the firm's productivity shocks (multiplied by a constant):

$$Var(FE_{n,n+1}) - Cov(FE_{n-1,n}, FE_{n,n+1}) = \frac{(\sigma-1)^2\sigma^2_{\nu_n}}{\sigma^2}. \quad (13)$$

Note that our decomposition is "nonparametric" in the sense that we do not impose any structure on σ_{ν_n}.

Table 3: How Learning and Age-Dependent Volatility Contribute to the Declining Variance of Forecast Errors

	(1)	(2)	(3)	(4)	(5)
Age n	$Var(FE_n)$	$Cov(FE_{n-1}, FE_n)$	$\frac{Cov(FE_{n-1},FE_n)}{Var(FE_n)}$	$\frac{Cov(FE_{n-1},FE_n)-Cov(FE_1,FE_2)}{Var(FE_n)-Var(FE_2)}$	$Var(FE_n) - Cov(FE_{n-1}, FE_n)$
1	0.242	–	–	–	–
2	0.174	0.034	19.8%	–	0.139
3	0.135	0.019	14.5%	38.3%	0.115
4	0.110	0.020	18.5%	22.0%	0.089
5	0.098	0.013	12.9%	28.8%	0.086
6	0.097	0.014	14.5%	26.5%	0.083
7	0.088	0.014	16.0%	23.7%	0.074
8	0.087	0.008	9.1%	30.5%	0.079
9	0.081	0.009	10.9%	27.6%	0.072
10	0.069	0.008	11.9%	25.0%	0.061
11	0.069	0.008	11.3%	25.4%	0.061

Notes: We have simplified the notation in this table so that $FE_{n,n+1} \equiv FE_n$. Columns 1 and 2 report the variance and covariance of the log forecast errors of firms at different ages in our data. Column 3 reports the ratio, $\frac{Cov(FE_{n-1},FE_n)}{Var(FE_n)}$, in percentage terms. The ratio indicates how much the covariance component (driven by learning) contributes to the level of $Var(FE_n)$. Column (4) reports the share contributed by the reduction in $Cov(FE_{n-1}, FE_n)$ in the overall reduction in $Var(FE_n)$. Mathematically, it equals $\frac{Cov(FE_{n-1},FE_n)-Cov(FE_1,FE_2)}{Var(FE_n)-Var(FE_2)}$. Column (5) reports the difference between Columns 1 and 2. According to the equation (13), this term is driven by age-dependent volatility and equals $\frac{(\sigma-1)^2\sigma^2_{\nu_{n+1}}}{\sigma^2}$. Note that all these decompositions are made under the assumption that the fixed cost is zero (no selection). There are empty cells (indicated by dashes) because we do not observe firms' sales expectations upon their entry, i.e., $E_0(R_1)$, and $FE_{0,1}$ cannot be measured from the data.

Following this logic, we use Table 3 to implement the decomposition exercise. Columns 1 and 2 of the table indicate the variance and covariance of forecast errors at age n in the data, whereas Column 5 is the difference between the two, capturing age-dependent volatility. In terms of levels, in general, the learning component (covariance terms) is small, explaining about 10% to 20% of the variance of the forecast errors (see Column 3, the ratio of Column 2 to Column 1). However, they have a larger contribution to the *change* in the variance of forecast errors over the firm's life cycle, ranging between 20% to 40% (Column 4). This is because the variance of shocks to labor productivity does not diminish to zero when firms are sufficiently old, which levels up the overall variance of forecast errors and makes the ratios in Column 3 small. In summary, both learning and age-dependent volatility are important to account for the life-cycle dynamics of firms' forecast errors.

5 Quantitative Analysis

In this section, we quantitatively assess the aggregate implications of imperfect information. In contrast with Section 4, we now allow for selection at entry, which renders it infeasible to derive a sharp mapping from structural parameters to autocovariance and variances of forecast errors. We calibrate the full model in Section 5.1 and analyze the gains from information improvements in Sections 5.2 to 5.4. We find that selection amplifies the gains from removing information frictions.

5.1 Calibration

We use data moments taken from the foreign activities survey. We normalize the aggregate demand shifter Y and wage rate w to one and the mean of the logarithm of the permanent demand $\bar{\theta}$ to zero. We set the elasticity of substitution between varieties σ to four and the discount factor β to 0.96 (assuming a real interest rate of 4% per annum).[21] The exogenous death rate η is set to 0.03 to match the exit rate of the largest 5% of firms above age 10. We impose an age threshold to avoid considering learning and age-dependent volatility for these firms, and only extremely negative shocks to labor productivity and the exogenous death shock lead to exits (Panel A of Table 4).

In our calibration, learning is parameterized by the two parameters, σ_θ and σ_ε. Guided by the decomposition exercise in Section 4.1, two natural candidate moments are the covariance

[21] The gains from information increase in σ as in David et al. (2016), who set $\sigma = 6$ as their baseline value but show more robustness checks with two other elasticity values 4 and 10. Our results are thus more comparable to their results under a conservative choice of an elasticity.

Table 4: Parameters Calibrated Without/By Solving the Model

Parameters	Value	Description	Source/Target	Moments	
				Data	Model
Panel A: Calibrated without solving the model					
σ	4	elasticity of substitution between different varieties	Bernard et al. (2003)		
β	0.96	discount factor	4% real interest rate		
η	0.03	exogenous death rate	exit rate of the largest 5% of firms above age 10		
Panel B: Calibrated by solving the model and matching moments					
f_m	0.0093	fixed cost	average exit rate of incumbents	0.093	0.093
σ_θ	0.96	std of θ	$Cov(FE_{t-1}, FE_t)$ at age one	0.034	0.034
σ_ε	1.36	std of ε	$Cov(FE_{t-1}, FE_t)$ above age ten	0.008	0.008
κ_0	0.33	$\sigma_{\nu_n} = \kappa_0 + \kappa_1(1 - n/10)^2$	Var(FE) above age ten	0.069	0.069
κ_1	0.28	$\sigma_{\nu_n} = \kappa_0 + \kappa_1(1 - n/10)^2$	Var(FE) above at age one	0.242	0.241
ρ	0.67	persistence in productivity	$\frac{Var[\log(\check{A}_{n+1}/\check{A}_{n-1})]}{Var[\log(\check{A}_{n+1}/\check{A}_n)]} - 1$	0.664	0.666

of fixed effects for the youngest firms and the oldest firms. Loosely speaking, conditional on other parameters, we calibrate σ_θ and σ_ε so that the model can match the autocovariance of fixed effects at ages one and two, and the autocovariance of fixed effects above age 10.

Learning contributes to the age-declining variance of fixed effects but only partially, as discussed in Section 4.1. We let the age-dependent volatility reproduce the rest of the age-declining variance of fixed effects. Following Atkeson and Kehoe (2005), we parameterize the age-dependent volatility using a quadratic function:

$$\sigma_{\nu_n} = \begin{cases} \kappa_0 + \kappa_1 \left(\frac{10-n}{10}\right)^2 & \text{if } n < 10 \\ \kappa_0 & \text{if } n \geq 10. \end{cases}$$

Therefore, σ_{ν_n} starts from a value of $\kappa_0 + \kappa_1$, then drops to and stays at κ_0 after age 10. We calibrate the two parameters so that the model can match the variance of forecast errors above age 10 and the variance of forecast errors at age one.[22]

We are left with the choices for the two other remaining parameters: the per-period

[22] When mapping the model to the data, we use a mix of age-one and age-two firms to mimic age-one firms in the data. Firms established in any month of the current fiscal year are considered age-one firms in the data. Late entrants with little information about θ make their predictions in the same manner as an age-one firm in the model, whereas early entrants behave like an age-two firm. We match the variance and covariance of fixed effects of age-one firms in the data by using a mix of age-one and age-two firms in the model, with shares close to 50% each, with the latter being slightly smaller due to exits. We use the same strategy for other firm ages.

fixed cost, f, and the AR(1) coefficient of the labor productivity process, ρ. For the former, we target the average exit rate of incumbent firms. For the latter, we first compute the "adjusted labor productivity" as:

$$\log \breve{A}_n = \log R_n - \frac{\sigma - 1}{\sigma} \log l_n = \frac{\theta}{\sigma} + \frac{\sigma - 1}{\sigma} \log \varphi_n + \frac{1}{\sigma} \log(Y) + \frac{\sigma - 1}{\sigma} \log(P),$$

where θ is firm-specific but time-invariant and Y and P are aggregate variables that do not vary across firms. The coefficient before $\log l_n$ is important—with this adjustment, the term related to expectation, $b(\varphi_{n-1}, \bar{s}_{n-1}, n-1)$, drops out from the labor productivity measure. Then, to calibrate ρ, we use the following data moment:

$$\frac{Var[\log(\breve{A}_{n+1}/\breve{A}_{n-1})]}{Var[\log(\breve{A}_{n+1}/\breve{A}_n)]} - 1, \quad n \geq 10. \tag{14}$$

[23] Note that without selection, this formula provides an unbiased estimate for the persistence parameter in a stationary AR(1) process, even in small samples (Lo and MacKinlay, 1988). In our modified setting, taking the one- and two-period differences in $\breve{A}_n$ removes the permanent demand shock θ. In addition, focusing on old firms ensures that σ_{ν_n} is constant, and we can apply the same argument as in Lo and MacKinlay (1988). Endogenous selection breaks the one-to-one mapping between this moment and ρ. However, we find that selection creates a very small bias, and that this moment tightly pins down ρ.

In Panel B of Table 4, we list the parameters and moments in an order such that, loosely, the moment provides the most information on the parameter in the same row. All moments are matched precisely. The calibrated σ_θ and σ_ε are 0.96 and 1.36, respectively, implying a signal-to-noise ratio of 0.50. We find the value of ρ to be 0.67, very close to the data counterpart of equation (14).

We show that the calibrated model closely matches the evolution of the variance and covariance of forecast errors over firms' life cycles, despite the fact that we are only targeting these moments at age one and above age 10. We show that our model also captures the

[23]As P and Y do not vary across firms, they drop out from the variance. Moreover, θ drops out from the difference in the logarithm of "adjusted labor productivity," as it is time-invariant.

increase in average firm sales and the decline in the standard deviation of firm growth as firms become older. We refer the reader to Appendix C.1 for details.

5.2 Comparative Statics: Intensive and Extensive Margins

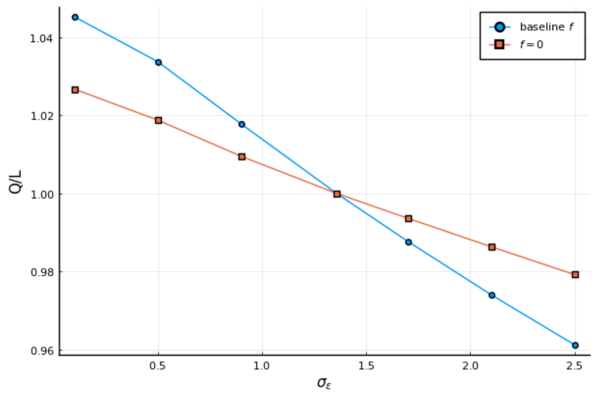
We first consider a change in the information environment by changing the value of σ_ε, holding other parameters fixed at the values described above. Our baseline σ_ε is 1.36, and we vary it between 0.10 and 2.50, with the highest value corresponding to the region with the highest σ_ε, as we show in our by-region calibration in Section 5.4. In addition, we consider a case where information about θ is perfect in that entrants know the true value of θ.

Figure 3 plots the impact of information frictions on aggregate outcomes. We compare our baseline model to a version of our model without selection. In Figure 3, the blue curves with dots summarize the comparative statistics with respect to σ_ε in the baseline model. The red curves with square markers indicate the same comparative statistics with respect to σ_ε in the model where we set the per-period fixed costs f to zero. In both models, the price index increases with σ_ε (top left panel), whereas labor productivity decreases with σ_ε (top right panel), with the slope being steeper in the baseline model.

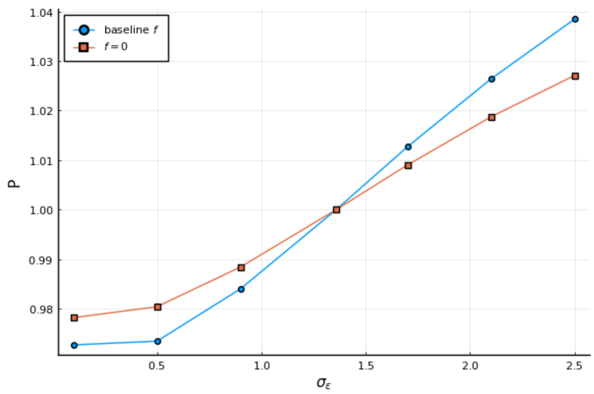
These productivity losses stem from the effects that operate through both intensive and extensive margins. For the intensive margin, we show that the correlation between firm capability ($\log \phi \equiv (\sigma - 1) \log \varphi + \theta$) and production scale ($\log b$) decreases with σ_ε (middle left panel).[24] This is because more severe informational imperfections tend to make firms with low demand θ produce too much, and vice versa for firms with high demand. Imprecise knowledge about demand θ makes output choice far from the optimal level at the intensive margin, which can be seen by the fact that the correlation between the true demand θ and the average of past noisy signals $\bar{s}$ decreases with σ_ε (middle right panel).

[24] Firm capability term $\log \phi$ is a combination of firm labor productivity φ and its permanent demand shifter θ. Scaling $\log \varphi$ by the coefficient $\sigma - 1$ ensures that this term solely determines firm-level output in a perfect information static model. In our dynamic imperfect information model, b is the only firm-level variable that determines expected profit (see equations (8) and (9)). In a perfect information static model, $\log b$ is linear in $\log \phi$ and thus the correlation is one.

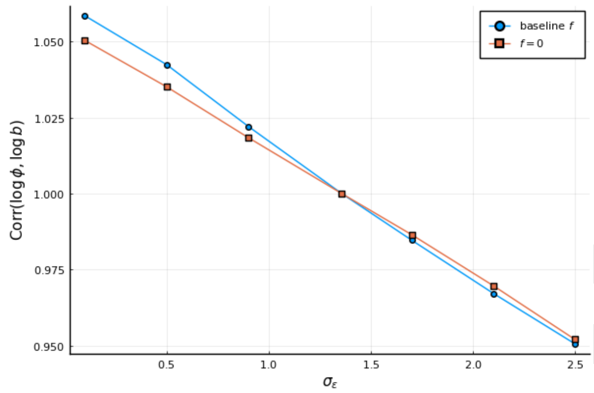
Figure 3: The Impact of σ_ε on Aggregate Outcomes



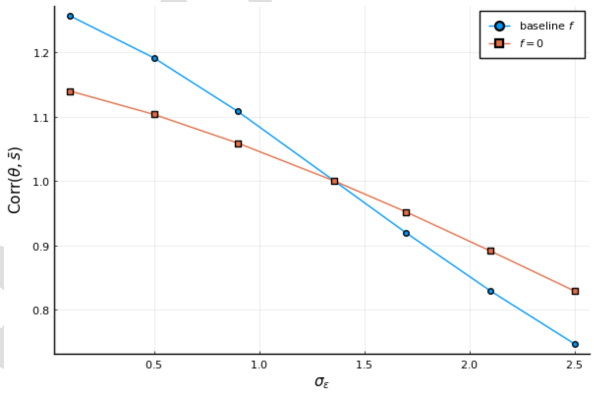
(a) Q/L



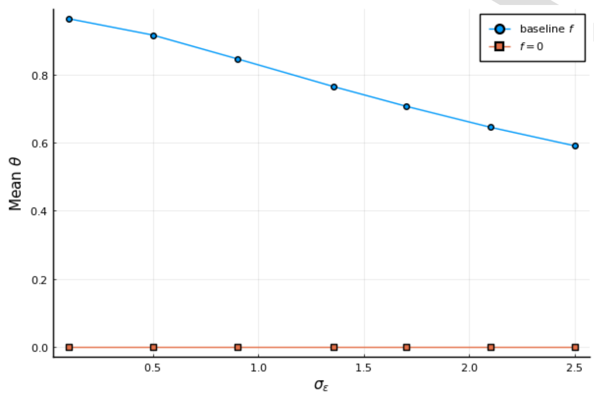
(b) P



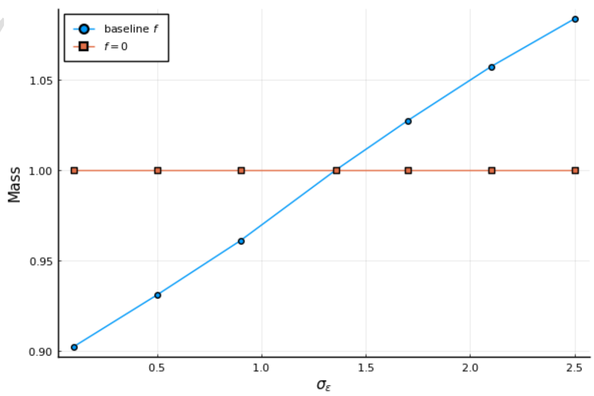
(c) Corr ($\log \phi$, $\log b$)



(d) Corr (θ, $\bar{s}$)



(e) Average Incumbent θ



(f) Relative Mass of Firms: M

Notes: Panels (a) to (f) show the aggregate industry price index, labor productivity, the correlation between $\log(\phi)$ and $\log(b)$ among incumbents, the correlation between θ and $\bar{s}$ among incumbents, the average incumbent's θ, and the equilibrium mass of firms under different fixed costs f and different values of σ_ε. All variables are normalized to one in the case that $\sigma_\varepsilon = 1.36$, the calibrated value in our baseline model. $\log \phi$ is defined as the combination of labor productivity and demand, $(\sigma - 1)\log \varphi + \theta$, which determines the size of the firm in a static model. b is defined as in equation (8). The blue dotted line indicates a model with fixed costs f at the value in the baseline calibration, whereas the red line with squares indicates a model with zero fixed costs (no selection).

For the extensive margin, the average demand shifter (of active firms) decreases with σ_ε (bottom left). More severe information imperfection causes firms with low demand but high values for noises to enter (and stay), whereas firms with high demand but low values for noises exit. This selection effect is similar to the one studied in Sager and Timoshenko (2021). Active firm mass increases with σ_ε (bottom right), as less tough competition (induced by more severe information imperfection) makes more firms stay. These effects are absent without selection (red curves with square markers). In this alternative model, the price index and labor productivity are only affected through the intensive margin, as all potential entrants (other than those that exit exogenously) are active in production.

Table 5: Aggregate Outcomes under Different σ_ε

Panel A: $f = 0.0093$ (benchmark)	(1)	(2)	(3)
Statistics	High Info. Friction $\sigma_\varepsilon = 2.50$	Baseline Info. Friction $\sigma_\varepsilon = 1.36$	Perfect Info.
Mass of Active Firms	11.224	10.359	9.046
Incumbents Average θ	0.591	0.764	1.046
Incumbents Average $\theta + (\sigma - 1) \log \varphi$	0.187	0.231	0.315
Q/L	3.482	3.623	3.853
Δ% Q/L	-3.88		6.36
Panel B: $f = 0$	(1)	(2)	(3)
Statistics	High Info. Friction $\sigma_\varepsilon = 2.50$	Baseline Info. Friction $\sigma_\varepsilon = 1.36$	Perfect Info.
Mass of Active Firms	32.333	32.333	32.333
Incumbents Average θ	0	0	0
Incumbents Average $\theta + (\sigma - 1) \log \varphi$	0	0	0
Q/L	4.528	4.624	4.794
Δ% Q/L	-2.08		3.66

Notes: This table reports the equilibrium outcomes under a high level of information frictions ($\sigma_\varepsilon = 2.50$), the baseline model ($\sigma_\varepsilon = 1.36$), and perfect information, with different values of fixed costs (baseline value, 0.0093, and alternative value, 0). As explained in footnote 24, the term $\theta + (\sigma - 1) \log \varphi$ can be interpreted as "firm capability," which uniquely determines a firm's size in a perfect information static model.

Table 5 shows the quantitative implications and highlights the role of selection. Labor productivity increases by 6.35% in our baseline model with selection, whereas it increases by 3.66% in the alternative model where the extensive margin does not play a role. Our comparative statistics show not only a substantial gain in overall productivity from eliminating the informational frictions over the firm's life cycle, but also the role of firm entry and exit in driving it.

5.3 Heterogeneous Effects Across Different Age Groups of Firms

One feature of our model is the gradual resolution of uncertainty over the life cycle of firms. Entrants and young firms face more severe informational imperfections and learn the true values of their demand shifters over time, while deciding in each period whether to stay or exit from the market. We proceed with the analysis to see how firms in different age groups are affected differently by the elimination of the information frictions and how much each age group's productivity change contributes to the overall productivity gains in the economy.

Consistent with the expression of aggregate output in equation (3), we define the average productivity of age-n ($n \geq 1$) firms as:

$$A_n \equiv \frac{Q_n}{L_n^{prod}} = \frac{\left(\int_{\omega \in \Omega_n} e^{\frac{\theta(\omega)}{\sigma}} q_n(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}}{L_n^{prod}}, \tag{15}$$

where L_n^{prod} is the number of workers used in production of all age-n firms, excluding workers used to pay for the fixed cost. Ω_n is the set of active age-n firms and $q_n(\omega)$ is the output of the firm that produces variety ω. Note that entrants are age-one firms, whereas incumbents are older than one. We define the average productivity of firms of all ages as:

$$A \equiv \frac{\left(\sum_n \int_{\omega \in \Omega_n} e^{\frac{\theta(\omega)}{\sigma}} q_n(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}}{L^{prod}} = \left(\sum_{n=1}^{N} A_n^{\frac{\sigma-1}{\sigma}} \left(\frac{\bar{L}_n^{prod} M_n}{\bar{L}^{prod} \sum_{n=1}^{N} M_n}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},$$

where $L^{prod} = \sum_{n=1}^{N} L_i^{prod}$, and N is the maximum age that we consider in the simulation. In addition, $\bar{L}_n^{prod}$ is the average employment of production workers of age-n firms, and $\bar{L}^{prod}$ is the average employment of production workers of all firms. $M_n \equiv \int_{\omega \in \Omega_n} d\omega$ is the measure of age-n firms that are active. Then, we define the normalized productivity $\tilde{A}_n = A_n M_n^{\frac{1}{1-\sigma}}$. Note that the difference between $\tilde{A}_n$ and A_n is that the former does not take into account the variety effect, reflected by the number of active firms in our model.

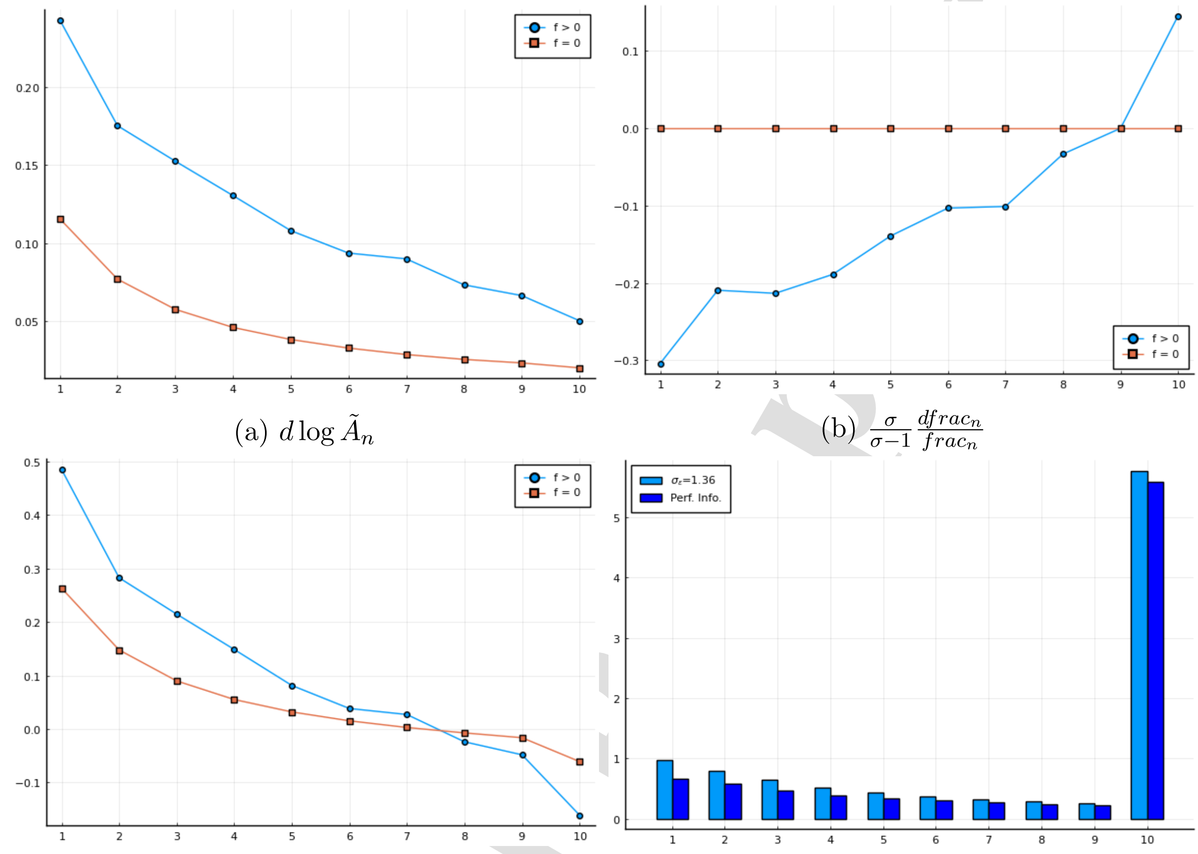
Finally, the log (or percentage) change in average labor productivity can be decomposed as:

$$\frac{dA}{A} = \sum_{n=1}^{N} \left[contri_n \left(d\log(\tilde{A}_n) + \frac{\sigma}{\sigma-1}\frac{dfrac_n}{frac_n} + d\log\left(\frac{\bar{L}_n^{prod}}{\bar{L}^{prod}}\right)\right)\right] + \frac{1}{\sigma-1}\frac{dM}{M}, \tag{16}$$

where the weight is defined as $contri_n \equiv frac_n \tilde{A}_n^{\frac{\sigma-1}{\sigma}} \left(\frac{\bar{L}_n^{prod}}{\bar{L}^{prod}}\right)^{\frac{\sigma-1}{\sigma}} / \sum_{n=1}^{N} frac_n \tilde{A}_n^{\frac{\sigma-1}{\sigma}} \left(\frac{\bar{L}_n^{prod}}{\bar{L}^{prod}}\right)^{\frac{\sigma-1}{\sigma}}$ and $frac_n$ is the fraction of active firms that are n years old among all active firms. The total mass of active firms is simply denoted by $M = \sum_{n=1}^{N} M_n$.

There are four terms related to the change in average productivity in equation (16). First, the term $d\log(\tilde{A}_n)$ is the change in normalized productivity for each age group. Second, $\frac{\sigma}{\sigma-1}\frac{dfrac_n}{frac_n}$ reflects the change in population shares for different age groups. Third, $d\log\left(\frac{\bar{L}_n^{prod}}{\bar{L}^{prod}}\right)$ is the change of the average size of age-n firms (relative to the overall mean). The final term, $\frac{1}{\sigma-1}\frac{dM}{M}$, reflects the variety effect. Figure 4 plots these terms when we move from our baseline level of imperfect information (with $\sigma_\varepsilon = 1.36$) to perfect information wherein all entrants know the true value of θ. In Figure 4, the blue curves with dots show the results for our baseline model, where only some firms enter and stay active, whereas the red curves with square markers indicate the results for an alternative model without selection, in which the per-period fixed costs f are set to zero.

Figure 4: Decomposing the Impact of σ_ε Across Age Groups: $\sigma_\varepsilon = 1.36 \rightarrow$ Perfect Information



(a) $d \log \tilde{A}_n$ (b) $\frac{\sigma}{\sigma-1} \frac{dfrac_n}{frac_n}$

(c) $d \log(\bar{L}_n^{prod}/\bar{L}^{prod})$ (d) Firm Mass

Notes: Panels (a) to (c) plot the three key components in the change in normalized industry labor productivity according to equation (16), contributed by firms of different ages n (capped at 10 years), when changing the model from the baseline of imperfect information ($\sigma_\varepsilon = 1.36$) to a dynamic model in which firms have perfect information about θ. The blue dotted line represents the case in which the fixed costs f are kept at the baseline value, 0.0093. The red line with squares represents the case where $f = 0$, i.e., the case without endogenous selection. Panel (d) shows the mass of firms at different ages in the imperfect and perfect information model, respectively.

Panel (a) shows that the (normalized) productivity gains are larger among young firms than old firms in both our models, but more so in our baseline model with selection (blue curves with dots). In Panel (b), the population shares of young and old firms change significantly in the baseline model but not in the alternative model (red curves with square markers). Selection becomes tougher when the information frictions become less severe, and this leads to a "better" selected group of firms operating in the economy. We discussed this in the previous section, but we now observe that this extensive margin effect operates more prominently among young firms. Note that selection is more relevant for young firms (given the information environment), as the exit rate flattens (with respect to age) for sufficiently old firms. Therefore, this *age-specific* selection effect makes the (normalized) age group-specific productivities increase and the population shares decrease more for young firms than old firms. These findings highlight the importance of post-entry selection especially among young firms. Relatedly, the average size of firms increases for young firms but decreases for old firms in Panel (c). Panel (d) shows a decline in the mass of firms for each age group after σ_ε declines, but only in the baseline model with selection.

5.4 Cross-Regional Analysis

As suggested by Fact 3 in Section 2.2, firms may face different levels of information frictions due to varying management practices and communication barriers in different countries. We use our model to quantify the degree of information frictions across countries/regions and demonstrate the potential gains from eliminating them. This allows us to use economies with smaller information frictions instead of an economy with perfect information as the benchmark, following a long tradition in the misallocation literature (Hsieh and Klenow (2009); David et al. (2016)).

We use data from eight major regions/countries of the world: Africa, the Middle East, Latin America, Eastern Europe, the Association of Southeast Asian Nations (ASEAN) countries, China, Western Europe, and the United States.[25] Similar to the baseline calibration, we target the covariance and variance of the youngest and oldest firms, together with the incumbent exit rates in each region. Panel A of Table 6 presents the calibrated parameters by region and the corresponding model moments. Each set of parameters enables us to precisely match the data moments, so we omit them from the table to save space, and report them in Appendix C-27. We find that Africa, the Middle East, Latin America, and

[25]These eight regions do not exhaust all foreign countries in which Japanese multinational firms operate, but they cover the majority of firm sales in our data. In addition, they display significant differences in income levels and business environments, and contain countries that are relatively homogeneous within each region. Appendix Table C-26 provides the full list of countries in each region.

Eastern Europe have higher values of σ_ε than the other regions, which are driven by their higher covariance of forecast errors targeted in the calibration. Firms in Latin America and Eastern Europe are revealed to have higher values of σ_θ than the other regions.

Table 6: Calibration and Gains from Eliminating Information Frictions by Region

	Parameters						Model Moments					Info. Gains
Region	σ_θ	σ_ε	$\sigma_\theta^2/\sigma_\varepsilon^2$	σ_{ν_1}	$\sigma_{\nu_{10}}$	f	Cov_1	Cov_{10}	Var_1	Var_{10}	exit rate	%Δ Q/L
Panel A: Change five parameters												
Africa	0.86	2.57	0.11	0.51	0.37	0.0152	0.040	0.020	0.186	0.100	0.105	12.56
Middle East	0.83	2.64	0.10	0.58	0.45	0.0142	0.038	0.019	0.226	0.134	0.102	12.16
Eastern Europe	1.41	1.80	0.62	0.58	0.32	0.0079	0.068	0.014	0.283	0.072	0.101	9.44
Latin America	1.62	1.66	0.95	0.39	0.39	0.0070	0.073	0.013	0.218	0.097	0.103	7.13
ASEAN	0.44	1.61	0.08	0.70	0.34	0.0074	0.011	0.006	0.264	0.073	0.078	3.68
China	1.12	1.48	0.57	0.64	0.31	0.0074	0.044	0.010	0.276	0.065	0.089	7.16
Western Europe	0.91	1.47	0.39	0.50	0.31	0.0131	0.034	0.009	0.179	0.065	0.106	6.95
United States	0.78	1.49	0.27	0.52	0.31	0.0147	0.028	0.009	0.180	0.063	0.110	7.11
Panel B: Change four parameters, fix f												
Africa	0.86	2.57	0.11	0.51	0.37	0.0093	0.039	0.020	0.184	0.098	0.073	9.77
Middle East	0.83	2.64	0.10	0.58	0.45	0.0093	0.037	0.018	0.225	0.130	0.074	9.83
Eastern Europe	1.41	1.80	0.62	0.58	0.32	0.0093	0.068	0.015	0.281	0.072	0.112	9.86
Latin America	1.62	1.66	0.95	0.39	0.39	0.0093	0.071	0.014	0.219	0.098	0.125	7.84
ASEAN	0.44	1.61	0.08	0.70	0.34	0.0093	0.012	0.006	0.268	0.072	0.103	4.04
China	1.12	1.48	0.57	0.64	0.31	0.0093	0.042	0.010	0.280	0.065	0.104	7.70
Western Europe	0.91	1.47	0.39	0.50	0.31	0.0093	0.034	0.010	0.177	0.066	0.081	5.93
United States	0.78	1.49	0.27	0.52	0.31	0.0093	0.029	0.008	0.177	0.063	0.077	5.55

Notes: Panel (A) shows the results when we re-calibrate five parameters for each region $(\sigma_\theta, \sigma_\varepsilon, \kappa_1, \kappa_0, f)$. We present age-dependent volatility $\sigma_{\nu_1}, \sigma_{\nu_{10}}$ instead of κ_1, κ_0 to facilitate interpretation. We target five moments in this calibration, $Cov(FE_{n-1,n}, FE_{n,n+1})$ for $n = 1$ and $n \geq 10$, $Var(FE_{n,n+1})$ for $n = 1$ and $n \geq 10$ and incumbent exit rates. $\%\Delta Q/L$ is the percentage change in labor productivity when we change the model from the calibrated imperfect information case to perfect information. Panel (B) reports the results when we re-calibrate the learning and uncertainty related parameters but keep the fixed costs at the baseline value of $f = 0.0093$. We target the first four moments but do not attempt to match the exit rates in the data. The model matches the data moments well (other than the untargeted exit rates in Panel B). To save space, we report the data moments in Appendix Table C-27. A full list of countries in each region can be found in Appendix Table C-26.

We use calibrated economies to assess productivity gains from eliminating informational imperfection. Moving from the calibrated economy to perfect information, $\sigma_\varepsilon = 0$, we report the increase in labor productivity in percentage terms in the last column of Panel A of Table 6. Regions with a larger σ_ε and σ_θ tend to have larger gains. A high σ_ε leads to noisier signals and potentially more misallocation at both the intensive and extensive margins.[26] A higher σ_θ increases the benefit of eliminating the information friction, as there is much more to learn over the life cycle. For instance, Africa and the Middle East feature the noisiest

[26] Firms in Latin America and Eastern Europe have higher values of σ_θ than the other regions, and their signal-to-noise ratios are the highest among the eight regions. This is broadly consistent with the view that firms acquire information optimally by paying a cost, which makes σ_ε (or equivalently, the signal-to-noise ratio, $\sigma_\theta^2/\sigma_\varepsilon^2$) endogenous to the level of σ_θ (see Sims (2003); Luo (2008); Mackowiak and Wiederholt (2009)).

signals, and their gains from eliminating information frictions are as large as 12.56% and 12.16%, respectively. ASEAN countries and China have low gains owing to their low σ_θ and σ_ε. This makes sense, as ASEAN countries and China are close to Japan geographically, which facilitates their Japanese affiliates' communication with the parent firms and reduces the forecast errors. Latin America and China have lower fixed costs, reducing their efficiency losses due to extensive margin misallocation.[27]

Overall, we show that the degree of imperfect information and the associated aggregate implications vary across regions/countries. For example, the productivity loss due to information frictions in Africa is almost twice as large as that in Western Europe (5.6 percentage points larger). Our results align with the notion that firms in developing economies face more information frictions, which are likely to arise from poorer management practices, but they also suggest that communication barriers, such as time zone differences, may lead to information frictions.[28] The first finding is in line with David et al. (2016), although learning in our model is dynamic rather than static and our results are based on a broader set of countries. The second finding is consistent with previous studies showing that time zone differences become barriers to international business (in a different context, see Gumpert (2018)). Our findings support the argument for improving management practices or communication efficiency (e.g., having more nonstop flights between cities). As highlighted by Hsieh and Rossi-Hansberg (2023), communication barriers exist even within a country such as the United States, and the reduction of such barriers can improve productivities.

[27] As discussed in Section 5.2, a lower fixed cost reduces the efficiency loss due to the extensive margin misallocation. Indeed, as is reported in Panel B of Table 6, when we keep the fixed cost at the baseline level for all regions, the gains from eliminating informational imperfections in Latin America and China increase, becoming about 1.9% to 2.2% higher than the gains in Western Europe and the United States, instead of being 0.05% to 0.2% higher in Panel A.

[28] As our findings and associated productivity gains are based on a Japanese firm sample, we caution against generalization. Domestic firms may face fewer information constraints because of lower communication barriers, whereas foreign-owned firms may have better forecasting accuracy owing to superior management practices, as extensively documented in the literature (Bloom and Van Reenen, 2007, 2010).

6 Conclusion

We analyze firm-level panel data on sales forecasts to identify imperfect information and its gradual resolution. The variance of forecast errors decreases with firms' experience, and the covariance of forecast errors is tightly linked to learning. We develop a model of heterogeneous firms' learning about their demand over their life cycle and show that learning contributes to a 20%–40% decline in the variance of forecast errors. We use this model to calibrate our cross-country data and measure potential gains from eliminating imperfect information. We believe there are at least two avenues for future research. First, causal evidence on how an improved information environment affects firm entry and exit would strengthen our understanding of the effects of information frictions on resource allocation. Second, given that the literature has identified different types of information frictions (static, life-cycle, and dynamic information frictions of maturing firms, such as rational inattention), it is crucial to propose a unified framework that can be used to quantify multiple sources of information frictions jointly, as different types of information frictions can have different policy implications.

References

ALTIG, D., J. M. BARRERO, N. BLOOM, S. J. DAVIS, B. MEYER, AND N. PARKER (2020): "Surveying business uncertainty," *Journal of Econometrics*.

ARKOLAKIS, C., T. PAPAGEORGIOU, AND O. A. TIMOSHENKO (2018): "Firm learning and growth," *Review of Economic Dynamics*, 27, 146–168.

ATKESON, A. AND P. J. KEHOE (2005): "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 113, 1026–1053.

BACHMANN, R., K. CARSTENSEN, S. LAUTENBACHER, AND M. SCHNEIDER (2021): "Uncertainty and change: survey evidence of firms' subjective beliefs," Tech. rep., National Bureau of Economic Research.

BAHAR, D. (2020): "The hardships of long distance relationships: time zone proximity and the location of MNC's knowledge-intensive activities," *Journal of International Economics*, 125, 103311.

BAKER, S. R., T. S. MCELROY, AND X. S. SHENG (2020): "Expectation formation following large, unexpected shocks," *Review of Economics and Statistics*, 102, 287–303.

BALEY, I. AND A. BLANCO (2019): "Firm Uncertainty Cycles and the Propagation of Nominal Shocks," *American Economic Journal: Macroeconomics*, 11, 276–337.

Baley, I., A. Figueiredo, and R. Ulbricht (2022): "Mismatch Cycles," *Journal of Political Economy.*

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

Bloom, N., S. J. Davis, L. Foster, B. Lucking, S. Ohlmacher, and I. Saporta-Eksten (2020): "Business-Level Expectations and Uncertainty," Working Paper 28259, National Bureau of Economic Research.

Bloom, N., T. Kawakubo, C. Meng, P. Mizen, R. Riley, T. Senga, and J. Van Reenen (2021): "Do Well Managed Firms Make Better Forecasts?" Working Paper 29591, National Bureau of Economic Research.

Bloom, N., R. Lemos, R. Sadun, D. Scur, and J. Van Reenen (2014): "JEEA-FBBVA Lecture 2013: The new empirical economics of management," *Journal of the European Economic Association*, 12, 835–876.

Bloom, N. and J. Van Reenen (2007): "Measuring and explaining management practices across firms and countries," *The quarterly journal of Economics*, 122, 1351–1408.

——— (2010): "Why do management practices differ across firms and countries?" *Journal of economic perspectives*, 24, 203–24.

Born, B., Z. Enders, G. J. Müller, and K. Niemann (2022): "Firm expectations about production and prices: Facts, determinants, and effects," *Handbook of Economic Expectations.*

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

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

Coibion, O., Y. Gorodnichenko, and S. Kumar (2015): "How Do Firms Form Their Expectations? New Survey Evidence," NBER Working Paper.

——— (2018): "How Do Firms Form Their Expectations? New Survey Evidence," *American Economic Review*, 108, 2671–2713.

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

David, J. M. and V. Venkateswaran (2018): "The Sources of Capital Misallocation," Working Paper.

Enders, Z., F. Hünnekes, and G. J. Müller (2022): "Firm expectations and economic activity," *Journal of the European Economic Association.*

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

Gravelle, J. G. (2009): "Tax Havens: International Tax Avoidance and Evasion," *National Tax Journal*, 62, 727–753.

Gumpert, A. (2018): "The Organization of Knowledge in Multinational Firms," *Journal of the European Economic Association*, 16, 1929–1976.

Hjort, J., H. Malmberg, and T. Schoellman (2021): "The missing middle managers: Labor costs, firm structure, and development," .

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

——— (2014): "The life cycle of plants in India and Mexico," *The Quarterly Journal of Economics*, 129, 1035–1084.

Hsieh, C.-T. and E. Rossi-Hansberg (2023): "The industrial revolution in services," *Journal of Political Economy Macroeconomics*, 1, 000–000.

Ilut, C. and H. Saijo (2021): "Learning, confidence, and business cycles," *Journal of Monetary Economics*, 117, 354–376.

Ilut, C., R. Valchev, and N. Vincent (2020): "Paralyzed by Fear: Rigid and Discrete Pricing Under Demand Uncertainty," *Econometrica*, 88, 1899–1938.

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

Lo, A. W. and A. C. MacKinlay (1988): "Stock market prices do not follow random walks: Evidence from a simple specification test," *The review of financial studies*, 1, 41–66.

Luo, Y. (2008): "Consumption dynamics under information processing constraints," *Review of Economic Dynamics*, 11, 366–385.

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

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

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

Muth, J. F. (1960): "Optimal properties of exponentially weighted forecasts," *Journal of the american statistical association*, 55, 299–306.

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

Sager, E. and O. A. Timoshenko (2021): "Demand Uncertainty, Selection, and Trade," Working Paper.

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

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

Highlights for the revised manuscript "Uncertainty, Imperfect Information, and Expectation Formation over the Firm's Life Cycle"

1. Firms gradually increase the precision of forecasts over their life cycle.
2. Firm learning plays a prominent role, among other reasons why firms make better forecasts as they become older.
3. The impact of information frictions underlying the gradual resolution of uncertainty on aggregate productivity is substantial and amplified in the presence of firm selection.