Firm Size and Complementarity between Geography and Products

Yoko Shibuya*

Duke University, Fuqua School of Business

December 3, 2022
Preliminary

Abstract

I develop a structural model of heterogeneous multi-product firms operating in multiple markets. In the model, firms choose: how many products to develop (product scope), how many markets to enter (geographic scope), and which products to sell in the markets they enter. Their decisions on product and geographic scopes interact in two ways: (1) the more products firms develop and sell in each market, the more likely they are to enter more markets, and (2) the more markets they enter, the more likely they are to develop more products. I estimate the model to match moments in the Japanese barcode-level transaction data using the simulated method of moments. The counterfactual exercise suggests that eliminating either product or geographic scope underestimates firm size heterogeneity by 64% to 96% due to the complementarity between the two margins. Furthermore, I use the model to estimate the welfare implication of size-dependent policies, taking an actual Small and Medium Enterprise (SME) subsidy in Japan as an example.

*I am deeply indebted to Pete Klenow, Kyle Bagwell, Matthew Grant, and Takeo Hoshi for their invaluable guidance and support. I am grateful to Adrien Auclert, Nick Bloom, Chad Jones, Martin Schneider, Monika Piazzesi, Andrés Rodríguez-Clare, Christopher Tonetti, and participants in IO, international trade, and macro seminars at Stanford University, Bank of Canada, the Federal Reserve Board, and Duke University for their helpful discussions. I am grateful to Freeman Spogli Institute for International Studies and SIEPR at Stanford University for financial support for this project.

1. Introduction

In many countries, markets consist of a few large firms and many small firms. The same is true for the Japanese manufacturing industry, where small and medium enterprises (SMEs) account for 98.5% of the total number of firms, and the remaining 1.5% of large firms command over half of the market share.[1] What is the economic mechanism that generates this enormous firm size heterogeneity?[2] Answering this question is essential for quantifying models of heterogeneous firms and for understanding the welfare implication of size-dependent policies.

Early research on firm size heterogeneity focused on productivity differences across firms; higher-productivity firms could charge lower prices and occupy larger market shares.[3,4] Recently, the availability of firm- or product-level sales data allows researchers to point to two extensive margins that amplify productivity differences: product scope (the number of products firms sell) and geographic scope (the number of markets in which firms sell their products to consumers).[5] The mechanism of how each scope amplifies

[1]The numbers are from the 2016 Economic Census for Business Activity. In the manufacturing industry, small and medium enterprises are defined as firms with fewer than 300 employees.

[2]Throughout the paper, my focus is on manufacturing firms, not retailers. I use the term "firm size" to mean the total sales of the manufacturing firm.

[3]For example, see Jovanovic (1982), Hopenhayn (1992), Melitz (2003), Luttmer (2007), and Bartelsman et al. (2013).

[4]More broadly speaking, there are models in which firm productivity differences emerge as the equilibrium outcome of ex ante homogeneous firms. For example, Bagwell and Ramey (1994) feature a model where homogeneous firms choose the degree of advertisement as a mixed strategy. Ex post productivity of firms depends on the degree of advertising. Such equilibria can be purified if firms have small cost differences about which they are privately informed, leading to a framework close to the typical models of heterogeneous productivity firms. For example, see Bagwell and Wolinsky (2002).

[5]The role of product scope in firm size heterogeneity is documented in Arkolakis et al. (2010), Bernard et al. (2010), Eckel and Neary (2010), Bernard et al. (2011), Mayer et al. (2014), and Hottman et al. (2016). The role of geographic scope in determining firm size heterogeneity is documented in recent trade literature such as Eaton et al. (2011). More recently, Bernard et al. (2019) finds the number of connections to buyers carries explanatory power for firm size, which has a similar role for geographic scope.

firm size heterogeneity is intuitive. As higher-productivity firms sell more products—or sell their products in more markets—they become larger.

Despite the extent to which previous literature has considered the role of the extensive margins, it has typically examined only *one* extensive margin at a time.[6] These models tend to leave one-third to two-thirds of the firm size variation unexplained.[7] Furthermore, these models do not account for the ways in which firms' product and geographic scopes interact with each other.

In this paper, I develop a structural model that incorporates both extensive margins. This model highlights the quantitative importance of having *both* geographic and product scope to explain firm size heterogeneity and understand the welfare implications of size-dependent policies. Counterfactual exercises show that missing either of the extensive margins in the model results in a dramatic underestimation of firm size heterogeneity due to the absence of complementarity between the two margins.

I begin by documenting empirical patterns that motivate the structure of the model. I use barcode-level retail scanner data from the Japanese consumer packaged goods industry. Unlike the household scanner or export data, which most existing literature uses, the retail scanner data allows me to observe barcode-level sales and the sales locations for each transaction at once. The visibility of both barcode-level sales and sales location enables

[6]An exception is Bernard et al. (2011) which models multi-product firms exporting to multiple markets. While their focus is on explaining firms' exporting behavior using U.S. firm-level export data, my focus is to quantify the interaction between product and geographic scope in generating firm size heterogeneity in a domestic economy context using Japanese product-level transaction data. In particular, this paper departs from Bernard et al. (2011) in two ways: First, I utilize product-level transaction data to show empirical patterns of product and geographic scopes. Second, I utilize the data to structurally estimate model parameters, which I use to quantify how eliminating either of the extensive margins affects firm size heterogeneity and the welfare implications of size-dependent policies.

[7]For example, Eaton et al. (2011) incorporate geographic scope into the model by considering single-product firms exporting to multiple countries. Although the model fits important geographical patterns of firm exporting behaviors, "it leaves the vastly different performance of the same firm in different markets as a residual."

me to explore the roles of geographic and product scope in firm size heterogeneity.

First, I show that both product scope (the number of products) and geographic scope (the number of markets to enter) are two important extensive margins to account for the firm size heterogeneity. Approximately 43% and 31% of firm size heterogeneity are accounted for by product and geographic scope in the accounting decomposition.

Next, I present three empirical patterns. The first pattern suggests that larger firms sell more products in a given market. The second and last patterns suggest that firms' decisions regarding market entry and the number of products to sell in a market are associated with the market size; namely, larger firms enter both large and small markets, whereas smaller firms enter only large markets. Moreover, firms sell more products in larger markets.

Motivated by the empirical patterns, I develop a model of heterogeneous-productivity firms. There are multiple markets with various market sizes in the economy. Firms make decisions on how many product to develop, which markets to enter, and which products to sell in which markets. When making such decisions, firms incur fixed costs from developing products, entering a market, and selling a product in each market.

The model offers insight into how geographic and product scope jointly amplifies initial productivity differences. Higher-productivity firms develop more products because they enter markets and sell more products in the entered markets, generating larger variable profits to cover the fixed cost of developing products. Higher-productivity firms enter more markets because they develop and sell more products and generate larger profits in each market to cover the fixed cost of entering a market. In total, higher-productivity firms develop and sell more products in more markets, gaining a larger market share in the economy.

I quantify the model and evaluate how well it fits the data. I structurally estimate the main model parameters using the simulated method of moments. The quantified model fits the three empirical patterns and successfully replicates the targeted and non-targeted moments in firm size, product scope, and geographic scope distributions found in the Japanese data.

Using the model, I run two counterfactual exercises to see how the absence of either extensive margin affects firm size heterogeneity in the model. I first eliminate the geographic scope margin by imposing a single market restriction on the model; this reduces product scope variation by 42% due to the lack of complementarity between product and geographic scope. In total, the absence of geographic scope margin and the reduction in product scope variation results in reducing firm size variation by 64%. Next, instead of eliminating geographic scope, I eliminate product scope margin by imposing single-product firm restriction on the model, i.e., I eliminate product development process. This reduces geographic scope variation by 77% because the model lacks complementarity such that developing more products allows firms to enter more markets. In total, eliminating the product scope margin results in reducing firm size variation by 96%. These exercises show that the complementarity between the two extensive margins is quantitatively large and important for generating firm size heterogeneity.

I conclude by using the parameterized model to examine the welfare implications of size-dependent policies. I solve the social planner's problem and find that the market equilibrium is efficient (i.e., any policy interventions are sub-optimal).[8] Despite this finding, size-dependent policies, especially SME subsidies, are very common in Japan, perhaps for political reasons.[9] What is the welfare cost of such sub-optimal subsidies? I use an actual

[8]As I will describe in Section 6, the discussion of the efficiency is related to Dhingra and Morrow (2019), where they analyze allocational efficiency of heterogeneous firm models with various market structure and demand specifications.

[9]One of the reasons why the Japanese government protects SMEs is their voting power: Japanese SMEs employ 34 million individuals, or approximately 70.1% of the private sector labor force, and thus wield a large number of votes in elec-

Japanese SME subsidy—given to expand their product scope—as an example and estimate the welfare cost of such subsidies under the baseline and single-economy models. When the geographic scope margin is absent, the welfare cost is underestimated by more than half: the Japanese SME subsidy covering two-thirds of the fixed costs of product development reduces the real consumption index by 2.19% in the baseline model, and by only 1.47% in the single-market economy model.

2. Three empirical facts in Japanese barcode-level transaction data

I use Japanese barcode-level transaction data to show the importance and empirical patterns of product and geographic scopes. The accounting decomposition shows that product and geographic scopes combined account for approximately 75% of sales variations, while the intensive margin explains the other 25%. Thereafter, four empirical facts about the product and geographic scopes are documented to motivate my model structure.

Data

The data source is Nikkei Point of Sales (POS) data, which enables me to observe the product-level prices and sales, location of sales, and manufacturing firms' information for millions of products with marked a barcode. Nikkei collects the sales, quantity, date, and time of each transaction made at a registered retail store, as well as the retail store information (name and address). I use the barcode's first seven (or nine) digits to identify the manufacturing firm that produces the product. GS1 provides the concordance

tions (please see https://www.oecd-ilibrary.org/sites/5989eb3a-en/index.html?itemId=/content/component/5989eb3a-en for more information).

table between barcodes and firm identifiers.[10] The GS1 concordance table is only available for Japanese firms, and thus the barcodes produced by foreign manufacturers comprising only 1% of overall sales are eliminated.[11] Combining the Nikkei POS data and the GS1 concordance table allows me to observe who produced which products and when and where those products are sold to consumers.

Using retail-scanned barcode-level transaction data has many advantages over the household scanner or firm-level export data, which most existing literature relies on. First, compared to the household scanner data, I can observe the sales location of each transaction as an address of the retail store that the barcode got scanned, which enables me to analyze firms' geographic scope. Furthermore, compared to the export data in which products are defined only at the industry category level, products are defined at the barcode level in my data. Barcode-level sales data allows me to analyze the role of product scope more accurately.

I use the Nikkei POS data from 2014 to 2018. Table 1 presents the number of retail stores, retail chains, prefectures, manufacturing firms, and barcodes in the sample. The data covers on average 350 retail stores a year. These stores represent a geographically balanced sample of 43 out of 47 prefectures located in Japan. The sales volume in each prefecture in the data is correlated with the actual market size of the prefecture.[12] For example, the three largest prefectures in terms of their populations (Tokyo, Osaka, and

[10]GS1 is a non-profit organization that develops global standards for business communication, including barcodes. For more detail, see https://www.gs1.org/.

[11]I identify foreign and domestic firms by the first two digits of barcodes; Barcodes produced by Japanese firms start with 45 or 49 (e.g., 45100000001), and those produced by foreign manufacturers start with different numbers. It was not possible to obtain the firm identifiers for under 2% of the barcodes of the products produced in Japan, so these products were dropped from the sample.

[12]The market size of each prefecture can be measured by prefecture-level employment (https://www.stat.go.jp/data/roudou/pref/index.html) or GDP (https://www.esri.cao.go.jp/jp/sna/data/data_list/kenmin/files/files_kenmin.html). Both measures show a positive correlation with the prefecture sale volume in the Nikkei POS from 2014 to 2018.

Kanagawa) are the three largest prefectures in terms of the number of stores and sales in the sample.

I observe the wide range of goods purchased by consumers at retail stores in the data. The database covers approximately 420 billion Japanese yen (≈ 4 billion USD) worth of transactions and $400,000$ barcodes spanning across $16,000$ manufacturing firms per year. Because Nikkei mainly samples supermarkets and convenience stores, the products in the data comprise mostly consumer packaged goods, such as processed foods, beverages, and household goods. Overall, the data covers around 20% of all expenditure on goods in the Japanese official consumer price index.[13]

Nikkei selects retail stores from various chains that cover multiple prefectures. The data samples retail stores from approximately 40 chains a year.[14] Each retail chain covers, on average, four prefectures a year. The largest chain in my sample is a supermarket chain called AEON and covers 22 prefectures a year. All retail stores in the database are relatively large, with average annual sales of 1.3 billion Japanese yen (≈ 11 million USD). Brick-and-mortar grocery or liquor stores are not included in the sample.

Product category information is available for each barcode, which allows me to analyze the role of product and geographic scopes in firm size heterogeneity within categories. Nikkei organizes barcodes into 3-digit product categories; there are 217 categories. The five largest of the 217 product categories in terms of total sales in my sample are Bento (packaged meal in a box for take-out), rice, frozen food, yogurt, and pastry. The product categories are used to generate the below empirical facts depending on the appropriate level of aggregation.

[13]The Nikkei Point of Sales data is used to construct a consumer price index provided by the University of Tokyo called UTokyo Daily Price Index (https://www.cmdlab.co.jp/price_u-tokyo/media_e/).

[14]The retail chains are identified by the first word of the store names (e.g., Safeway Menlo park).

Table 1: Sample statistics

Year	Retail stores	Prefectures	Retail chains	Firms	Barcodes
2014	336	43	39	15,590	393,313
2015	348	43	40	16,289	407,497
2016	351	44	40	16,718	411,694
2017	348	43	39	16,628	407,497
2018	345	43	39	15,393	326,407

Note: The fifth and sixth columns are the numbers of *domestic* manufacturing firms and the barcodes produced by the domestic firms. The total number of prefectures in Japan throughout the sample period is 47.

Firm size, product and geographic scopes

Table 2 presents the market share, product scope, and geographic scope by ranked decile in the sample. I calculate the decile values for each quarter and 3-digit product category and weigh the values by the product category-quarter sales. Three striking features emerge. First, the markets are skewed toward a few large firms — the largest 10% of firms occupy approximately 90% of the market share. Second, larger firms sell more products than smaller firms — the largest 10% of firms sell 77 barcodes, while the median firms sell four barcodes. Finally, larger firms sell their products in more locations — the largest 10% of firms sell their products in 28 out of 43 prefectures, while the median firms sell in six.[15]

How important are product and geographic scopes, compared with intensive margin, to account for the observed firm size distribution? Figure 1 presents an accounting decomposition of the firm sales into the number

[15] A common concern regarding product scope is that larger firms might be connected one large retail chain and sell their products in many stores and prefectures within the chain. As shown in Section A.1, this is not the case: larger firms sell in more retail chains, and these retail chains are not systematically larger than the retail chains that smaller firms sell their products in.

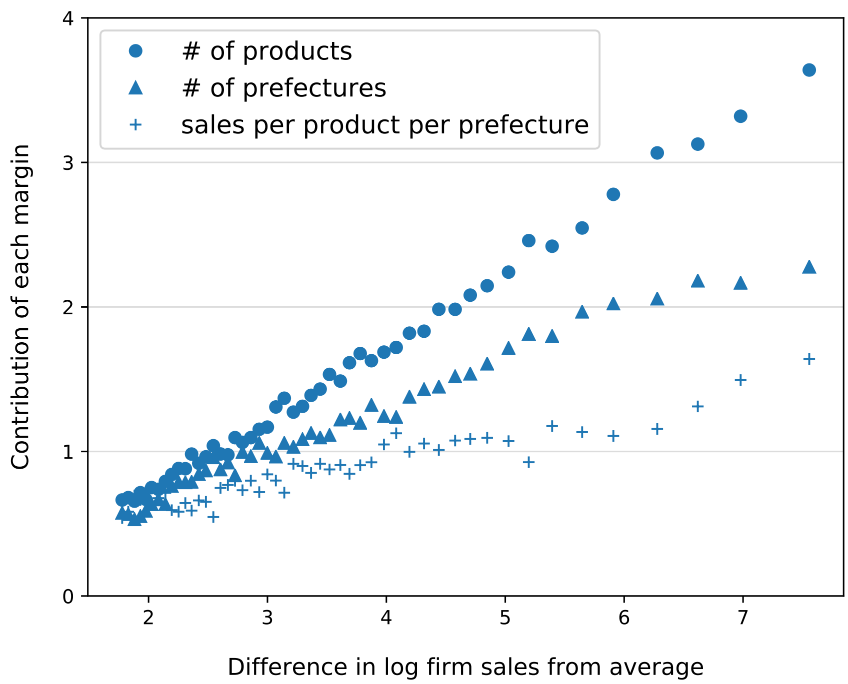
Table 2: Firm size, number of barcodes and prefectures by decile

Ranked decile	Decile market share (%)	Avg. no. barcodes	Median no. barcodes	Avg. no. prefectures	Median no. prefectures
1	89.95	73.38	51.76	28.08	29.56
2	6.93	19.74	15.56	17.30	16.54
3	2.47	10.24	8.08	12.46	11.16
4	1.09	6.86	5.30	9.50	7.98
5	0.55	4.82	3.78	7.40	6.02
6	0.29	3.58	2.78	5.50	4.40
7	0.14	2.76	2.14	4.04	3.06
8	0.07	2.18	1.60	2.86	2.10
9	0.03	1.72	1.24	2.10	1.46
10	0.01	1.30	1.02	1.42	1.08

Note: The largest decile is ranked first. The values are weighed by 3-digit product category sales per quarter. Decile market share does not perfectly adds up to 100% because it is calculated as an weighted average decile market share in each product category. Number of prefectures is calculated as the number of prefectures in which each manufacturing firm sells at least one product.

of products firms sell (product scope), the number of prefectures in which firms sell their products (geographic scope), and the sales per product per prefecture (intensive margin) for the 50 firms with the largest average market share in each product category. Each point represents an average contribution of a particular margin. For example, the column of points at the right of the figure indicates that the largest firm in a product group is, on average, almost eight log units larger than an average firm. Of this 8 log unit difference, the dotted point indicates that, on average, approximately 3.5 log units (43%) are attributed to the product scope differences. The triangles reveal that approximately 2.3 log units (31%) are, on average, attributable to the geographic scope differences, whereas the intensive margin is responsible for the rest (25%). Thus, approximately 75% of sales variations across the top 50 firms are attributed to product and geographic scope variations.

Figure 1: Accounting decomposition of the firm sales by rank



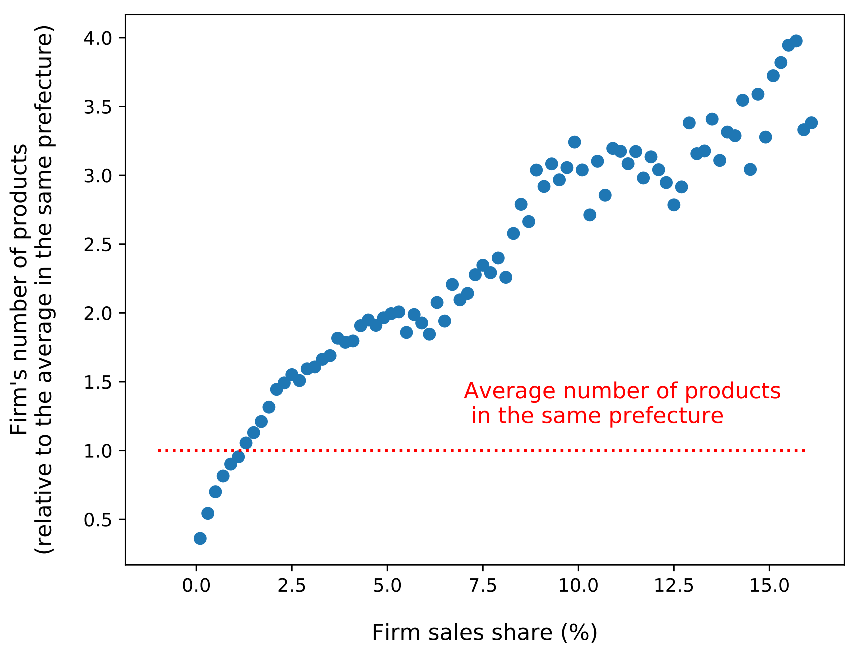
The sales decomposition confirms that geographic scope, in addition to product scope, is a quantitatively important margin for firm size heterogeneity. Therefore, in §3, I construct a theoretical framework to analyze firms' decisions on product and geographic scope. Before moving on to the model, I present three empirical facts that motivate my model structure. These empirical facts reveal regularities in firms' product scope decisions within and across markets and firms' entering decisions across markets.

Fact 1: Larger firms sell more products in a given market

According to Figure 2, larger firms not only sell more products in the economy (see Table 2), but also they sell more products in each particular market. The figure plots firms' number of products in each prefecture on the y-axis and the firm sales share on the x-axis. The y-axis is normalized by the average number of products per firm in the same prefecture. I calculated firms' number of products in each prefecture relative to the average number of products per firm in the same prefecture (y-axis) and firm sales share (x-axis) in each 3-digit product category and created the binned scatter plot where each bin is defined as 0.1% of firm sales share. For example, the figure suggests that firms with 2.5% of sales share in the economy sells about 50% more products than the average firms in the same prefecture in the prefectures that the firm sells their products in. The positive slope indicates that the firms with a larger sales share in the economy sell more products in each market.

Fact 1 might seem obvious, but not necessarily. It is possible that larger firms sell only a small subset of their products in each market, but because they sell in more markets, they sell more products in total at the country level. Fact 1 implies that this is not the case.

Figure 2: Firms' number of products in each prefecture



Fact 2: Firms sell more products in larger markets

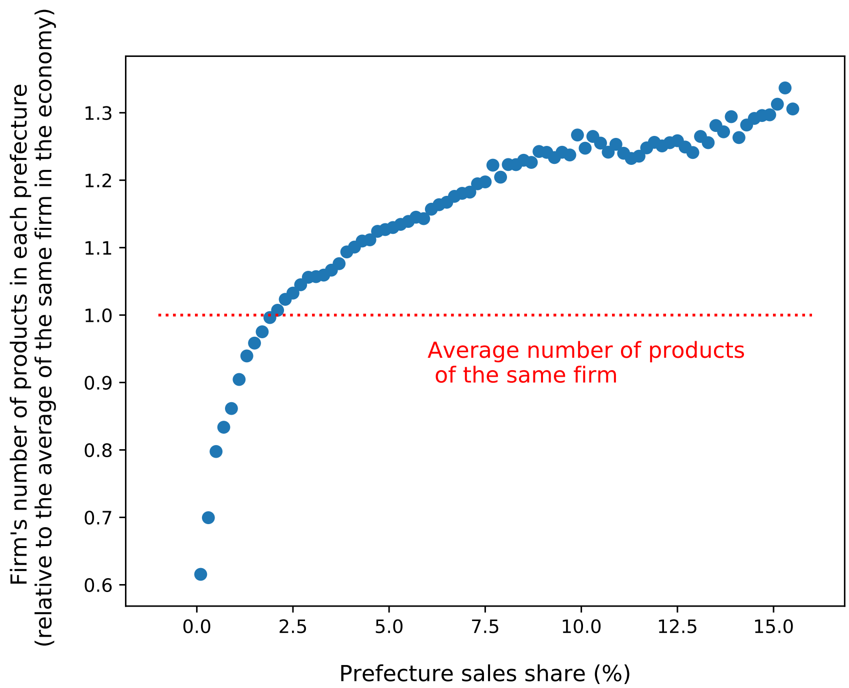
How do firms' decisions on how many products to sell in a market depend on the market size? Figure 3 shows that the number of products firms sell increases with the market size. The figure plots firms' number of products in each prefecture (relative to the average number of products of the same firm across prefectures) on the y-axis and prefecture market size (as a share of total sales in all the prefectures) on the x-axis. I calculated the average number of products of firms in each prefecture (y-axis) and prefecture sales share (x-axis) in each 3-digit product category and created the binned scatter plot, where each bin is defined as 0.1% of prefecture sales share. For example, the figure suggests that if firms sell their products in prefectures that have 2.5% and 5% of sales shares, then the same firms sell 10% more products in the prefecture with 5% of sales share than in the prefecture with 2.5% of sales share. The positive slope in the figure suggests that the same firms sell more products in larger markets.

Fact 3: Larger firms penetrate smaller markets

Fact 2 shows the firm's decision on how many products to sell in a market depends on the market size/ Fact 3 asks whether firms' market entering decision also depends on the market size.

Figure 4 shows that larger firms enter both small and large markets, while smaller firms enter only large markets. The figure plots average firm sales in a particular prefecture relative to the average firm sales in the whole country on the y-axis and the prefecture market size (as a share of total sales in all the prefectures) on the x-axis. I calculated the average firm size in each prefecture (y-axis) and prefecture sales share (x-axis) in each 3-digit product category and then created the binned scatter plot where each bin is defined as 0.1% of prefecture sales share. Because the y-axis is normalized by the

Figure 3: Firm's number of products in each prefecture

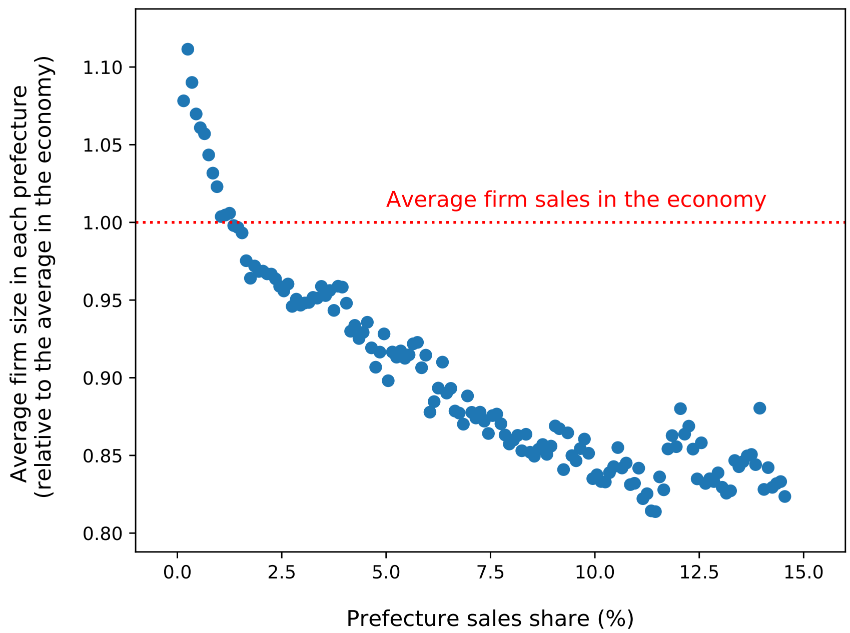


average firm size in the country, if the value is above one, it means that the firms in the prefecture are larger than the average firms in the country. For example, the point on the left end shows that the average firm size in the prefectures that have 0.0 to 0.1% of prefecture sales shares is 1.12 times larger than the average firm size in the economy The negative slope in the figure indicates that larger firms are more likely to penetrate smaller prefectures.

3. A heterogeneous-firm model

To understand mechanisms that generate firm size, product scope, and geographic scope variations, I build a model of heterogeneous firms, where firms choose which products to develop, which markets to enter, and which products to sell in the markets they enter. The economy consists of multiple markets with heterogeneous market sizes. Products are differentiated, and

Figure 4: Average firm size in each prefecture



firms are monopolistically competitive in each market. Firms are subject to three fixed costs: the fixed costs of developing a product, entering a market, and selling a product in each market.

3.1 Consumer preference

There are R number of markets in the economy, each indexed by r. A representative consumer in market r is endowed with L_r amount of labor. The representative consumer supplies the labor endowment inelastically with no disutility coming from labor. I assume that firms are mobile across markets to hire labor, and therefore, wage is equalized across markets. I denote w as wage per unit labor supply and normalize wage $w = 1$.

The representative consumer chooses consumption amount (C_{kr}) of each product k in market r to maximize the following constant elasticity of substitution (CES) utility function subject to the budget constraint:

$$U_r = \left[\int_{f \in \Omega_r} \int_{k \in \Omega_{fr}} (\phi_k C_{kr})^{\frac{\sigma-1}{\sigma}} \; \mathrm{d}k \; \mathrm{d}f \right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

$$\text{s.t.} \int_{f \in \Omega^r} \int_{k \in \Omega_{fr}} p_{kr} C_{kr} \, \mathrm{d}k \; \mathrm{d}f \leq w L_r, \tag{2}$$

where Ω_r is the set of firms entering market r, Ω_{fr} is the set of products that firm f sells in market r. p_{kr} is the price of product k in market r, and ϕ_k is taste shifter of product k. The parameter $\sigma > 1$ is the elasticity of substitution across products.

The consumer demand on product k in market r is given by

$$C_{kr} = (\phi_k)^{\sigma-1} \left(\frac{p_{kr}}{P_r} \right)^{-\sigma} \frac{L_r}{P_r}. \tag{3}$$

The taste shifter ϕ_k captures a popularity of product k that is common across all markets. Given the consumption amount of the product, a product with a higher taste shifter generates a higher utility. Therefore, given price, consumers consume a product with a higher taste shifter more (see expression (3)).[16, 17]

The corresponding price index for the CES utility function in market r is given by

$$P_r \equiv \left[\int_{f \in \Omega_r} \int_{k \in \Omega_{fr}} \left(\frac{p_{kr}}{\phi_k} \right)^{1-\sigma} \; \mathrm{d}k \; \mathrm{d}f \right]^{\frac{1}{1-\sigma}}. \tag{4}$$

[16] I acknowledge that consumer taste for the same product can differ across markets. For example, households in wealthier cities might have different tastes for products compared to other cities (CITATION). This model prioritizes simplicity and does not capture these heterogeneity by assuming common taste shifter for same products across markets.

[17] Using a log-normal distribution for consumer tastes has an advantage in fitting the data. With the consumer tastes following log-normal and productivity following Pareto distributions, sales per product per region increases with productivity, which fits the empirical finding in the Japanese data (see Figure 1). This is in contrast with assuming Pareto distribution for consumer tastes and productivity, where sales per product per region becomes independent of productivity. See Fernandes et al. (2019) for more discussion.

3.2 Firm technology

There is an unbounded measure of potential firms who are identical prior to their entry into the economy. To enter the economy, firms must incur a sunk cost of entry, F_e. Once the sunk entry cost is incurred, each firm f draws productivity φ_f from the continuous distribution $g(\varphi)$, with cumulative distribution $G(\varphi)$.

After observing its productivity, firms make two simultaneous decisions: How many products to develop and which markets to enter. Firms incur the fixed cost of developing products, which increase with the total number of products firms develop, N_f. Specifically, the fixed costs of developing N_f number of products equals $F_d \cdot (N_f)^\theta$ where $F_d > 0$ and $\theta > 1$. Firms also incur the fixed cost of entering each market, F. Both fixed costs are paid in unit of labor.

After incurring the fixed costs of developing products and entering markets, firms draw taste shifter of N_f products. The taste shifter of each product is drawn from the continuous distribution $h(\phi)$, with a cumulative distribution $H(\phi)$.[18]

Finally, after taste shifters for all the developed products have been observed, the firm decides which products to sell in the markets they choose to enter. Upon choosing which products to sell in which markets, firms incur the fixed cost of selling per product in a given market, F_p. All the fixed costs are paid in units of labor.

In addition to these fixed costs, there is a constant marginal cost for each product, which depends on the firm's productivity. The marginal cost of a product is the same across products and markets within a firm due to the following assumptions. First, the production function is the same across

[18]To exploit the low of large numbers' feature, it is assumed that the productivity and taste shifter are independent of one another and independently distributed across firms. Similarly, it is assumed that the taste shifter is independently distributed across the products.

products and firms and is linear in labor, the sole factor of production. The amount of labor needed to produce one unit of output is $\frac{1}{\varphi_f}$. Second, I assume the free mobility of firms in the economy, which equates wage across markets.[19] Third, I assume there are no variable costs of trade across markets. In total, firm f's marginal cost of selling a product is given by φ_f^{-1}.

With the production technology explained above, total labor employed by firm f is

$$l_f = \sum_{r \in \Omega_f} \left\{ \int_{k \in \Omega_{fr}} \frac{q_{kr}}{\varphi_f} + F_p \, \mathrm{d}k + F \right\} + F_d (N_f)^\theta,$$

where q_{kr} denotes the output of product k (sold by firm f) in market r.

Each market is monopolistically competitive. Under the assumption of CES preferences and monopolistic competition, optimal prices are a constant markup over marginal costs:

$$p_{kr} = p_f = \frac{\sigma}{\sigma - 1} \varphi_f^{-1}. \tag{5}$$

Since a firm's marginal cost is common across products and markets, the optimal pricing is common across products and markets within a firm as well.

3.3 Equilibrium production, entry, and product development decisions

Firms optimization problem involves two steps. First, given productivity draw, firms choose how many products to develop and which markets to enter. Second, given productivity and product appeal draws, firms choose which products to sell in the markets they enter. We solve the optimization

[19]The single wage in the economy is normalized to one.

problem backward.

Production decision in each market. Using firm f's pricing rule defined by expression (5), the equilibrium revenue received by a firm f in market r from selling a product k is:

$$r_{kr} = \left(\frac{\sigma}{\sigma - 1}\right)^{1-\sigma} (\phi_k \varphi_f P_r)^{\sigma-1} L_r. \tag{6}$$

The corresponding equilibrium variable profits from selling product k to market r are:

$$\pi_{kr} = \frac{r_{kr}}{\sigma} - F_p. \tag{7}$$

Note that firm's decision on which products to sell in the markets they enter happen after product development and market entry decision. Thus, the fixed costs of developing products and entering a market are not in consideration here.

Firm f sells product k if and only if the profit from selling the product is positive, i.e., $\pi_{kr} > 0$. Because the demand for a product in a given market is monotonically increasing in the taste shifter, and $\frac{r_{kr}}{\sigma}|_{\phi_k=\infty} > F_p$ and $\frac{r_{kr}}{\sigma}|_{\phi_k=0} < F_p$ hold for any given productivity level φ_f, a *cutoff taste shifter* for each firm f in market r, ϕ^*_{fr}, is defined as

$$\pi_{kr}|_{\phi_k=\phi^*_{fr}} = 0. \tag{8}$$

Given entering market r, firm f sells products with taste shifters higher than the cutoff taste shifters. Solving equation (8) to obtain the following analytical expression for the cutoff taste shifter:

$$\phi^*_{fr} = \frac{1}{\varphi_f P_r} \left\{ \frac{\sigma F_p}{L_r} \left(\frac{\sigma}{\sigma - 1}\right)^{\sigma-1} \right\}^{\frac{1}{\sigma-1}}. \tag{9}$$

In a given market, higher-productivity firms have lower cutoff taste shifer, which means higher-productivity firms sell a larger share of their developed products. The total number of products in market r equals the number of developed products times the share of developed products with taste shifter above the cutoff taste shifter, i.e., $N_f(1 - H(\phi^*_{fr}))$.

Joint decision on product development and market entry. Given the cutoff taste shifter in each market, firm f makes a joint decision on whether to enter each market and how many products to develop. Firm f chooses how many products to develop (N_f) and the set of markets to enter (Ω_f) to maximize the following profits:

$$\max_{N_f,\Omega_f} \pi_f = \sum_{r\in\Omega_f} \left\{ N_f \int_{\phi^*_{fr}}^{\infty} \pi_{kr} \, \mathrm{d}H(\phi) - F \right\} - F_d \left(N_f\right)^{\theta}. \tag{10}$$

Firm f's profits equal the sum of profits in each market minus the fixed cost of developing products. Firm f's profits in each market depend on two terms. The first term is the variable profit from selling the developed products with taste shifters above the cutoff taste shifter, $N_f \int_{\phi^*_{fr}}^{\infty} \pi_{kr} \, \mathrm{d}H(\phi)$. The second term is the fixed cost of entering the market, F.

The decisions on how many products to develop and which markets to enter are interdependent. The increase in profits by entering an additional market depends on the number of products sold in the market, which depends on the number of developed products. Moreover, the increase in profits by developing an additional product depends on in how many markets the additional product to be sold. This interdependence of firms' decisions is the source of complementarity between product scope and geographic scope, which will be discussed in the next subsection.

We can define a cutoff productivity, φ^*_r, in market r such that firms with productivity higher than the cutoff value enter the market. The profit of entering market r increases with productivity by the following three reasons. Given the number of products developed, firm f's profit of entering market r increases with productivity because $\frac{\delta \pi_f}{\delta \varphi_f} \geq 0$ and $\frac{\delta \phi^*_{fr}}{\delta \varphi_f} \geq 0$. Similarly, given the set of markets entered, firm f's optimal number of developed products increases with productivity, i.e., $\frac{\delta N^*_f}{\delta \varphi_f} \geq 0$. Finally, firm f's profit of entering market r increases with the number of developed products, i.e., $\frac{\delta \pi_{fr}}{\delta N_f} \geq 0$.

Because the profit entering market r increases with productivity, and $\pi_{fr}|_{\varphi_f=\infty} > 0$ and $\pi_{fr}|_{\varphi_f=0} < 0$ hold for any market r, we can define *cutoff productivity* in market r, φ^*_r, as

$$\pi_{fr}|_{\varphi_f=\varphi^*_r} = 0. \tag{11}$$

Firms with productivity higher than the cutoff productivity makes a positive profit by entering the market, and thus, enter the market.

Complementarity between product scope and geographic scope. Firm's joint decision on product development and market entry is the source of two-way complementarity between product scope and geographic scope, which amplifies the firm size heterogeneity. This subsection shows the analytical presentation of the complementarity.

First, there is a complementarity such that developing more products increase the profit of entering more markets. Define $\pi_f(N_f, M_f)$ as the profit of firm f when developing N_f products and entering M_f markets. Compare the increase in firm f's profits from entering one more market (namely, market r') when the firm develops N_f products and $N_f + 1$ products:

$$\pi_f(N_f, M_f+1) - \pi_f(N_f, M_f) = N_f \int_{\phi^*_{fr'}}^{\infty} \pi_{kr'} \, \mathrm{d}H(\phi) - F,$$

$$\pi_f(N_f+1, M_f+1) - \pi_f(N_f+1, M_f) = (N_f+1) \int_{\phi^*_{fr'}}^{\infty} \pi_{kr'} \, \mathrm{d}H(\phi) - F.$$

The first terms on the right-hand sides capture the increase in variable profits by entering one more market and the second term captures the increase in the fixed cost of entering one more market. By comparing the the two expressions, the increase in profit of entering one more market is higher when firms develop one more products, i.e.,

$$\pi_f(N_f+1, M_f+1) - \pi_f(N_f+1, M_f) \geq \pi_f(N_f, M_f+1) - \pi_f(N_f, M_f). \quad (12)$$

This first complementarity comes from the fact that, by developing more products, the firm sells more products in a given market and, thus, makes larger variable profits by entering one more market. The fixed cost of entering one more market is independent of the number of products developed. The complementarity implies that higher-productivity firms that develop more products are more likely to enter more markets.

Second, there is a complementarity such that entering more markets increase the profit of developing more products. Compare the increase in profit of developing one more product when the firm enters one more market:

$$\pi_f(N_f+1, M_f) - \pi_f(N_f, M_f) = \sum_{r=1}^{M_f} \int_{\phi^*_{fr}}^{\infty} \pi_{kr} \, \mathrm{d}H(\phi) - F_d \left((N_f+1)^\theta - (N_f)^\theta \right),$$
$$\pi_f(N_f+1, M_f+1) - \pi_f(N_f, M_f+1) = \sum_{r=1}^{M_f+1} \int_{\phi^*_{fr}}^{\infty} \pi_{kr} \, \mathrm{d}H(\phi) - F_d \left((N_f+1)^\theta - (N_f)^\theta \right).$$

The first terms on the right-hand sides capture the increase in the variable profits by developing (and selling) one more products, and the second terms capture the increase in the fixed costs of developing.

The increase in profit of developing one more product is higher when firms enter one more market , i.e.,

$$\pi_f(N_f+1, M_f+1) - \pi_f(N_f, M_f+1) \geq \pi_f(N_f+1, M_f) - \pi_f(N_f, M_f). \quad (13)$$

The second complementarity comes from the fact that developing one more product increases expected variable profits of entering markets, and, thus, the firms are likely to enter more markets. The fixed costs of developing products do not depend on how many markets the firms enter. The complementarity implies that higher-productivity firms that enter more markets are likely to develop more products.

3.4 Free-entry

In an equilibrium with positive entry, the expected value of entry must equal the sunk entry cost, which requires the following free entry condition to hold:

$$\frac{\Sigma_r \Pi_r}{M} = F_e, \tag{14}$$

where M is the mass of firms entering the economy and $\Pi_r \equiv \int_{f \in \Omega^r} \pi_{fr} \, \mathrm{d}f$ is the sum of profits of all surviving firms in each consumption market r. Equation (14) pins down the equilibrium mass of firms in the economy.

3.5 Aggregation

Using the cutoff taste shifter condition for each firm (8), the cutoff productivity condition (11), and the pricing strategy for each firm (5), the price index (4) in market r can be expressed as

$$P_r = \left[M \int_{\varphi_r^*}^{\infty} (\varphi_f)^{\sigma-1} N_f^* \int_{\phi_{fr}^*}^{\infty} (\phi_k)^{\sigma-1} \mathrm{d}H(\phi) \mathrm{d}G(\varphi) \right]^{\frac{1}{1-\sigma}}. \tag{15}$$

The price index of each market depends on four elements—mass of firms in the economy, the average productivity in the market, the number of developed products of each firm, and the average taste shifter of each firm. A decrease in the cutoff productivity, φ_r^*, has three effects on the price index.

First, the decrease in the cutoff productivity means more firms are entering the market, which decreases the price index due to the love-of-variety effect of CES preference. Moreover, the entering firms (develop and) sell multiple products in the market, further decreasing the price index. Lastly, the inflow of marginal firms decreases the average productivity in the market, which increases the price index. These insights will be argued further when discussing the model fit to the three empirical patterns in Section 4.4.

4. Estimation and model evaluation

I structurally estimate model parameters by the simulated method of moments using the Nikkei POS data. With the quantified model, I evaluate model performance by checking whether the model fits three empirical patterns found in Section 2, and targeted and non-targeted moments in the Japanese data.

4.1 Structural estimation

The model parameters are divided into four categories. The first category is the parameters that can be directly determined by the data. The number of markets (R) and the labor endowment of each market (L_r) are included in this category. The second category comprises the parameters that are estimated from the data. The parameters for taste shifter distributions $\log(\phi_k) \sim N(\mu, \sigma^2)$ are included in this category. The third category is the parameters that are estimated with GMM to match some moments in the data. The fixed cost of entering the economy (F_e) is set to match the mass of firms in the model to the number of firms in the data. In addition, there are three types of fixed costs $\{F_d, F, F_p\}$ set to match the top 10% vs. median firms' stats of firm size, product, and geographic scopes. Specifically, compared to the median firms, the top 10% firms generates 139.2 times more sales, produce 15.9 times more products, and enter 4.5 times more markets, and the three fixed costs are estimated with GMM to match the three targeted values. The final category is the parameters that are estimated using SMM. The elasticity of substitution in the CES utility function (σ), shape parameter for productivity distribution (α), and a parameter in the fixed cost of developing products (θ) are in this category. I define $\Theta_1 = \{F_d, F, F_p\}$ as a set of parameters that are estimated using GMM, and $\Theta_2 = \{\sigma, \alpha, \theta\}$ as a set of parameters that are estimated using SMM.

4.1.1 Moments in the SMM

To estimate the set of model parameters $\Theta_2 = \{\sigma, \alpha, \theta\}$, we target three moments.

The first moment is the elasticity of firm's sales share per product in each market with respect to price. Define firm f's sales share in the market r as S_{fr} and the number of products firm f sells in market r as N_{fr}. Then, firm f's sales share per product in market r is given by

$$\log \frac{S_{fr}}{N_{fr}} = (1-\sigma)\log p_f + (\sigma - 1)\log P_r + \frac{1}{N_{fr}} \int_{\phi^*_{fr}}^{\infty} (\phi_k)^{\sigma-1}\, \mathrm{d}H(\phi), \qquad (16)$$

It is possible to estimate σ from using equation (16), i.e., by regressing the firm sales share per product on firm price and market fixed effect.[20] However, such an estimation suffers from selection bias. The assumption required for the identification is firm prices and the average taste shifter are orthogonal, i.e., $p_f \perp \frac{1}{N_{fr}} \int_{\phi^*_{fr}}^{\infty} (\phi_k)^{\sigma-1}\, \mathrm{d}H(\phi)$, which is likely to be violated. Firms with a higher productivity (and thus a lower price) can sell products with a lower taste shifter, which leads to a positive correlation between firms'

[20] Upon estimating σ, I compute firm-level price as the average of product-level prices in each market weighted by expenditure share of each product within firm.

prices and the average taste shifter, underestimating σ. Thus, we target the *biased* estimate of σ estimated by expression (16) as a targeted moment in the SMM.

The second moment is the sales share of each product sold in each market. Define S_{kfr} as the expenditure share of product k sold by firm f in consumption market r. Taking the log of the expenditure share gives the following expression:

$$\log S_{kfr} = \log(\frac{\sigma}{\sigma - 1})^{1-\sigma} + (\sigma - 1)(\log(\phi_k) + \log(\varphi_f) + \log(P_r)). \tag{17}$$

Given the known value of the elasticity of substitution (σ), I can compute productivity (φ_f) for each firm by (1) regressing the log of product sales in each consumption market onto product, firm, and consumption market fixed effects, and (2) dividing the coefficients of the firm fixed effects by $(\sigma - 1)$ and take an exponential of it. However, if firms have higher productivity, then the products with a smaller taste shifter (less popular products) are still likely to be sold. The selection bias underestimates the coefficient of firm fixed effects and thus the productivity of higher-productivity firms. Thus, I use the *biased* estimate of the shape parameter of productivity distribution estimated by expression (17) as a targeted moment in the SMM.

The third moment is the elasticity of the number of developed products with respect to the firm's total sales. The parameter θ controls the elasticity of number of developed products with respect to the firm's productivity. A higher-productivity firm develops more products, but the parameter θ controls the degree of how many more products a higher-productivity firm develops compared to a lower-productivity firm.

By solving the expression (10), firm f's optimal number of developed products given the set of markets to enter (Ω_f), N_f^*, is given by

$$N_f^* = \left[\frac{1}{\theta F_d}\left\{\sum_{r\in\Omega_f}\int_{\phi_{fr}^*}^{\infty}(p_f - \mathrm{MC}_f)C_{fr} - F_p \,\mathrm{d}H(\phi)\right\}\right]^{\frac{1}{\theta-1}}. \tag{18}$$

From the expression (18), the number of products a firm develops is positively correlated with the firm's total sales. Therefore, I use the elasticity of the number of developed products with respect to the firm total sales in the economy as a target moment to estimate θ in the model.

I use the simulated method of moments (SMM) to help find unbiased estimates of these parameter values. The key purpose of using SMM is to find true parameter values $\{\sigma, \alpha, \theta\}$ so that the targeted moments in the actual data match the targeted moments in the model-generated data.

4.1.2 Estimation algorithm

The estimation of the model parameters involves the following 6 steps. Throughout the steps, the same data as in Section 2 is used , the Nikkei POS data from 2014 to 2018.

Preparation. Pin down the values of the number of markets (R) and the size of each market r (L_r) from the data. The number of market is set to match the median number of prefectures in Nikkei POS data from 2014-2018. The size of each market is set to match the average total sales volumes in each prefecture in the Nikkei POS data from 2014 to 2018.

Step 1. Assume true values of the elasticity in CES utility function (σ), productivity distribution $\varphi_f \sim \text{Pareto}(\alpha, 1)$, and the elasticity of the number of developed products with respect to productivity θ in the model.

Step 2. Given the value of σ, compute the values of taste shifter (ϕ_k) for each firm in the data. Then, compute the parameter values of the taste shifter distribution $\log(\phi_k) \sim N(\mu, \sigma^2)$. To obtain unique values for the taste shifter for each product k for firm f (ϕ_k), I use the variation of sales across products within each firm. The relative expenditure shares between two products within a firm in the same market is given by

$$\frac{p_f C_{kr}}{p_f C_{k'r}} = \left(\frac{\phi_k}{\phi_{k'}}\right)^{\sigma-1}. \tag{19}$$

For each firm f, I choose a product that is sold in the largest number of prefectures in the firm as a base product such that the taste shifter of the product is normalized as one, i.e., $\phi_k = 1$. I then estimate the relative taste shifter of other products using expression (19).[21]

Step 3. Compute the equilibrium in the model given the values of σ and productivity distribution set in Step 1 and the taste shifter distribution estimated in Step 2. Set the fixed cost of entering the economy F_e to match the number of firms in the model to the data.

Step 4. Find a set of parameter values of three fixed costs $\{F_d, F, F_p\}$ that matches top 10% vs. median firms' statistics in the data. Specifically, I match between top 10% vs. median firms' sales, product scope (number of products), and geographic scope (number of markets) in the equilibrium and data. Define $m(\Theta_1)$ as a vector of the top 10% vs. median moments in the model and $m(1)$ as the corresponding moments in the data. Then, I seek a set of parameter values $\hat{\Theta}_1$ to minimize the distance between the moments

[21] Upon estimating the taste shifter, I restrict the sample to the firms that sell multiple products in more than five markets(prefectures).

in the data and model using the criterion function:

$$\hat{\Theta}_1 = \text{argmin}\,(m(1) - m(\Theta_1))W_1(m(1) - m(\Theta_1))', \tag{20}$$

where W_1 is the variance-covariance matrix of the error function, $m(1) - m(\Theta_1)$.

Step 5. Simulate 10,000 artificial firms and compute product- and firm-level expenditure shares in each market, firm-level prices, and the number of products for each firm (see Appendix A.2.1 for a detailed simulation algorithm). Using the model-generated data, compute the three targeted moments.

I iterate these steps until the targeted moments in the model-generated data converges to the targeted moments of these parameter values estimated in the actual data. The SMM selects a vector of parameter values $\hat{\Theta}_2$ that minimizes the distance between the moments in the actual data and the model-generated data, using the criterion function:

$$\hat{\Theta}_2 = \text{argmin}\,(m(2) - m(\Theta_2))W_2(m(2) - m(\Theta_2))', \tag{21}$$

where $m(\Theta_2)$ is a vector of moments, $\{\hat{\sigma}, \hat{\alpha}, \hat{\theta}\}$, estimated in the model-generated data, $m(2)$ is a vector of corresponding moments in the actual data, and W_2 is the variance-covariance matrix of the error function, $m(2) - m(\Theta_2)$.

4.2 Estimation results

Table 3 presents the parameter values. Our estimate of σ confirms the upward bias from the selection bias. The biased estimate of σ in the actual data using expression (16) is 1.32, and the estimated value of σ that generates the same value of biased estimate in the model generated data is 2.40, which confirms the point that selection bias underestimates elasticity σ. The estimated value ($\sigma = 2.40$) is consistent with the literature in trade and macro; for example, the estimates in Ossa (2014) range from 2 to 4 for most categories including non-food consumer goods.

The biased estimates of shape and scale parameters for productivity distribution in the actual data were $\alpha = 3.52$ and the parameter values that produces the same value of biased estimates in the model-generated data are $\alpha = 2.30$. The biased estimate of scale parameter is larger than the true value, which confirms that the selection bias underestimates high-productivity firms' productivity. The results are consistent with the literature (for example, see Eaton et al., 2011) and satisfy the technical restriction for the distribution of firm revenue to have a finite mean, that is $\alpha > \sigma - 1$.

The interpretation of the magnitude of the fixed costs are as follows. The model features a free-entry condition, which ensures that firms' variable profits finance three types of fixed costs exactly. Given that the estimated σ is 2.40 and firms' profit margin is given by $\frac{1}{\sigma-1} = 0.71$, it means that, on average, 71% of firms' sales are used to pay the three types of fixed costs. In this economy, firms have a large market power and use large variable profits to pay the fixed costs.

4.3 Qualitative fit: Three empirical patterns

I start by showing important equilibrium characteristics of the model: the price index, cutoff productivity, and cutoff taste shifter (of each firm) decrease with the market size. These equilibrium characteristics are then connected to the observed empirical patterns in Facts 1-3.

Figure 5 plots the price index and cutoff productivity of a firm in each consumption market on the y-axis, and prefecture market size as a share of the aggregate market size in the economy on the x-axis. Price index and

Table 3: Parameter values

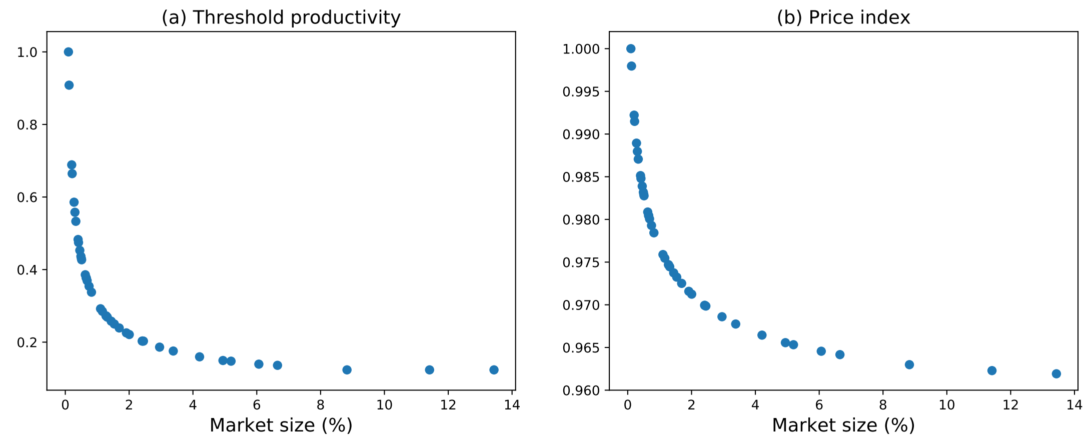
Parameter	Definition	Value	Target
R	Number of s	43	Data (number of prefectures)
L_r	Labor endowment for each region		Data (sales in each prefecture)
$\log(\phi) \sim N(\mu, \sigma^2)$	Distribution of common component	$\mu = 0.47, \sigma = 0.92$	Estimated from the data
$\varphi \sim$ Pareto$(\alpha, \bar{\varphi})$	Distribution of productivity	$\alpha = 2.3, \bar{\varphi} = 1(set)$	Estimated from the data (SMM)
σ	Elasticity of substitution in CES	2.4	Estimated from the data (SMM)
θ	Parameter in the fixed costs of developing products	2.6	Estimated from the data (SMM)
F_d	Fixed cost of developing products	0.079	Parameters to fit top10/median stats
F	Fixed cost of entering a region	0.05	
F_p	Fixed cost of selling per product in each region	0.35	
F_e	Fixed cost of entering the production economy	1.25	Fit the # of firms in the economy

cutoff productivity in the figure are normalized by the value in the smallest market.

Panel (a) of Figure 5 shows that the cutoff productivity decreases with the market size. A larger market size has two offsetting effects on the cutoff productivity. Given the price index, a larger market size means demand for each good, allowing the lower-productivity firms to profit sufficiently to cover the fixed cost of entering the market. A larger market size also induces more competition across entering firms, which decreases the firms' profits. Overall, the first effect dominates, and the cutoff productivity is lower in larger markets, i.e., more firms enter larger markets.

Panel (b) of Figure 5 shows that the price index also decreases with the market size. As described in Section 3.3, when more lower-productivity firms enter larger markets, there are two offsetting effects on the price index: the love-of-variety effect decreases the price index, while lower average productivity increases the price index. In total, the love-of-variety effect exceeds the effect of lower average productivity; thus the price index falls with the market size. It also means that the welfare in a market increases with the market size; since the wage is equalized across markets, the real wage increases as

Figure 5: Cross-sectional results



the price index decreases with the market size. Consumers in larger markets enjoy higher welfare generated by more varieties because more firms enter and sell more products in larger markets.[22] These cross-sectional results are used to explain the observed patterns in Facts 1-3.

Fact 1: Larger firms sell more products in a given market. Fact 1 is implied by the cutoff taste shifter condition (8). The cutoff taste shifter condition for firm f in market r can be rewritten as

$$\phi^*_{fr} = \left(\frac{F_p(\sigma - 1)\frac{\sigma}{\sigma - 1}}{L_r} \right)^{\frac{1}{\sigma-1}} \frac{1}{\varphi_f P_r}. \tag{22}$$

Given price index P^r and market size L^r, higher productivity lowers the cutoff taste shifter. Therefore, a higher-productivity firms sell more products upon entering a market. In the model, higher-productivity firms have larger sales. In total, larger firms sell more products in a given market.

Fact 2: Firms sell more products in larger markets. From expression (22),

[22]Literature in urban economics points to welfare differences across geographic locations due to different varieties available in each location. For example, Handbury (2019) shows that wealthier consumers enjoy a lower price index in wealthier cities because more varieties that are catered to wealthier households are available in these cities.

we can see that the larger market size has two opposite impacts on the cutoff taste shifter. First, a larger market size decreases (a higher L_r) the cutoff taste shifter because there is a larger pie of consumer demand. Second, a larger market induces more competition, which means a lower price index, and the cutoff taste shifter increases with the price index. All in all, the first effect dominates, and the cutoff taste shifter decreases with the market size, i.e., firms sell more products in larger markets.

Fact 3: Larger firms penetrate smaller markets. Because the cutoff productivity decreases with market size, only high-productivity firms enter small markets. This follows the empirical pattern, according to which the average firm size decreases with the market size. The model also replicates the convex decreasing shape of an average firm size in the data.

4.4 Quantitative fit: Targeted and non-targeted moments

To evaluate the model fit to the data quantitatively, I compare firm size, product scope, and geographic scope distributions in the model vs. the data (Nikkei POS). When comparing the model and data, I plot and visually compare the firm size, product scope, and geographic scope distributions in the model and data.

As stated in Section 4.1, I estimate three fixed costs $\{F_d, F, F_p\}$ by GMM to match the largest 10% vs. the median firm statistics in firm size, product scope, and geographic scope. The first row of Table 4 confirms that the model fits these three moments in the data.

Next, I show that not only these targeted moments in the data, but the model also replicates the non-targeted distributions of firm size, product scope, and geographic scope. The top panel of Figure 6 plots the log of average percentile sales normalized by top percentile sales. The x-axis of the figure shows the percentile from the largest to the smallest, from right to left.

The model performs well in replicating the sales distribution. The middle and bottom panels of Figure 6 plot the number of products and prefectures in each percentile normalized by the top percentile value, respectively. The model replicates entire distributions of firms' product scope and geographic scope.

Overall, the model replicated the observed firm size, product scope, and geographic scope heterogeneity well. The differences between the actual and the predicted distributions come from the fact that I approximate the actual productivity and taste shifter distributions by the best fitted distributions (Pareto and log-normal distributions). However, the best fitted distributions still do not match the actual distributions perfectly.[23]

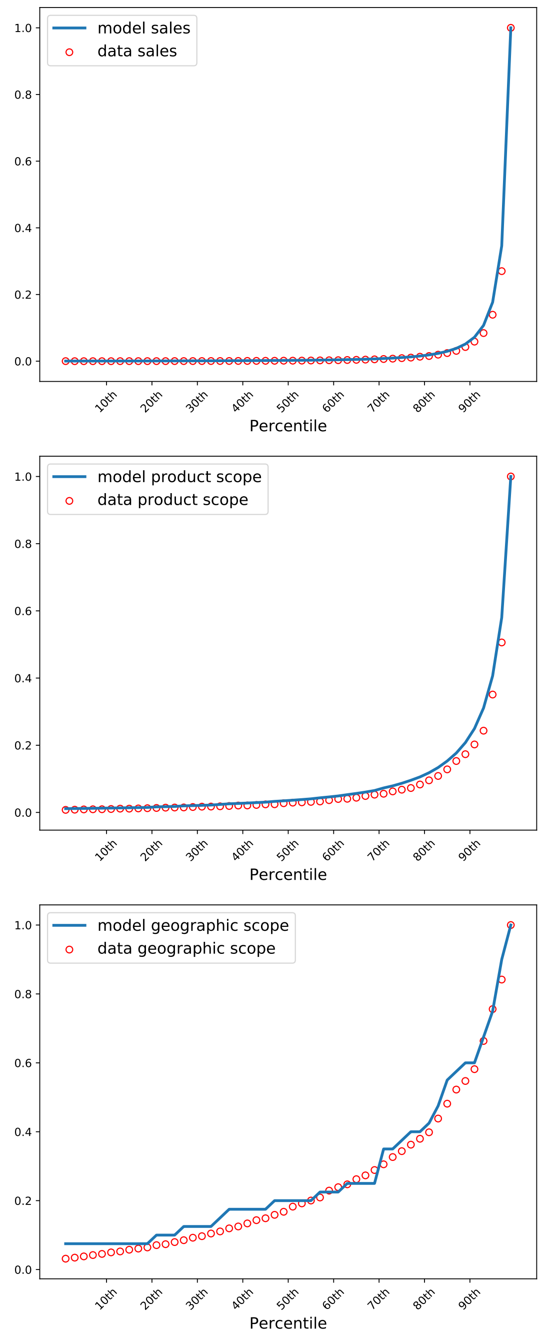
Table 4: Top 10% vs. median firms' statistics in the data and model

$\frac{\text{Top 10\% firms}}{\text{median firms}}$	Firm size		Product scope		Geographic scope	
	Data	Model	Data	Model	Data	Model
Baseline	139.2	139.0	15.9	15.1	4.5	4.5
CF1	139.2	49.8	15.9	8.7	4.5	1
CF2	139.2	4.78	15.9	1	4.5	1.01

Note: The model contains top 10% firms vs. median firms' statistics for firm size, product and geographic scopes in the Nikkei POS data from 2014 to 2018 and in the model. CF1 corresponds to the counterfactual equilibrium where geographic scope is eliminated by imposing single-market in the model. CF2 corresseponds to the counterfactual equilibrium where the product scope is eliminated by imposing single-product firms in the model.

[23]For example, it is well known that the Pareto distribution overestimates productivity of low-productivity firms. This is reflected as the overestimation of the product and geographic scope in the smallest decile in Figure 6.

Figure 6: Non-targeted moments: Model vs. data



Note: In all panels, the x-axis is the number of decile, from the smallest firms on the left to the largest firms on the right. The top panel shows the sales distribution in the data and the model. The sales are normalized by the top percentile sales. The middle panel shows the average number of products of the firms in each percentile in the data and the model. The average number of products in each percentile is normalized by the largest percentile value. The bottom panel shows the average number of markets that firms in each percentile sell at least one product in. Again, the values are normalized by the largest percentile value.

5. Counterfactual Exercises

To see the quantitative importance of the complementarity between geographic and product scopes to account for the firm size variations, I run two counterfactual exercises that eliminate each scope and see how it affects the model fit to the data. The counterfactual exercises show that eliminating either geographic or product scopes reduces not only a significant fraction of firm size variations but also the variations in the other extensive margin, which indicates the significance of the complementarity between the two extensive margins.

CF1: Eliminating geographic scope

To eliminate geographic margin from the model, I solve a single-market version of the model. Specifically, I assume that there is a single market in the economy in which the labor endowment is equal to the average labor endowment in the original economy, i.e., $L = \frac{1}{R}\Sigma_r L_r$.

Moving from the original to the single-market model shuts off geographic scope variations across firms because all the surviving firms enter one market. Therefore, compared with the original equilibrium, where higher-productivity firms enter more markets and acquire larger revenues, the firm size variations are expected to be smaller. In addition, unlike the original equilibrium where firms sell different sets of products across markets due to consumer taste heterogeneity, the product scope variations are expected to be smaller in the single-market model.

The second column of Table 4 reports the result from eliminating geographic scope. First, because I eliminate geographic scope, the top 10% firms vs. median firms' value of the geographic scope is reduced from 4.8 to 1 in terms of the top 10% vs. median firms' values. Second, product scope heterogeneity is also reduced by 42% (from 15.1 to 8.7). This is explained by a lack of complementarity between geographic and product scope through consumer taste heterogeneity across markets. In total, moving from the multiple-market to the single-market models reduces firm size heterogeneity by 64% (from 139.0 to 49.8)

CF2: Eliminating product scope

To eliminate product scope from the model, I solve a single-product firm version of the model. Specifically, each firm is endowed with one product for which the taste shifter equals one in any consumption market, i.e., product development process is eliminated from the model. The demand for the product of firm f in market r equals

$$C_{fr} = \left(\frac{p_f}{P_r}\right)^{-\sigma} \frac{L_r}{P_r}.$$

The model is now similar to Melitz (2003) but with asymmetric market size.

Moving from the original to the single-product firm model shuts off product scope variations across firms because all the surviving firms produce only one product. Therefore, compared with the original equilibrium where higher-productivity firms produce more products and acquire larger revenues, the firm size variations are expected to be smaller. In addition, unlike the original equilibrium, where higher-productivity firms produce more products in a given market, when all the firms produce only one product, lower-productivity firms have a higher chance to enter the market. Therefore, geographic scope variations are expected to be smaller as well.

The third column of Table 4 reports the result of eliminating product scope. When product scope heterogeneity is completely eliminated (15.1 to 1), geographic scope heterogeneity is also reduced by 77% (from 4.5 to 1.01) in terms of the top 10% versus median firms' values. This is due to the lack of complementarity between product and geographic scopes described above. In total, moving from the multi-product firm model to the single-product firm model reduces firm size heterogeneity by 96% (from 139.0 to 4.78).

The two counterfactuals show that complementarity between product and geographic scope is quantitatively important. Existing models of firm heterogeneity with only either one of the extensive margins miss firm size heterogeneity due to a lack of variation in missed margin but also the variation in the other margin coming from complementarity between the two margins. It should be noted that these counterfactual exercises are conducted with all the parameter values fixed at the baseline values presented in Table 3.

6. Welfare implication of size-dependent policies

I solve the social planner's problem and show that the market equilibrium is efficient. Given the result, I discuss the welfare implication of sub-optimal size-dependent policies, taking actual policy in Japan as an example. I compare the welfare cost between the baseline and single-product firm model and find that the missing product scope margin largely underestimates the welfare cost of such policies.

6.1 Social planner's problem

It is well known that the market equilibrium in my baseline model, Melitz (2003), is efficient.[24] To understand whether this efficiency is carried over

[24] For example, Melitz and Redding (2015) show that with CES preferences and monopolistic competition, the open economy equilibrium in the heterogeneous firm model is efficient. Dhingra and Morrow (2019) analyze the efficiency of heterogeneous firm models in a more generalized setting.

when multi-product firms and multiple regions with heterogeneous market sizes are added, I solve the social planner's problem, whose objective is to maximize the real consumption index of the representative consumers in the economy.

The real consumption index for the representative consumer in the economy is defined as

$$Q = \Sigma_r \beta_r^s \left[M^s \int_{\varphi_r^s}^{\infty} N_f^s \int_{\phi_{fr}^s}^{\infty} (\phi_k C_{kr}^s)^{\frac{\sigma-1}{\sigma}} \, \mathrm{d}H(\phi) \, \mathrm{d}G(\varphi) \right]^{\frac{\sigma}{\sigma-1}}, \tag{23}$$

where β_r^s is the social planner's welfare weight on the real consumption index of each market r. The welfare weight is set as proportional to the labor endowment, i.e., $\beta_r^s = \frac{L_r}{\sum_r L_r}$. For each firm that enters the economy, the social planner chooses the number developed products N_f^s. For each consumption market r, the social planner chooses the productivity cutoff φ_r^s, the taste shifter cutoff ϕ_{fr}^s for all producing firms $\varphi_f \geq \varphi_r^s$, and the output level C_{kr}^s for all produced goods $\phi_k \geq \phi_{fr}^s$ for all producing firms $\varphi_f \geq \varphi_r^s$. Lastly, the planner chooses the mass of firms M^s subject to the aggregate labor constraint:

$$\Sigma_r L_r = \Sigma_r M^s \left\{ \int_{\varphi_r^s}^{\infty} N_f^s \int_{\phi_{fr}^s}^{\infty} \frac{C_{kr}^s}{\varphi_f} + F_p \, \mathrm{d}G(\phi) \, \mathrm{d}G(\varphi_f) + \int_{\varphi_r^s}^{\infty} F \, \mathrm{d}G(\varphi) \right\} + M^s \left\{ \int_{\varphi^*}^{\infty} F_d (N_f^s)^{\theta} \, \mathrm{d}G(\varphi_f) + F_e \right\},$$

where φ^* is the smallest cutoff productivity in the economy. The social planner faces the same productivity distribution $G(\varphi)$, taste shifter distribution $H(\phi)$, and fixed costs. The following proposition summarizes the social planner's choice with general distributions of productivity and taste shifter.

Proposition 1 *The social planner's choice of the output level, number of developed products, taste shifter cutoff, productivity cutoff, and mass of firms* *entering the economy coincides with the market allocation, i.e., the market equilibrium is efficient.*

The proof of this proposition is presented in Appendix A.3.

The intuition behind the efficiency of the market equilibrium is the following. First, the CES utility function and monopolistic competition leads to constant markups across firms, $\frac{\sigma}{\sigma-1}$. Therefore, there is no markup distortion across products or firms; hence the output level in the market equilibrium is optimal. Next, as for the taste shifter cutoff and productivity cutoff, there are two sources of potential inefficiency. When introducing another product or firm, firms do not internalize the positive effect on consumer surplus (the love-of-variety effect) and the negative effect on the sales of other products and firms (the business-stealing effect). The CES utility function uniquely ensures that these two externalities exactly offset each other. Lastly, as for the mass of firms in the economy, a potential inefficiency comes from the fact that, with heterogeneous productivity, some firms make profits and some others make a loss. A free-entry condition ensures that the profits and losses exactly finance the fixed costs of entering, which means that the mass of firms entering the economy is at an efficient level. Therefore, the market implements the efficient allocation.

6.2 Welfare cost of sub-optimal SME subsidies

The efficiency of the market equilibrium implies that any policy distorting the market allocation is sub-optimal. In this section, I estimate the social cost of the subsidies for Small and Medium Enterprises (SMEs) to expand their product scope, a policy tool frequently used in Japan. I consider two cases: the baseline and the single-product firms. Comparison between the welfare costs of the subsidies under the baseline model and the single-product firm model allows us to measure the importance of having both product and geographic scopes for quantifying effects of size-dependent policies.

6.2.1 SME subsidies in Japan

Although my model suggests that the market equilibrium with a vast difference in firm performance is efficient, the Japanese government has conducted various policies to promote SME growth.[25] One of the main policy tools to promote SME growth is subsidy for SMEs to expand their product varieties.

There are many types of subsidies for SMEs to expand their product varieties.[26] Coverage and financial sources of the subsidies vary across different types of subsidies, but most of them share the eligibility requirement and the types of costs that are subsidized. The condition for the companies to be eligible for subsidies is to have equity or employment less than a cutoff value.[27] The types of costs covered by subsidies include the costs of purchasing production equipment for the new products, hiring experts/consultants to develop new products, and the costs of market research for the new products.

Among these SME subsidies, I take the subsidy provided by the National Federation of Small Business Associations (a government owned financial agency) in 2020-2021 financial year as an example to analyze the welfare cost of the subsidy. This particular SME subsidy aims to promote growth of the SMEs located in Japan. The firms eligible for the subsidy are manufacturing firms with less than 300 employees. Whether a particular firm is eligible for

[25]The rationale for the Japanese government to support SMEs growth that is not captured in my model includes the political power of SMEs. Since SMEs represent 70% of all employment in Japan, the voting power of SMEs is larger than their economic presence.

[26]For example, a list of subsidies available as of June 4th, 2021, can be found here https://hojyokin-portal.jp/subsidies/list. Note that the website is available only in Japanese.

[27]For the manufacturing firms, the usual cutoff values are 300 million Japanese yen (approximately 3 million USD) for equity and 300 people for employment. Firms in the service and other industries have lower cutoff values.

the subsidy is determined based on the number of employees listed in the previous year's tax return documents of the enterprise tax. The subsidy covers the costs of purchasing production equipment for the new products, fees of using intellectual property rights for the technology necessary for the new products, costs of hiring experts/consultants for the product development, the raw ingredient and outsourcing costs for the new products. The coverage of the subsidy is up to two-thirds of the total costs. The costs are first paid by the companies, and the agency reviews the receipts and determines whether these costs are eligible for a subsidy. After getting it approved by the agency, up to two-thirds of the costs are reimbursed back to the companies.[28]

6.2.2 Modeling the SME subsidy

The eligibility, types of the costs covered by the subsidy, and the coverage of the subsidy are modeled in the following ways. In the model, firms' employments have a one-to-one relationship with their productivity, i.e., higher-productivity firms use more labor. Thus, imposing an eligibility threshold for the subsidy by employment (less than 300 employess) is identical to imposing an eligibility threshold on productivity in the model.[29] Define $\hat{\varphi}$ as the threshold productivity allowing the firms with the productivity below the threshold to be considered eligible for a subsidy. Because the firms who have less than 300 employees comprise 98.5% of all firms in the manufacturing industry in Japan, I set $\hat{\varphi}$ such that surviving firms with productivity less than

[28]See more detail about the subsidy in https://portal.monodukuri-hojo.jp/common/bunsho/ippan/13th/reiwakoubo_20221025.pdf. Note that the document is available only in Japanese.

[29]One caveat is that, in reality, firms may modify their number of employees to become eligible for a subsidy. However, a firm's eligibility is determined based on the previous year's tax return documents, and the announcement of the subsidy was made after the deadline for the tax document. Therefore, I conclude that firms did not decrease their number of employees to become eligible for the subsidy.

$\hat{\varphi}$ consist of 98.5% of all surviving firms:

$$\frac{G(\hat{\varphi}) - G(\varphi^*)}{1 - G(\varphi^*)} = 0.985, \tag{24}$$

where φ^* is the smallest cutoff productivity in the market equilibrium (before the subsidy is imposed).

Lastly, the costs covered by the actual subsidy correspond to the fixed cost of developing products in the model. Define s as a fraction of subsidy for the fixed cost of entering a market. Then, the government provides subsidy for the firms whose productivity is less than the threshold productivity $\hat{\varphi}$, i.e., the fixed cost of developing products = $(1 - s)F_d(N_f)^\theta$ for the firms with productivity $\varphi_f \leq \hat{\varphi}$. Lastly, to be consistent with the actual coverage of the subsidy, I set $s = \frac{2}{3}$ in the model.

I assume a balanced budget in the economy by imposing the uniform consumption tax on consumers in each market. The representative consumer's budget constraint in market r becomes

$$\int_{f \in \Omega_r} \int_{k \in \Omega_{fr}} (1 + \tau) p_{kr} C_{kr} \, \mathrm{d}k \, \mathrm{d}f \leq L_r. \tag{25}$$

The consumption tax τ is given by the following balanced budget:

$$\sum_r \int_{f \in \Omega_r} \int_{k \in \Omega_{fr}} \tau p_{kr} C_{kr} \, \mathrm{d}k \, \mathrm{d}f = M \int_{\varphi^*}^{\hat{\varphi}} s F_d(N_f)^\theta \, \mathrm{d}G(\varphi), \tag{26}$$

where the subsidy is granted for the firms entering the economy ($\varphi_f \geq \varphi^*$) but with the productivity less than the threshold ($\varphi_f \leq \hat{\varphi}$). In the single-market model, the balanced budget equation becomes

$$\int_{f \in \Omega_r} \int_{k \in \Omega_{fr}} \tau p_{kr} C_{kr} \, \mathrm{d}k \, \mathrm{d}f = M \int_{\varphi_r^*}^{\hat{\varphi}} s F_d(N_f)^\theta \, \mathrm{d}G(\varphi),$$

where market r is the only market that exists in the single-market economy.

6.2.3 Results: baseline and single-product market model

Moving from the market equilibrium to the equilibrium under the subsidy affects the real consumption index of each market in the following way. The firms with productivity less than the threshold productivity receive a subsidy to cover the fixed cost of developing more products; thus they are more likely to develop more products. The subsidy is financed by an uniform consumption tax on consumers in all the markets, which is borne by all the entering firms as a reduction in total consumer demand, including the high-productivity firms that do not receive the subsidy. Furthermore, because of the entry of more firms due to the subsidy, the competition within a market becomes harsher. In total, the subsidy helps the low-productivity firms develop more products and enter more markets but hurts the high-productivity firms' profits.

In the baseline model, the subsidy results in the loss of the real consumption index in the economy by 2.19%. In the model with single-market economy, the welfare costs of the subsidy are reduced to 1.47% of the real consumption index. There are two reasons why the single-market model underestimates the cost of the subsidy. First, the subsidy helps the lower-productivity firms to develop more products, but these firms only enter one market. Thus, the entry of lower-productivity firms induces a minor competition increase compared to the baseline model where subsidized lower-productivity firms enter multiple markets. Second, the high-productivity firms that suffer from the reduction in total consumer demand enter only one market with and without subsidy. Therefore, higher-productivity firms' geographic scope margin is not distorted, and the only effect is reflected on their product scope and intensive margins. Taken together, the single-market model underestimates the welfare cost of the SME subsidies.

7. Conclusion

The recent literature in firm heterogeneity has documented the importance of two extensive margins to account for firm size heterogeneity: product and geographic scopes. Although much attention has been paid to the role of each extensive margin separately, I argue in this paper that the two extensive margins interact in a major way and thus, for the need to have both extensive margins in one model. I develop a heterogeneous-firm model with multi-product firms and multiple markets with various market sizes.

In the model, firms' decisions on product and geographic scopes interact in two ways. First, as firms expect to develop and sell more products in given markets, they are more likely to enter more markets. Second, as firms enter more markets, they develop greater varieties of products because the firms make larger profits per product. The calibrated model reveals that each complementarity is quantitatively important. Eliminating either geographic or product scope from the model reduces firm size variations by 64 to 96%, failing to incorporate the complementarity between two margins. I further use the model to quantify the welfare implication of the actual size-dependent subsidy in Japan. Notably, the model with only a single extensive margin underestimates the welfare implication of such subsidies.

There are a couple of additional directions for research suggested by the paper. First, my analysis focuses on explaining cross-sectional firm size variations, and it is left unknown how the firms' performance and product and geographic scopes evolve. Exploring these aspects requires access to panel data that would allow analyzing observations made over some course of time. Second, my framework does not incorporate distance. Including distance in the model lets us explore what role shipping costs and taste heterogeneity, which are correlated with the distance between the markets, play in generating firm size heterogeneity.

References

Arkolakis, Costas, Sharat Ganapati, and Marc-Andreas Muendler, “The extensive margin of exporting products: A firm-level analysis,” Technical Report, National Bureau of Economic Research 2010.

Bagwell, Kyle and Asher Wolinsky, “Game theory and industrial organization,” *Handbook of Game Theory with Economic Applications*, 2002, *3*, 1851–1895.

__ and Garey Ramey, “Coordination economies, advertising, and search behavior in retail markets,” *The American Economic Review*, 1994, pp. 498–517.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, “Cross-country differences in productivity: The role of allocation and selection,” *American economic review*, 2013, *103* (1), 305–34.

Bernard, Andrew B, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, “The origins of firm heterogeneity: A production network approach,” Technical Report, National Bureau of Economic Research 2019.

__ , Stephen J Redding, and Peter K Schott, “Multiple-product firms and product switching,” *American Economic Review*, 2010, *100* (1), 70–97.

__ , __ , and __ , “Multiproduct firms and trade liberalization,” *The Quarterly journal of economics*, 2011, *126* (3), 1271–1318.

Dhingra, Swati and John Morrow, “Monopolistic competition and optimum product diversity under firm heterogeneity,” *Journal of Political Economy*, 2019, *127* (1), 196–232.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, “An anatomy of international trade: Evidence from French firms,” *Econometrica*, 2011, *79* (5), 1453–1498.

Eckel, Carsten and J Peter Neary, “Multi-product firms and flexible manufacturing in the global economy,” *The Review of Economic Studies*, 2010, *77* (1), 188–217.

Fernandes, Ana M, Peter J Klenow, Sergii Meleshchuk, Martha Denisse Pierola, and Andres Rodriguez-Clare, “‘The Intensive Margin in Trade: How Big and How Important?,” *Unpublished manuscript*, 2019.

Handbury, Jessie, “Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities,” Technical Report, National Bureau of Economic Research 2019.

Hopenhayn, Hugo A, “Entry, exit, and firm dynamics in long run equilibrium,” *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, “Quantifying the sources of firm heterogeneity,” *The Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Jovanovic, Boyan, “Selection and the Evolution of Industry,” *Econometrica: Journal of the Econometric Society*, 1982, pp. 649–670.

Luttmer, Erzo GJ, “Selection, growth, and the size distribution of firms,” *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

Mayer, Thierry, Marc J Melitz, and Gianmarco IP Ottaviano, “Market size, competition, and the product mix of exporters,” *American Economic Review*, 2014, *104* (2), 495–536.

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

__ and Stephen J Redding, “New trade models, new welfare implications,” *American Economic Review*, 2015, *105* (3), 1105–46.

Ossa, Ralph, “Trade wars and trade talks with data,” *American Economic Review*, 2014, *104* (12), 4104–46.

A. Appendix

A.1 Retail chains and firm sizes

This section answers whether larger firms are connected to a few larger retail chains and sell their products in many stores within the few retail chains. My data shows that this is not the case; larger firms sell in more retail chains, and these retail chains are not systematically larger than the retail chains that smaller firms are connected.

The first column of Table 5 shows the average number of prefectures that firms sell their products in increases with firm size. If larger firms have a wider geographic scope because they sell to a few retail chains with many stores in many prefectures, then I would expect the average number of retail chains to not increase with firm size. The second column of Table 5 shows that the average number of retail chains monotonically increases with firm size. It means that larger firms sell not only in many prefectures but also in many retail chains.

I also check whether the retail chains that larger firms connect have more stores than those that smaller firms connect. The last column of Table 5 shows that it is not the case; the average number of stores per retail chain does not increase monotonically with firm size. The smallest 10% firms sell to retail chains that have, on average, the largest number of stores. The findings in Table 5 suggest that the reason why larger firms sell their products in more prefectures is not that they sell to large retail chains, but they sell to many retail chains and stores.

A.2 More explanation on structural estimation

Section A.2.1 explains simulation algorithm used in the SMM.

Table 5: Number and size of retail chains by decile

Ranked decile	Avg. no. prefectures	Avg. no. retail chains	Avg. no. stores per chain
1	28.08	14.99	17.39
2	17.30	10.11	14.24
3	12.46	7.59	14.63
4	9.50	5.68	14.38
5	7.40	4.53	16.90
6	5.50	3.65	14.90
7	4.04	2.94	16.67
8	2.86	2.39	16.74
9	2.10	1.93	17.60
10	1.42	1.43	18.33

Note: The largest decile is ranked first. The values are weighed by 3-digit product category sales per quarter. Number of prefectures is calculated as the number of prefectures in which each manufacturing firm sells at least one product. Number of retail chains is calculated as the total number of retail chains firms sell their products to. Finally, average number of stores per chain in each decile is calculated as the average number of stores of each retail chain that firms in each decile sell their products to.

A.2.1 Simulation algorithm

A simulation of the model requires a set of parameters $\Theta_2 = \{\sigma, \alpha, \theta\}$, the data for product- and firm-level sales, (firm-level) prices, and number of products firms sell in each market. It involves the following steps:

Preparation: Assume a value of α, a shape parameter of the Pareto distribution. Draw productivity of 100,000 artificial firms from the productivity distribution: $\varphi_f \sim \text{Pareto}(\alpha, 1)$, where the scale parameter of the distribution is normalized as one.

Step 1. Among these simulated firms, randomly sample 10,000 firms in each simulation. Among the 10,000 artificial firms, use only the surviving firms in the model equilibrium, i.e., the firms with the productivity higher than the lowest cutoff productivity, for the remaining simulation.

Step 2. Calculate the prices of the surviving firms by expression 5. Note that, in my model, the firms' pricing depends only on productivity (φ_f) and the elasticity of substitution (σ).

Step 3. For each surviving firm f, compute the optimal number of developed products, N_f^* from expression (18). Draw the taste shifter for each product for N_f^* products from the distribution: $\log(\phi_k) \sim N(\mu, \sigma^2)$. Compute the firm's entering decision to each market.

Step 4. Compute the number of products firms sell in each market by the cutoff taste-shifter condition (8).

Step 5. Compute the product-level sales for each firm in each market by the following expression:

$$p_{fr} C_{kr} = (\phi_k)^{\sigma-1} \left(\frac{p_f}{P_r}\right)^{1-\sigma} L_r. \tag{27}$$

Compute the firm-level sales in each market by the following expression:

$$\int_{k \in \Omega_{fr}} p_f C_{kr} \, \mathrm{d}k = \int_{k \in \Omega_{fr}} (\phi_f)^{\sigma-1} \left(\frac{p_f}{P_r}\right)^{1-\sigma} L_r \, \mathrm{d}k. \tag{28}$$

Step 6. Estimate σ using expression (16) and φ_f for each surviving firm from expression (17). Compute the best-fitted shape parameter for the estimated productivity distribution: $\varphi_f \sim$ Pareto$(\alpha, 1)$.

Step 7. Compute the elasticity of the number of developed products with respect to the total sale of the firm.

These seven steps generate the three targeted moments: biased estimates of $\{\hat{\sigma}, \hat{\alpha}\}$ and the elasticity of the number of developed products with respect to the total sales of the firm. The SMM finds a set of the parameter values (σ, α, θ) such that the targeted moments in the model and data converge the same values. The weighing matrix is calculated as a variance-covariance matrix of the error function using the following bootstrap procedure: (i) re-sample, with replacement, 10,000 firms from our initial data set of the 100,000 firms 1000 times. (ii) For each re-sampling b, calculate moments $m(\Theta_2)$. (iii) Calculate

$$W_2 = \frac{1}{1000}\sum_{b=1}^{1000}(m_2^b - m^b(\Theta_2))(m_2^b - m^b(\Theta_2))'. \tag{29}$$

A.3 Social planner's problem

Proof of Proposition 1. First, I derive the planner's choice on intensive margin. The planner chooses the output levels C_{kr} to equate the the marginal rates of transformations (MRT) and the marginal rates of substitution (MRS) for the products of the firms with different productivity and for the products of the same firm in each market. The MRT between the products for any two different firms with productivity φ_1 and φ_2 in market r is:

$$MRS = \left(\frac{\phi_{k_1,r}}{\phi_{k_2,r}}\right)^{\frac{1-\sigma}{\sigma}} \left(\frac{C_{k_1,r}}{C_{k_2,r}}\right)^{\frac{1}{\sigma}}. \tag{30}$$

The MRT for the products for any two different firms with productivity φ_1 and φ_2 in market r is:

$$MRT = \frac{\varphi_1}{\varphi_2}. \tag{31}$$

The social planner chooses an output level to equate the MRS and MRT:

$$MRS = MRT \implies \frac{C_{k_1 r}}{C_{k_2 r}} = \left(\frac{\phi_{k_1 r}}{\phi_{k_1 r}}\right)^{\sigma-1}\left(\frac{\varphi_1}{\varphi_2}\right)^{\sigma}, \tag{32}$$

which yields the same relationship between relative quantities, relative productivity, and relative taste shifters as those in the market equilibrium. Similarly, I can show that the planner's choice of the output level for two goods within a same firm also coincides with that in the market equilibrium. Therefore, the output level in the market equilibrium is efficient.

Next, I derive the social planner's choice of the taste shifter cutoff. For a given firm f with productivity φ_f in market r, the social planner sets a taste shifter cutoff ϕ^s_{fr} that satisfies the first order conditions with respect to ϕ^s_{fr} and the $C(\phi^s_{fr})$, where $C(\phi^s_{fr})$ is the consumption quantity of the product with the cutoff taste shifter. Combining the two first order conditions, I obtain

$$\left(\frac{1}{\sigma-1}\right)\varphi_f^{-1}C(\phi^s_{fr}) = F_p. \tag{33}$$

The condition coincides with the cutoff taste shifter condition in the market equilibrium (8). Therefore, the value of the cutoff taste shifter in the market equilibrium is efficient.

Next, I derive the planner's choice of the productivity cutoff in each market. Rewrite the real consumption index (23) as

$$Q = \Sigma_r \beta_r (M)^{\frac{\sigma}{\sigma-1}} \left[\int_{\varphi^r_s}^{\infty} (C(\phi^s_{fr})\phi^s_{fr})^{\frac{\sigma-1}{\sigma}} N^s_f \int_{\phi^s_{fr}}^{\infty} (\phi_k)^{\frac{\sigma-1}{\sigma}} \, \mathrm{d}H(\phi)\, \mathrm{d}G(\varphi)\right]^{\frac{\sigma}{\sigma-1}}, \tag{34}$$

and the aggregate labor constraint as

$$\Sigma_r L_r = \Sigma_r M \left\{\int_{\varphi^s_r}^{\infty} \frac{C(\phi^s_{fr})\phi^s_{fr}}{\varphi_f} N^s_f \int_{\phi^s_{fr}}^{\infty} \phi_k^{\frac{\sigma-1}{\sigma}} + F_p \mathrm{d}H(\phi)\, \mathrm{d}G(\varphi) + \int_{\varphi^s_r}^{\infty} F \, \mathrm{d}G(\varphi)\right\} + \int_{\bar{\phi}^s_r}^{\infty} F_d (N^s_f)^{\theta} \, \mathrm{d}G(\tag{35}$$

Combining the first order conditions with respect to the productivity cutoff φ^s_r and the consumption quantity of products with the taste shifter cutoff sold by a firm with the productivity cutoff $C(\phi^r_{f,s}(\varphi^r))$ I obtain

$$\left(\frac{1}{\sigma-1}\right)(\varphi^s_r)^{-1} K \int_{\phi^s_{fr}}^{\infty} C(\phi^s_{fr}(\varphi^s_r))\mathrm{d}H(\phi) = K \int_{\phi^s_{fr}}^{\infty} F_p \mathrm{d}H(\phi) + F, \tag{36}$$

which corresponds to the cutoff productivity condition in the market equilibrium (11). Therefore, the productivity cutoff in the market equilibrium is efficient.

Lastly, I derive the planner's choice of the mass of firms entering the economy. To do so, I take the first order condition with respect to the mass of firms M, combine the condition with the first order condition with respect to $C(\phi^s_{fr}(\varphi^s_r))$, and take summation Σ_r:

$$\frac{\frac{1}{\sigma}\sum_r L_r - \sum_r \left\{ M\int_{\varphi^s_r}^{\infty} N^s_f \int_{\phi^r_{f,s}}^{\infty} F_p \,\mathrm{d}H(\phi) + F\mathrm{d}G(\varphi)\right\} - \int_{\bar{\varphi}^s_r}^{\infty} F_d (N^s_f)^{\theta}\,\mathrm{d}H(\phi)}{M} = F_e. \tag{37}$$

The first term in the numerator is equal to the total variable profit made in the market equilibrium. The second term in the numerator is equal to the total fixed costs paid by the firms in the market equilibrium. Thus, the numerator on the left-hand side of (37) equals the total profit made by the firms in the market equilibrium. Therefore, the condition (37) coincides with the free-entry condition in the market equilibrium; hence the mass of firms in the market equilibrium is efficient. ■