RIETI Discussion Paper Series 24-E-019

Firms' Internationalization Decisions and Demand Learning

TANI, Naoki
Kyoto University / Ministry of Finance

OGAWA, Eiji
RIETI

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 24-E-019
February 2024

Firms' Internationalization Decisions and Demand Learning[1]

Naoki TANI
Institute of Economic Research, Kyoto University and Ministry of Finance
Eiji OGAWA
Faculty of Economics, Tokyo Keizai University and RIETI

Abstract

This study investigates the rationale for firms to establish foreign affiliates without export experience in the affiliate destinations despite facing demand uncertainty. As over half of Japanese firms start their first internationalization process through foreign direct investment (FDI), we attempt to explain this pattern by developing a statistical decision-making model in which firms make decisions about entering foreign markets via FDI or exports under demand uncertainty. Our model incorporates the proximity-concentration tradeoff and the demand-learning mechanism through which firms predict their demand level based on information about the number and average productivity of their neighbors. We analytically show that demand-learning from neighbors affects firms' entry and exit decisions regarding FDI. We provide supporting evidence for the main predictions of the model by using a rich micro-level dataset of Japanese multinational firms.

Keywords: demand uncertainty, experimentation, export, FDI, learning, productivity
JEL classification: F1, F2, D8, L1

[1]This study was conducted as a part of the project "Exchange Rates and International Currency" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). We would like to thank Arata Ito, Shujiro Urata, Keiichiro Kobayashi, Masato Mizuno, Masayuki Morikawa, and participants of the RIETI DP Seminar where the draft of this paper was presented for their helpful comments. The authors are also grateful for the helpful comments and suggestions provided by Yuki Masujima, Kiyotaka Sato, Etsuro Shioji, and Yushi Yoshida. This study utilizes the micro data of the questionnaire information based on the "Basic Survey of Japanese Business Structure and Activities" and the "Basic Survey on Overseas Business Activities" which are conducted by the Ministry of Economy, Trade and Industry (METI), and the Kikatsu-Kaiji converter, which is provided by RIETI. We gratefully acknowledge financial support from the JSPS KAKENHI Grant (grant numbers:21K01567). All remaining errors are our own.

1 Introduction

There is a growing need to study the dynamics of multinational firms because they are important actors in the global economy and international trade. In fact, approximately 90% of U.S. exports and imports flow through multinational firms (Bernard, Jensen, and Schott (2009)). Analyzing how domestic firms expand their businesses to foreign markets and become multinationals by participating in international supply chain networks yields useful implications for policy makers concerned with promoting international competitiveness in global markets.

Our study focuses on the process through which firms internationalize. In particular, when a firm advances into a foreign market, it can do so through two channels: via exports or via foreign direct investment (FDI). Recent studies have documented that most firms enter into FDI with prior export experience in the destination market, with about 90% of Belgian firms doing so as reported by Conconi, Sapir, and Zanardi (2016) and about 70% of Japanese firms doing so as reported by Deseatnicov and Kucheryavyy (2017). According to Conconi et al. (2016), most FDI entrants conduct export experiments in order to learn about the demand level and reduce demand uncertainty in the destination market. The larger sunk entry cost of FDI compared to that of export makes firms test demand level in destination markets before investment in order to avoid the risk of failure of FDI. However, when we analyze whether a specific firm conducts export experiments before FDI entry, we must first identify whether the firm has existing affiliated companies in the destination market. If so, the firm does not need to test the market demand since it already knows the demand through its existing affiliates; if the firm has export experience before the FDI entry, we should consider it as the intra-firm trade to operate its existing affiliated companies. We report a novel fact that over half of firms start FDI without export experience when we analyze firms that entered specific foreign regions for the first time, which are destination regions where the firms do not have existing affiliated firms. However, why do many firms start FDI and pay the large sunk cost of entry without learning about the demand level through export experiments? This is the main research question this study seeks to answer.

This study develops a model of firms' demand learning from neighbors by extending the exporters' social learning model of Fernandes and Tang (2014) and explains how the information obtained from neighboring firms operating in destination markets affects firms' internationalization decisions regarding FDI and export. In particular, we relax Fernandes and Tang (2014)'s assumption to restrict neighbors' information firms can obtain from a more realistic perspective and further incorporate the setup of the dynamics of firms' internationalization choices between export and FDI.

Our model analytically demonstrates that the higher level of neighbors' productivity negatively affect firms' decision of FDI and exports. This also increases the likelihood of export experiments before FDI entry. This result indicates that if a firm learns that the mean productivity of its neighbors in a specific market is low, the firm is more likely to start FDI without conducting export experiments. This is because a lower mean productivity of neighbors indicates that the level of the market demand is so high that unproductive firms can operate. Firms that predict high demand in markets can start FDI without export experiments. Our model also shows that the number of neighboring firms in the destination market itself has a direct impact on firms' entry decision, depending on the difference between the mean of the prior distribution of the market demand and the demand level inferred from the neighbors' information. Additionally, regarding firms' exit decision from FDI market, our model indicates that the mean productivity of neighbors has a negative impact on exit decisions.

We provide supporting evidence for the main prediction of the model using a rich micro-level dataset of Japanese multinational firms. We first test our theoretical assumption that firms can learn the level of market demand from information about their neighbors. We confirm that firms can make better predictions about the market demand in which they have many neighbors. Then, using the fixed-ordered logit model, we show that the mean of neighbors' productivity has a significantly negative effect on firms' decision regarding entry especially in the manufacturing sector. Moreover, the number of neighbors promotes manufacturing firms to start the internationalization process and begin FDI without conducting export experiments. Thus, demand-learning from neighbors affects firms' entry decisions, especially in the manufacturing sector as predicted by our model. Finally, we also report that the mean productivity of neighbors negatively affects firms' exit decisions.

The first contribution of this study is to unveil a new empirical finding that over half of firms start their first-time internationalization process via FDI, by considering firms' prior experience of FDI in destination regions. This finding contributes to previous works such as Conconi et al. (2016) by suggesting the importance of considering firms' previous FDI experience; firms should be able to reduce demand uncertainty in the destination markets and predict their demand levels without conducting export experiments.

The second contribution of this study is that it develops a social learning model on the basis of Fernandes and Tang (2014). Recently, many studies have incorporated the Jovanovic (1982)'s firm learning setup into their models (e.g., Albornoz, Calvo-Pardo, Corcos, and Ornelas (2012), Timoshenko (2015), Conconi et al. (2016), Arkolakis, Papageorgiou, and Timoshenko (2018), Li (2018), and Chen, Senga, Sun, and Zhang (2018)). Most of these models focus on firms' demand-learning after entry into destination markets. In particular, they assume that firms do not know their true demand in foreign markets before entry. Firms ex-ante know only the distribution from which the demand parameter is drawn and do not obtain information about the parameter before entry. An exception to this literature is the work by Fernandes and Tang (2014), which assumes that firms can predict whether they have sufficient demand in destination markets, given their productivity, by learning from neighbors' average sales. Our model contributes to the literature to rationalize firms' ex-ante demand-learning behavior from information about the number of neighbors and their average productivity. We do not assume that firms can observe their neighbors' average sales from a realistic perspective, which is an extension of Fernandes and Tang (2014).

In addition to demand-learning, we formulate firms' choices between export and FDI. The point of the firm-choice model is the proximity-concentration tradeoff (e.g., Markusen (1984), Brainard (1997), Helpman, Melitz, and Yeaple (2004), Ramondo, Rappoport, and Ruhl (2013), and Conconi et al. (2016)). In particular, firms entering foreign markets face a trade-off: exports incur per-unit transportation costs, but sunk entry costs are small. By contrast, opening a foreign subsidiary enables them to save transportation costs, but they face larger sunk entry costs. Helpman et al. (2004) show that the higher fixed cost of establishing foreign subsidiaries causes a selection effect which only allows the most productive firms to start FDI. Conconi et al. (2016) study the pattern of multinationals that serve foreign markets before FDI. They explain that firms experiment through exporting before starting FDI because larger sunk entry costs of FDI compared with that of exporting make firms "wait and see" whether their true demands are large enough to operate in the markets. Their model justifies export experiments to learn the true demand in foreign markets before starting FDI. However, as mentioned before, in reality, many firms start FDI without export experiments. Thus, we develop a model to explain that ex-ante demand-learning from neighbors reduces the market uncertainty and promotes firms' entry FDI decisions without export experiments. In particular, a firm's entry decision depends on whether its productivity is higher than the productivity threshold of entry, which is determined endogenously by the firm's prediction of demand. We contribute to the literature by incorporating the mechanism of demand-learning from neighbors into the standard firms' internationalization choice model.

We also contribute to the literature by providing empirical evidence using data from Japanese multinationals, including affiliate firms' sales forecast data. Using this forecast data, Chen et al. (2018) justify the reliability of sales forecasts by showing that the affiliates do not use naive rules to make their forecasts. Chen, , Sun, and Zhang (2022) also provide evidence that multinational firms learn about their profitability in a particular market by observing their performance in nearby markets. Using this direct measure of firms' demand forecast, we show that firms with more neighbors in destination markets make better predictions of their demand levels. We also show that the number and mean productivity of neighbors affects firms' entry and exit decisions regarding FDI and exports.

The remainder of this paper is organized as follows. Section 2 documents the pattern that motivates our research. Section 3 presents our theoretical model to explain our novel finding. Section 4 introduces the data and Section 5 presents the results of our empirical analysis. Section

6 concludes the paper.

2 Experience of FDI entrants

Table 1: Previous export and FDI experience of Japanese FDI entrants before entry

	Export experience > 0 (to destination region)	Export experience = 0 (to destination region)	Total
Existing affiliated firm > 0 (in destination region)	5,443	1,699	7,142
	76%	24%	100%
Existing affiliated firm = 0 (in destination region)	2,466	3,387	5,853
	42%	58%	100%
Total	7,909	5,086	12,995
	61%	39%	100%

Source: Authors' calculation based on the Basic Survey on Overseas Business Activities and the Basic Survey of Japanese Business Structure and Activities from 2003 to 2016, which are taken from the Ministry of Economy, Trade, and Industry (METI) in Japan.

Note: The table presents the number and share of Japanese firms that established foreign affiliates by whether they have previous export or FDI experience in destination regions. Firms which entered Central and South America, Africa, and Oceania after 2010 are excluded because of data limitations.

We document a novel fact about the previous experience of FDI entrants. Table 1 shows that 61% of Japanese FDI entrants have previous export experience to destination regions in total, which is almost the same with Deseatnicov and Kucheryavyy (2017). Conconi et al. (2016) report that about 90% of Belgian firms have export experience before establishing subsidiaries. Conconi et al. (2016) and Deseatnicov and Kucheryavyy (2017) state that firms start the internationalization process by exporting in order to learn about foreign market demand. Specifically, Conconi et al. (2016) theoretically shows that firms conduct export experiments before making investment decisions because they face demand uncertainty in foreign markets.

However, we need to reanalyze the percentage of FDI entrants that have previous export experience. This is because when we analyze whether a specific firm conducts export experiments before FDI entry, we must first identify whether the firm has existing affiliated companies in the destination market. If so, the firm does not need to test the market demand since it already knows the demand through its existing affiliates. In fact, among firms that already have affiliated firms in destination regions, 76% have export experience before additional FDI entry. We should not consider them as export experiments undertaken to test potential demand in destination markets but consider them as the intra-firm trade to operate its existing affiliated companies. Meanwhile, among those that do not have existing affiliated firms in destination regions, only 42% have previous export experience. This is a much lower share than that stated in the literature[2].

Thus, in contrast to the literature, over half of firms start FDI without export experience when we analyze firms that entered specific foreign regions for the first time, which are destination regions where the firms do not have existing affiliated firms. However, why do many firms start their internationalization process via FDI and pay the large sunk costs of entry without learning about the demand levels through export experiments? Section 3 theoretically explains why these firms can start FDI without export experience even though they face demand uncertainty.

3 Model

This section develops a tractable model of firms' internationalization decisions under demand uncertainty. The structure of firms' demand-learning behavior is based on Fernandes and Tang (2014), in which firms ex-ante learn their demand levels by observing their neighbors' export performance. The definition of neighbors in Fernandes and Tang (2014) is firms that operate in the same city and export to the same country. They assume that firms can observe three signals from their neighbors: (1) the number of neighbors; (2) the conditional mean of neighbors' productivity; (3) the average neighbors' export revenue to the destinations. These three signals from neighbors are used to ex-ante learn the demand level in the destination market.

Our model extends the social learning mechanisms of Fernandes and Tang (2014) by relaxing the assumption about the signals from neighbors in their model. In particular, we do not assume that firms can observe the average neighbors' export revenue to the destination markets. This is because in reality, it is difficult to obtain information about the sales of export and affiliates of other firms in a specific foreign market. The firm-destination level export value is not accessible since the transaction data of customs is confidential. Moreover, in general, firms do not even know whether other firms export to a specific foreign market or not. Hence, our model does not assume that firms can observe other firms' sales of exports or sales of their affiliated companies in foreign markets.

Instead, we make two assumptions about the information about neighbors: (1) firms know the number of neighbors operating affiliated firms in destination markets. (2) Firms also know the conditional mean of neighbors' productivity in the home country. Note that we do not assume that firms can observe the average sales of neighbors' affiliates in destinations. The fact that firms usually publish lists of the countries in which they have foreign affiliates on their websites makes our first assumption more reliable. Moreover, in order to justify our second assumption, firms need to know the productivity level of their neighbors. Thus, the definition of neighbors in our model is restricted to firms that operate in the same industry of the home

[2] As reported in Table A.1 and A.2 in Appendix A, restricting our sample to FDI in specific regions or industries makes minimal change in our results.

country and have affiliates in destination markets. Thus, a firm's neighbors are its competitors in the same industry in the home market. As firms often gather information about competitors to survive in the home market, it is natural to assume that firms with knowledge about their own markets know the average productivity of competitors in their home market. The detail of the demand-learning system is discussed in Section 3.4.

Furthermore, another extension of Fernandes and Tang (2014) is that we incorporate the dynamics of firms' internationalization choices between exports and FDI to our model. This allows us to demonstrate how demand-learning from neighbors affects firms' entry and exit decisions of FDI and explain why many firms start serving a foreign market via FDI without export experience. In what follows, we describe the details of the model.

3.1 Setup

Our model focuses on a simple, two-period structure. Without a loss of generality, we assume that the firm does not discount the future. Firms have heterogeneous productivity and constant-elasticity-of-substitution (CES) preferences and operate under monopolistic competition, following Melitz (2003).

We consider firm i that produces a variety in the domestic market. Firm i knows its own productivity ρ_i, which is drawn from a cumulative distribution function $G(\rho)$. Each firm decides whether to advance in a foreign market m and whether to do so via exports or establishing a foreign affiliate firm.

3.2 Demand

As in Fernandes and Tang (2014), firm i's demand in market m is given by

$$q_{im}^{f} = (p_{im}^{f})^{-\sigma} \exp(d_m + z_{im}), \qquad f \in \{I, E\}, \tag{1}$$

where p_{im}^{f} is the price of the product of firm i supplied through f in market m. The set of f has I and E, which corresponds to foreign direct investment (FDI) and exports, respectively. $\sigma > 1$ is the elasticity of substitution between varieties.

The market-specific demand parameter is $d_m = \ln(P_m^{\sigma} Y_m)$, where P_m and Y_m are the ideal price index and total expenditure in market m, respectively. However, firm i does not know the true value of d_m. The prior belief of d_m is normally distributed with mean $\bar{d}_m$ and variance v_{dm}: $d_m \sim \mathcal{N}(\bar{d}_m, v_{dm})$. For example, a large $\bar{d}_m$ means that there is a large demand in the market on average, and a large v_{dm} means large uncertainty about the market demand.

The firm-market-specific demand parameter is z_{im}, which is not known to a firm that enters a market. As the firm does not know the true value of z_{im}, its prior belief follows a normal distribution with mean zero and variance v_{zm}: $z_{im} \sim \mathcal{N}(0, v_{zm})$.

Before entering market m, firm i faces uncertainty about both d_m and z_{im} and knows only their prior distributions. Thus, the firm must make its internationalization decision without knowing its true values of d_m and z_{im}. If the firm enters the market m via exports or FDI, the firm can learn the true values of its demand parameters. After entry, the firm makes post-entry decisions based on the information about their true values. For example, a firm that entered via exports in the first period makes a choice between continuing exports, switching to FDI, and exiting from the market in the second period, depending on the true values of its demand.

3.3 Profit function and entry thresholds

Each firm is a monopolistic competitor and chooses from establishing an affiliated firm in market m, exporting to the market, or staying in the home market. When it decides to start the internationalization process, it incurs two types of per-unit cost of production: C^I_{im} when it serves via FDI and produces goods in market m and C^E_i when it produces goods domestically and serves via exports. The per-unit cost of production via FDI is written as $C^I_{im} = \zeta_m c^I_{im}$, where $\zeta_m \geq 1$ is the market-specific penalty in terms of lost productivity associated with the transfer of operational methods from parent to the affiliate's country[3]. c^I_{im} is firm i's production cost in a foreign country m.

The firm also faces an iceberg transportation cost $\tau_m > 1$ only when it exports goods to market m. The CES preferences and the familiar constant mark-up rule derive the prices of goods.

The price of goods of firm i when it is produced in market m is,

$$p^I_{im} = \frac{\sigma}{\sigma - 1} \frac{C^I_{im}}{\rho_i}, \tag{2}$$

and the price of goods produced domestically and exported to market m is,

$$p^E_{im} = \frac{\sigma}{\sigma - 1} \frac{\tau_m C^E_i}{\rho_i}, \tag{3}$$

where ρ_i is firm i's productivity in the home market.

Firm i also incurs a sunk cost of entry K^I_{im} when it establishes an affiliate in market m, but the firm can bypass transportation cost τ_m. Similarly, the firm bears a sunk cost of export K^E_{im}

[3]Recent literature, including Tintelnot (2017) and Head and Mayer (2019), report that parents are on average more productive than their foreign affiliates. Our setup follows Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple (2018). ζ_m includes lack of technology transfers and information frictions between workers in the home country and the affiliate's country. Regarding Japanese multinational firms, Lu, Tomiura, and Zhu (2020) report that there is a substantial difference in affiliates' productivity across destination countries and parents' technology transfers and distance from headquarters' country are important factors to determine their affiliates' productivity.

at the initial period of export entry. As with the literature, we impose a standard assumption of the relative magnitudes of these costs[4] for the proximity-concentration tradeoff to arise.

Assumption 1. $\frac{K_{im}^E}{K_{im}^I} \leq \left(\frac{\tau_m C_i^E}{C_{im}^I}\right)^{1-\sigma} \leq 1,$ *for all i and m*

Assumption 1 ensures that FDI is not always more profitable than exports because of larger sunk entry costs even though the marginal production cost of FDI $\frac{C_{im}^I}{\rho_i}$ is lower than that of exports $\frac{\tau_m C_i^E}{\rho_i}$.

When firm i starts FDI in market m, its initial profit function is

$$\pi_{im}^I(\rho_i, d_m, z_{im}) = \left[\frac{\sigma}{\sigma-1}\frac{C_{im}^I}{\rho_i}\right]^{1-\sigma} \exp(d_m + z_{im}) - K_{im}^I. \tag{4}$$

Similarly, the profit function of starting exports is

$$\pi_{im}^E(\rho_i, d_m, z_{im}) = \left[\frac{\sigma}{\sigma-1}\frac{\tau_m C_i^E}{\rho_i}\right]^{1-\sigma} \exp(d_m + z_{im}) - K_{im}^E. \tag{5}$$

As firm i faces demand uncertainty before entry and only has knowledge about the prior distribution of d_m and z_{im}, its expected profits of FDI and exports are

$$\begin{aligned} E[\pi_{im}^I(\rho_i)] &= \left[\frac{\sigma}{\sigma-1}\frac{C_{im}^I}{\rho_i}\right]^{1-\sigma} E[\exp(d_m + z_{im})] - K_{im}^I \\ &= \left[\frac{\sigma}{\sigma-1}\frac{C_{im}^I}{\rho_i}\right]^{1-\sigma} \exp(\bar{d}_m + \frac{v_m}{2}) - K_{im}^I, \end{aligned} \tag{6}$$

and

$$E[\pi_{im}^E(\rho_i)] = \left[\frac{\sigma}{\sigma-1}\frac{\tau_m C_i^E}{\rho_i}\right]^{1-\sigma} \exp(\bar{d}_m + \frac{v_m}{2}) - K_{im}^E, \tag{7}$$

where $v_m = v_{dm} + v_{zm}$.

From (6) and (7), the cutoff productivity of starting FDI and export is

$$\tilde{\rho}_{im}^I(\bar{d}_m, v_m) = \left[\frac{K_{im}^I - K_{im}^E}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left[(C_{im}^I)^{1-\sigma} - (\tau_m C_i^E)^{1-\sigma}\right]\exp\left(\bar{d}_m + \frac{v_m}{2}\right)}\right]^{\frac{1}{\sigma-1}}, \tag{8}$$

[4]For example, Horstmann and Markusen (1992), Brainard (1993), Brainard (1997), Helpman et al. (2004), and Ramondo et al. (2013).

and

$$\tilde{\rho}_{im}^{E}(\bar{d}_m, v_m) = \left[\frac{K_{im}^{E}}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (\tau_m C_i^E)^{1-\sigma} \exp\left(\bar{d}_m + \frac{v_m}{2}\right)} \right]^{\frac{1}{\sigma-1}}. \tag{9}$$

Under the assumption, we have $\tilde{\rho}_{im}^{E} \leq \tilde{\rho}_{im}^{I}$ for any i and m. Firms whose productivity ρ_i is $\tilde{\rho}_{im}^{E} \leq \rho_i < \tilde{\rho}_{im}^{I}$ start export and those with $\tilde{\rho}_{im}^{I} \leq \rho_i$ start FDI.

3.4 Learning from neighbors

An important feature of our model is firms' demand-learning from its neighbors. The pioneering work of this mechanism is Fernandes and Tang (2014). As discussed above, our model relaxes the assumption of Fernandes and Tang (2014) in order to make the model more realistic.

We assume that firm i knows the number of neighbors operating affiliated firms in market m at $t-1$, $n_{m,t-1}$[5]. In addition, firm i is assumed to know the time-varying conditional mean of neighbors' productivity in the home market, $\hat{\rho}_{m,t-1}^{nb} = E[\rho | \rho_i \geq \tilde{\rho}_{im,t-1}^{I}]$, where $\tilde{\rho}_{im,t-1}^{I}$ is the productivity threshold for FDI entry at $t-1$. Using information about $n_{m,t-1}$ and $\hat{\rho}_{m,t-1}^{nb}$, the firm can update its belief about the market demand.

The objective of the learning process is d_m. As the firm ex-ante knows the distribution of d_m, the distribution serves as a prior belief about its market demand. The learning process begins with calculating the minimum expected demand level $\underline{d}_{m,t-1}^{nb}$ using the time-varying conditional mean of neighbors' productivity in the home market, $\hat{\rho}_{m,t-1}^{nb}$. Deriving the minimum expected demand level enables the firm to truncate the prior distribution with the truncation point $\underline{d}_{m,t-1}^{nb}$. Then, the firm obtains the mean of the left-sided truncated normal distribution $\hat{d}_{m,t-1}^{nb}$, which is the expected demand level inferred from $\hat{\rho}_{m,t-1}^{nb}$.

Finally, the firm updates its prior belief $\bar{d}_m$ based on the weighted $\hat{d}_{m,t-1}^{nb}$. The firm puts more weight on $\hat{d}_{m,t-1}^{nb}$ when there are more neighbors. This is because the reliability of the inferred expected demand level $\hat{d}_{m,t-1}^{nb}$ depends on the number of signals the firm obtains from its neighbors, which is the number of neighbors. The learning process described so far yields the posterior belief $\bar{d}_{mt}^{post}(n_{m,t-1}, \hat{d}_{m,t-1}^{nb})$.

The detailed process of belief updating is described in the following subsections.

[5]Though the location of firms' affiliates is usually published on the websites, the destination of export is very rarely published. Thus, we do not assume that firms can get information about neighbors' exports.

3.4.1 Demand level inferred from neighbors' information

Firm i's belief updating starts from inferring the minimum expected demand level $\underline{d}^{nb}_{m,t-1}$. Using $\hat{\rho}^{nb}_{m,t-1}$, firm i infers $\exp(\underline{d}^{nb}_{m,t-1})$ such that $E[\pi^I_{im}(\hat{\rho}^{nb}_{m,t-1})] = 0$:

$$\exp(\underline{d}^{nb}_{m,t-1}) = \frac{K^I_{im}}{\left(\frac{\sigma}{\sigma-1}\frac{C^I_{im}}{\hat{\rho}^{nb}_{m,t-1}}\right)^{1-\sigma}\exp\left(\frac{v_{zm}}{2}\right)}. \tag{10}$$

Taking the natural log of (10) yields

$$\underline{d}^{nb}_{m,t-1} = \ln K^I_{im} - (1-\sigma)\ln\left(\frac{\sigma}{\sigma-1}\frac{C^I_{im}}{\hat{\rho}^{nb}_{m,t-1}}\right) - \frac{v_{zm}}{2}. \tag{11}$$

Then, using the minimum expected demand level $\underline{d}^{nb}_{m,t-1}$, firm i truncates the range of the prior distribution of d_m. It derives the mean of the left-sided truncated normal distribution from the prior belief of d_m and the truncation point $\underline{d}^{nb}_{m,t-1}$:

$$\begin{aligned}\hat{d}^{nb}_{m,t-1} &= E\left[d_m | d_{m,t-1} \geq \underline{d}^{nb}_{m,t-1}\right] \\ &= \bar{d}_m + \frac{\sqrt{v_{d_m}}\phi(X)}{1-\Phi(X)},\end{aligned} \tag{12}$$

where $X = \frac{\underline{d}^{nb}_{m,t-1} - \bar{d}_m}{\sqrt{v_{d_m}}}$ and ϕ denotes the probability density function of the standard normal distribution and Φ the corresponding cumulative distribution function.

(12) is the demand level inferred from the conditional mean of the neighbors' productivity $\hat{\rho}^{nb}_{m,t-1}$. In order to demonstrate this first learning process more concretely, we show the derivations of $\underline{d}^{nb}_{m,t-1}$ and $\hat{d}^{nb}_{m,t-1}$:

$$\frac{\partial \underline{d}^{nb}_{m,t-1}}{\partial \hat{\rho}^{nb}_{m,t-1}} < 0; \qquad \frac{\partial \hat{d}^{nb}_{m,t-1}}{\partial \hat{\rho}^{nb}_{m,t-1}} < 0.$$

The higher mean of the neighbors' productivity $\hat{\rho}^{nb}_{m,t-1}$ implies that firms with lower productivity do not operate in market m. It further indicates that the market-specific demand d_m is more likely to be lower. This is because, if d_m was high enough for unproductive firms to be able to operate in the market, there would be more unproductive firms and the mean productivity of neighbors would be lower in market m. Thus, the higher $\hat{\rho}^{nb}_{m,t-1}$ leads to the lower $\underline{d}^{nb}_{m,t-1}$ and $\hat{d}^{nb}_{m,t-1}$.

3.4.2 Posterior demand

Based on (12), firm i updates its prior belief about the demand parameter in the manner proposed by DeGroot (2005):

$$\begin{aligned} \bar{d}_{mt}^{post}(n_{m,t-1}, \hat{d}_{m,t-1}^{nb}) &= E\left[d_{mt}|n_{m,t-1}, \hat{d}_{m,t-1}^{nb}\right] \\ &= \delta_t \hat{d}_{m,t-1}^{nb} + (1-\delta_t)\bar{d}_m, \end{aligned} \tag{13}$$

where δ_t is the weight the firm places on information from neighbors $\bar{d}_{m,t-1}^{nb}$ when updating its belief. Again, following DeGroot (2005), δ_t is derived as follows:

$$\begin{aligned} \delta_t(n_{m,t-1}, v_{dm}, v_{zm}) &= \frac{n_{m,t-1}v_{dm}}{v_{zm} + n_{m,t-1}v_{dm}} \\ &= \left(1 + \frac{1}{n_{m,t-1}}\frac{v_{zm}}{v_{dm}}\right)^{-1}. \end{aligned} \tag{14}$$

Given $n_{m,t-1}$, v_{dm}, v_{zm}, and $\bar{d}_{mt}^{post}$, the conditional variance of posterior belief d_{mt}^{post} is,

$$\begin{aligned} v_{mt}^{post}(n_{m,t-1}, v_{dm}, v_{zm}) &= \frac{v_{zm}v_{dm}}{v_{zm} + n_{m,t-1}v_{dm}} \\ &= \left(\frac{1}{v_{dm}} + \frac{n_{m,t-1}}{v_{zm}}\right)^{-1}. \end{aligned} \tag{15}$$

As with Fernandes and Tang (2014), we can interpret how the number of neighbors affects the belief updating and the posterior belief using the partial differentiation of (14) and (15):

$$\frac{\partial \delta_t}{\partial n_{m,t-1}} > 0; \quad \frac{\partial v_{mt}^{post}}{\partial n_{m,t-1}} < 0. \tag{16}$$

The first inequality implies that a firm places larger weight on information from its neighbors if there are more neighbors in the destination market. According to the second inequality, the firm obtains a more precise posterior belief when it has more neighbors operating in the market.

3.5 Effect of the mean of neighbors' productivity on entry thresholds

Given the posterior mean $\bar{d}_{mt}^{post}$ and variance v_{mt}^{post} defined in (14) and (15), the posterior productivity thresholds of entry for FDI and exports are

$$\tilde{\rho}_{im}^{I}(\bar{d}_{mt}^{post}, v_{mt}^{post}) = \left[\frac{K_{im}^{I} - K_{im}^{E}}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left[(C_{im}^{I})^{1-\sigma} - (\tau_m C_i^E)^{1-\sigma}\right]\exp\left(\bar{d}_{mt}^{post} + \frac{v_{mt}^{post}}{2}\right)}\right]^{\frac{1}{\sigma-1}}, \tag{17}$$

and

$$\tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) = \left[\frac{K^E_{im}}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (\tau_m C^E_i)^{1-\sigma} \exp\left(\bar{d}^{post}_{mt} + \frac{v^{post}_{mt}}{2}\right)} \right]^{\frac{1}{\sigma-1}}. \tag{18}$$

Before analyzing the impact of neighbors' average productivity on entry thresholds, we begin by taking the derivative of posterior demand (13) with respect to the conditional mean of neighbors' productivity $\hat{\rho}^{nb}_{m,t-1}$ as follows:

$$\frac{\partial \bar{d}^{post}_{mt}}{\partial \hat{\rho}^{nb}_{m,t-1}} = (1-\sigma)\delta_t \frac{\phi(X)}{\hat{\rho}^{nb}_{m,t-1}} \frac{\phi(X) - X[1-\Phi(X)]}{[1-\Phi(X)]^2} < 0, \tag{19}$$

where $X = \frac{d^{nb}_{m,t-1} - \bar{d}_m}{\sqrt{v_{dm}}}$ as defined above. The higher mean of the neighbors' productivity reduces the posterior demand level[6]. This is because the higher mean productivity implies that the demand level inferred from the signals of neighbors, $\hat{d}^{nb}_{m,t-1}$, is lower, as discussed in Section 3.4.1. The lower $\hat{d}^{nb}_{m,t-1}$ results in lower $\bar{d}^{post}_{mt}$ because of the belief updating equation (13).

We now analyze how learning from neighbors' $\hat{\rho}^{nb}_{m,t-1}$ affects firms' entry decisions. First, we derive the partial derivatives of (17) and (18) with respect to neighbors' mean productivity $\hat{\rho}^{nb}_{m,t-1}$:

$$\frac{\partial \tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial \hat{\rho}^{nb}_{m,t-1}} = \frac{\partial \tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial \hat{\rho}^{nb}_{m,t-1}} = \delta_t \frac{\phi(X)}{\hat{\rho}^{nb}_{m,t-1}} \frac{\phi(X) - X[1-\Phi(X)]}{[1-\Phi(X)]^2} > 0. \tag{20}$$

The inequality shows that a higher conditional mean of neighbors' productivity increases the cutoff productivity of FDI and exports by reducing the posterior belief of the market-specific demand in (19). In the inverse,, a decrease of the mean of neighbors' productivity lowers the threshold of starting FDI and exports.

To be more specific, consider a firm i that has $n_{m,t-2}$ neighbors in a destination market m and the productivity of which is lower than the threshold of market m at $t-1$: $\rho_i < \tilde{\rho}^I_{im}(\bar{d}^{post}_{m,t-1}, v^{post}_{m,t-1})$. In other words, firm i does not start FDI in market m at $t-1$ and chooses to wait and see[7]. If some unproductive firms start FDI in market m at $t-1$, the conditional mean of neighbors' productivity $\hat{\rho}^{nb}_{m,t-1}$ decreases at $t-1$, which results in the decline of the entry threshold at t: $\tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) < \tilde{\rho}^I_{im}(\bar{d}^{post}_{m,t-1}, v^{post}_{m,t-1})$. As a result, when the productivity of firm i is higher than

[6]In (19), $\phi(X) - X[1-\Phi(X)] > 0$ using the property of Mills' ratio proved by Gordon (1941).

[7]For simplicity, this example ignores the possibility that firm i starts exporting instead of FDI. The likelihood of starting exports instead of FDI is discussed in Section 3.7.

$\tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$, firm i starts FDI in market m at t. Thus, a decline of the mean of neighbors' productivity at $t-1$ promotes FDI entry of firms at t, which did not enter at $t-1$[8].

Additionally, we examine the effect of the number of neighbors in market m on the neighbors' effect of (20). The derivative of (20) with respect to the number of neighbors is

$$\frac{\frac{\partial \tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial \hat{\rho}^{nb}_{m,t-1}}}{\partial n_{mt-1}} = \frac{\frac{\partial \tilde{\rho}^{E}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial \hat{\rho}^{nb}_{m,t-1}}}{\partial n_{mt-1}} = \frac{\frac{\partial \tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial \hat{\rho}^{nb}_{m,t-1}}}{\partial \delta_t} \frac{\partial \delta_t}{\partial n_{m,t-1}} > 0, \tag{21}$$

using (16). The inequality (21) indicates that the increment of neighbors strengthens the effect of the mean of neighbors' productivity on the entry threshold only through changing the weights a firm places on its neighbors' information in its belief updating. That is to say, the increase of neighbors improves the reliability of neighbors' information, which allows firms to place greater weight on the mean of neighbors' productivity and thus strengthens the neighbors' effect of (20).

Notice that the number of neighbors itself can affect not only the weight δ_t and the posterior mean $\bar{d}^{post}_{mt}$, but also the posterior variance v^{post}_{mt}. Although the effect on the posterior variance is not included in (21), it is discussed in Section 3.6.

These theoretical results about the effect of the mean of neighbors' productivity are summarized by the following proposition:

Proposition 1. *The likelihood of a firm's starting FDI and exports decreases with the strength of the mean of neighbors' productivity, and more so if there are more neighbors operating in the market. Conversely, the decline of the mean of neighbors' productivity motivates a firm to enter the foreign market and that effect is strengthened by the number of neighbors.*

3.6 Effect of the number of neighbors on entry thresholds

As discussed in Section 3.5, the number of neighbors has an indirect influence on the entry thresholds through strengthening the effect of the conditional mean of neighbors' productivity on the one hand. On the other hand, the prevalence of neighbors also has a direct effect on the productivity cutoff of FDI entry.

The elasticity of the entry threshold with respect to the number of neighbors is

$$\frac{\partial \tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial n_{mt-1}} = \frac{\partial \tilde{\rho}^{E}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})}{\partial n_{mt-1}} = \frac{v_{dm} v_{zm}}{1-\sigma} \left(\frac{1}{n_{m,t-1} v_{dm} + v_{zm}} \right)^2 \left(\hat{d}^{nb}_{m,t-1} - \bar{d}_m - \frac{v_{dm}}{2} \right). \tag{22}$$

[8]The productivity range of those firms is $\tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) \le \rho_i < \tilde{\rho}^{I}_{im}(\bar{d}^{post}_{m,t-1}, v^{post}_{m,t-1})$.

This elasticity can be positive or negative depending on the sign of $\hat{d}^{nb}_{m,t-1} - \bar{d}_m - \frac{v_{dm}}{2}$.

If the demand level inferred from the conditional mean of the neighbors' productivity, $\hat{d}^{nb}_{m,t-1}$, is higher than the mean of the prior belief, $\bar{d}_m$[9], and the difference is greater than or equal to $\frac{v_{dm}}{2}$: $\hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2}$; then (22) becomes negative[10]. The number of neighbors reduces the productivity cutoff of FDI entry. The underlying intuition is that when a firm obtains a more positive signal of the demand level $\hat{d}^{nb}_{m,t-1}$ inferred from its neighbors compared to the mean of its prior knowledge $\bar{d}_m$, the prevalence of neighbors makes the signal more reliable, which enables the firm to reduce the entry thresholds of FDI and exports.

By contrast, when the difference between $\hat{d}^{nb}_{m,t-1}$ and $\bar{d}_m$ is smaller than $\frac{v_{dm}}{2}$: $\hat{d}^{nb}_{m,t-1} - \bar{d}_m < \frac{v_{dm}}{2}$; then (22) becomes positive. The number of neighbors increases the entry thresholds of FDI and exports. That is, when the mean of the demand level inferred from neighbors' information is not much greater than the prior mean, improving the reliability of the neighbors' information does not have much impact on the cutoff productivity but the prevalence of neighbors reduces the variance of the posterior belief at the same time, which increases the entry thresholds of FDI.

Note that the number of neighbors reduces the variance of the posterior belief in both cases. Thus, the prevalence of neighbors decreases the risk of exits after over-optimistic entry. We discuss the effect of the number of neighbors on exit decisions later in Section 3.8.

The theoretical predictions of the effect of the number of neighbors are summarized as follows:

Proposition 2. *The number of neighbors in a destination market has different effects on the likelihood of entry via FDI and exports depending on the difference between the mean of the prior belief and the demand level inferred from the neighbors' information: (1) if $\hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2}$, then the number of neighbors has a positive direct effect on the likelihood of starting FDI and exports; (2) the number of neighbors has a negative direct effect on the likelihood of starting FDI and exports otherwise.*

3.7 Effect of demand-learning from neighbors on export experiments

This subsection explains the export experiment that a firm starts FDI after learning its demand via exports. According to Conconi et al. (2016) and Deseatnicov and Kucheryavyy (2017), export experiments are widely used by firms that start FDI. However, as we report in Section 2, over half of Japanese FDI entrants do not conduct export experiments when they first establish affiliated firms in the destination market. The effect of demand-learning from

[9] We always have $\hat{d}^{nb}_{m,t-1} \geq \bar{d}_m$. This is because belief updating in our model enables firms to truncate the left-tail of the prior distribution.

[10] Notice that $1 - \sigma < 0$ in (22).

neighbors provides an explanation for why the majority of firms start FDI without conducting export experiments.

Suppose that firm i knows the true values of d_m and z_{im} after entry via FDI or exports and it incurs the fixed operation cost K_{im} at the second period of entry, which is assumed to be equal for FDI and exports. We further assume that $K_{im} \leq K^E_{im} \leq K^I_{im}$.

The second-period profit function of continuing FDI for firms starting FDI in the first period is

$$\pi^{II}_{im}(\rho_i, d_m, z_{im}) = \left[\frac{\sigma}{\sigma - 1}\frac{C^I_{im}}{\rho_i}\right]^{1-\sigma} \exp(d_m + z_{im}) - K_{im}. \tag{23}$$

Similarly, the second-period profit function of continuing exports for firms starting exports in the first period is

$$\pi^{EE}_{im}(\rho_i, d_m, z_{im}) = \left[\frac{\sigma}{\sigma - 1}\frac{\tau_m C^E_i}{\rho_i}\right]^{1-\sigma} \exp(d_m + z_{im}) - K_{im}. \tag{24}$$

The productivity threshold of firms starting FDI after export entry is $\tilde{\rho}^{EI}_{im}(d_m, z_{im})$ such that $\pi^I_{im} = \pi^{EE}_{im}$ from (4) and (24):

$$\tilde{\rho}^{EI}_{im}(d_m, z_{im}) = \left[\frac{K^I_{im} - K_{im}}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left[(C^I_{im})^{1-\sigma} - (\tau_m C^E_i)^{1-\sigma}\right]\exp\left(d_m + z_{im}\right)}\right]^{\frac{1}{\sigma-1}}, \tag{25}$$

The productivity range of firms conducting export experiments before FDI entry depends on the relationship of the values between $\tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$, $\tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$, and $\tilde{\rho}^{EI}_{im}(d_m, z_{im})$ in (17), (18), and (25):

(1) If $\tilde{\rho}^{EI}_{im}(d_m, z_{im}) < \tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$, the productivity range of firms that switch from exports to FDI is $\tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) \leq \rho_i < \tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$;

(2) If $\tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) \leq \tilde{\rho}^{EI}_{im}(d_m, z_{im}) < \tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$, the productivity range of starting FDI after export entry is $\tilde{\rho}^{EI}_{im}(d_m, z_{im}) \leq \rho_i < \tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$;

(3) if $\tilde{\rho}^I_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) \leq \tilde{\rho}^{EI}_{im}(d_m, z_{im})$, no firm starts FDI after export entry.

In the following subsubsections, we analyze the effect of demand-learning from neighbors on the productivity range discussed in cases (1) and (2) above.

3.7.1 The case of $\tilde{\rho}^{EI}_{im}(d_m, z_{im}) < \tilde{\rho}^E_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt})$

This case corresponds to the situation in which the true demand is much higher than the expected demand predicted before export entry. Then, the export entrant can start FDI regardless of the value of its productivity. The productivity range of firms that switch from exports to FDI is defined as follows:

$$\tilde{D}_{im}(\bar{d}_{mt}^{post}, v_{mt}^{post}) = \tilde{\rho}_{im}^{I}(\bar{d}_{mt}^{post}, v_{mt}^{post}) - \tilde{\rho}_{im}^{E}(\bar{d}_{mt}^{post}, v_{mt}^{post})$$
$$= \left[\frac{1}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\exp\left(\bar{d}_{mt}^{post} + \frac{v_{m}^{post}}{2}\right)}\right]^{\frac{1}{\sigma-1}} \left[\left(\frac{K_{im}^{I} - K_{im}^{E}}{(C_{im}^{I})^{1-\sigma} - (\tau_m C_i^E)^{1-\sigma}}\right)^{\frac{1}{\sigma-1}} - \left(\frac{K_{im}^{E}}{(\tau_m C_i^E)^{1-\sigma}}\right)^{\frac{1}{\sigma-1}}\right]. \quad (26)$$

The semi-elasticities of $\tilde{D}_{im}(\bar{d}_{mt}^{post}, v_{mt}^{post})$ with respect to the conditional mean of neighbors' productivity and the number of neighbors are expressed as follows:

$$\frac{\partial \ln \tilde{D}_{im}(\bar{d}_{mt}^{post}, v_{mt}^{post})}{\partial \hat{\rho}_{m,t-1}^{nb}} > 0; \quad (27)$$

$$\frac{\partial \ln \tilde{D}_{im}(\bar{d}_{mt}^{post}, v_{mt}^{post})}{\partial n_{m,t-1}} = \begin{cases} \leq 0 & \text{if } \hat{d}_{m,t-1}^{nb} - \bar{d}_m \geq \frac{v_{dm}}{2} \\ > 0 & \text{otherwise.} \end{cases} \quad (28)$$

The effect of the conditional mean of neighbors' productivity on the productivity range of export experiments is positive. As stated in Proposition 1, the increase of the mean of neighbors' productivity is a negative signal about the market-demand. Thus, firms are more likely to conduct export experiments before FDI in order to avoid the risk of failure of FDI.

Two inequalities of (28) mean that the productivity range of conducting export experiments decreases with the number of neighbors if the mean of the demand level inferred from neighbors' information is much greater than the prior mean: $\hat{d}_{m,t-1}^{nb} - \bar{d}_m \geq \frac{v_{dm}}{2}$. According to Proposition 2, in this case, the number of neighbors strengthens the reliability of the positive signal from neighbors and promotes firms' entry via FDI, which results in the decline of export experiments. In contrast, if $\hat{d}_{m,t-1}^{nb} - \bar{d}_m < \frac{v_{dm}}{2}$, the number of neighbors prevents firms from making optimistic predictions about the market demands and then promotes export experiments to avoid the risk of FDI failure.

3.7.2 The case of $\tilde{\rho}_{im}^{E}(\bar{d}_{mt}^{post}, v_{mt}^{post}) \leq \tilde{\rho}_{im}^{EI}(d_m, z_{im}) < \tilde{\rho}_{im}^{I}(\bar{d}_{mt}^{post}, v_{mt}^{post})$

We next consider the case that some of the export entrants switch to FDI because their true demands are higher than their prediction. This is the most likely scenario. The analysis is similar to that of the case in Section 3.7.1.

The productivity range of switching from exports to FDI is

$$\tilde{D}_{im}^{EI}(\bar{d}_{mt}^{post}, d_m, v_{mt}^{post}, z_{im}) = \tilde{\rho}_{im}^{I}(\bar{d}_{mt}^{post}, v_{mt}^{post}) - \tilde{\rho}_{im}^{EI}(d_m, z_{im}). \tag{29}$$

The elasticities of $\tilde{D}_{im}^{EI}(\bar{d}_{mt}^{post}, d_m, v_{mt}^{post}, z_{im})$ with respect to $\hat{\rho}_{m,t-1}^{nb}$ and $n_{m,t-1}$ are derived as follows:

$$\frac{\partial \tilde{D}_{im}^{EI}(\bar{d}_{mt}^{post}, d_m, v_{mt}^{post}, z_{im})}{\partial \hat{\rho}_{m,t-1}^{nb}} > 0; \tag{30}$$

$$\frac{\partial \tilde{D}_{im}^{EI}(\bar{d}_{mt}^{post}, d_m, v_{mt}^{post}, z_{im})}{\partial n_{m,t-1}} = \begin{cases} \leq 0 & \text{if } \hat{d}_{m,t-1}^{nb} - \bar{d}_m \geq \frac{v_{dm}}{2} \\ > 0 & \text{otherwise.} \end{cases} \tag{31}$$

Thus, as in (27) and (28), the increase of the conditional mean of neighbors' productivity has a positive effect on the productivity range of the export experiment. In short, a firm is more likely to start its internationalization process by exports when the mean of its neighbors' productivity is higher in the destination market.

The inequality (31) also means that the prevalence of neighbors has a negative impact on the productivity range of the export experiment if $\hat{d}_{m,t-1}^{nb} - \bar{d}_m \geq \frac{v_{dm}}{2}$ and otherwise has a positive impact on that range.

Thus, we have the following proposition:

Proposition 3. *The strength of the conditional mean productivity of neighbors that operate affiliate firms in a destination market encourages a new entrant to conduct an export experiment before establishing its affiliate in the market. The number of neighbors decreases the export experiment before FDI entry if* $\hat{d}_{m,t-1}^{nb} - \bar{d}_m \geq \frac{v_{dm}}{2}$*. Otherwise, it increases the export experiment.*

3.8 Effect of demand-learning from neighbors on survival of firms' affiliates

Finally, we consider the effect of demand-learning from neighbors on the productivity range of firms that exit after entry.

As we always have $\pi_{im}^{E}(\rho_i, d_m, z_{im}) \leq \pi_{im}^{II}(\rho_i, d_m, z_{im})$ from Assumption 1 and the assumption of the second-period fixed operation cost, firms that started FDI never switch to export.

Thus, FDI entrants exit only when $\pi_{im}^{II}(\rho_i, d_m, z_{im}) < 0$. This is the case that the true demand is extremely lower than the expected demand and the operation cost is higher than the revenue from operating in the foreign market. The productivity threshold of continuing FDI at the second period is

$$\tilde{\rho}^{II}_{im}(d_m, z_{im}) = \left[\frac{K_{im}}{\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (C^{I}_{im})^{1-\sigma} \exp(d_m + z_{im})} \right]^{\frac{1}{\sigma-1}}. \tag{32}$$

From (17) and (32), a firm starts FDI at the first period and exits at the second period only when its productivity is $\tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}) \leq \rho_i < \tilde{\rho}^{II}_{im}(d_m, z_{im})$. The productivity range of firms' exit from FDI is defined as follows:

$$\tilde{D}^{Iexit}_{im}(\bar{d}^{post}_{mt}, d_m, v^{post}_{mt}, z_{im}) = \tilde{\rho}^{II}_{im}(d_m, z_{im}) - \tilde{\rho}^{I}_{im}(\bar{d}^{post}_{mt}, v^{post}_{mt}). \tag{33}$$

The elasticities of (33) with respect to the conditional mean of neighbors' productivity and the number of neighbors yield

$$\frac{\partial \tilde{D}^{Iexit}_{im}(\bar{d}^{post}_{mt}, d_m, v^{post}_{mt}, z_{im})}{\partial \hat{\rho}^{nb}_{m,t-1}} < 0, \tag{34}$$

and

$$\frac{\partial \tilde{D}^{Iexit}_{im}(\bar{d}^{post}_{mt}, d_m, v^{post}_{mt}, z_{im})}{\partial n_{m,t-1}} = \begin{cases} \geq 0 & \text{if } \hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2} \\ < 0 & \text{otherwise.} \end{cases} \tag{35}$$

The effect of the conditional mean of neighbors' productivity on the productivity range of exit is negative. As stated in Propositions 1 and 3, the increase of the mean of neighbors' productivity makes a firm more pessimistic about its market demand, which encourages the firm to conduct export experiments before establishing an affiliate. Thus, the decline of firms that started FDI based on optimistic expectations about the market demand results in the decrease of exits.

Two inequalities of (35) means that the productivity range of exits increases with the number of neighbors if $\hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2}$. From Proposition 2, in this case, the number of neighbors promotes firms' optimistic entry via FDI and thus increases firms' exits. In contrast, if $\hat{d}^{nb}_{m,t-1} - \bar{d}_m < \frac{v_{dm}}{2}$, the number of neighbors prevents firms from making optimistic predictions about the market demands and then reduces exits as discussed in Section 3.6.

The effects of demand-learning on exit decisions are summarized as follows:

Proposition 4. *The conditional mean of neighbors' productivity increases the likelihood of the survival of a firm that entered a foreign market via FDI. The number of neighbors reduces the likelihood of the firm's survival if* $\hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2}$. *In contrast, if* $\hat{d}^{nb}_{m,t-1} - \bar{d}_m < \frac{v_{dm}}{2}$, *the prevalence of neighbors increases the likelihood of the firm's survival.*

3.9 Discussion on the effects of demand-learning from neighbors

Our model aims to provide insights into the dynamics of firms' internationalization decisions from the perspective of demand uncertainty and learning from neighbors. The main theoretical predictions of our model are the effect of demand-learning from the conditional mean of neighbors' productivity on decisions related to FDI and export entries, export experiments, and exit, which are stated in Propositions 1, 3, and 4. Even in situations in which firms cannot observe profits of other firms in foreign markets, they can learn market-specific demands from the mean of the neighbors' productivity in the home market.

In addition, we argue that the number of neighbors itself has opposite effects on firms' decisions, as stated in Propositions 2, 3, and 4, depending on the signs of $\hat{d}^{nb}_{m,t-1} - \bar{d}_m - \frac{v_{dm}}{2}$. Although we can not specify the values of $\hat{d}^{nb}_{m,t-1} - \bar{d}_m - \frac{v_{dm}}{2}$ of each country-industry pair precisely, we can state that the number of neighbors has a direct impact on the firms' entry decisions in foreign markets.

We test these theoretical predictions in Section 5.

4 Data

In our empirical analysis, we use a rich data set taken from the Ministry of Economy, Trade, and Industry (METI) in Japan[11]. The main data source is the Basic Survey on Overseas Business Activities ("foreign survey" hereafter), which includes information about foreign affiliates of Japanese parent companies. Our data set records the information about affiliate firms that began operating from 2003 to 2016. The condition of foreign affiliates recorded in the foreign activities survey is (1) an overseas firm in which a Japanese parent corporation has invested capital of 10% or more or (2) an overseas corporation funded over 50% by a subsidiary that is funded over 50% by a Japanese parent. The response rate is around 70%, and we assume no systematic difference between respondents and non-respondents. The foreign survey contains affiliates' information such as location, year of establishment, operation status, sales, employment, and intermediate inputs. More importantly, it records affiliate firms' sales forecasts, which enables us to analyze the sales forecast errors of affiliates firms. Chen et al. (2018) justify the reliability of the sales forecasts in the foreign survey by showing that the affiliates do not use naive rules to make their forecasts. In Section 5, we first test whether this sales forecast depends on the number of neighbors in the markets to justify our theoretical assumption that firms can learn from their neighbors to update posterior demand belief.

In our dataset, firm i's neighbors' affiliates are defined as Japanese-affiliated firms working in the same industry and destination country with firm i. We calculate the number of neighbors

[11]The dataset is provided by the Research Institute of Economy, Trade, and Industry (RIETI) as a part of the project "Exchange Rates and International Currency" led by Professor Eiji Ogawa.

by industry and destination country annually. Then, we calculate the time-varying conditional mean of neighbors' productivity $\hat{\rho}^{nb}_{m,t-1}$ by using parent-firm-level productivity, estimated by Levinsohn and Petrin (2003), and take its average by year, industry, and destination country.

In addition, we combine the foreign survey with the Basic Survey of Japanese Business Structure and Activities ("domestic survey" hereafter) by using the converter prepared by RI-ETI[12]. The domestic survey contains Japanese parents' regional-level exports to North America, Central and South America, Asia, Europe, the Middle East, Oceania, and Africa. We use this information to capture parent firms' export experience before starting FDI.

We also use the World Bank national account data to obtain the real gross domestic product (GDP) of host countries. The area of the destination country is obtained from Mayer and Zignago (2011).

As we focus on firms' internationalization process through which they expand their real business to foreign countries, we drop "tax havens" from our sample[13]. We also exclude affiliate firms established in Central and South America, Africa, or Oceania after 2010 because the domestic survey does not contain the data about exports to those regions after 2009 and we can not identify whether new FDI entrants in those regions have export experience before investment.

Table 2 provides the number of FDI entrants for each year. As mentioned in Section 2, over half of firms without FDI experience in the destination region started FDI without export experience to the destination region.

[12]We use Kikatsu-Kaiji converter, which enables us to combine parent firms in the domestic survey and their affiliate firms in the foreign survey by each year.

[13]Following Deseatnicov, Fujii, Saito, and Kucheryavyy (2020), we exclude Panama, Liberia, Cayman Islands, Virgin Islands, British Bermuda, Bahamas, Northern Mariana Islands, and the Netherlands Antilles, in the order of the number of affiliates.

Table 2: Statistics of FDI entrants

Year of establishment	FDI Experience> 0		FDI Experience= 0	
	Export Exp.> 0	Export Exp.= 0	Export Exp.> 0	Export Exp.= 0
2003	521	128	351	463
2004	561	153	324	428
2005	540	158	250	359
2006	450	145	231	291
2007	387	142	187	257
2008	334	115	164	226
2009	284	100	100	162
2010	396	133	178	234
2011	567	147	219	357
2012	512	158	206	303
2013	355	122	129	164
2014	285	78	73	98
2015	178	86	38	36
2016	73	34	16	9
Total	5,443	1,699	2,466	3,387

Note: This table summarizes the number of FDI entrants by their export experience and FDI experience in the destination region. The number of entrants in 2015 and 2016 is smaller than that in other years as most new entrants respond to the foreign survey two years after entry.

5 Empirical evidence

Before testing our theoretical propositions stated in Section 3, we provide direct evidence to justify our assumption that firms can learn their demand level to update posterior demand belief from the information of neighboring firms operating in the destination market.

In the foreign survey, each affiliate firm reports its sales forecast for the next year, which enables us to examine whether firms can predict their demand level more accurately when they have more neighbors. Based on Chen et al. (2018)'s econometric specification, we regress firms' absolute forecast errors on the number of neighbors per area. The specification is:

$$|FE^{log}_{ict}| = \beta_1 \ln(n_{cm,t-1}/Area_c) + \sum_{n=2}^{6} \beta_n Age_{ict} + \boldsymbol{\beta X_{ict}} + \alpha_i + \delta_m + \delta_c + \delta_t + \epsilon_{ict}, \quad (36)$$

where $FE^{log}_{ict} \equiv log(\frac{Sales_{ict}}{E_{t-1}[Sales_{ict}]})$. $Sales_{ict}$ and $E_{t-1}[Sales_{ict}]$ are firms' affiliates' realized sales and their sales forecasts made at $t-1$, respectively. We use the density of neighbors in destination market $n_{cm,t-1}/Area_c$, which is the number of neighbors in the destination industry and country divided by the area of the country, as our measure of the number of neighbors. Our model also contains affiliate firms' age Age_{ict}, affiliates' characteristics X_{ict}, and fixed effects of firm α_i, industry δ_m, destination country δ_c, and year δ_t.

Table 3 reports the results of the estimation. Columns (1)-(3) include all firms and columns (4)-(5) restrict the sample to firms without FDI experience in the destination country. The significant negative effect of Age_{ict} on $|FE^{log}_{ict}|$ is consistent with Chen et al. (2018). The variables of our interest are $\ln(n_{cm,t-1}/Area_c)$ and a dummy for export experience. The coefficient of the number of neighbors has a negative effect on firm's forecast error and the effect is stronger for firms without FDI experience. This means that firms, especially inexperienced firms, can predict their sales more accurately when they have more neighbors in destination markets. In addition, the absolute effect of firms' previous export experience on forecast error is insignificant for firms without previous FDI experience. As this specification focuses on firms that have entered destination markets for the first time, the insignificant coefficient of export experience may indicate that that learning demand level from export experience is not as strong as stated in the literature such as Conconi et al. (2016).

In short, we confirm that new entrants learn their demand level from their neighbors' affiliates in destination markets, which directly justify our theoretical assumption that firms can learn their demand level to update posterior demand belief from the information of neighboring firms operating in the destination market.

We now examine Propositions 1, 2 and 3 regarding the effects of demand-learning from neighbors on firms' internationalization decisions between FDI and exports. We restrict our sample to firms which established affiliated firms in specific destination regions for the first time. This is because if a firm has existing affiliated companies in the destination region, we can not identify whether the firm starts its internationalization process via FDI or exports. Thus, our analysis focuses on whether firms' first internationalization processes in specific regions start through FDI or exports.

Table 3: Neighbors' effect on forecast

Dep. Var: $\lvert FE_{ict}^{log} \rvert$	All Firms			Firms without FDI experience in c		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(n_{cm,t-1}/Area_c)$	-0.006	-0.008*	-0.007	-0.008*	-0.009*	-0.009*
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
$1(\text{Age}_{ict} = 2)$	-0.159***	-0.165***	-0.162***	-0.155***	-0.161***	-0.148***
	(0.019)	(0.022)	(0.022)	(0.026)	(0.029)	(0.029)
$1(\text{Age}_{ict} = 3)$	-0.240***	-0.249***	-0.242***	-0.234***	-0.246***	-0.228***
	(0.019)	(0.022)	(0.022)	(0.026)	(0.030)	(0.029)
$1(\text{Age}_{ict} = 4)$	-0.297***	-0.308***	-0.297***	-0.299***	-0.311***	-0.289***
	(0.019)	(0.022)	(0.021)	(0.026)	(0.030)	(0.029)
$1(\text{Age}_{ict} = 5)$	-0.315***	-0.332***	-0.315***	-0.318***	-0.342***	-0.315***
	(0.019)	(0.022)	(0.021)	(0.026)	(0.030)	(0.029)
$1(\text{Age}_{ict} \geq 6)$	-0.352***	-0.369***	-0.340***	-0.361***	-0.385***	-0.341***
	(0.019)	(0.021)	(0.021)	(0.026)	(0.029)	(0.028)
Dummy for FDI	-0.016**	-0.009	0.007			
experience in c	(0.007)	(0.008)	(0.008)			
Dummy for export		-0.022***	-0.021***		-0.011	-0.011
experience		(0.007)	(0.007)		(0.008)	(0.008)
$\ln(\text{GDP}_{ct})$		-0.122***	-0.123***		-0.157***	-0.148***
		(0.036)	(0.035)		(0.041)	(0.041)
$\ln(\text{Parent Employment}_t)$			0.007***			0.003
			(0.003)			(0.003)
$\ln(\text{Employment}_t)$			-0.040***			-0.039***
			(0.003)			(0.003)
Industry FE	Y	Y	Y	Y	Y	Y
Country FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
N	37830	31459	30364	26504	21493	20845
R^2	0.057	0.061	0.075	0.065	0.067	0.078

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Regressions in column 1-4 includes all firms. Those in columns 4-6 include only firms that have no existing affiliates in the destination country before entry. "Dummy for FDI experience in c" is a dummy variable indicating whether firms have existing affiliates in the destination country before entry. "Dummy for export experience" indicates whether firms have export experience in the region of the destination country before entry.

Although we include three choices ("Start FDI", "Start export", and "Stay in home market") in our discrete choice model, we cannot use the multinominal logit model because of the violation of the "independence of irrelevant alternatives" (IIA) assumption. We use the fixed-effects ordered logit model (Baetschmann, Staub, and Winkelmann (2015) and Baetschmann, Ballantyne, Staub, and Winkelmann (2020)), which controls the unobserved time-invariant heterogeneity of firms.

The ordered dependent variable is

$$Entry_{ict} = \begin{cases} 2 & \text{if } \kappa_{ic2} < y^*_{ict} \quad \text{(to start FDI)}, \\ 1 & \text{if } \kappa_{ic1} < y^*_{ict} \leq \kappa_{ic2} \quad \text{(to start export)}, \\ 0 & \text{if } y^*_{ict} \leq \kappa_{ic1} \quad \text{(to stay in home market)}. \end{cases} \tag{37}$$

where κ_{ic}s are the thresholds and the latent variable y^*_{ict} is

$$y^*_{ict} = \beta_1 \bar{\rho}^{nb}_{cm,t-1} + \beta_2 \ln(n_{cm,t-1}/Area_c) + \beta_3 \bar{\rho}^{nb}_{cm,t-1} \times \ln(n_{cm,t-1}/Area_c) + \boldsymbol{\beta X_{ict}} + \alpha_i + \delta_t + \varepsilon_{it}, \tag{38}$$

where $\bar{\rho}^{nb}_{cm,t-1}$ is the mean of the productivity of neighbors operating in destination market m at $t-1$. We estimate the productivity of firms from the parent data of the domestic survey based on the method of Levinsohn and Petrin (2003) and then compute the mean of those operating affiliates in foreign countries by each country, industry, and year. As mentioned in Section 3, we assume that firms can observe the number of neighbors and the mean productivity of neighbors in the destination country before entering into new foreign markets. In column (1) of Table 4, the coefficient of $\bar{\rho}^{nb}_{cm,t-1}$ is negative and significant, which is consistent with Proposition 1. Although this effect becomes insignificant when we include year fixed effects in column (2), it still has a negative and significant effect for manufacturing firms in columns (3) and (4). The higher mean productivity of neighbors has a negative effect on the entry of both FDI and exports as expected in Proposition 1. The marginal probability effect of $\bar{\rho}^{nb}_{cm,t-1}$ in column (4) at the average implies that 1% increase in the mean productivity of neighbors reduces the probability of starting FDI by 0.03% points and export by 0.01% points, which increases the likelihood of export experiment and is consistent with Proposition 3.

The absolute impact of the number of neighbors is much larger than that of the mean productivity in columns (3) and (4). This confirms the positive direct effect of the number of neighbors on the likelihood of entry. This result corresponds to the case of $\hat{d}^{nb}_{m,t-1} - \bar{d}_m \geq \frac{v_{dm}}{2}$ in Proposition 2. From column (4), 1% increases in $\ln(n_{cm,t-1}/Area_c)$ increase the probability of starting FDI by 2.1% points and exports by 1.2% points on average, which reduce the likelihood of export experiments as expected in Proposition 3.

The indirect effect of the number of neighbors through strengthening the mean productivity effect is also confirmed in columns (3) and (4). Although the sign of the coefficient is consistent with Proposition 1, the impact is small.

In sum, results from the fixed-effects ordered logit model support Proposition 1, 2, and 3 for manufacturing firms[14]. The reason that the coefficient of $\ln(n_{cm,t-1}/Area_c)$ is insignificant in columns (1) and (2) is that non-manufacturing firms may avoid competition with neighbors.

Table 4: Learning effect on entry

Dep. Var: $Entry_t$	All Firms		Manufacturing	
	(1)	(2)	(3)	(4)
mean of neighbors' productivity $\bar{\rho}^{nb}_{cm,t-1}$	-0.002*	0.001	-0.006***	-0.003*
	(0.001)	(0.001)	(0.002)	(0.002)
$\ln(n_{cm,t-1}/Area_c)$	0.054	0.000	0.228***	0.209***
	(0.041)	(0.054)	(0.066)	(0.078)
$\bar{\rho}^{nb}_{cm,t-1} \times \ln(n_{cm,t-1}/Area_c)$	-0.000	0.000	-0.001***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
ln(Productivity$_t$)	0.000**	0.000*	0.001	0.001
	(0.000)	(0.000)	(0.002)	(0.002)
ln(Parent Employment$_t$)	0.798***	0.337**	0.710***	0.312
	(0.175)	(0.168)	(0.251)	(0.265)
ln(GDP$_{ct}$)	8.663***	2.135***	7.758***	1.623*
	(0.442)	(0.669)	(0.535)	(0.829)
S.D. of neighbors' sales in m	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Firm-Country FE	Y	Y	Y	Y
Year FE		Y		Y
N	19694	19694	11296	11296
Log Likelihood	-4210.4461	-3887.1769	-2517.0203	-2347.1251

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Regressions in column 1-2 includes all firms. Those in column 3-4 include only manufacturing firms.

Finally, we test Proposition 4 using the conditional logit model. Our specification is

[14] As reported in Table A.3 in Appendix A, we estimate our specification in the manufacturing sector by regions. Although the coefficients of $\bar{\rho}^{nb}_{cm,t-1}$ and $\ln(n_{cm,t-1}/Area_c)$ are not significant in regions other than Asia because of its small sample size, the effects are consistent with our theoretical predictions.

$$\Pr(y_{ict}^{exit}) = \beta_1^{exit} \bar{\rho}_{cm,t-1}^{nb} + \beta_2^{exit} \ln(n_{cm,t-1}/Area_c) + \boldsymbol{\beta X_{ict}} + \alpha_i + \varepsilon_{it}, \quad (39)$$

where y_{ict}^{exit} is equal to one when an affiliate firm i exits from country c at year t and equal to zero otherwise. The result is reported in Table 5. The coefficient of the mean of the neighbors' productivity is negative and significant as expected in Proposition 4. When the mean productivity of neighbors is higher in the destination market, firms are less likely to exit. In contrast, the coefficient of the number of neighbors is significantly positive, which means that firms are more likely to exit from markets where there are many neighbors.

Firms have more optimistic predictions of their demand levels in destination markets in which there are more neighbors due to the direct effect stated in Proposition 2. This has a positive effect on the exit decision. On the contrary, the higher mean of neighbors' productivity makes FDI entrants more pessimistic about their demand. Thus, firms that have entered the market despite the pessimistic prediction of demand are less likely to exit.

Table 5: Learning effect on exit

Dep. Var: $Exit_t$	All Firms		Firms without FDI experience in c
	(1)	(2)	(3)
mean of neighbors' productivity $\bar{\rho}^{nb}_{m,t-1}$	-0.003***	-0.001*	-0.002**
	(0.000)	(0.000)	(0.001)
$\ln(n_{cm,t-1}/Area_c)$	0.431***	0.255**	0.301 **
	(0.052)	(0.100)	(0.124)
ln(Productivity$_t$)		-0.007***	-0.006*
		(0.001)	(0.003)
ln(Parent Employment$_t$)		-0.268	0.483
		(0.575)	(0.968)
ln(Employment$_t$)		-0.550***	-0.658***
		(0.171)	(0.241)
ln(GDP$_{ct}$)		82.620***	73.316***
		(6.683)	(7.292)
N	11537	6609	3803
Log Likelihood	-3438.2577	-316.2632	-200.11263

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Regressions in column 1-2 includes all firms. Those in column 3 include only firms that have no existing affiliates in the destination country before entry.

6 Conclusion

The demand-learning literature explains that most FDI entrants conduct export experiments in order to reduce demand uncertainty and learn about the demand level in the destination market. The larger sunk entry cost of FDI compared to that of exports makes firms test demand level in the destination market before making investments.

In contrast to the discussions present in the demand-learning literature, this study provides a novel fact that over half of firms start their first internationalization process through FDI, not exports. To explain this pattern in a manner consistent with the demand-learning theory, we develop a statistical decision-making model in which firms make decisions about entering foreign markets via exports or FDI under demand uncertainty. Our model incorporates the proximity-concentration tradeoff and the demand-learning mechanism through which firms predict their demand level based on information about the number and average productivity of their neighbors. We analytically show that a higher level of neighbors' productivity negatively affects firms' decisions of exports and FDI. This also increases the likelihood of export experiments before FDI entry and indicates that if a firm learns that the mean productivity of its neighbors in a specific market is low, the firm is more likely to start FDI without conducting export experiments. Our model also shows that the number of neighboring firms in the destination market itself has a positive impact on firms' entry decisions. Additionally, regarding firms' decisions to exit from FDI, the mean productivity of its neighbors has a negative impact on exit decisions according to our model.

We provide supporting evidence for the main predictions of the model using a rich micro-level dataset for Japanese multinational firms. We first test our theoretical assumption that firms can learn about the market demand level from information about their neighbors. We confirm that firms can make better predictions in markets in which firms have many neighbors. Then, using the fixed-ordered logit model, we show that the mean of the neighbors' productivity has a significantly negative effect on firms' decision regarding entry. Moreover, the number of neighbors encourages manufacturing firms to start the internationalization process and start FDI without export experiments. Thus, demand learning from firms' neighbors affects firms' entry decisions, especially in the manufacturing sector, as predicted by our model. Finally, we also report that the mean productivity of neighbors negatively affects firms' exit decision.

This study identifies the importance of ex-ante demand-learning in the firms' internationalization processes. The results indicate that in markets with a sufficient number of neighbors, firms can make appropriate demand forecast and entry decisions without governmental support. By contrast, when there are not a sufficient number of neighbors' affiliates in a destination market, firms can not make precise demand forecasts and inefficiencies may occur, such as missed opportunities due to the inability to appropriately forecast demand. When policy makers aim to encourage domestic firms to expand their business into foreign markets, policy makers need to support firms to invest in countries in which there is a small number of neighboring firms. This is because potential entrants cannot predict the level of demand in such markets well, which prevents the entry of firms with a sufficiently high true level of demand. Policy makers can support firms in collecting information about the destination country. On the contrary, in countries in which there are many existing neighboring firms, entrants can make entry decisions based on more accurate predictions of their demand level without the government's support. Thus, governments' support for overseas expansion should focus on markets in which firms do not have a sufficient amount of neighbors' information.

References

Albornoz, F., Calvo-Pardo, H. F., Corcos, G., & Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, *88*, 17–31.

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

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., & Yeaple, S. (2018). Innovation and production in the global economy. *American Economic Review*, *108*(8), 2128–2173.

Baetschmann, G., Ballantyne, A., Staub, K., & Winkelmann, R. (2020). feologit: A new command for fitting fixed-effects ordered logit models. *Stata Journal*, *20*(2), 253-275.

Baetschmann, G., Staub, K. E., & Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *178*(3), 685–703.

Bernard, A. B., Jensen, J. B., & Schott, P. K. (2009). Importers, exporters and multinationals: a portrait of firms in the us that trade goods. In *Producer dynamics: New evidence from micro data* (pp. 513–552). University of Chicago Press.

Brainard, S. L. (1993). *A simple theory of multinational corporations and trade with a trade-off between proximity and concentration.* National Bureau of Economic Research Cambridge, Mass., USA.

Brainard, S. L. (1997). An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *American Economic Review*, *87*(4), 520–544.

Chen, C., , Sun, C., & Zhang, H. (2022). Learning and information transmission within multinational corporations. , *European Economic Review*(143), 104016.

Chen, C., Senga, T., Sun, C., & Zhang, H. (2018). Uncertainty, imperfect information, and expectation formation over the firm's life cycle. (RIETI Discussion Paper Series 18-E-010).

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

DeGroot, M. H. (2005). *Optimal statistical decisions*. John Wiley & Sons.

Deseatnicov, I., Fujii, D., Saito, Y. U., & Kucheryavyy, K. (2020). Why do japanese mnes enter and exit foreign markets? (RIETI Discussion Paper Series 20-E-055).

Deseatnicov, I., & Kucheryavyy, K. (2017). Exports and fdi entry decision: Evidence from japanese foreign-affiliated firms. (RIETI Discussion Paper Series 17-E-036).

Fernandes, A. P., & Tang, H. (2014). Learning to export from neighbors. *Journal of International Economics*, *94*, 67–84.

Gordon, R. D. (1941). Values of mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *The Annals of Mathematical Statistics*, *12*(3), 364–366.

Head, K., & Mayer, T. (2019). Brands in motion: How frictions shape multinational production. *American Economic Review*, *109*(9), 3073–3124.

Helpman, E., Melitz, M. J., & Yeaple, S. R. (2004). Export versus fdi with heterogeneous firms. *American Economic Review*, *94*(1), 300–316.

Horstmann, I. J., & Markusen, J. R. (1992). Endogenous market structures in international trade (natura facit saltum). *Journal of International Economics*, *32*(1), 109-129.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, *50*(3), 649–670.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, *70*(2), 317–341.

Li, S. (2018). A structural model of productivity, uncertain demand, and export dynamics. *Journal of International Economics*, *115*, 1–15.

Lu, Y., Tomiura, E., & Zhu, L. (2020). Productivity of offshore affiliates and distance from headquarters: Evidence from affiliate-level data of japanese multinationals. (RIETI Discussion Paper Series 20-E-047).

Markusen, J. (1984). Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics*, *16*, 205–226.

Mayer, T., & Zignago, S. (2011). Notes on cepii's distances measures: The geodist database.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695–1725.

Ramondo, N., Rappoport, V., & Ruhl, K. J. (2013). The proximity-concentration tradeoff under uncertainty. *Review of Economic Studies*, *80*, 1582–1621.

Timoshenko, O. A. (2015). Product switching in a model of learning. *Journal of International Economics*, *95*, 233–249.

Tintelnot, F. (2017). Global production with export platforms. *The Quarterly Journal of Economics*, *132*(1), 157–209.

Appendix A

Table A.1: Statistics of FDI entrants by region of destination

Region of destination	FDI Experience> 0		FDI Experience= 0	
	Export Exp.> 0	Export Exp.= 0	Export Exp.> 0	Export Exp.= 0
Asia	4094	1151	1907	2653
	(42%)	(12%)	(19%)	(27%)
North America	506	223	239	367
	(38%)	(17%)	(18%)	(27%)
Europe	699	253	254	238
	(48%)	(18%)	(18%)	(16%)
Others	144	72	66	129
	(35%)	(18%)	(16%)	(31%)
Total	5,443	1,699	2,466	3,387

Note: This table summarizes the number of FDI entrants by their export experience and FDI experience in destination region. The percentage of FDI entrants of each group to whole FDI entrants in each region is shown in parentheses.

Table A.2: Statistics of FDI entrants in the manufacturing and non-manufacturing industry

Industry	FDI Experience> 0		FDI Experience= 0	
	Export Exp.> 0	Export Exp.= 0	Export Exp.> 0	Export Exp.= 0
Manufacturing	2665	555	1274	1465
	(45%)	(9%)	(21%)	(25%)
Non-manufacturing	2778	1144	1192	1922
	(39%)	(16%)	(17%)	(27%)
Total	5,443	1,699	2,466	3,387

Note: This table summarizes the number of FDI entrants by their export experience and FDI experience in destination regions. The percentage of FDI entrants of each group to whole FDI entrants in each industry category is shown in parentheses.

Table A.3: Learning effect on entry in the manufacturing sector by region

Dep. Var: $Entry_t$	Asia		Other regions	
	(1)	(2)	(3)	(4)
mean of neighbors' productivity $\bar{\rho}^{nb}_{cm,t-1}$	-0.006***	-0.003*	-0.003	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)
$\ln(n_{cm,t-1}/Area_c)$	0.220***	0.213**	0.256	0.207
	(0.074)	(0.088)	(0.165)	(0.237)
$\bar{\rho}^{nb}_{cm,t-1} \times \ln(n_{cm,t-1}/Area_c)$	-0.001***	-0.000**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
ln(Productivity$_t$)	-0.002	-0.003	0.001	0.003*
	(0.002))	(0.002)	(0.002)	(0.002)
ln(Parent Employment$_t$)	0.292	0.211	1.044**	0.711
	(0.301))	((0.301)	(0.528)	(0.538)
ln(GDP$_{ct}$)	7.503***	2.026**	19.577***	-2.796
	(0.538)	(1.028)	(2.860)	(5.913)
S.D. of neighbors' sales in m	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Firm-Country FE	Y	Y	Y	Y
Year FE		Y		Y
N	9211	9211	2085	2085
Log Likelihood	-1988.7614	-1865.8505	-498.16502	-453.53252

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.