Uncertainty, Imperfect Information, and Learning in the International Market*

Cheng Chen Tatsuro Senga Chang Sun Hongyong Zhang†

May, 2019

Abstract

Using a dataset of Japanese multinational firms that contains firm-level sales forecasts, we provide new evidence on imperfect information and learning over the firm's life cycle. We find that firms make non-negligible and positively correlated forecast errors over time. However, they make more precise forecasts and less correlated forecast errors, as they become more experienced. We then build a model with heterogeneous firms that gradually learn about their demand. We quantify the learning and real options channels along the age dimension, through which greater micro-level uncertainty adversely affects resource allocation at the extensive margin and thus depresses productivity.

Keywords. imperfect information, learning, uncertainty, firm expectations

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, Olga Timoshenko, 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 and University of Hong Kong, 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 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 about future outcomes, such as sales and profits. However, as we seldom observe firms' expectations directly, how firms respond to and resolve uncertainty over time remains unknown. This makes it difficult to quantitatively isolate the degree of uncertainty and imperfect information from the data 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. By using data of Japanese firms that serve the foreign market, we show that firms become better informed about future sales over their life cycle. We begin by showing that the precision of forecasts increases with firms' market experience, which is accumulated either through multinational production or via exporting prior to multinational production. Moreover, although each firm's forecast errors are autocorrelated, the auto-correlation declines with market experience. To account for these facts, we extend a dynamic industry equilibrium model with heterogeneous firms to allow firms to serve the foreign market via exporting or multinational production. We then embed Jovanovic-type learning into this framework. As in Jovanovic (1982), with uncertainty and imperfect information, firms' dynamic decisions on the mode of service may be distorted, which leads to productivity losses at the industry level. Our quantitative exercises substantiate that these productivity losses through extensive margin dynamics are substantial, showing the quantitative importance of learning and the real options effects in driving misallocation in our model.

We begin by constructing our new dataset—a parent-affiliate matched 20-year panel

[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), and market entry (Dixit, 1989).

dataset on Japanese multinational firms—taken from business surveys conducted by the Japanese government on a yearly basis. The distinctive features of our dataset include the following: (1) it contains quantitative forecasts for future sales at the affiliate level—a direct measure of firms' subjective expectations; (2) it covers both expectations and realized outcomes, which allows us to calculate forecast errors for each firm; and (3) its panel structure and the inclusion of many young affiliates enable the analysis of within-firm variation of forecast errors over the firm's life cycle.[3]

Exploring our dataset, we show the following features of forecast errors made by an individual firm regarding its sales. First, firms are making small forecast errors on average, and firm-level components explain most of the variations in forecast errors, while aggregate components such as country-year and industry-year fixed effects explain a tiny fraction of these errors. Second, firms make more precise forecasts as they become more experienced in the destination market, either through multinational production or via exporting prior to it. Third, past forecast errors are positively correlated with current and future forecast errors. Moreover, this positive correlation declines with the firms' experience in the destination market. All things considered, we find that firms become better informed as they operate longer in the destination market, as captured by the declining variance and serial correlation of forecast errors as firms becomes more experienced.

In light of this evidence, we build a model that integrates Jovanovic's (1982) model of learning into a standard monopolistically competition model with heterogeneous firms, as in Arkolakis et al. (2017). We extend this model with the following three key ingredients. First, firms can serve the foreign market by exporting or multinational production, as in Helpman et al. (2004). Exporting to the foreign market requires firms pay the iceberg trade costs, while multinational production in the foreign market requires sunk entry costs but no such iceberg trade costs. This implies that a firm's optimal behavior of entering multinational

[3]For analysis on business cycles, see, for example, Bachmann et al. (2017), who study firms' expected investment over the business cycles using panel data of German firms.

production is described by the threshold rules.[4]

With our second model ingredient—learning, these threshold rules are belief-driven[5]. In our model, firms face a downward sloping demand curve in a setting where each firm gradually learns about its demand, which is heterogeneous across firms. The firm-specific demand is the sum of a time-invariant, permanent component, and a transitory component. Crucially, firms only observe the sum of the two components, not each of them separately, and thus have to gradually learn about the permanent component over time using past sales. As a result, the threshold rules involve the firms' *belief* about their permanent demand, and the firms start multinational production when the expected permanent demand is above a certain threshold. Different from a perfect information benchmark where multinational affiliates and exporters sort by the permanent component of demand perfectly, such *sorting* is imperfect in the learning model. In other words, the pecking order result for the permanent demand in the prefect information benchmark does not hold in our learning model. As a result, it causes losses in the allocative efficiency and lowers industry productivity. We call this mechanism the learning effect. Moreover, uncertainty implies a negative impact of real options on multinational production. In our model, young firms face high uncertainty because of their lack of experience, which is captured by a large posterior variance of the (permanent) demand distribution. For those firms, the option value of waiting is high, and they adopt a "wait-and-see" rule for multinational production (and for exiting) by exporting when they are young. The inaction of entering multinational production (and of exiting) by young exporters with high permanent demand draws (and by young exporters with low permanent demand draws) leads to productivity losses at the industry level.

Finally, to match the autocorrelation of forecast errors, we integrate sticky information into the model in the spirit of Mankiw and Reis (2002). This allows us to reproduce the positive but declining auto-correlation of forecast errors over the firm's life cycles.[6] Our

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

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

[6] The serial correlation of forecast errors is zero in a Bayesian learning model, as in Jovanovic (1982), since Bayesian updating with the unbiased prior yields the best linear unbiased estimator (BLUE) for the

economy is populated by informed firms and uninformed firms. Informed firms update their beliefs using Bayes' rule as in Jovanovic (1982), while uninformed firms keep using the prior belief. All entrants are uninformed and thus use their prior belief to forecast. In each period, a fraction of uninformed firms stochastically become informed and never become uninformed again.[7]

Our quantitative exercise involves solving the model numerically and parameterizing it using our dataset. A defining feature of our parameterization is that we use cross-sectional moments of forecast errors to pin down the key parameters of the model. In our model, when a firm is old enough, the forecast errors are caused almost all by the transitory shocks since the firm has discovered its permanent demand component. In contrast, for entrants (or young firms) with little experience, both the permanent demand and the transitory shock contribute to the forecast errors. Therefore, the variance of forecast errors of young and old firms is informative about the variance of the permanent shocks and the transitory shocks. Finally, as the serial correlation of forecast errors is only caused by firms that are uninformed in adjacent periods, we use the autocovariance of forecast errors to discipline the probability that an uniformed firm keeps being uninformed in the next period. The model is able to capture the decline in the absolute value and the autocorrelation of forecast errors as firms become more experienced. It can also capture other salient features of the data, such as the decline of sales growth volatility as affiliates become older.

Our model has quantitative implications for industry-level productivity. Our data show substantial cross-country differences in the variance of firm-level transitory shocks, potentially reflecting different business environments affected by government policies.[8] Motivated by this finding, we vary the variance of transitory shocks and examine its impacts on industry-level productivity, using a benchmark of the perfect information model with the same set

permanent demand.

[7]We integrate sticky information into our model to reproduce the positively correlated forecast errors, although there are other ways, such as introducing rational inattention or noisy information into the model to match the result. See, for example, Sims (2003), Luo (2008), and Mackowiak and Wiederholt (2009).

[8]For instance, Japanese affiliates in Argentina and Venezuela receive transitory shocks that are three to four times more volatile (in terms of the variance) than Japanese affiliates in the U.S.

of parameters. We prove that the variance of transitory shocks (with a constant mean) has no impact on resource allocation and industry-level productivity in the perfect information benchmark, as the transitory shock is unexpected and the firm cannot make decisions based on it.

Industry productivity decreases with the variance of transitory shocks in the imperfect information model for two reasons. First, as the signal-to-noise ratio declines when the transitory shocks become more volatile, learning becomes less effective. This leads to a *worse* sorting of firms into service modes, which we invent a rank test to show. We further show that the worsened sorting results in lower allocative efficiency and productivity at the industry level.[9] Second, firms become more cautious when learning becomes less effective. Consequently, more exporting firms will wait and see before switching to multinational production or to exit, verified by the expanding inaction region of exporters over their market experience. This also results in lower allocative efficiency and industry productivity. Both effects combined reduce the average productivity substantially when we move from the perfect information benchmark to the imperfect information world, with the losses ranging from 2.6% (for countries with low uncertainty) to 7.4% (for countries with high uncertainty). These numbers are comparable to the findings from David et al. (2016). We then implement comparative statics exercises in our imperfect information model and find that average productivity falls by 4.3%, when the variance of transitory shocks increases from its lower bound (i.e., Poland) to its upper bound (i.e., Argentina and Venezuela) in our data. In short, losses in industry-level productivity due to imperfect information or more volatile transitory shocks are sizable in our model.

[9]The essence of the rank test is to quantify the overall overlap of two distributions which should not overlap in the perfect information benchmark. We prove several appealing features of this test in the online appendix.

Related Literature

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—assuming a particular way of forming expectations as discussed by Manski (2018). A growing literature breaks with this tradition by collecting and analyzing direct expectations data. The seminal works by Coibion and Gorodnichenko (2012), Coibion and Gorodnichenko (2015), and Coibion et al. (2018) conducted a diagnostic study on how agents form expectations and how they respond to shocks. Indeed, these studies have informed us how to best model and calibrate a theoretical framework, highlighting the usefulness of such a direct-measure-oriented approach.[10] Our paper differs from these studies, as we study firms' expectations of their own future circumstances, while previous research has relied on macroeconomic expectations. Moreover, we focus on the life cycle properties of firms' expectations.[11]

Our paper contributes to the recent literature on uncertainty. For example, Bloom (2009), Bachmann et al. (2013), Vavra (2014), Kehrig (2015), Senga (2016), Bloom et al. (2018) and others show that various dispersion measures of uncertainty fluctuate in a countercyclical fashion. For instance, Bloom et al. (2018) find that a variety of cross-sectional dispersion measures are correlated with stock market volatility. Moreover, they show that countercyclical volatility coupled with sunk costs of investment and hiring can trigger large drops in output over the business cycles because of firms' increasing inactions. Regarding firm-level uncertainty, Leahy and Whited (1996) and Bloom et al. (2007) show a negative relationship between uncertainty and investment. Our contributions are that (1) we show firms are gradually better informed and make more precise forecasts over their life cycle, (2) and we uncovers a new dimension, firm age, along which increasing uncertainty triggers firms'

[10] See, for example, Mankiw and Reis (2002), Sims (2003), Mackowiak and Wiederholt (2009), and Tanaka et al. (2018).

[11] Other papers that study micro-level expectations include Bloom et al. (2017) for American firms, Bachmann et al. (2017) for German firms, and Boneva et al. (2018) for firms in the U.K.

inactions under imperfect information.

Our paper shows that productivity losses through extensive margin dynamics—firms' entries and exits—are substantial, highlighting the importance of uncertainty as a key driver of industry-level productivity under imperfect information.[12] The analysis of the real options effect of uncertainty on productivity is related to Bloom (2009) and Bloom et al. (2018).[13] While Bloom (2009) and Bloom et al. (2018) emphasize variations of uncertainty over the business cycles, we focus on how uncertainty evolves over the firm's life cycle. Moreover, different from Bloom et al. (2018), our paper shows that increasing uncertainty can lead to allocative losses not only through the real options effect, but also through the learning effect.

The importance of uncertainty and informational imperfection in the international market has been studied by Impullitti et al. (2013), Handley (2014), Novy and Taylor (2014), Alessandria et al. (2015), Handley and Limão (2015), Timoshenko (2015), Handley and Limão (2017), among others. For instance, Handley and Limão (2017) quantify the impact of policy uncertainty in the trade context using a model of sunk costs of exporting.[14] Different from these studies, we study uncertainty associated with multinational production.[15] We complement their work by focusing on learning as a mechanism of reducing uncertainty.

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 multinational firms and their foreign affiliates to study properties of the forecast errors made by each firm and their relationship with firm age and export experience. First, the forecast errors made by firms become smaller

[12]Similar to the proof contained in Appendix B of Arkolakis et al. (2017), we can show that the social planner's solution is the same as the equilibrium allocation of a decentralized economy, *if* both of them have the same allocation at the extensive margin. Therefore, there is no distortion at the intensive margin in our model (in terms of relative quantities produced by various firms). The proof is available upon request.

[13]This also links with the earlier work of Abel (1983), Bernanke (1983), and Dixit and Pindyck (1994).

[14]See, Roberts and Tybout (1997), for evidence of sunk costs to export market entry.

[15]Related to our paper, Gumpert et al. (2016), Conconi et al. (2016), and Deseatnicov and Kucheryavyy (2017) examine the interaction between exporting and multinational production.

as they grow older. Second, export experience reduces the size of the forecast errors made by entrant firms. Finally, the forecast errors are autocorrelated, but the serial correlation declines over the affiliates' life cycles. Overall, the results presented in this section indicate that firms become better informed as they operate longer in the market, hence obtaining more experience.

2.1 Data

This subsection describes our panel data of Japanese multinational firms and their foreign affiliates. We combine two firm-level surveys executed by the Ministry of Economy, Trade and Industry (METI): the Basic Survey of Japanese Business Structure and Activities ("domestic activities survey" hereafter) and the Basic Survey on Overseas Business Activities ("foreign activities survey" hereafter).

The domestic activities survey provides information about domestic business activities of Japanese firms, including multinational parent firms. It covers firms with more than 50 workers and 30 million Japanese yen in paid-in capital from the following industries: mining, manufacturing, wholesale trade, retail, and hospitality. A key variable for our study is their export to seven regions: North America, Latin America, Asia, Europe, the Middle East, Oceania, and Africa. Combined with the foreign activities survey, we can measure each Japanese multinational firm's previous export experience in a region before an affiliate is established there.

The foreign activities survey contains information about overseas affiliates of Japanese multinational firms (hereafter called "multinational affiliates"), including affiliates' location, industry, sales, employment, and investment. It covers two types of overseas businesses: (1) direct (first-tier) affiliates with more than 10% of the equity share capital owned by Japanese multinational firms, and (2) second-tier affiliates with more than 50% of the equity share capital owned by Japanese multinational firms' affiliates. Combining these two surveys yields a panel of 2,300 parent firms and 14,000 affiliates each year from 1995 to 2014.[16]

2.2 Forecast Errors

Importantly, the foreign activities survey asks not only about the realized sales in the previous fiscal year, but also about the *projected sales* for the next fiscal year. Using this variable as firms' expectations of future sales, we define the deviation of the realized sales from the projected sales as the forecast error.

First, our leading measure for the forecast error used in this paper is the log point deviation of the realized sales from the projected sales 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 prediction about sales in period $t+1$ from period t.[17] A positive (negative) forecast error means that the firm under-predicts (over-predicts) its sales.

Second, we define the percentage deviation of the projected sales from the realized sales as

$$FE_{t,t+1}^{pct} = \frac{R_{t+1}}{E_t\left(R_{t+1}\right)} - 1.$$

As forecast errors calculated using the above methods contain extreme values, we trim the top and bottom one percent of observations of the forecast errors.

Third, we construct a measure for the "residual forecast error" measure in an effort to isolate the firm-level idiosyncratic components reflected in the forecast errors. To exclude systemic components, such as aggregate business cycles, from the forecast errors, we project our two measures of the forecast error, $FE_{t,t+1}^{\log}$ and $FE_{t,t+1}^{pct}$, onto country-year and industry-

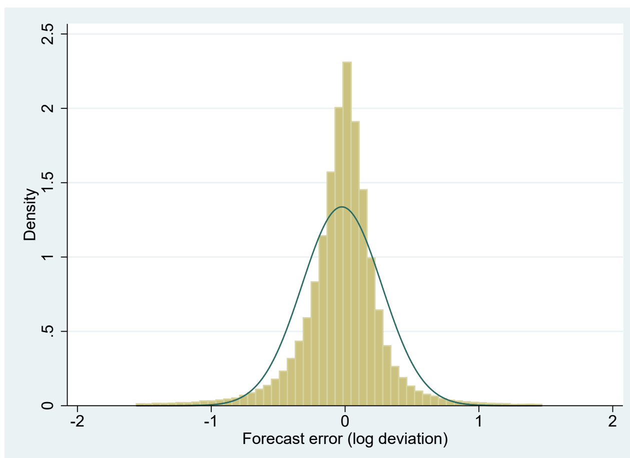
[16] We have approximately 3,200 parent firms and 17,000 affiliates (per year) in the foreign activities survey. Affiliates with relatively small parent firms are lost in this process, as there are minimum thresholds (on firm size) for parent firms to be included into the domestic activities survey.

[17] We omit the time subscriptions whenever there is no confusion.

year fixed effects and obtain the residuals, $\hat{\epsilon}_{FE^{\log}}$ and $\hat{\epsilon}_{FE^{pct}}$. As it turns out, the fixed effects only account for about 11% of the variation, which indicates that firm-level uncertainty plays a dominant role in generating the firms' forecast errors.

In Figure 1, we plot the distribution of our leading measure of forecast errors, $FE^{\log}$, across all affiliates 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). In Table 1, we report the summary statistics regarding the forecast errors.

Figure 1: Distribution of the forecast errors



Notes: Histogram of FE^{log} with fitted normal density (solid line).

The first four rows are about forecast errors and residual forecast errors. While the mean of the residual forecast errors, $\hat{\epsilon}_{FE^{\log}}$ and $\hat{\epsilon}_{FE^{pct}}$, is zero by construction, the mean of FE^{log} and FE^{pct} is also close to zero. In the middle four rows, we report the summary statistics of the absolute value of various constructed forecast errors. Since the country-year and industry-year fixed effects account for a small fraction of the variation, the mean, median, and standard deviation of $|\hat{\epsilon}_{FE^{log}}|$ (and $|\hat{\epsilon}_{FE^{pct}}|$) are similar to those of $\left|FE^{\log}\right|$ (and $|FE^{pct}|$). The patterns of manufacturing firms' forecast errors are similar to the overall patterns, as shown by the last four rows of the table.

Overall, these results show that the forecasts are unbiased on average. In the following

Table 1: Summary statistics of the forecast errors

	Obs.	mean	std. dev.	median
FE^{log}	132050	-0.024	0.298	-0.005
FE^{pct}	132589	0.017	0.333	-0.006
$\hat{\epsilon}_{FE^{log}}$	131754	0.000	0.281	0.011
$\hat{\epsilon}_{FE^{pct}}$	132293	0.000	0.315	-0.023
$\lvert FE^{log}\rvert$	132050	0.200	0.223	0.130
$\lvert FE^{pct}\rvert$	132589	0.204	0.264	0.130
$\lvert\hat{\epsilon}_{FE^{log}}\rvert$	131754	0.184	0.212	0.116
$\lvert\hat{\epsilon}_{FE^{pct}}\rvert$	132293	0.189	0.252	0.117
FE^{log} - Manufacturing	91574	-0.022	0.278	-0.003
FE^{pct} - Manufacturing	91858	0.014	0.307	-0.004
$\lvert FE^{log}\rvert$ - Manufacturing	91574	0.186	0.208	0.123
$\lvert FE^{pct}\rvert$ - Manufacturing	91858	0.189	0.243	0.123

Notes: FE^{log} is the log deviation of the realized sales from the projected sales, while FE^{pct} is the percentage deviation of the realized sales from the projected sales. $\hat{\epsilon}_{FE^{log}}$ is the residual log forecast error, which we obtain by regressing FE^{log} on a set of industry-year and country-year fixed effects. Similarly, $\hat{\epsilon}_{FE^{pct}}$ is the residual percentage forecast error, which we obtain by regressing FE^{pct} on a set of industry-year and country-year fixed effects. The manufacturing subsample refers to affiliates in manufacturing or the wholesale/retail sector whose parent firm is in the manufacturing sector.

sections, we use these measures of forecast errors to present a set of stylized facts that are key for understanding our dataset. The stylized facts presented below are robust across all these measures of forecast errors.

Fact 1: Precision of Forecasts Increases over Affiliates' Life Cycles

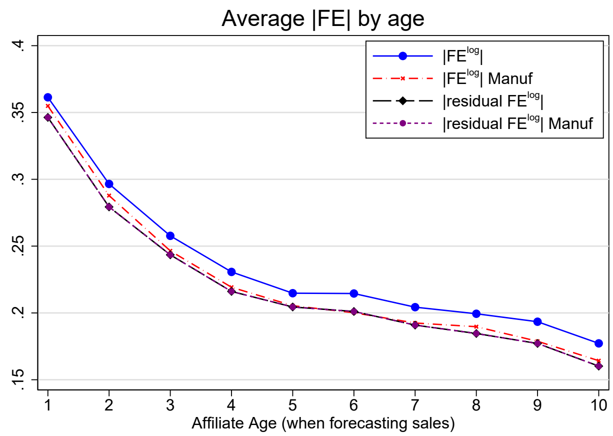
Figure 2 presents the average absolute value of the log forecast errors (and residual log forecast errors) by age cohorts.[18] It is clearly that the precision of forecasting future sales increases as the firm becomes older. Specifically, as foreign affiliates grow from age one to age ten, the absolute value of their log forecast errors decline from 36% to 18% on average. Moreover, the decline of the absolute value of the forecast errors happens mainly in the first five years after market entry. Similar patterns emerge when we use the absolute value of the residual forecast errors. We further confirm these patterns formally by estimating an OLS regression of affiliate i's absolute value of forecast error in year t:

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

[18]Firm age has been truncated at the age of ten.

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. We also control for affiliate or parent sales X_{it} in some regressions. We use age one as the base category; therefore, the age fixed effects represent the difference in the absolute value of the forecast errors between age n and age one. To further control for heterogeneity across affiliates, we also run a regression with affiliate fixed effects δ_i instead of the industry fixed effects δ_s. Column 1 in Table 2 shows the baseline specification

Figure 2: Distribution of the forecast errors



Note: Average absolute value of FE^{log} by age cohorts.

with industry and country-year fixed effects. It is clear that as affiliates become older, the absolute value of their forecast errors declines. On average, affiliates that are at least ten years old have absolute forecast errors that are 16 log points lower. In columns 2 and 3, we add affiliates' sales and their parent firms' sales in Japan into the specification. While larger affiliates tend to have smaller forecast errors, affiliates with larger parent firms tend to have larger forecast errors. This may be because larger affiliates tend to diversify their products or because these affiliates have better planning and thus more precise forecasts; at the same time, larger parent firms may choose to enter riskier and more volatile markets. Once we control for the affiliates' fixed effects, as reported in column 3, the effect of parent firm size disappears. All in all, the result that the forecast errors decline over the affiliates' life cycles survives, even after we have controlled for affiliates' sales and their parent firms' domestic sales.

To further evaluate the robustness of our results, we restrict our sample to (1) surviving entrants and (2) manufacturing firms.[19] Column 4 reports the result for a subsample of affiliates that have survived and continuously appeared in the data from age one to age seven, while column 5 reports the result for a subsample of manufacturing affiliates. The age-dependent declines in the forecast errors robustly show up in both sample.

In the online appendix, we provide additional robustness checks of the above regressions using other measures for the absolute forecast errors. We also address the concern that potential biases in firms' forecast errors may change over their life cycles, which can also contribute to changes in $|FE|$. Specifically, we design a two-step procedure to test whether the conditional variance of the forecast errors depends on age.[20] We consistently find that firms become better at forecasting their sales as they become older.

Fact 2: Increasing Precision of Forecasts through Exporting

As seen above, in a foreign market the precision of forecasts increases over the affiliates' life cycles. It is commonly known that many firms serve a foreign market via exporting before setting up their foreign affiliates. Therefore, the previous export experience of parent firms may have effects on foreign affiliates' forecasts, which is the phenomenon investigated by this subsection. In particular, we calculate the affiliates' absolute forecast errors at age one and regress this measure on various measures of previous export experience, controlling for industry fixed effects and country-year fixed effects.[21]

We restrict our sample to first-time entrants into countries or regions that we identify using the founding year of the affiliates. As export data at the firm-destination country level

[19] Endogenous exits affect our estimates of the age effects in two ways. First, affiliates with higher uncertainty may exit early as they are more likely to be hit by unexpectedly bad shocks. They may also delay their exits, since they have already paid the sunk costs (of doing multinational production), and there is an option value of remaining in the market, as in Bloom (2009).

[20] The basic idea is to 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.

[21] Specifically, the affiliates' absolute forecast errors at age one is the log deviation of the realized sales at age two from the projected sales at age one.

Table 2: Age effects on the absolute forecast errors

Dep.Var: $\lvert FE^{log}_{t,t+1}\rvert$	(1)	(2)	(3)	(4)	(5)
Sample:		All Affiliates		Entrants that have survived >= 7 years	Manufacturing
Age=2	-0.068***	-0.034***	-0.029***	-0.041***	-0.026**
	(0.007)	(0.008)	(0.008)	(0.012)	(0.010)
Age=3	-0.104***	-0.050***	-0.040***	-0.045***	-0.036***
	(0.007)	(0.008)	(0.008)	(0.012)	(0.010)
Age=4	-0.130***	-0.068***	-0.055***	-0.052***	-0.053***
	(0.007)	(0.007)	(0.009)	(0.012)	(0.010)
Age=5	-0.144***	-0.078***	-0.060***	-0.067***	-0.053***
	(0.007)	(0.008)	(0.009)	(0.012)	(0.010)
Age=6	-0.143***	-0.078***	-0.057***	-0.066***	-0.054***
	(0.007)	(0.007)	(0.009)	(0.012)	(0.010)
Age=7	-0.154***	-0.085***	-0.062***	-0.081***	-0.059***
	(0.007)	(0.007)	(0.009)	(0.012)	(0.011)
Age=8	-0.157***	-0.085***	-0.061***	-0.072***	-0.055***
	(0.007)	(0.007)	(0.009)	(0.013)	(0.011)
Age=9	-0.161***	-0.090***	-0.064***	-0.069***	-0.057***
	(0.007)	(0.007)	(0.009)	(0.014)	(0.011)
Age=10	-0.173***	-0.093***	-0.062***	-0.075***	-0.054***
	(0.007)	(0.007)	(0.009)	(0.012)	(0.011)
log(Parent Domestic Sales)$_t$		0.006***	0.002	0.009***	0.002
		(0.001)	(0.002)	(0.001)	(0.002)
log(Affiliate Sales)$_t$		-0.021***	-0.026***	-0.026***	-0.028***
		(0.001)	(0.002)	(0.002)	(0.002)
Industry FE	Yes	Yes	No	Yes	No
Country-year FE	Yes	Yes	Yes	Yes	Yes
Affiliate FE	No	No	Yes	No	Yes
N	131447	116362	111002	16750	82143
R^2	0.096	0.117	0.369	0.128	0.362

Notes: Standard errors are clustered at parent firm level, * 0.10 ** 0.05 *** 0.01. All coefficients are significant at 1% level, except for the log of parent firm's domestic sales in column 3. The dependent variable is the absolute value of forecast errors in all regressions. Age is the age of the affiliate when making the forecasts. Regressions in columns 1, 2 and 3 include all affiliates, while the regression in column 4 only includes entering affiliates that have continuously appeared in the sample from age 1 to age 7.

are not available, we obtain information on parent firms' previous export experience at the regional level using the domestic activities survey data. Finally, we focus on affiliates in either the manufacturing sector or the wholesale and retail sector whose parent firms are in the manufacturing sector.[22]

In Table 3, we provide evidence that previous export experience reduces the forecast errors made by the foreign affiliates that enter a market for the first time. In columns 1 and 2, we use dummy variables that equal one only if the parent firm of the affiliate exported to the same region in the year (or in one of the two years) prior to the entry of multinational production. In columns 3 and 4, we use the more sophisticated definition of export experience. We follow the literature (Conconi et al. (2016) and Deseatnicov and Kucheryavyy (2017)) and define export experience as the number of years since export entry and then reset the experience when the firm stops exporting for two consecutive years.[23] Columns 1-3 show show that having previous export experience reduces the absolute forecast errors by 13-16 log points, and column 4 shows that one additional year of export experience reduces the forecast errors by 1.3 log points.

In the online appendix, we provide a battery of robustness checks for Table 3. The above results are robust to (1) using the sample of first-time entrants into each region, (2) controlling for parent firm and affiliate size, (3) and using the subsample of affiliates with most of the sales in the local market. Taken together, we show that previous export experience is associated with lower initial uncertainty for entering affiliates. This shows that the existence of information value provided by exporting activities.

[22]Following Conconi et al. (2016), we include distribution-oriented foreign direct investments such as wholesale and retail affiliates in our analysis, as affiliates in these industries may sell the same products as what the parent firms had previously exported. As a result, previous export experience helps reduce demand uncertainty for these affiliates as well.

[23]We present the distribution of export experience across first-time entrants in the online appendix.

Table 3: Forecast error and previous exporting

Dep.Var: $\lvert FE^{log}_{1,2}\rvert$	(1)	(2)	(3)	(4)
$Exp_{-1} > 0$	-0.159**			
	(0.065)			
$Exp_{-1} > 0$ or $Exp_{-2} > 0$		-0.151**		
		(0.064)		
$Exp\quad Expe. > 0$			-0.132*	
			(0.070)	
$Exp\quad Expe.$				-0.013**
				(0.006)
Industry FE	Yes	Yes	Yes	Yes
Country-year FE	Yes	Yes	Yes	Yes
N	553	561	658	658
R^2	0.486	0.499	0.472	0.472

Notes: Standard errors are clustered at parent firm level, * 0.10 ** 0.05 *** 0.01. Dependent variable is affiliates' initial forecast error, which is calculated as the absolute log deviation of the realized sales at age = 2 from the projected sales (predicted by an affiliate at age = 1). We only include affiliates that are first-time entrants into a particular host country. Exporting experience (Exp Expe.) is defined at the continent level for each parent firm. Each column head indicates the different measure of exporting experience used in the regression.

Fact 3: Correlated Forecast Errors over Affiliates' Life Cycles

A number of studies have investigated the time-series properties of expectational errors in various contexts. Among them, a growing literature has highlighted the serial correlation of forecast errors, interpreted as evidence against the assumption of full information made in economic models. For example, Ryngaert (2017) shows that professional forecasters' expectational errors of future inflation rates are autocorrelated, indicating their imperfect information about macroeconomic conditions.[24] Instead of inflation expectations, we use data on the sales expectations made by individual firms and show that their forecast errors are positively autocorrelated over time. Importantly, the serial correlation of forecast errors declines with the affiliate's age, a similar age dependent profile of the precision of forecasts as presented above.

First, we present the summary statistics on the correlation of forecast errors over time, which refers to the serial correlation between $FE^{\text{log}}_{t-1,t}$ and $FE^{\text{log}}_{t,t+1}$, where $FE^{\text{log}}_{t,t+1}$ refers to the error in period $t+1$ made by the forecast in period t. Table 4 shows that the forecast errors made by the same firm in two consecutive years are positively and significantly correlated,

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

Table 4: Serial correlation of the forecast errors made in two consecutive years

Sample	All Firms	Manufacturing	Survivors	Manufacturing & Survivors
corr. $(FE^{log}_{t-1,t}, FE^{log}_{t,t+1})$	0.137	0.136	0.171	0.168
	96868	68429	15556	11111
corr. $(FE^{pct}_{t-1,t}, FE^{pct}_{t,t+1})$	0.105	0.103	0.136	0.124
	97393	68694	15658	11174
corr. $(\hat{\epsilon}^{log}_{FE,t-1,t}, \hat{\epsilon}^{log}_{FE,t,t+1})$	0.114	0.116	0.150	0.151
	96589	68346	15484	11086
corr. $(\hat{\epsilon}^{pct}_{FE,t-1,t}, \hat{\epsilon}^{pct}_{FE,t,t+1})$	0.087	0.086	0.120	0.113
	97111	68611	15584	11149

Notes: $FE^{log}_{t-1,t}$ is the log deviation of the realized sales in year t from the projected sales in year $t-1$, while $FE^{log}_{t,t+1}$ is the log deviation of the realized sales in year $t+1$ from the projected sales in year t. $FE^{pct}_{t-1,t}$ is the percentage deviation. The other two measures, $\hat{\epsilon}^{log}_{FE,t-1,t}$ and $\hat{\epsilon}^{pct}_{FE,t-1,t}$, are the residual terms of the first three measures, which we obtain by regressing FE^{log} and FE^{pct} on a set of industry-year and country-year fixed effects. The top and bottom one percent observations of forecast errors are trimmed. The manufacturing sample includes affiliates in manufacturing, wholesale, or retail whose parent firms are in manufacturing. The survivor sample includes entrants that have continuously appeared in the sample from age 1 to age 7. All correlation coefficients are positively significant at the 1% level. The integers below the correlation coefficients are the numbers of observations.

irrespective of the measure of forecast errors we look at and the sample we use.[25] This evidence indicates that firms tend to make the same systematic mistake in forecasting their own sales overtime. Moreover, we also find that the positive serial correlation of the forecast

Table 5: Serial correlation of the forecast errors for different age groups

Sample	Age 2-4	Age 5-7	Age $\geq$ 8
corr. $(FE^{log}_{t-1,t}, FE^{log}_{t,t+1})$	0.175	0.131	0.122
	6708	9313	52265
corr. $(FE^{pct}_{t-1,t}, FE^{pct}_{t,t+1})$	0.141	0.096	0.092
	6799	9350	52395
corr. $(\hat{\epsilon}^{log}_{FE,t-1,t}, \hat{\epsilon}^{log}_{FE,t,t+1})$	0.160	0.120	0.097
	6695	9296	52213
corr. $(\hat{\epsilon}^{pct}_{FE,t-1,t}, \hat{\epsilon}^{pct}_{FE,t,t+1})$	0.129	0.088	0.072
	6786	9333	52343

Notes: $FE^{log}_{t-1,t}$ is the log deviation of the realized sales in year t from the projected sales in year $t-1$, while $FE^{log}_{t,t+1}$ is the log deviation of the realized sales in year $t+1$ from the projected sales in year t. Firm age refers to the age in year t. The top and bottom one percent of observations of the forecast errors are trimmed. The sample only includes affiliates in manufacturing, wholesale, or retail whose parent firms are in manufacturing, i.e., the manufacturing sample. All correlation coefficients are positively significant at the 1% level. The integers below the correlation coefficients are the numbers of observations.

errors declines with affiliate age, as shown by Table 5. When firms become more experienced, the tendency of making systematically the same forecast errors is reduced, hinting that firms learn and thus become more informed about their own information environment over their

[25] The regression analysis in the online appendix further confirms this finding.

life cycle.

We use regressions to further show that the positive serial correlation of forecast errors is attenuated, when the affiliate becomes older and when its parent firm has previous export experience prior to the entry of multinational production. The regression equation we run is

$$\mathbf{1}\Big(Sign(FE^{pct}_{i,t,t+1}) = Sign(FE^{pct}_{i,t-1,t})\Big) = age_{i,t} + X_{i,t-1} + \delta_i + \delta_{ct} + \varepsilon_{it}, \tag{2}$$

where δ_i and δ_{ct} represent affiliate and country-year fixed effects, respectively. The indicator function, $\mathbf{1}\Big(Sign(FE^{pct}_{i,t,t+1}) = Sign(FE^{pct}_{i,t-1,t})\Big)$, equals one if the forecast errors made in two consecutive years have the same sign and their absolute values are all bigger than 1%. Otherwise, it takes the value of zero. In other words, the firm forecasts its sales precisely, if its forecast error is small enough.[26] We focus on the coefficients of the age dummies or a continuous age variable top-coded at age ten. In some of the regressions, we also control for affiliate size and parent size, which have little effect on the age effects.

The results are summarized in Table 6. As seen in Table 6, older firms tend to make less systematic forecast errors, as shown by the negatively significant coefficients in front of the age dummies and the age of the affiliate. Moreover, this finding is robust to the inclusion of affiliate size and parent size. In the online appendix, we show that using the residual forecast errors to run the above regression does not change our findings. We also show that positive export experience reduces the correlation of the forecast errors for first-time entrants into the destination markets. These empirical findings substantiate that market experience helps firms learn about their business conditions.

[26] We use 1% as the cutoff to exclude cases when $FE_{t-1,t}$ and $FE_{t,t+1}$ are of the same signs but only deviate from zero by a small margin. Note that our findings are robust to the different thresholds we use to define this indicator function, such as 0.5%.

Table 6: Age and serial correlation of the forecast errors

Dep.Var: $Sign(FE^{pct}_{t,t+1}) = Sign(FE^{pct}_{t-1,t})$	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		All Affiliates			Manufacturing	
Age=3	0.013	0.006		0.015	0.009	
	(0.012)	(0.013)		(0.014)	(0.016)	
Age=4	-0.011	-0.022		-0.015	-0.021	
	(0.013)	(0.014)		(0.015)	(0.017)	
Age=5	-0.007	-0.018		-0.021	-0.026	
	(0.014)	(0.016)		(0.017)	(0.019)	
Age=6	-0.019	-0.025		-0.013	-0.015	
	(0.014)	(0.015)		(0.017)	(0.019)	
Age=7	-0.037***	-0.049***		-0.040**	-0.047**	
	(0.014)	(0.016)		(0.017)	(0.019)	
Age=8	-0.037**	-0.049***		-0.037*	-0.048**	
	(0.015)	(0.017)		(0.019)	(0.020)	
Age=9	-0.018	-0.037**		-0.019	-0.033	
	(0.015)	(0.017)		(0.019)	(0.020)	
Age=10	-0.035**	-0.053***		-0.039**	-0.051**	
	(0.016)	(0.017)		(0.019)	(0.021)	
Age (cut at ten)			-0.007***			-0.006***
			(0.002)			(0.002)
log(Parent Domestic Sales)$_{t-1}$		-0.009*	-0.009*		-0.008	-0.008
		(0.005)	(0.005)		(0.006)	(0.006)
log(Affiliate Sales)$_{t-1}$		0.005	0.004		0.003	0.003
		(0.004)	(0.004)		(0.004)	(0.004)
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Affiliate FE	Yes	Yes	Yes	Yes	Yes	Yes
N	90653	80873	80873	64579	60697	60697
R^2	0.206	0.214	0.214	0.205	0.210	0.210

Notes: Standard errors are clustered at parent firm level. * 0.10 ** 0.05 *** 0.01. Dependent variable equals 1, if forecast errors made in two consecutive years have the same sign and their absolute values are all bigger than 1%. Otherwise, it takes the value of zero. Forecast error is calculated as the percentage deviation of the realized sales from the projected sales. Age is cut at ten. Note that as the dependent variable can only be defined for affiliates which are at least two years old, the age dummies can be estimated from three (all relative to age-two firms).

3 Model

To study how uncertainty and learning affect resource allocation and allocative efficiency, this section proposes a dynamic model that integrates Jovanovic's (1982) model of learning into a standard monopolistically competition model with heterogeneous firms. We extend it allow firms to serve the foreign market via exporting or multinational production. The model features firm learning (as in Arkolakis et al. (2017)) and information rigidity (similar to Mankiw and Reis (2002)). After describing the setup and equilibrium of the model, we show that both mechanisms are needed to match all the facts documented above.

3.1 Setup: Demand and Supply

In our model, there are two countries, Japan and the foreign country. Each Japanese firm produces a differentiated variety and has to decide whether to serve the foreign market, and if so, whether through exporting or multinational production. We do not explicitly model Japanese domestic firms and ignore domestic demand for two reasons. First, it helps us highlight the trade-off between trade and multinational production. Second, since we do not have a representative sample of Japanese domestic firms, we lack relevant moments to calibrate parameters specific to domestic production.[27]

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

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

[27]The domestic activities survey we use does cover some firms that do not export or produce abroad. However, since the threshold for the survey is quite high (50 workers and 30 million yen of paid-in capital), it misses a large number of small domestic firms. We think it is more representative for exporters and multinational affiliates.

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

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

In the first nest, χ_i is the demand shifter for country i goods, 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:

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

where $\tilde{P}_t$ is the aggregate price index for all goods, and $P_{jp,t}$ is the ideal price index for Japanese goods. 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 Eaton and Kortum (2002) and Melitz (2003). In our calibration, δ is set to be between one and σ.

In our model, we assume that the Japanese varieties make up a small fraction of foreign consumers' consumption and treat $\tilde{Y}_t$ and $\tilde{P}_t$ as exogenous.[28] 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 call it the aggregate demand shifter. In addition, as we only focus on Japanese firms, we suppress the subscript jp in the following analysis and derive the ideal price index of Japanese goods as

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

We use the term, $A_t \equiv Y_t P_t^{\sigma-\delta}$ to denote the aggregate demand condition faced by all

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

Japanese firms in period t and rewrite the firm-level demand function as

$$q_t(\omega) = A_t e^{a_t(\omega)} p_t(\omega)^{-\sigma}. \tag{6}$$

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 shock $\varepsilon_t(\omega)$ as in Arkolakis et al. (2017):

$$a_t(\omega) = \theta(\omega) + \varepsilon_t(\omega),\ \varepsilon_t(\omega) \overset{i.i.d.}{\sim} N\left(0, \sigma_\varepsilon^2\right). \tag{7}$$

Firms understand that $\theta(\omega)$ is drawn from a normal distribution $N\left(\bar{\theta}, \sigma_\theta^2\right)$, and the independently and identically distributed (i.i.d.) transitory shock, $\varepsilon_t(\omega)$, is drawn from another normal distribution $N(0, \sigma_\varepsilon^2)$. We assume that the firm observes the sum of the two demand components, $a_t(\omega)$, at the end of each period, not each of them separately. Thus, the firm needs to learn about its permanent demand every period by forming an posterior belief about the distribution of θ.

In addition, we assume that there are two types of firms in the economy in the same spirit as in sticky information models a la Mankiw and Reis (2002): the informed firms and the uninformed firms. All entrants are uninformed in the sense that they use the prior distribution of θ (i.e., $N(\bar{\theta}, \sigma_\theta^2)$) to form the expectation. At the end of each period, $1-\alpha$ fraction of the remaining uninformed firms become informed. From the time when the uninformed firms become informed, they begin to update the beliefs by utilizing realized demand shifters and will never become uninformed again. The remaining uninformed firms still use the prior distribution of θ to form posterior beliefs.

To ensure that experienced and unexperienced multinationals coexist, we need to introduce ex-ante heterogeneity across firms. We assume that firms are heterogeneous in their entry costs of multinational production, following Das et al. (2007). Therefore, firms with low entry costs may become multinationals without exporting experience. On the other hand, we assume that firms are homogeneous in labor productivity. In order to produce q units of output, the firm has to employ the same amount of workers. We make this assumption as our data show that experienced multinational affiliates are larger and more productive than the unexperienced ones (see online appendix for the evidence). Heterogeneity in ex-ante labor productivity would imply the opposite pattern, as only the most productive firms enter multinational production without export experience in such a world.

Firms hire workers in a perfectly competitive labor market. Exporters can employ domestic workers to produce at a constant wage w, while multinational affiliates employ foreign workers at a constant wage w^*. 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 28) and the export-to-domestic (gross) output ratio is also small for Japan, ranging from 6% to 8% from 2005 to 2009.[29]

The industry structure features monopolistically competition. There is an exogenous mass of potential entrants J (from Japan) that decide whether or not to enter the foreign market each period. Each entrant draws a permanent demand shifter θ from a normal distribution, $N\left(\bar{\theta}, \sigma_\theta^2\right)$, and a sunk entry cost f_m^e of multinational production from a log-normal distribution, $logN(\mu_{f_m^e}, \sigma^2_{f_m^e})$. 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 entry cost of f_x^e, and setting up an affiliate 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 have to 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). In addition, they can also choose to permanently exit. We assume that incumbent multinational affiliates can switch to exporting without paying the

[29]Ideally, instead of output shares, one will want to evaluate the assumptions based on labor shares of exporters (and multinational firms), which are not easily unavailable. However, given that larger firms tend to have lower labor shares (Autor et al., 2017; Sun, 2018), the labor shares of Japanese exporters and multinationals are likely to be even lower.

sunk entry cost of exporting, as they have already established their appearance in the foreign market. Each period, firms also have to pay a fixed cost, f_x, in order to export or, f_m, in order to do multinational production.

For firms that serve the foreign market, they decide how much to produce in period t before the overall demand shifter, a_t, is realized. After the demand shifter in period t is realized, they choose the price p_t to sell all the products produced, as we assume there is no storage technology and firms cannot accumulate inventories. For informed incumbent firms, they update their beliefs about the permanent demand after observing the demand shifter in period t. Additionally, a randomly selected $1-\alpha$ fraction of uninformed firms become informed at the end of each period.

3.2 Belief Updating

In this subsection, we discuss how a firm forms the ex post belief for its permanent demand. For an informed firm, it has observed the realized demand shifters in the past t periods: $a_1, a_2, \ldots, a_t$. Since both the prior and the realized demand shifters are normally distributed, the Bayes' rule implies that the posterior belief about θ is normally distributed with mean μ_t and variance σ_t^2 where

$$\mu_t = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2}\bar{\theta} + \frac{t\sigma_\theta^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2}\bar{a}_t, \tag{8}$$

and

$$\sigma_t^2 = \frac{\sigma_\varepsilon^2 \sigma_\theta^2}{\sigma_\varepsilon^2 + t\sigma_\theta^2}. \tag{9}$$

The history of signals $(a_1, a_{2,}, \ldots, a_t)$ is summarized by age t and the average demand shifter:

$$\bar{a}_t \equiv \frac{1}{t}\sum_{i=1}^{t} a_i \text{ for } t \geq 1; \quad \bar{a}_0 \equiv \bar{\theta}.$$

Therefore, the firm believes that the overall demand shifter in period $t+1$, $a_{t+1} = \theta + \varepsilon_{t+1}$, has a normal distribution with mean μ_t and variance $\sigma_t^2 + \sigma_\varepsilon^2$. For an uninformed firm, its belief for the mean and variance of θ is the same as the prior belief. For future use, we define the signal-to-noise ratio as

$$\lambda \equiv \frac{\sigma_\theta^2}{\sigma_\varepsilon^2}.$$

3.3 Static Optimization of Per-Period Profit

We study the firm's static optimization problem in the steady state in this subsection. As all aggregate variables such as wages, the ideal price index and expenditure on Japanese goods are constant in the steady state, 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, the firm hires labor and produces q_t quantity of output to maximize its expected per-period profit, $E_{a_t|\bar{a}_{t-1},t}(\pi_{o,t})$, and the realized per-period profit for a multinational affiliate ($o = m$) 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.[30] Firms set the price after observing the realized demand a_t to sell all the output. Maximizing $E_{a_t|\bar{a}_{t-1},t}(\pi_{o,t})$, the optimal output choice is

$$q_{o,t} = \left(\frac{\sigma - 1}{\sigma}\right)^{\sigma} \left(\frac{b(\bar{a}_{t-1}, t-1)}{MC_o}\right)^{\sigma} \frac{Y}{P^{\delta-\sigma}}, \tag{10}$$

where

$$b(\bar{a}_{t-1}, t-1) = E_{a_t|\bar{a}_{t-1},t-1}\left(e^{a_t/\sigma}\right) = \exp\left\{\frac{\mu_{t-1}}{\sigma} + \frac{1}{2}\left(\frac{\sigma_{t-1}^2 + \sigma_\varepsilon^2}{\sigma^2}\right)\right\}, \tag{11}$$

and t is the firm's effective market experience ($t = 1$ for uninformed firms).

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

The resulting price and per-period profit function are

$$p_{o,t}\left(a_t\right) = \frac{\sigma}{\sigma - 1} e^{a_t/\sigma} \frac{MC_o}{b\left(\bar{a}_{t-1}, t-1\right)}; \tag{12}$$

$$E\pi_{o,t} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}} \frac{b\left(\bar{a}_{t-1}, t-1\right)^{\sigma}}{{MC_o}^{\sigma-1}} \frac{Y}{P^{\delta-\sigma}} - wf_o. \tag{13}$$

3.4 Dynamic Optimization and Equilibrium Definition

In each period, an entrant or incumbent firm chooses among three different 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, t, the history of demand shocks, $\bar{a}_{t-1}$. [31] 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 an informed firm ($t \geq 2$) as

$$v(o', o, f_m^e, t, \bar{a}_{t-1}) = \begin{cases} E_{t-1}\pi_{x,t} + \beta(1-\eta)E_{t-1}V\left(x, f_m^e, t+1, \bar{a}_t\right) & \text{if } o' = x, \\ E_{t-1}\pi_{m,t} - wf_m^e \mathbf{1}(o = x) + \beta(1-\eta)E_{t-1}V\left(m, f_m^e, t+1, \bar{a}_t\right) & \text{if } o' = m, \\ 0 & \text{if } o' = exit, \end{cases} \tag{14}$$

where o is the mode of service in the previous period, and o' is the choice of mode this period. The expectation is taken based on information available at the end of period $t-1$, and $V(\cdot)$ is the value function. Note that previous service mode cannot be *entry* for informed firms in the above expression. For an uninformed firm, the choice-specific value function is more complicated, as the firm stays uninformed with probability α next period. In particular, we

[31] For uninformed firms, their current effective market experience is one and they make decisions based on the prior belief $\bar{a}_0 = \bar{\theta}$.

can write

$$v(o', o, f^e_m, 1, \bar{a}_0) = \begin{cases} E_0\pi_{x,1} - wf^e_x\mathbf{1}(o = ent) \\ +\beta(1-\eta)E_0\left(\alpha V\left(x, f^e_m, 1, \bar{a}_0\right) + (1-\alpha)V\left(x, f^e_m, 2, \bar{a}_1\right) \right) & \text{if } o' = x, \\ E_0\pi_{m,1} - wf^e_m\mathbf{1}(o = ent, x) \\ +\beta(1-\eta)E_0\left(\alpha V\left(m, f^e_m, 1, \bar{a}_0\right) + (1-\alpha)V\left(m, f^e_m, 2, \bar{a}_1\right) \right) & \text{if } o' = m, \\ 0 & \text{if } o' = exit. \end{cases} \tag{15}$$

With these choice-specific value functions in hand, we define the value functions as

$$V(o, f^e_m, t, \bar{a}_{t-1}) = \max_{o' \in \{x,m,exit\}} \left\{v(o', o, f^e_m, t, \bar{a}_{t-1})\right\}, \tag{16}$$

and we denote the corresponding policy function as $o'(o, f^e_m, t, \bar{a}_{t-1})$. The definition of equilibrium is contained in the Appendix.

3.5 Intuition on Matching Facts about Forecasts and forecast errors

In this subsection, we show how our model is able to match facts 1-3 presented in Section 2. We illustrate the intuition using a special case in which there is no endogenous switching of production modes. In the online appendix, we also show that the perfect information model cannot be used to rationalize these facts especially the serially correlated forecast errors, even when firms endogenously exit the market.

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

1. *Variance of forecast errors declines with years of experience.*

2. Forecast errors made in two consecutive periods by the same firm are positively correlated when $\alpha > 0$ but are uncorrelated when $\alpha = 0$. When $\alpha > 0$, the positive correlation between these two terms declines with years of experience.

Proof. See online appendix. ■

Both learning and the reduction in information rigidity over time contribute to the first property. Thanks to learning, informed firms accumulate more experience and have clearer information about their permanent demand, when they operate in the market for a longer period of time. Second, as more firms become informed over time and informed firms make more accurate forecasts, the variance of forecast errors goes down with firm's market experience. Since exporting helps firms accumulate sales experience, it is a natural result the learning mechanism also rationalizes fact 2.

The above proposition also rationalizes the finding of serially correlated forecast errors presented in Section 2.2. The positive correlation is triggered by firms that are uninformed in two consecutive periods only, as their forecasts do not change over time. In the proof, we show that firms that are informed in both periods and firms that switch from being uninformed to being informed within the two periods have uncorrelated forecast errors. Therefore, information rigidity is needed to match the serial correlation of forecast errors, while Jovanovic-type learning alone (i.e., setting $\alpha = 0$ in our model) cannot do so. As the share of uninformed firms declines with market experience, the positive autocorrelation also decreases.[32]

4 Quantitative Analysis

In this section, we describe the procedures used for calibrating our model. The calibrated model is able to capture the dynamics of the affiliates' forecast errors observed in Section 2.

[32] In the online appendix, 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 Jovanovic-type learning model.

In the online appendix, we show that the calibrated model is able to capture other salient features of the data, such as exporters' sales growth and declining exit rates over their life cycles, which we do not directly target in the calibration. We illustrate the model's aggregate implications at the end of this section.

4.1 Calibration

We first normalize a set of parameters that are 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.[33]

Next, we calibrate a set of parameters without solving our model, as listed in 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; Arkolakis et al., 2018). 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).[34] We set the discount factor, β, to 0.96, which implies a real interest rate of 4%.

The exogenous death rate η and the per-period fixed costs of multinational production f_m are crucial for the exit rates of multinational affiliates. As there is strong selection in the model, the affiliates' exit rates are likely to decline over their life cycles if the per-period fixed costs are positive. However, we did not find a significant decline for the affiliates' exit rates over their life cycle, even for the affiliates without export experience.[35] Therefore, we postulate that $f_m = 0$ and set η to 0.03 so that the model can match the average exit rate

[33]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*.

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

[35]See the online appendix for more details.

of the affiliates (3%). As a result, multinational affiliates exit only due to exogenous death shocks.

Table 7: Parameters calibrated without solving the model

Parameters	Description	Value	Source
σ	Elasticity of substitution between 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 multinational affiliates
f_m	Per-period fixed costs of multinational production	0	Flat profile of affiliates' exit rate over their life cycles

Three parameters that govern information rigidity and learning can also be backed out without calibrating the model. Since we have shut down the endogenous exits for multinational affiliates, there is no selection on the permanent demand draw among multinational affiliates after entry. As all entrants have the same prior belief for their permanent demand, they choose to become multinational affiliates based on the entry costs of multinational production, 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 in the online appendix assuming no selection, and these formulas can be directly applied to the *unexperienced multinational affiliates.* In particular, we target the variance of forecast errors of age-one unexperienced affiliates and of the unexperienced affiliates older than ten, as they are the most informative about σ_θ and σ_ε, respectively. We also target the covariance of forecast errors at age one and two, since the positive auto-covariance is caused by information rigidity, governed by parameter α.[36] The calibrated value of α is 0.29, which implies that 71% of uninformed firms become informed each period and less than three percent of firms are still uninformed after three years.

The remaining four parameters are jointly calibrated by solving the equilibrium and

[36]In practice, due to the partial-year effects (firms have not completed their first year of operation at the time of the survey), age-one firms in the data may have less information than age-one firms in the model. We therefore assume that age-one firms in the data correspond to a mixture of age-zero and age-one firms in the model (with equal weights), and age-two firms in the data correspond to a mixture of age-one and age-two firms in the model (with equal weights). We then adjust the formulas derived in equations (4) and (5) of the online appendix accordingly.

Table 8: Parameters related to the forecast errors and moments

Parameters	Value	Description	Moments	Data	Model
σ_θ	2.15	Std of time-invariant shock	Var. of FE at age 1	0.52	0.52
σ_ϵ	0.98	Std of transitory shock	Var. of FE at age 10	0.26	0.26
α	0.29	prob of sleeping	Cov of $FE_{1,2}$ and $FE_{2,3}$	0.055	0.055

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

matching four moments. The parameters are as follows: the 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 affiliates at age one and the share of exports in total sales (i.e., total exports plus total sales of multinational affiliates).

In Table 9, we list the moments in an order such that, loosely speaking, the moment is the most informative about the parameter in the same row. A higher export per-period fixed cost raises the exporter exit rate, while a higher average entry cost of multinational production increases the fraction of exporters among all firms selling in the foreign market. In the model, only firms with small enough entry costs of multinational production become unexperienced multinational affiliates. Thus, a higher $\sigma_{f_m^e}$ raises the share of unexperienced multinational affiliates. Finally, the iceberg trade costs have a large impact on the intensive margin of export, so we include the sales share of exporters among all firms as a targeted moment.

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

Parameters	Value	Description	Moments	Data	Model
f_x	0.0062	export fixed cost	average exit rate of exporters	0.08	0.08
$\mu_{f_m^e}$	-0.06	mean of log FDI entry cost	fraction of exporters among active firms	0.69	0.71
$\sigma_{f_m^e}$	1.81	Std of log FDI entry cost	fraction of experienced MNEs at age 1	0.72	0.74
τ	1.10	iceberg trade cost	Exporter sales share	0.32	0.31

4.2 Dynamics of Forecast Errors

We examine the changes in $\left|FE^{\log}\right|$ over the affiliates' life cycles in Figure 3.[37] We first estimate the age effects on $\left|FE^{\log}\right|$ for the affiliates that enter a foreign market for the first time. In the same regression, we interact the age fixed effects with a dummy variable that indicates whether the parent firm has previous export experience in the same region. We plot the estimated fixed effects for experienced and unexperienced multinational affiliates in the left panel of Figure 3, using the age-one unexperienced multinational affiliates as the base group. In the right panel, we plot the average $\left|FE^{\log}\right|$ by affiliate age predicted by the calibrated model, again by normalizing the average $\left|FE^{\log}\right|$ to zero for unexperienced multinational affiliates at age one.[38]

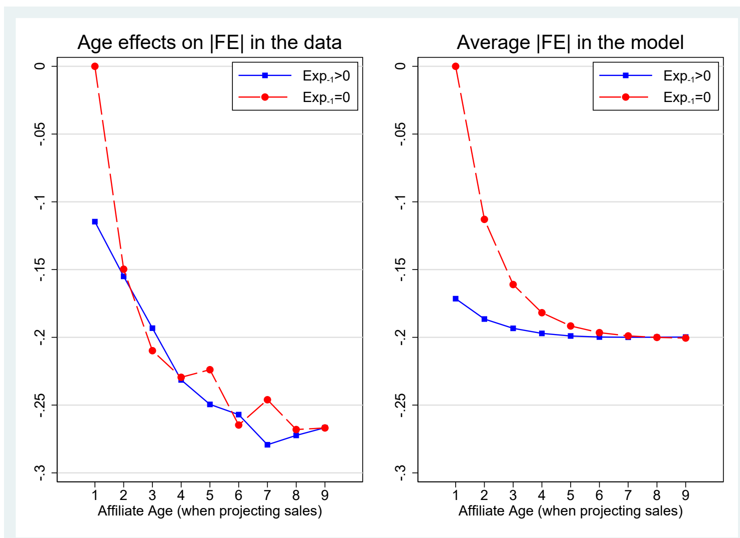
Consistent with the data, the model predicts that the average $|FE^{\log}|$ declines over the affiliates' life cycle and that the initial $\left|FE^{\log}\right|$ is lower for affiliates whose parent firms have export experience. However, the model predicts a much smaller decline in $\left|FE^{\log}\right|$ for experienced multinational affiliates and a larger difference between experienced multinational affiliates and unexperienced multinational affiliates at age one.[39] The reason for this is that firms in the model learn very fast — there is almost no uncertainty about θ after four periods. This implies experienced firms start with very precise forecasts about 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.

[37] Other untargeted moments such as the dynamics of exporters and affiliates in terms of sales growth, exit rates, and sales growth volatility are discussed in the online appendix. In particular, we show that our model endogenously generates sales growth volatility that declines with market experience, which is consistent with the findings from the data.

[38] Note that as we target at the variance of forecast errors at age one and ten in the calibration, the absolute value of forecast errors at age one and ten are untargeted moments.

[39] For example, in the model, the average $\left|FE^{\log}\right|$ of experienced multinational affiliates drops by 0.03 over their life cycles, while the corresponding empirical moment is 0.16.

Figure 3: Forecast error — age profile: data v.s. model



Notes: The left panel shows the estimated age effects on average $|FE^{log}|$ for affiliates in the data, while the right panel shows the average $|FE^{log}|$ by affiliate age 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 affiliates at age $t-1$ and age t, since most affiliates enter into the destination market in the middle of each fiscal year. The solid line shows the estimated age effects for the affiliates whose parent firms have previous export experience, that is, $Exp_{-1} > 0$, while the dashed line shows the estimated age effects for the affiliates without previous export experience, that is, $Exp_{-1} = 0$. The age effect of age-one affiliates without previous export experience is normalized to zero.

4.3 Implications for Allocation and Productivity

In this subsection, we study the impact of imperfect information and learning on industry-level productivity. First, we use our dataset to highlight the substantial heterogeneity of micro-level uncertainty across countries. Second, we use our model to quantify by how much such a cross-country variation of uncertainty drives the variation of productivity across countries. Finally, we investigate the various effects through which uncertainty affects productivity, which include the learning effect and the real options effect. We then decompose these effects step by step to highlight the quantitative importance of imperfect information and learning in driving industry-level productivity.

4.3.1 Uncertainty and Losses in Productivity: A Cross-country Analysis

As reported in Figure 4, we highlight the substantial variation of micro-level uncertainty across countries. The standard deviation of the forecast errors for firms above age ten ranges from 0.2 to 0.3 for most of the countries, with some higher numbers for Argentina (0.41) and for Venezuela (0.46). Through the lens of our model, this implies that Japanese affiliates in Argentina and Venezuela receive transitory shocks that are three to four times more volatile than Japanese affiliates in the E.U. and the U.S.

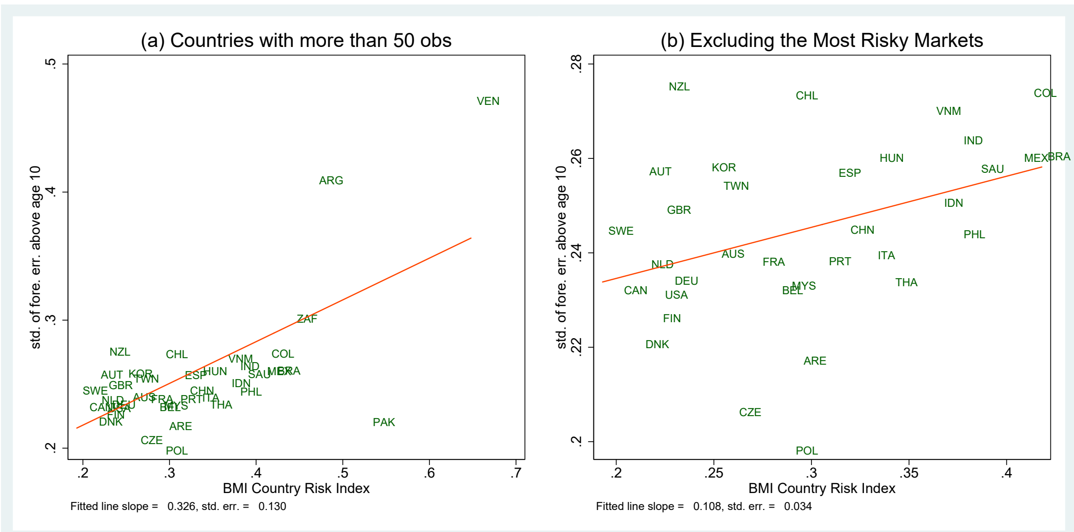
As discussed before, the standard deviation of the forecast errors made by old enough firms is close to $\sigma_\varepsilon/\sigma$. Therefore, with the standard deviation of the forecast errors of firms older than ten for each country, we can retrieve country-specific σ_ε, by applying $\sigma = 4$ to all countries. From this, we can back out σ_ε for each country, which ranges from 0.8 for Poland to 1.8 for Venezuela, with an average of 0.98 across countries.

Motivated by this finding, we ask how much this cross-country variation of uncertainty transforms into differences in productivity across these countries. Specifically, we vary σ_ε from 0.8 to 1.8, with a constant value of $\bar{\theta} + \sigma_\varepsilon^2/2\sigma$.[40] As seen in Figure 5, the average

[40] As both demand shocks follow the log normal distribution, a larger variance of the shock leads to a higher mean. A mean-preserving-spread (MPS) requires us to deduct $\frac{\Delta\sigma_\varepsilon^2}{2\sigma}$ from each firm's θ draw when the change in the variance of ε is $\Delta\sigma_\varepsilon^2$. In the online appendix, we prove that both firm-level and aggregate-level

labor productivity, Q/L, decreases monotonically with σ_ε. All other things being equal, the productivity can vary by 5% when uncertainty varies from the level of Poland to that of Argentina and Venezuela. As learning becomes less effective when σ_ε increases, there are more entrants that choose multinational production instead of exporting immediately after entry, as can be seen in the top right panel of Figure 5.

Figure 4: Firm-level uncertainty and country-level aggregate uncertainty



Notes: Each dot in the graph represents a country. The value of the vertical axis represents the standard deviation of the forecast errors of affiliates above 10 years old, which is a proxy for $\sigma_\varepsilon/\sigma$ in our model. The horizontal axis corresponds to the aggregate uncertainty at the country level, measured by the Business Monitor International (BMI) country risk index. Panel (a) includes all countries with at least 50 observations. Panel (b) excludes the four most risky markets according to the BMI country index: Venezuela, Argentina, South Africa, and Pakistan. The line in each panel represents the fitted line. The corresponding slopes and standard errors of the slopes are presented below the panels.

To measure "misallocation" at the extensive margin due to imperfect information, we invent an allocation efficiency index inspired by Bowen et al. (1987).[41] With the index, we can then examine how allocative efficiency varies with the model parameters. Conditional on a particular value of the multinational production's entry cost, we denote the cumulative density function (CDF) of the permanent demand draw of multinational affiliates and exporters conditional as $G(\theta)$ and $F(\theta)$, respectively. We then define the allocative efficiency

outcomes are unchanged, when we implement such a MPS of $e^{a_t/\sigma}$ by increasing the variance of transitory shocks.

[41]Bowen et al. (1987) calculate a similar statistic to test the predictions of the Heckscher-Ohlin model.

index using the following integral:

$$A_{m,x}(G,F) \equiv \int \mathbf{1}(\theta_m \geq \theta_x) dG(\theta_m) dF(\theta_x) = \int \mathbf{1}(\theta_m \geq \theta_x) g(\theta_m) f(\theta_x) d\theta_m d\theta_x, \qquad (17)$$

where $g(\theta)$ and $f(\theta)$ are the corresponding probability density functions (PDF) of $G(\theta)$ and $F(\theta)$. Economically, what this measure does is to calculate the fraction of exporting firms with lower permanent demand draws than multinational affiliates with a given value of θ_m, and to integrate this fraction over the distribution of θ_m for multinational affiliates.[42] We call $A_{m,x}(G,F)$ the allocation efficiency index along the margin of exporting and multinational production. Similarly, we can use the same equation as above to derive another index for the comparison between the (permanent demand) distribution of active firms (i.e., exporters or multinational affiliates) and that of exiting firms. As the bottom panels of Figure 5 show, both indices deteriorate as σ_ε increases.

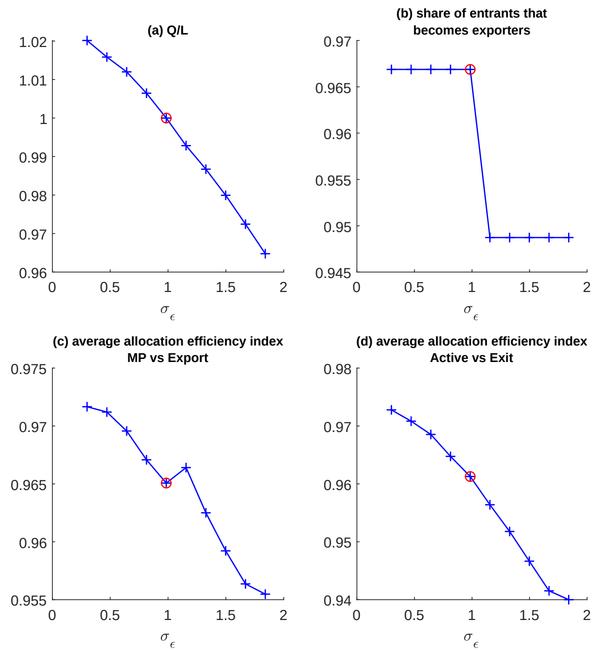
4.3.2 Learning Effect and Real Options Effect

In our model, uncertainty impacts industry-level productivity via affecting resource allocation at the extensive margin. Here, we isolate the various effects that operate in our model; and, among them, we highlight two major ones by which greater uncertainty adversely affects productivity: (1) the learning effect and (2) the real options effect.

First, the pecking order result that exists in the perfect information benchmark does not hold in the imperfect information model, as firms make extensive margin choices before knowing their permanent demand perfectly. Specifically, firms with good permanent demand might exit the market because of a series of bad transitory shocks at young ages, while firms with bad permanent demand might conduct multinational production thanks to a series of

[42] If the distribution of θ were discrete, we would be doing (pairwise) comparisons for each possible pair of (θ_x, θ_m) from the permanent demand distribution of the exporters and of the multinational affiliates. Then, we calculate the total probability with which exporters have lower θ draws than multinational affiliates. This is similar to the rank test used in Bowen et al. (1987). In the online appendix, we prove that this measure has several attractive properties. For instance, this index achieves its maximum, one, in the perfect information benchmark for any f_m^e, since firms' decision of exporting or multinational production follows a threshold rule strictly.

Figure 5: Industry-level productivity falls with σ_ε^2 with a constant $\bar{\theta} + \sigma_\varepsilon^2/2\sigma$



Note: We keep $\bar{\theta} + \sigma_\varepsilon^2/2\sigma$ constant.

good transitory shocks at young ages. This creates a source of "misallocation" and reduces industry-level productivity. This effect depends on the signal-to-noise ratio, λ, and we call this (first) effect the learning effect.

Second, even if firms with good permanent demand draws eventually become multinational affiliates, they might export for long time because of the high uncertainty concerning their θ's at young ages. Similarly, even if firms with bad permanent demand exit eventually, they might stay in the market for a long period of time due to the same reason as well. This creates an inaction region (for entry into multinational production and exiting from the market) over exporting firms' average market experience, which shrinks as the firm ages. This second effect is the real options effect which exists due to gradual learning in our model and lowers industry-level productivity. In order to illustrate the two new channels, we take the following steps.

Model (1) → Model (2)

First, we start from a version of our model with $\sigma_\varepsilon = 0.8$ and $\sigma_\theta = 1.750$ and increase σ_ε by 50% from 0.8 (Poland) to 1.2 and leave the other parameters unchanged. Two effects emerging from this exercise are (1) the learning effect due to the smaller signal-to-noise ratio $\lambda = \sigma_\theta^2/\sigma_\varepsilon^2$ (4.738 → 2.126) and (2) the real options effect due to a higher dispersion of the transitory shocks, σ_ε (0.8 → 1.2).

Table 10: Average allocation efficiency index and industry-level productivity in four models

Model	σ_ε	σ_θ	λ	A(G,F)	Q/L(Imperfect info.)	Q/L(Perfect info.)
1	0.800	1.750	4.783	0.973	6.452	6.623
2	1.200	1.750	2.126	0.966	6.378	6.623
3	1.200	2.624	4.783	0.968	6.529	6.863
4	1.800	2.624	2.126	0.961	6.352	6.863

Notes: This table reports the average labor productivity in the four models in both the perfect information world and imperfect information world. In addition, it reports the weighted average of the allocation efficiency index (for a comparison between exporting and multinational production) in each model, where the weights are the probabilities of f_m^e of the entrants. Note that we always keep $\bar{\theta} + \sigma_\theta^2/2 + \sigma_\varepsilon^2/2\sigma$ constant in our four models.

Table 11: Effects of uncertainty on average productivity: decomposition of channels

Each channel:	Learning effect	Real options effect	Selection effect	Loss in Q/L
Model (1) to (2)	negative	negative	—	-1.1 %
Model (1) to (3)	—	negative	positive	1.2 %
Model (1) to (4)	negative	negative	positive	-1.5 %

Notes: We decompose the overall change in industry-level productivity into various channels. Note that we always keep $\bar{\theta}+\sigma_\theta^2/2+\sigma_\varepsilon^2/2\sigma$ constant to control the Jensen's effect (coming from variations in both $e^{\sigma_\theta^2/\sigma}$ and $e^{\sigma_\varepsilon^2/\sigma}$) and the Oi-Hartman-Abel effect (coming from variations in $e^{\sigma_\theta^2/\sigma}$). We discuss these effects in detail in the online appendix.

As the signal-to-noise ratio decreases, the distribution of the allocation efficiency index over the possible range of multinational production entry costs moves downward *monotonically* from Model (1) to Model (2), as shown in Figure 6. That is, for each given realization of f_m^e, Model (2) yields a smaller allocation efficiency index than Model (1). This downward shift implies a worse allocation of firms at the extensive margin. Moreover, this learning effect also has an indirect effect on the expansion of the inaction region. That is, exporting firms want to learn about their own fundamental demand for a longer period of time to make more informed extensive margin choices as learning becomes less effective. This can be seen from the expansion of the inaction region shown in panel (a) of Figure 7.[43] In total, the two negative effects hinder the efficient allocation of firms and reduce the average labor productivity, as shown in Table 10.

Finally, productivity is unchanged when we increase σ_ε in a perfect information model where firms know θ upon entry, as shown by the last column of Table 10.[44] However, we can investigate how the average productivity changes, when we move from the perfect information benchmark to the imperfect information model in Models (1) and (2). For instance, even for a country where the variance of transitory shocks is small (e.g., Poland) as in Model (1), the average productivity, Q/L, falls by 2.6% if we move from the perfect information world (6.623) to the imperfect information world (6.452). Such a productivity loss is more pronounced for a country where the variance of transitory shocks is large (e.g., New Zealand), as in Model (2), since the same exercise yields a 3.7% drop in the average productivity (from 6.623 to 6.378).

Model (1) → Model (3)

Next, to see the learning effect and real options effect separately, we start from Model (1) and increase σ_ε by 50% from 0.8 to 1.2. We also increase σ_θ proportionately, leaving the signal-to-noise ratio unchanged ($\lambda = 4.783$). As in the previous exercise, the real options effect emerges due to the more dispersed σ_ε ($0.8 \rightarrow 1.2$). This can be seen in panel (b) of Figure 7, showing the expansion of the inaction region. However, there is a crucial difference here as there is no learning effect when we move from Model (1) to (3) with a fixed λ.[45]

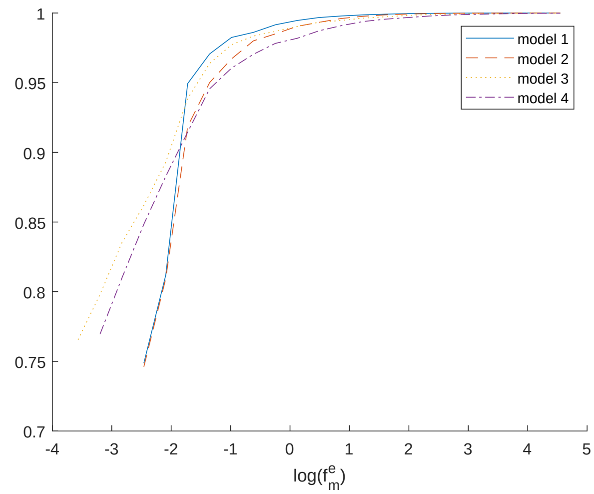
Without the learning effect at work in this exercise, the size of the productivity losses

[43] We show the same pattern for different values of the multinational production's entry cost and pick up one value for the purpose of illustration only.

[44] We prove this result in the online appendix.

[45] Note that in Figure 6, the distributions of the allocation efficiency index in Models (1) and (3) do not have a clear ordering. For some values of f_m^e, Model (1) yields a higher allocation efficiency index than Model (3), while the opposite pattern is true for the other values of f_m^e.

Figure 6: The allocation efficiency index under different values of f_m^e



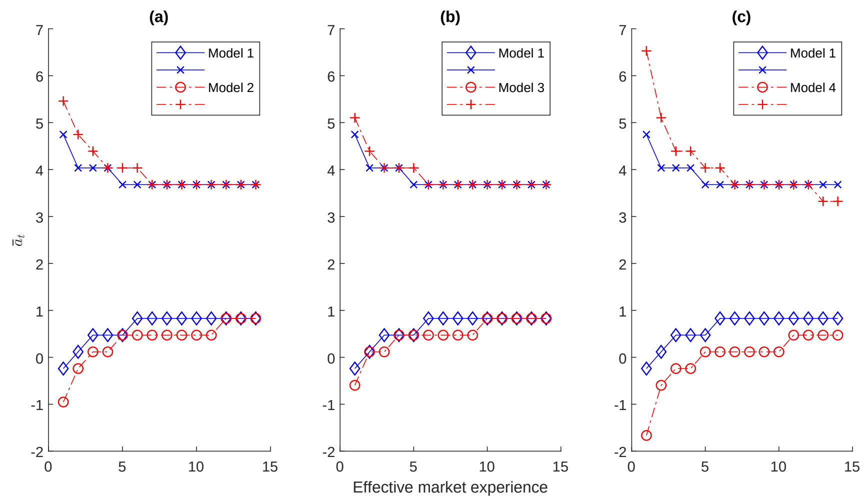
Notes: This figure plots the allocation efficiency index (for a comparison between exporting and multinational production) over the possible range of entry costs of multinational production in four models. Note that we always keep $\bar{\theta} + \sigma_\theta^2/2 + \sigma_\varepsilon^2/2\sigma$ constant in our four models.

when we move from Model (1) to (3) should be mitigated. However, as seen in the fifth column of Table 10, productivity in Model (3) is even higher than in Model (1). This is because an increase in σ_θ^2 generates an additional effect, which has nothing to do with imperfect information and learning. In fact, the average productivity also increases under perfect information, as seen in the last column of Table 10. We label this as *a selection effect* in that an increase in σ_θ^2 leads to more extremely productive and unproductive firms. As the most productive firms which choose multinational production replace the least productive firms which exit in the new equilibrium, market competition intensifies, leading to a tougher selection into survival. Average productivity increases as a result of this selection effect. In total, the selection effect dominates the real options effect, leading to a higher value of average productivity in Model (3) than in Model (1).[46] When we move from the perfect information world to the imperfect information world, the comparison between Model (2) and (3) reveals another role of σ_θ^2: while average productivity falls by 3.7% in Model (2) with

[46]Note that the selection effect appears only when we have multiple extensive margin choices. In the online appendix, we prove that both firm-level and aggregate-level outcomes are unchanged in the perfect information benchmark with *one* production mode (e.g., multinational production).

$\sigma_\theta = 1.750$, it falls by 4.9% in Model (3) with $\sigma_\theta = 2.624$. In short, productivity losses due to informational imperfection are more pronounced when the variance of permanent demand draws is larger.

Figure 7: Inaction region by age in different models



Notes: This figure plots the inaction region of exporters (to continue exporting in the next period) over their life cycles. The inaction region is defined over $\bar{a}_{t-1}$ (i.e., average past demand shifters) and conditional on a particular value of the entry cost of multinational production, $f_m^e = 0.781$. It is obtained from the policy functions under different parameter values but assuming constant aggregate demand and price as in Model (1). Note that we always keep $\bar{\theta} + \sigma_\theta^2/2 + \sigma_\varepsilon^2/2\sigma$ the same across all the models.

Model (1) $\rightarrow$ Model (4)

Finally, to highlight the quantitative importance of the learning effect, we start from Model (3) and increase σ_ε^2 by 50% from 1.2 to 1.8 (Venezuela) and leave the other parameters unchanged. By doing so, we have the learning effect back and make the other two effects, the real options effect and the selection effect, at work as well (see the summary of these exercises in Table 11). By comparing Models (1) and (4), one can see the learning effect reduces the average productivity despite the fact that the selection effect is at work, raising the average productivity. The panel (c) of Figure 7 illustrates the expansion of the inaction region, and Table 10 shows that the average allocation efficiency index drops from Model (1) to (4). In total, the two negative effects dominate the positive selection effect, which leads to a lower average productivity in Model (4) compared with Model (1), as shown by the fifth column of Table 10 (Q/L: 6.452 $\rightarrow$ 6.352). When we move from the perfect information world to the imperfect information world in Model (4), the average productivity falls by 7.4% (6.863 $\rightarrow$ 6.352), which is the biggest productivity loss among the four models.[47]

The volatility of the transitory shocks has a sizable and negative impact on industry-level productivity in our imperfect information model, when the variance of the transitory shocks increases from its lower bound to its upper bound in our data. Specifically, we start from Model (4) and reduce σ_ε from 1.8 (for Argentina and Venezuela) to 0.8 (for Poland). We find that the average productivity, Q/L, drops by 4.3% (from 6.633 to 6.352). In short, the losses in industry-level productivity due to imperfect information or volatile transitory shocks are sizable in our model.

Remarks

Although we focus on industry-level analysis, we conjecture that the two new channels (along the age dimension) through which uncertainty affects resource allocation and productivity survive in a general equilibrium multi-country model (with endogenous wages and income) for two reasons. First, the driving force for the two channels is the signal-to-noise ratio, which is not affected by general equilibrium forces. Second, firms' inactions and imperfect sorting at the extensive margin appear along the age dimension. As long as adjustments in wages and income do not affect firms with different ages differently, our new channels will not be impacted.[48]

Before ending this subsection, we discuss the similarities and differences between the real options effect in our model and the classical real options effect studied in Bloom (2009)

[47] In our model, the existence of imperfect information per se does not contribute to the productivity losses substantially. If we shut down learning in our model (i.e., $\alpha = 1$), average productivity, Q/L, would fall by 46.9% from our calibrated model with $\alpha = 0.29$. Therefore, it is the learning that matters for productivity losses.

[48] We can also vary the fraction of firms that switch from being uninformed to being informed every period, as it is related to the level of sticky information in our model. Simulation results show that an increase in this fraction generates the same two negative effects as above: the learning effect and the real options effect. Results are available upon request.

and Bloom et al. (2018). In our model, the real options effect disappears when the firm is old enough and thus knows its permanent demand perfectly. In Bloom (2009) and Bloom et al. (2018), the real options effect disappears if we shut down the innovations to the firm's productivity/demand over time. Put differently, in our model, the real options effect only depends on the existence of imperfect information and is *not* caused by time-varying permanent demand/productivity shocks.[49] To the contrary, the real options effect in Bloom (2009) and Bloom et al. (2018) is caused by the increasing variance of innovations to the firm's productivity/demand over time and is *not* driven by informational imperfection. As a result, the two models yield the same real options effect, but along different dimensions. Relatedly, both models point out that the delayed actions at the extensive margin that are triggered by increasing uncertainty cause "misallocation" and result in losses in productivity.

5 Conclusion

In this paper, we use a unique dataset of Japanese multinational firms, which contains information on sales forecasts at the affiliate level, to detect information imperfection and learning in the international market. We document several new and important stylized facts concerning affiliates' forecasts and forecast errors. We view these facts as direct evidence for the existence of age-dependent firm-level uncertainty, imperfect information and learning. We then build up a dynamic industry equilibrium model of trade and multinational production, which features sticky information and learning, in order to explain the documented facts.

The quantitative model implies substantial information value of exporting. We also find that firm-level demand uncertainty due to more volatile transitory shocks exacerbates the productivity losses from imperfect information via the real options channel and the learning channel along the age dimension. We view understanding the cross-country variation in firm-level uncertainty as a fruitful venue for future research.

[49]In fact, the firm's permanent demand is assumed to be fixed over time in our model.

6 Appendix

6.1 Definition of equilibrium

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

1. *value functions* $V(o, f_m^e, t, \bar{a}_{t-1})$*, choice-specific value functions* $v(o', o, f_m^e, t, \bar{a}_{t-1})$ *and policy functions* $o'(o, f_m^e, t, \bar{a}_{t-1})$ *that satisfy equations (14), (15) and (16);*

2. *policy functions of optimal output* q_o *that satisfy equation (10);*

3. *prices in the current period* $p_o(a_t), o \in \{m, x\}$ *that satisfy equation (12);*

4. *a measure function of firms* $\lambda(o, f_m^e, t, \bar{a}_{t-1}, \theta)$*,* $o \in \{x, m, ent\}$ *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*

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

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 affiliates should be a fixed point of the aggregate law of motion, i.e., given any Borel set of* $\bar{a}_t$*,* Δ_a*, measures of informed firms with* $t \geq 2$ *satisfy*

$$\lambda(o', f_m^e, t+1, \Delta_a, \theta) = \sum_{o \in \{x,m\}} \int \begin{array}{c} \mathbf{1}\left(\bar{a}_t \in \Delta_a, o'(o, f_m^e, t, \bar{a}_{t-1}) = o'\right) \\ \times (1-\eta) \Pr(\bar{a}_t | \bar{a}_{t-1}, \theta)\, \lambda(o, f_m^e, t, d\bar{a}_{t-1}, \theta) \end{array}.$$

Informed firms with $t = 1$ were uninformed in the previous period:

$$\lambda\left(o', f_m^e, 2, \Delta_a, \theta\right) = (1-\alpha) \sum_{o \in \{x,m,ent\}} \int \begin{array}{c} \mathbf{1}\left(\bar{a}_1 \in \Delta_a, o'\left(o, f_m^e, 1, \bar{a}_0\right) = o'\right) \\ \times (1-\eta) \Pr\left(\bar{a}_1 | \bar{a}_0, \theta\right) \lambda\left(o, f_m^e, 1, d\bar{a}_0, \theta\right) \end{array}.$$

Finally, measures of uninformed exporters and multinationals ($t = 0$) satisfy the following law of motion:

$$\lambda\left(o', f_m^e, 1, \Delta_a, \theta\right) = \alpha \sum_{o \in \{x,m,ent\}} \int \begin{array}{c} \mathbf{1}\left(\bar{a}_1 \in \Delta_a, o'\left(o, f_m^e, 1, \bar{a}_0\right) = o'\right) \\ \times (1-\eta) \Pr\left(\bar{a}_1 | \bar{a}_0, \theta\right) \lambda\left(o, f_m^e, 1, d\bar{a}_0, \theta\right) \end{array}.$$

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

$$P^{1-\sigma} = \sum_{t \geq 1} \sum_{o \in \{x,m,ent\}} \int_{f_m^e, \bar{a}_{t-1}, \theta, a_t} \begin{array}{c} e^{a_t} p_o\left(a_t, q_o\left(b\left(\bar{a}_{t-1}, t-1\right)\right)\right)^{1-\sigma} \\ \times \lambda\left(o, df_m^e, t, d\bar{a}_{t-1}, d\theta\right) d\Pr\left(a_t | \theta\right) \end{array}.$$

References

Abel, Andrew B., "Optimal Investment Under Uncertainty," *The American Economic Review*, 1983, *73* (1), 228–233.

— **and Janice C. Eberly**, "Optimal Investment with Costly Reversibility," *The Review of Economic Studies*, 1996, *63* (4), 581–593.

Alessandria, George, Horag Choi, Joseph P Kaboski, and Virgiliu Midrigan, "Microeconomic Uncertainty, International Trade, and Aggregate Fluctuations," *Journal of Monetary Economics*, 2015, *69*, 20–38.

Andrade, Philippe and Hervé Le Bihan, "Inattentive Professional Forecasters," *Journal of Monetary Economics*, November 2013, *60* (8), 967–982.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and Production in the Global Economy," *American Economic Review*, August 2018, *108* (8), 2128–2173.

— **, Theodore Papageorgiou, and Olga Timoshenko**, "Firm Learning and Growth," Working Paper 2017.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research May 2017.

Bachmann, Rüdiger, Steffen Elstner, and Atanas Hristov, "Surprise, Surprise – Measuring Firm-Level Investment Innovations," *Journal of Economic Dynamics and Control*, October 2017, *83*, 107–148.

— **,** — **, and Eric R Sims**, "Uncertainty and Economic Activity: Evidence from Business Survey Data," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 217–249.

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

Bernanke, Ben S., "Irreversibility, Uncertainty, and Cyclical Investment," *The Quarterly Journal of Economics*, 1983, *98* (1), 85–106.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, 2003, *93* (4), 1268–1290.

Bertola, Guiseppe and Ricardo J. Caballero, "Irreversibility and Aggregate Investment," *The Review of Economic Studies*, 1994, *61* (2), 223–246.

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

— **, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry**, "Really Uncertain Business Cycles," *Econometrica*, May 2018, *86* (3), 1031–1065.

— **, Steven J. Davis, Lucia Foster, Brian Lucking, Scott Ohlmacher, and Itay Saporta-Eksten**, "Business-Level Expectations and Uncertainty," SSRN Scholarly Paper 3085377, Social Science Research Network, Rochester, NY December 2017.

Bloom, Nick, Stephen Bond, and John Van Reenen, "Uncertainty and Investment Dynamics," *The Review of Economic Studies*, 2007, *74* (2), 391–415.

Boneva, Lena, James Cloyne, Martin Weale, and Tomasz Wieladek, "Firms' Price , Cost and Activity Expectations: Evidence from Micro Data," Working Paper 2018.

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

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

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

— , — , and Saten Kumar, "How Do Firms Form Their Expectations? New Survey Evidence," *American Economic Review*, September 2018, *108* (9), 2671–2713.

Conconi, Paola, André Sapir, and Maurizio Zanardi, "The Internationalization Process of Firms: From Exports to FDI," *Journal of International Economics*, March 2016, *99*, 16–30.

Das, Sanghamitra, Mark J. Roberts, and James R. Tybout, "Market Entry Costs, Producer Heterogeneity, and Export Dynamics," *Econometrica*, May 2007, *75* (3), 837–873.

David, Joel M., Hugo A. Hopenhayn, and Venky Venkateswaran, "Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, May 2016, *131* (2), 943–1005.

Deseatnicov, Ivan and Konstantin Kucheryavyy, "Exports and FDI Entry Decision: Evidence from Japanese Foreign-affiliated firms," Working Paper 17-E-036, RIETI March 2017.

Dixit, Avinash, "Entry and Exit Decisions under Uncertainty," *Journal of Political Economy*, 1989, *97* (3), 620–638.

Dixit, Avinash K. and Robert S. Pindyck, *Investment Under Uncertainty*, Princeton University Press, 1994.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Feenstra, Robert C., Philip Luck, Maurice Obstfeld, and Katheryn N. Russ, "In Search of the Armington Elasticity," *The Review of Economics and Statistics*, August 2017, *100* (1), 135–150.

Guiso, Luigi and Giuseppe Parigi, "Investment and Demand Uncertainty," *The Quarterly Journal of Economics*, February 1999, *114* (1), 185–227.

Gumpert, Anna, Andreas Moxnes, Natalia Ramondo, and Felix Tintelnot, "Multinational Firms and Export Dynamics," Working Paper 2016.

Handley, Kyle, "Exporting under Trade Policy Uncertainty: Theory and Evidence," *Journal of International Economics*, 2014, *94* (1), 50–66.

— and Nuno Limão, "Trade and Investment under Policy Uncertainty: Theory and Firm Evidence," *American Economic Journal: Economic Policy*, November 2015, *7* (4), 189–222.

— and — , "Policy Uncertainty, Trade and Welfare: Theory and Evidence for China and the U.S," Working Paper 2017.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, "Export versus FDI with Heterogeneous Firms," *The American Economic Review*, March 2004, *94* (1), 300–316.

Impullitti, Giammario, Alfonso A. Irarrazabal, and Luca David Opromolla, "A Theory of Entry into and Exit from Export Markets," *Journal of International Economics*, May 2013, *90* (1), 75–90.

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

Kehrig, Matthias, "The Cyclical Nature of the Productivity Distribution," Working Paper 2015.

Leahy, John and Toni M. Whited, "The Effects of Uncertainty on Investment: Some Stylized Facts," *Journal of Money, Credit, and Banking*, 1996, *28* (1), 64–83.

Luo, Yulei, "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 2008, *11*, 366–385.

Mackowiak, Bartosz and Mirko Wiederholt, "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 2009, *99* (3), 769–803.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, November 2002, *117* (4), 1295–1328.

Manski, Charles F., "Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise," *NBER Macroeconomics Annual*, April 2018, *32* (1), 411–471.

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

Novy, Dennis and Alan M. Taylor, "Trade and Uncertainty," Working Paper 19941, National Bureau of Economic Research February 2014.

Roberts, Mark and James Tybout, "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 1997, *87* (4), 545–64.

Ryngaert, Jane, "What do (and don't) Forecasters Know about U.S. Inflation," Working Paper 2017.

Senga, Tatsuro, "A New Look at Uncertainty Shocks: Imperfect Information and Misallocation," Working Paper 2016.

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

Sun, Chang, "Multinational Production with Non-neutral Technologies," Working Paper, The University of Hong Kong 2018.

Tanaka, Mari, Nicholas Bloom, Joel M David, and Maiko Koga, "Firm Performance and Macro Forecast Accuracy," Working Paper 24776, National Bureau of Economic Research June 2018.

Timoshenko, Olga A., "Product Switching in a Model of Learning," *Journal of International Economics*, 2015, *95* (2), 233–249.

Vavra, Joseph, "Inflation Dynamics and Time-varying Volatility: New Evidence and an Ss Interpretation," *Quarterly Journal of Economics*, February 2014, *129* (1), 215–258.