The Demand Side of Firm Growth: Evidence from Mexico*

Louise Guillouët
Columbia University

JOB MARKET PAPER

April 3, 2022
[Click here for the latest version]

Abstract: In developing countries, there is uncertainty about product quality, leading consumers to seek credible signals of quality. One of them is the fact that a good is produced by a foreign firm. Combining barcode-level consumption data from Mexico with information about the origin of the producers of the good, I measure a precise foreign price premium of at least 16%. While the availability of foreign goods increases consumers' welfare, the dominance of foreign firms may also hinder the growth of domestic firms. I then document the following novel facts about the consumer packaged goods industry in Mexico: 1) domestic firm growth is driven by surviving goods rather than new goods; 2) domestic goods have slower and longer life-cycles than foreign goods; 3) the extensive customer margin is key to growth for both types of firms; 4) domestic firms, depend relatively more on the intensive margin for customer growth; and 5) new customers of older domestic goods are poorer than those of new goods. I estimate a demand model, showing that the price premium elicited in the raw data can be attributed to consumers' relative preference for foreign goods. Importantly, this preference fades over time. I show that this is consistent with consumers learning about product quality, and provide consumer-level empirical evidence for this mechanism.

JEL: D22; F23; L25

Keywords: growth; quality uncertainty; international competition; learning; consumer goods

*Columbia University Department of Economics (email: louise.guillouet@columbia.edu). Enrique Seira at ITAM was essential in obtaining access and understanding of the data. I am very grateful to Eric Verhoogen, Amit Khandelwal and Jack Willis for their continuous encouragement and guidance throughout this project. I wish to thank Vittorio Bassi, Michael Best, Laura Boudreau, Gautam Gowrisankaran, Morgan Hardy, Jonas Hjort, Lucas Husted, Rocco Macchiavello, David Martimort, Monica Morlacco, Suanna Oh, Golvine de Rochambeau, Matthieu Teachout, Pietro Tebaldi and seminar participants at Columbia University, Analysis Group (Paris), Oxford University's Global Priorities Institute, and Dusseldörf Institute for Competition Economics for helpful feedback. I gratefully acknowledge the financial support of the Program for Economic Research and the Center for Economic Development and Policy at Columbia University. I would also like to thank Carmen Castro Rios at ITAM for her help in accessing the data and manipulating it on ITAM's Beethoven server. The data agreement details are registered with the Columbia University Institutional Review Board as the protocol AAAT7865. The views expressed herein are those of the author and do not necessarily reflect the views of Kantar, Columbia University or ITAM. All errors are my own.

1 Introduction

A prevalent feature in developing countries is the lack of information about the quality of goods. In the absence of regulation, firms need to implement signaling strategies such as advertising, branding, or certifications. One such signal is stating that a good is being imported from, or exported to a high-income country.[1] The trade literature has documented that firms based in high-income countries produce higher quality goods on average (Schott (2004), Hummels and Klenow (2005)). In this paper, I show that consumers indeed use the origin of goods as a signal of quality.

Globalization therefore benefits consumers in developing countries by allowing them to import high-quality products or to buy high-quality products from foreign firms investing at home (multinational corporations or henceforth MNCs). In this setting: Mexico, twenty years after the signature of the North American Free Trade Agreement, MNCs headquartered in high-income countries dominate many product markets. While this quality channel, in addition to the variety and price channels, increases consumer welfare, globalization also increases the set of competitors domestic firms must contend with.[2] It may become very difficult for relatively unknown domestic firms to grow when competing with long-established, high-reputation MNCs, as analyzed by Schmalensee (1982), even if they do produce high-quality goods. This demand problem undermines firm-level upgrading policies, which often focus on the supply side (Verhoogen (2021)).

In this paper, I ask how uncertainty about product quality explains the foreign price premium in Mexico, and how it affects the growth patterns of domestic firms. I study this question in the context of the Mexican consumer packaged goods sector, which represent almost half of households' expenditure in at-home consumption. Uncertainty about quality is a critical issue in the consumer food sector because low quality means that people get sick. However, it is unverifiable for the consumer and costly to verify for the regulator. As a result, it is not enforced consistently in developing countries. The United States' Food and Drug Administration (FDA) often recalls products that were exported from Mexico to the United States, a lower bound for the number of potentially unreported issues with product that did not leave the domestic market.[3]

I answer this question by leveraging a rarely used barcode-level dataset covering the universe of consumer-packaged goods in Mexico from 2010 to 2015. I establish whether the firm is foreign or domestic by merging the names of the 3985 firms making the products with an administrative dataset listing the firms which have receive foreign direct investment (FDI).[4]

[1] See an example in the Washington Post, September 9th 2012.

[2] Globalization also increases the size of the market firms have access to. For example, Alfaro-Ureña et al. (2019) show that becoming a supplier to an MNC has a large positive impact on domestic firm performance in Costa Rica. The relevant policy interventions might be those that help domestic firms access these opportunities (Atkin et al. (2017a), Hjort et al. (2020)). But Goldberg and Reed (2020) argue that ultimately firm sales must also come from the middle-class consumers at home.

[3] The incidence of foodborne diseases is four times higher in Mexico than in the United States. Source: OMICS and CDC. For a list of recent examples of food contamination events, see this media article and scientific article. See also a recent statement by the US FDA and its Mexican counterpart, Cofepris.

[4] 90% of the imports of developing countries from high-income countries are intermediate goods (Faber (2014)). Since I study consumption goods, it therefore seems nelevant to focus on MNCs goods instead of imported goods. The comparison of the two is left for future research.

First, I quantify the foreign price premium. Controlling for a range of product characteristics, I show that global brands charge at least a 16% price premium. Despite this higher price, the aggregate market share of foreign firms is far above 50% for most product categories. I hypothesize that this reflects a belief that foreign firms produce high-quality products, in a context of uncertainty. While this means that the market has identified a signal, increasing welfare compared to the absence of signal, it also suggests that domestic firms may have trouble selling their goods even if they are of high-quality.

I investigate this dynamic question by establishing five novel facts on firm-level sales growth. First, using a data-driven definition of new goods, I decompose the sales of firms in each year between sales of "new goods" (products that were introduced in the current year) and sales of "surviving goods" (products that are sold in the current year but were already being sold in the previous year). By comparing the sales of new goods and the sales of surviving goods of the same firm over time, I find that for domestic firms, sales growth mostly comes from the growth of surviving goods as opposed to new goods. This is in contrast with firms in the United States (Argente et al. (2019)): in the Nielsen data, surviving goods appear to contribute negatively to sales growth. Although I find that for foreign firms operating in Mexico, the majority of sales growth comes from new goods, surviving goods still contribute positively to growth. This fact suggests that Mexican consumers may generally value novelty less than consumers in the United States, and that domestic firms have an even harder time selling new products than foreign firms in Mexico.

Second, at the product level, new goods introduced by Mexican firms start off by selling considerably less than comparable new foreign products. But, conditional on surviving, sales grow and stay higher than in the initial quarter for up to two years. This is in contrast with the pattern identified in the United States by Argente et al. (2019). My finding is not due to data limitations, as I do find that when foreign firms introduce new goods, sales grow for a short period after the introduction of the good and then decline for a long period. Argente et al. (2019) interpret this as a consequence of the high rate of product innovation in this market: new products generate demand, but this demand gets attracted to newer products after a while. While my finding confirms that at least initially, domestic products start with a disadvantage, it also points to the fact that domestic firms in Mexico may acquire a better image relative to foreign products over time.

Third, following Einav et al. (2021), I show that for both types of firms, the key driver of sales growth is growth in the number of customers, as opposed to growth in the number of units sold to each customer or growth in the monetary value generated per unit.[5] The importance of the customer margin emphasizes the importance of marketing and outreach activities for firms. Importantly, this particular result does not appear to differ much for domestic as opposed to foreign firms, suggesting that the tools that foreign firms have used to achieve such a large customer base and large sales may also work for domestic firms.

Fourth, I show that in order to grow their customer base, domestic firms depend relatively more on the intensive margin of product markets — growing the number of customers they sell

[5]This latter variable can be interpreted as the average price per unit at the firm level. It can increase if the firm increases the price of existing products, but it would also increase if the firm introduced higher priced products.

each product to —, as opposed to the extensive margin — growing the number of products they sell. This suggests that the first-order problem faced by domestic firms is not that their products do not match the taste of domestic customers, but potentially that they must overcome barriers to convince customers that they will appreciate each of their existing products.

Fifth and last, I show that the new customers of domestic products that have survived several quarters are poorer than the new customers buying domestic products which have just been released. By contrast, the new customers of foreign products that have survived several quarters are richer than the customers of these same foreign products when they were new. Together with the second fact about the domestic product life-cycle, this finding suggests that customers who face stricter constraints in their consumption decisions do not buy new domestic products immediately, but instead wait until these products survive a certain age before purchasing them.

Motivated by these facts, I estimate a model of demand to quantify the relative preference for foreign goods, and its evolution over time. I show that consumers value foreign goods more than domestic goods on average. However, this foreign premium declines as products age. I show that while the premium is quite homogeneous across income groups, the decline is mostly driven by richer households. This finding is consistent with two models. On the one hand, it is consistent with a world of uncertainty about domestic product quality where consumers learn over time as shown in Appendix C. While poor consumers are attracted to low prices, they cannot afford to waste money on bad products. Instead, richer consumers are attracted to novelty, they can afford to experiment and are therefore able to discover high-quality domestic products. On the other hand, the finding that the foreign preference declines over time could be attributed to imitation. If domestic firms imitate foreign firms, and lag behind because they introduce products systematically after foreign firms do, then foreign products would depreciate over time.

I present several pieces of evidence supporting the hypothesis of consumer learning. First, I show that consumer behavior exhibits state-dependence: at the product-consumer level, having consumed a product before increases the probability of consuming it again, *ceteris paribus*. This effect is higher for domestic products than foreign products, despite domestic products being less successful on average, underlining the relative uncertainty surrounding domestic products. Within a product-consumer cell, I also show that the quantity purchased increases with purchase experience, consistent with a learning story. This effect is again higher for domestic products. Moreover, I show that this effect stabilizes over time, consistent with state dependence converging over time when it is generated by learning (Dubé et al. (2010)). Third, I exploit product heterogeneity to show that for products for which quality is more salient, such as infant formula, consumers exhibit more state dependence. Since there is no reason to believe infant formula is more amenable to imitation, this supports the learning hypothesis. Finally, I show that, consistently with the model, richer households exhibit more state dependence and therefore *apriori* more learning.

The rest of the paper is organized as follows. Section 2 describes the literature this paper aims to contribute to. Section 3 describes the data and the setting. Section 4 establishes the foreign price premium. Section 5 discusses the five dynamic facts I establish. Section 6 estimates a demand model allowing for the interaction of foreign origin and time. Section 7 shows evidence of the mechanisms at work. Section 8 concludes. Additional figures are presented in Appendix A and additional tables in Appendix B. The conceptual model is formalized in Appendix C.

2 Literature

Akerlof (1970)'s seminal theoretical paper showed that when buyers cannot tell apart high-quality goods from low-quality goods, both types will be pooled under the same price, leading to markets unraveling and ultimately large welfare costs. Buyers and sellers must therefore rely on strategies to signal and identify high-quality goods. A well-studied strategy, relational contracts, is not really possible in the retail context (see Macchiavello et al. (2015) for example). Warranty schemes (see Shimp and Bearden (1982)'s series of experiments) may work for durable goods but not for non-durables, while certification schemes are not always welfare increasing (see Dranove and Jin (2010) for a review, and more recently Bai (2018)). This paper contributes to thinking about how reputations can address asymmetric information.

Shapiro (1983)'s elegant model shows that if consumers can learn about quality over time, it is possible for firms to establish a reputation for quality and sell high-quality goods for a premium. For example, Bennett and Yin (2019) show that an Indian pharmacy brand was successful in selling high-quality products and increasing the average quality of drugs in the markets where it operates. However, Bronnenberg et al. (2015) show that more informed shoppers (such as pharmacists for drugs, and chefs for groceries) are less likely to buy branded products, suggesting that the brand price premium is too high compared to the quality it delivers. The underlying hypothesis in my paper is that the MNC price premium may be too high, reflecting frictions.[6]

Reputations can also be collective, as shown theoretically by Tirole (1996), including at the national level. Researchers often rely on Armington (1969) preferences to incorporate the origin of goods in trade models. In this paper, I contribute to providing empirical evidence for this type of preferences. The marketing literature has studied the effect of Country of Origin on consumers' willingness to pay for goods, although few use prices as outcomes,[7] see Papadopoulos and Heslop (1993) for a thorough review. Generally, consumers project a country's historic advantage in a product line (for example, German firms make good cars and French firms make good wines). Cagé and Rouzet (2015)'s model shows that individual firms may benefit or suffer from a country's overall reputation, leading to multiple equilibria. Bai et al. (2019) have shown that being associated with a country whose reputation deteriorates, such as China's in the context of the melanine contamination scandal, can harm an innocent firm especially if it is trying to export to high-income countries.

On the other-hand, buyers in developing countries generally value products imported

[6] The premium could also be due to a form of status associated with foreign brands (see for example Orlove (1997)), Batra et al. (2000)). This would mean that there is no such thing as "too high" a premium, as long as it generates utility for consumers. I believe it is less likely in the context I study, which is at-home consumption.

[7] The closest I could find was Hulland et al. (1996)'s study. They collected the prices of goods illegally imported to the Philippines and regressed them on their country of manufacture, controlling for the multinational. They find a 55% to 65% import price premium, even for products imported from low-income countries, suggesting that part of their finding might be attributed to the peculiarity of the environment.

from or made by MNCs headquartered in high-income countries.[8] However, this is has not been well documented. The trade literature has documented that higher-income countries generally produce higher-price and higher quality goods (Hummels and Klenow (2005), Schott (2004)). Faber (2014) shows that firms that use more imported inputs sell higher priced goods in Mexico. Ge et al. (2015) use customs data from China to show that foreign MNCs enjoy a higher export price than Chinese firms, controlling for productivity, which they attribute to quality. However, they do not observe product-level characteristics. In this paper, I turn to the consumer side and have access to a very high level of detail. Also using scanner data, Bems and Di Giovanni (2016) find a 28% import premium for consumer goods in Latvia. Confirming the point made by Dubois et al. (2022) about the potential of scanner data in developing countries, this paper is to my knowledge the first to estimate a precise foreign MNC price premium for consumer packaged goods in a developing country.

I then propose to unpack this foreign premium, contributing to a large literature trying to understand the formation of consumer preferences (see Bronnenberg and Dubé (2017) for a review). My central hypothesis is that the foreign premium reflects a belief about quality, which may be matched by domestic products over time as consumers learn about them. Estimating consumer learning is difficult: using consumption data together with survey data, Shin et al. (2012) show that learning is imprecisely- and often over-estimated. Dubé et al. (2010) show that state-dependence can be mistaken for learning. However, Crawford and Shum (2005) and Ackerberg (2003) show that there can be learning for new products. Here, the data allow me to measure the novelty of products at the intensive margin, in terms of their age. I am therefore able to measure learning by comparing newer products to older products.

This paper also contributes to thinking about the role of of advertising and marketing in helping firms signal individual quality (see Bagwell (2007) for a review), including at the export margin (Arkolakis (2010)). Using survey data on Pakistani soccer ball manufacturers, Atkin et al. (2017b) show that that the firms who charge the highest markups are the ones that make the most marketing efforts and not necessarily the most productive ones. Using an RCT, Hjort et al. (2020) show that teaching firms marketing skills can expand the market they have access to, in particular towards large buyers, and thus enhance growth possibilities. My paper uses less direct evidence, following recent papers looking at firms in the United States. Using scanner data, Hottman et al. (2016) and Afrouzi et al. (2020) have shown that firms' reputation are an important driver of sales. Using credit card data, Einav et al. (2021)'s model suggests marketing is a key input towards building that reputation, although they see it as a trade-off with innovation efforts. In line with these papers, I don't have direct data on marketing, but the findings in this paper suggest that it can have a significant impact of firm performance through the demand side.[9]

I contribute to understanding the life-cycle of products. Argente et al. (2019) study the

[8]See the India Times's 2015 article about consumer packaged goods; See also IFPRI's note on rice in Ghana, Ragasa et al. (2014) to understand why this may be a concern.

[9]The internet may introduce cheaper alternatives for firms to market their products (see Chen and Wu (2020) for example). While e-commerce is still out of reach for many small firms in developing countries, it is an avenue for growth. In this paper, I study a market yet relatively untouched by the Internet, but which will be increasingly so: in 2020, 25% of Mexican consumers have bought groceries online up from 13% in 2017 and less than 3% in 2014, OECD, ICT Access and Usage by Households and Individuals.

product life-cycle of consumer goods in the United States. They show that product turnover is high and firms must constantly reinvent their product scope in order to avoid business stealing, even though this strategy increases cannibalization. Perla (2019) proposes an alternative model to explain these life-cycles. His central idea is that customers may be "aware" of some products and not others. This means that firms have more market power than what the nominal number of competitors suggests. As products age though, customers learn about their existence through social networks. A relevant review of empirical work estimating social learning effects is Mobius and Rosenblat (2014), with some of the most famous ones Foster and Rosenzweig (1995) and Duflo and Saez (2003). It is not possible to estimate social learning effects empirically here due to the structure of the data, but it would be an interesting follow-up as a survey or an experiment.

Last, this paper contributes to understanding how trade, through the presence of MNCs, affects the welfare of consumers in developing countries. Fajgelbaum and Khandelwal (2016) show that trade is generally pro-poor because it decreases the price of goods that poor consumers spend a relatively higher share of their budget on. However, using more detailed consumer survey data from India, Atkin (2013) shows that because preferences change slowly, trade-induced decreases in prices may not increase poor consumers' calorie intake as much as one would expect. Using barcode-equivalent data from Mexico, Atkin et al. (2018) further show that the arrival of Walmart in new areas is relatively more beneficial for higher-income consumers. However, they do not observe the origin of the goods consumed, which I do. Using information about the origin of goods, Atkin and Donaldson (2015) show that intra-national trade costs might limit the ability of consumers in remote area to benefit from trade. However, they are only able to look at about a dozen different goods across three countries. In this paper, I have access the precise origin of the firms supplying the goods consumed in the entire urban Mexico. As underlined by Dubois et al. (2022), the richness of this scanner data, rarely used in developing countries, offers new insights.

3 Data and Setting

In this section, I describe the three primary data sources containing information on the Mexican consumer goods sector, covering the period of January 2010 to December 2015.

3.1 Data

The main source of data is a rotating household panel shared by Kantar World Panel.[10] Households are visited weekly by surveyors, who scan and collect all receipts for the packaged goods purchased for at-home consumption. I observe[11] 7182 distinct households per month on average and a total of 15750 unique households from all 30 out of the 32 Mexican states.[12] The sample is designed to represent metropolitan areas (collection of municipalities) in Mexico with more

[10]Kantar World Panel is an international company that operates in more than 50 countries. They specialize in the collection of household consumption data for marketing and sales strategy purposes. For more information on the data-sharing agreement, see Aguilar et al. (2021).

[11]Given the observations I exclude because the purchases are not identified by their manufacturer.

[12]On average, households stay in the panel for 3.4 years. I observe 1191 households on all 72 months. The two states not represented are Baja California Sur and Zacatecas.

than 50,000 individuals.

The panel provides information on the universe of the purchases of packaged goods, except for liquor and tobacco. For each purchase, I observe the transaction date, several characteristics about the good purchased, the price, the number of units purchased, the type of the store where the purchase was made,[13] whether the product was subject to special promotions, and the payment method. Importantly for the rest of the paper, I also observe the name of the manufacturer of the good. For soft drinks purchases for example, the product characteristics would include whether the drink is diet or regular, the flavor (cherry, coke, etc.), the package (can, plastic bottle, etc.), potentially other information such as temporary marketing by artists or for big sports events, and the content size. The manufacturer variable would then indicate "Coca-Cola FEMSA" if the item was produced by the Coca-Cola subsidiary in Mexico. I mostly do not observe products purchased in bulk or by the weight such as fresh fruits and vegetables, meat and fish, etc. Products that for which the manufacturer is not identified are dropped.[14]

The panel also contains economic, demographic and geographic information about each household. I observe these variables at the yearly level. They include information about household members' age, gender and occupation. I also observe asset-like characteristics about households: I know whether they have a fridge, a TV, and other appliances. I observe a few characteristics about the dwelling where households live. Last, I observe a socioeconomic status (SES), which is computed based on households' assets, the dwelling characteristics, the head of household education and purchasing power, following the methodology developed by the the Mexican Association of Market Intelligence Agencies (AMAI). I observe six values for the SES.[15] Geographic variables include the state, city and neighborhood of residence for Mexico City residents.

The second source of data is the yearly updated directory of private establishments (Directorio Estadístico Nacional de Unidades Económicas or DENUE) conducted by the Mexican national statistical institute, INEGI. DENUE was first created in 2010 based on the 2009 Economic Census. Since then, DENUE has been used as a sampling frame for business surveys.[16] Although this dataset provides the exact addresses of all 5478689 establishments listed, the KWP dataset only lists one of the names of the firm, which may be multi-establishment. I therefore merge this name with either the given or the official name of an establishment in the administrative dataset, but cannot tell which establishment actually corresponds to each good. This dataset is important because it helps me narrow down the definition of the firms I study in this paper.

The third source of data is the yearly updated register of foreign investment (Registro Nacional de Inversiones Extranjeras or RNIE) which is maintained by the Mexican Economic Secretary (Secretaria de Economia). Firms based in Mexico that receive foreign capital directly (FDI, as opposed to through stock) are listed there. There are 65810 firms in total. I match

[13] I observe the name of the retail chain if the purchase was made at a retail chain. If not, I observe whether the purchase was made in a store that does not belong to a chain or in an open-air market.

[14] For this reason, I am unable to do the analysis on unbranded goods as unbranded goods are not identified by their manufacturer.

[15] AB, C+, C, D+, D and E. See AMAI's website for more explanation

[16] See a commentary from the American Statistical Association

the name of the manufacturers collected by Kantar with the names of firms listed in this list. Based on this, I define a firm as foreign if it has received a positive amount of foreign direct investment. When relevant, I also collect the country or countries from which the investment originate from.

3.2 Descriptive Statistics

Households I summarize the characteristics of the panelists surveyed in 2010 in Table B.1. To see how they compare to a representative data, I also describe these characteristics for the households surveyed in the expenditure and income survey of households, ENIGH, conducted by the national statistical office, INEGI, in 2010. Households in the Kantar data appear to be slightly younger and more educated than households in the ENIGH dataset, which is not surprising given that Kantar focuses on urban households. I do not observe the income of Kantar households. While it is difficult to compare expenditure since the two surveys report it using very different categories, I offer a comparison in the last line of the table, which suggests that spending on consumer packaged goods represent 46% of total spending on at-home food, personal care and household care.

Goods I observe 66059 different products across 82 categories. For each product, I observe between one and seven characteristics such as brand, flavor, color, size, number of units in the package etc. This description is similarly precise as the one available in better known datasets such as Nielsen which report at the barcode level. In the baseline specification of the empirical analysis, I will refer to these goods defined by Kantar under a product identifier as "barcodes".

A large part of the empirical analysis relies on the identification of "new" goods in the dataset. In this paper, the definition of new products is empirical. I define the date of birth of a good as the first date it appears in the dataset.[17] Of course, for goods that appear in the first month of data, I cannot tell whether they are 0 month old, 1 month old, or 12 months old. In the analysis that follows, I always control for cohort fixed effects and verify the robustness of my main results to excluding goods that appear before the first six months of the data.

Firms These barcodes are manufactured by 3985 different firms, 94% of which receive no foreign direct investment. The few firms who do are much larger as shown in Figure A.1, which plots the distribution of annual total expenditure panelists made on each firm. Among foreign firms, 49% receive FDI from the USA.[18] The rest of the foreign investment is from 30 different countries, which limits my ability to do separate analysis by country of origin.

4 Stylized Static Facts

In this section, I present novel evidence on the place occupied by foreign firms in the retail landscape faced by consumers in a middle-income country.

[17] I validate this data-driven definition by checking that new products referenced in the marketing firm The Market Think's review were indeed classified as new according to our definition. Although most products do not appear in our dataset, the ones that do were indeed classified as new. Examples include Coca-Cola's Life drink and Nestle's Oikos Greek-style yogurt.

[18] This is similar to the share of US investment in the total foreign investment inflows received by Mexico over the last 20 years. Author's calculation based on the data published by the Secretaria de Economia.

4.1 The Foreign Price Premium

In this section, I study prices. Although an obvious outcome of interest, it is not easy to compare prices of goods sold by foreign firms and sold by domestic firms in the same country. The reason is that most datasets which include the origin of the good are at relatively low level of disaggregation (see for example Ge et al. (2015)), while most datasets that precisely describe goods that are being purchased by consumers do not include the origin of the goods. This is the case for a widely used set of consumption surveys, the Living Standard Measurement Surveys designed by the World Bank and then conducted by the World Bank or the relevant national statistical offices. Very few questions in few questionnaires include expenditure questions for both a domestic item and its imported version. In Appendix Table B.2, I show the import price premium measured on rice. I find a positive and significant price premium for all five countries for which I was able to identify it. In four out of five countries, the premium is above 20%, a large dispersion. Of course, it might be attributed to unobservable quality differences, such as grade (Ragasa et al. (2014)). In the Kantar data, it is possible to control for this type of quality attribute.

However, even scanner datasets in both high-income and middle-income countries do not include the origin of the good except for Bems and Di Giovanni (2016) in Latvia. I provide to my knowledge the first estimate of an import price premium in developing countries with such a high level of precision. I run the following regression:

$$\ln\ p_{i,g,t} = \alpha + \beta\ \text{foreign}_i + \zeta_{gt} + \mu_k + \epsilon_{i,g,t} \tag{1}$$

where I regress barcode i's price p observed in city g in month t on a dummy for whether the manufacturer is foreign or not. I control for city by time fixed effects ζ_{gt} and for product fixed effects $(i \in k)$. Table 1 describes the results, varying the narrowness of the product fixed effects used. Looking at the most simple regression, I find a 16% price premium, which is not driven by a volume bias: controlling for the barcode volume does not move the MNC price premium (Column 2).

Using an increasingly narrow product definition and therefore increasing the number of product fixed effects, I find a consistent, close to 20% price premium, where the upper bound is obtained by controlling for the largest number of product fixed effects (the narrowest product definition). The R2 jumps between Column (1) and Column (2) as I include the unit volume, but is relatively stable after this. The fact that controlling for observables does not change the coefficient of interest or the R2 much is reassuring regarding the potential for unobservables to bias the estimate (Oster (2019)).

One can wonder whether this is different across source countries. I propose to separate the country of origin of the foreign MNCs into two groups: MNCs headquartered in high-income countries and the rest. I run the same specification as above, separating the foreign dummy along these two groups. I show in Appendix Table B.3 that across the five specifications including the volumeme control, the price premium is twice higher for MNCs based in high-income countries compared to MNCs based in any other country, consistent with expectations based on the literature (Schott (2004), Hummels and Klenow (2005)).

	Log price					
	(1)	(2)	(3)	(4)	(5)	(6)
MNC	0.164	0.164	0.158	0.169	0.189	0.205
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
City by month FEs	6011	6011	6011	6011	6011	6011
Product FEs	180	180	10912	16647	23635	26397
Volume control	No	Yes	Yes	Yes	Yes	Yes
R2	0.53	0.87	0.91	0.92	0.93	0.94
N	43759653	43759369	43758021	43697214	42445758	37548230

Table 1: MNC price premium

Notes: Table reports regressions of the log price on a dummy for whether the firm is foreign, city by month fixed effects, and various level of characteristics. Column (1) only controls for the product category by subcategory fixed effects. Columns (2) through (6) also control for the volume of the barcode. Columns (3) through (6) include narrower fixed effects defined by product characteristics. The row "product FEs" describes how many fixed effects are absorbed in the regression. The number of observation changes across specifications because the reghdfe command drops singleton observations.

Following Bems and Di Giovanni (2016), I then investigate heterogeneity among product categories. Focusing on the 20 largest product categories in the dataset, I run regression 8 separately for each category. For each, I show in Figure 1 the premium obtained in three different specifications. In the first specification, I only control for city by month fixed effects. In the second specification, I also include a control for one product characteristic. In the third specification, I also include a control for the unit volume of the barcode.

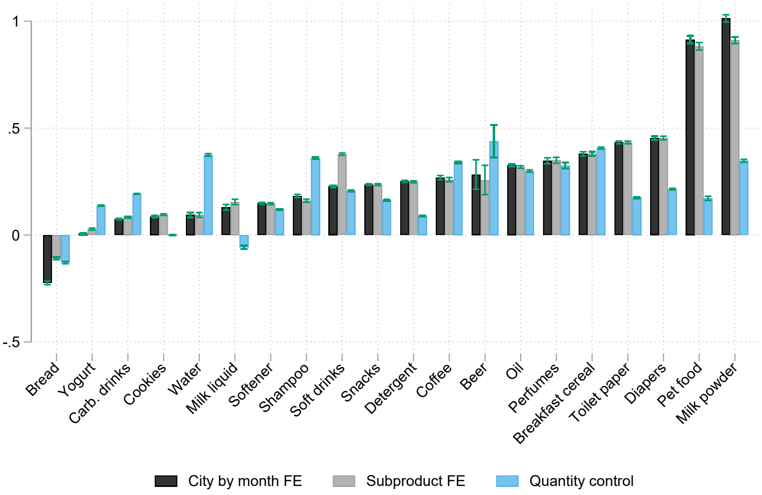


Figure 1: Foreign price premium

Notes: Figure reports the coefficient obtained from the purchase-level regression where the log price of the product is regressed on a dummy that turns on if the product is sold by a foreign firm and characteristics fixed effects. I repeat this regression three times for each of the top 20 categories in the dataset. Each regression uses incrementally additional characteristics as controls, which are described in the legend (gray and blue coefficients are obtained from regressions that also have city-by-month and and subproduct fixed effects, respectively). 95% confidence intervals are reported using the green bars. Products are sorted according to the size of the coefficient obtained in the first regression.

Almost all the coefficients are positive, but there is a lot of heterogeneity. I note that once again, controlling for the unit volume changes some coefficients, but adding additional controls does not move the coefficient much. The foreign price premium can be very large, up to 100% for the milk powder category, which is probably related to the subsidized price of the domestic products sold by public establishments such as LICONSA and the importance of quality of products such as infant formula. In the one category among the top 20 where the foreign price premium appears to be negative, bread, I hypothesize that this is because Bimbo, a Mexican company, has itself become a MNC operating in several countries including the United States.

4.2 The Foreign Share Premium

I have shown that foreign products are sold at a 16 to 20% price premium. One could still think that as a result these are elite items, sold only to the richest consumers. On the contrary, 55% of all sales in the dataset go to foreign firms and it does not appear to vary much across income groups. The fact that MNCs enjoy large market shares despite high prices suggests that consumers believe them to be of higher quality (Khandelwal (2010)). I run a similar regression as above, for firm-level market share in each product category:

$$y_{k,t} = \alpha + \beta \text{ foreign}_k + \vartheta_t + \epsilon_{k,t} \tag{2}$$

except now k is a firm and t is a year. I look at market shares defined in terms of sales, quantity and volume.[19]

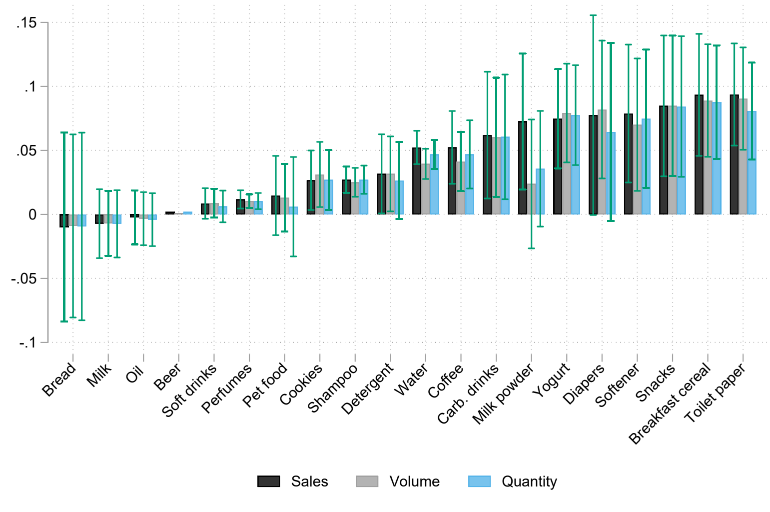


Figure 2: Foreign share premium

Notes: Figure reports the coefficient obtained from the firm-product-category-level regression where the market share of the firm in that product category is regressed on a dummy that turns on if the firm is foreign and year fixed effects. I repeat this regression three times for each of the top 20 categories in the dataset. Each regression uses a different definition of the market share, which are described in the legend. 95% confidence intervals are reported using the green bars. Products are sorted according to the size of the coefficient obtained in the first regression.

[19] To compute volumes, I use the information given by Kantar about the content of each barcode, usually in milliliters or grams depending on the product category. I drop observations that have a different unit from the rest of their category (about 1.5% of the data).

Figure 2 shows the Foreign coefficients β obtained for the 20 largest product categories. All but four coefficients are positive, which I interpret as the "foreign share premium". I point out the correlation between the noisy, negative coefficients observed for the sectors of milk, beer and bread with the existence of three very large Mexican firms in these sectors (Lala, Modelo and Bimbo respectively which are themselves MNCs, based in Mexico). The noisy negative coefficient obtained for milk is further linked to the importance of a public establishment, LICONSA,[20] which sells milk to 18% of urban households. Further, I remark that the negative share premia for bread, milk and beer are correlated with the smaller price premia for these product categories, emphasizing the idea that in the categories for which there is no domestic powerhouse that is trusted, consumers turn to foreign firms which have proven their ability to deliver quality in other markets.

5 Stylized Dynamic Facts

In this section, I leverage the rich data about consumer goods and establish a series of five novel facts about the dynamics of demand faced by domestic and multinational companies in Mexico.

5.1 Domestic Firms Grow Relatively More through Surviving Goods

I follow Argente et al. (2019) who use scanner data to study the life-cycle of products in the United States. I decompose a firm's growth rate into the sum of a "product innovation" component and a "product life-cycle" component. I use the following approximation:

$$\Delta S_{k,t} = \underbrace{\Delta S_{k,t}^{old,survive} - \overline{S}_{k,t-1}^{old,exit}}_{\text{life-cycle}} + \underbrace{n_{k,t}^{new} \times \overline{s}_{k,t}^{new}}_{\text{innovation}} \tag{3}$$

where $\Delta S_{k,t}$ is the firm's annual growth rate, and where the life-cycle component is further decomposed in the annual growth rate of the aggregate sales of "surviving products", products that sold a positive amount in both year $t-1$ and year t, $\Delta S_{k,t}^{old,survive}$, from which I subtract the share of sales that the products that exited between year $t-1$ and t (products that sold a positive amount in year $t-1$ and zero in year t) represented in the sales of the firm in year $t-1$, $\overline{S}_{k,t-1}^{old,exit}$. The innovation component is the product of the rate of introduction of new products in the firm's portfolio $n_{k,t}^{new}$ and the relative sales of new products compared to older products $\overline{s}_{k,t}^{new}$, obtained by taking the ratio of the average sales of a new product in year t to the average sales of a surviving product in year t.

I show in Figure 3 that domestic firms grow more than foreign firms. This may not seem surprising given that on average domestic firms are smaller than foreign firms. However, this finding is robust to binning firms into categories based on size as determined by sales. In Appendix Figure A.2, I show the comparison of domestic firms and foreign firms in the top quarter of the firm size distribution and in the bottom quarter of the firm size distribution. This is important because I showed above that looking at a snapshot of the Mexican economy between 2010 and 2015, MNCs are dominating the consumer goods market, but this domination may be fading over time if domestic firms continue to grow.

[20] See Jiménez-Hernández and Seira (2021) for an assessment of government's role in milk provision.

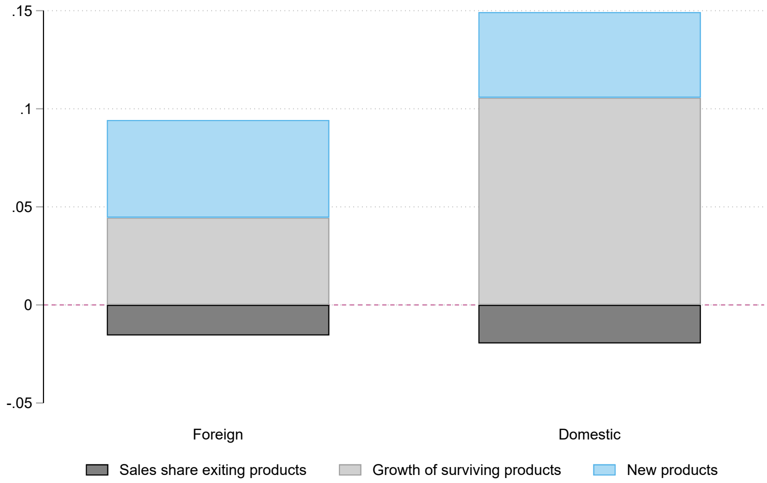


Figure 3: The extensive customer margin is key to firm growth

Notes: Figure represents the average firm-level year-to-year growth components, separated out by the origin of the firms. The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products. See Appendix Table B.5.

For foreign firms, growth appears to be driven equally by the introduction of new products and the growth of surviving goods. The "new products" component is comparable across foreign and domestic firms. The difference in average growth between foreign firms and domestic firms is almost entirely driven by the higher growth rate of surviving products for the latter. Appendix Figure A.3 shows the relative components, illustrating this point even more clearly.

One could wonder whether this is simply driven by the fact that foreign firms introduce new goods more frequently. I therefore study the rate of introduction of new goods at the firm level. Columns (1) to (3) of Appendix Table B.4 show that if anything, foreign firms appear to introduce slightly less new goods than foreign firms. However, this result may be driven by quality heterogeneity: if the new products introduced by domestic firms systematically sell poorly, either the firms go under or have to release a different product. Therefore, in Columns (4)-(6) of Appendix Table B.4 I focus on firms that sell at least one good that is not new in the year of observation. I find that the rate of introduction of new goods is smaller for all firms, at only 11% for domestic firms on average, and is larger for foreign firms by 4.6 percentage points. The first conclusion of this table is that both the rate of introduction and customers' appreciation of domestic firms' new goods must be playing a role here. Moreover, this foreshadows the implications of the foreign premium: if certain firms face difficulties finding customers for the same underlying level of quality, they will have less ability to invest in quality and less incentive to do so, reinforcing the prevalence of low quality products in the market.

Overall and including for foreign firms, I find that surviving goods contribute positively to the growth of firms from year to year. This is quite different from what Argente et al. (2019) observe: they find that goods that survive sell less overall from one year to the next, therefore contributing negatively to firm growth. There are two possible sources for this difference. One is that the life-cycle of products is different in Mexico. The second one is that because firms are newer on average in the Kantar data than in the Nielsen data exploited by Argente et al. (2019), the stock of surviving goods are also newer in the Kantar data, and therefore in any year there are fewer very old goods whose sales decrease dramatically in the following year. This hypothesis is partly confirmed by looking at the data separately depending on whether firms are four years or older, or younger than three years (based on the earliest day I observe the firm in the data). In Appendix Figure A.4, surviving goods contribute negatively to the growth of foreign firms which are four years or older. However, it is still not the case for domestic firms that are four years or older. For firms that are less than four years old, the pattern is very similar to the one shown above. This confirms the hypothesis that product life-cycles are different from Mexican firms in Mexico than for US firms in the US, although they may not be different for US firms operating in Mexico, as I will show below.

Another question of interest is how this decomposition varies depending on the time horizon chosen to look at the evolution of sales. Because many products survive only a few years, when looking at the decomposition of the two-year or five-year growth rates, a much higher share of growth comes from new products, both for foreign and for domestic firms. Further, exiting products take a much larger toll on growth, as shown in Appendix Figure A.5. However, the difference is still striking between foreign and domestic firms: even over five years, sales growth for domestic firm comes relatively more from the growth of sales of surviving products than new products compared to foreign firms.

5.2 Domestic Products Have a Slower Life-Cycle

I again follow Argente et al. (2019) and analyze the evolution of product-level sales over time. I estimate the following equation:

$$\log y_{i,t} = \alpha + \sum_{a=1}^{14} \beta_a D_a + \lambda_{jt} + \theta_c + \epsilon_{i,t} \tag{4}$$

where i is a barcode observed in a certain quarter t, a is a potential age of the good in quarters (between 0 and 14). j is a product category and c is a cohort-quarter. I regress a good u's log sales in a quarter t on dummies for the age of this product, product category interacted with quarter fixed effects, and cohort fixed effects (corrected following Deaton (1997)'s suggestion to avoid collinearity). I restrict the analysis to products born between 2011Q1 and 2012Q2, to maximize the change of measuring product age cleanly and to be able to follow products for 14 quarters, which is the maximum number of quarters products born in 2012Q2 can be observed for until 2015Q4. I only keep products which sold a positive amount in each quarter of their "life". I perform this regression separately for domestic and foreign products.

In Figure 4, I plot the coefficients obtained on each of the 14 quarter-age dummies. The coefficients obtained for the products introduced by foreign firms are very similar to the ones obtained by Argente et al. (2019) in the United States scanner data: products are born and sell, they exit in the next quarter or sell more, and then in each following quarter they exit or sell less and less until they exit. The authors attribute this to a constant arrival of products in the market, which means that after a while demand for the aging products is captured by

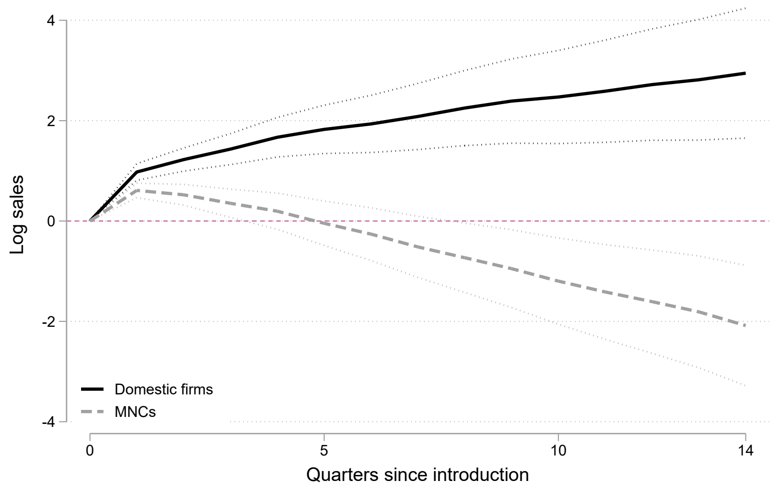


Figure 4: Product life-cycle

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression where I regress these dummies, quarter by product category fixed effects, and cohort fixed effects on the sales of a product in a given quarter. I run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines. See Appendix Table B.7.

newer products released by the same firm (cannibalization) or other firms' products (business stealing). By contrast, when domestic firms introduce new products, sales increase in the initial quarters and then remain high for a long period. While this may seem like an advantage for domestic products, they sell much less at any age, so overall the lifetime total sales of the average domestic product are much lower than the lifetime total sales of the average foreign product. Moreover for a given level of total lifetime sales, displacing sales later in time means that the present discounted values of domestic products is smaller.

A potential explanation for this phenomenon could be the endogenous timing of product retirement: if foreign firms decide to retire products later than domestic firms conditional on a sales trajectory, I would observe that conditional on surviving a set number of quarters, foreign average product sales decrease over time. Since the product duration is the same between foreign and domestic goods, it seems unlikely. I can further rule out this hypothesis by showing that the survival rate of foreign products is not different from that one of domestic products as they age, as shown in Appendix Figure A.6.

A potential concern is again that the difference between domestic firms and foreign firms be entirely driven by size heterogeneity. The findings are robust to controlling for firm size as measured by firm sales, as shown in Appendix Table B.7, Columns (3) and (4). When looking separately at four categories of firms, binned based on their sales, the difference between foreign and domestic firms only holds for the first two bins. Firms on the larger side of the spectrum do not exhibit this difference. In Appendix Figure A.7, I show similar figures as above for firms in the bottom 25% and the top 25% of the size distribution. This means that the average difference we observe is driven by differences between small domestic firms and small foreign firms. This is significant: it means that large Mexican firms such as Bimbo, Lala, Modelo etc. overcome the collective reputation of domestic firms. By contrast, foreign firms never face this problem, regardless of their size. Even if they are very small, consumers trust their new products to be of high quality. This suggests that some Mexican firms have the potential to grow and be as productive as MNCs operating in Mexico. However, only some of them are able to overcome the trust issue. It may be the case that the selection on the ability to gain consumers' trust is not efficient, and that some firms who have the potential to be very productive do not grow.

Because of the particular necessities of this regression, it is difficult to slice the data further between firms that are older and firms that are younger: 95% most of the products first appearing in the data between 2011Q1 and 2012Q2 and surviving 14 quarters were introduced by firms born in 2010 or before. The results are robust to excluding younger firms from the analysis.

5.3 The Extensive Customer Margin is Key to Firm Growth

I now decompose firm sales in a different manner. Here, I follow Einav et al. (2021) in studying the exact decomposition of a firm's sales:

$$\text{Sales} \equiv \text{Customers} \times \frac{\text{Quantity}}{\text{Customers}} \times \underbrace{\frac{\text{Sales}}{\text{Quantity}}}_{\text{Unit value}}$$

This equation can be interpreted in the following way: in order to double sales, firms may do one of three things, *ceteris paribus*:

(i) double their number of customers - the extensive margin

(ii) double the number of items they sell to each customer - one intensive margin

(iii) double the monetary value obtained from each unit sold - another intensive margin

The firm upgrading literature, and particularly the technology adoption literature, has mostly focused on (iii) (Verhoogen (2021)). The intuitive argument is that if a firm increases the quality of its output, it will be able to sell each unit for a higher price, and not only increase sales but also profits. Here, I remain agnostic and look at which margin seems to generate the most variation in sales. Of course, these three variables are endogenous as highlighted by the sentence above and I am not making any causal claim. I begin by taking logs of each element:

$$\log(\text{Sales}) = \log(\text{Customers}) + \log(\text{Quantity per Customer}) + \log(\text{Unit value})$$

I perform three regressions, regressing each element in turn on the log of Sales.

$$\log(\text{Customers}_{kt}) = \alpha + \beta_C \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5a}$$
$$\log(\text{Quantity per Customer}_{kt}) = \alpha + \beta_Q \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5b}$$
$$\log(\text{Unit value}_{kt}) = \alpha + \beta_U \log(\text{Sales})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{5c}$$

where i is a firm and t is a year. I control for firm fixed effects (η_k) and year fixed effects (ϑ_t) which means the coefficients are identified from looking at the years when firms grew faster, or slower, than their own average. By construction,

$$\beta_C + \beta_Q + \beta_U \equiv 1$$

The results are presented in Figure 5. The customer margin accounts for 60% of sales variation

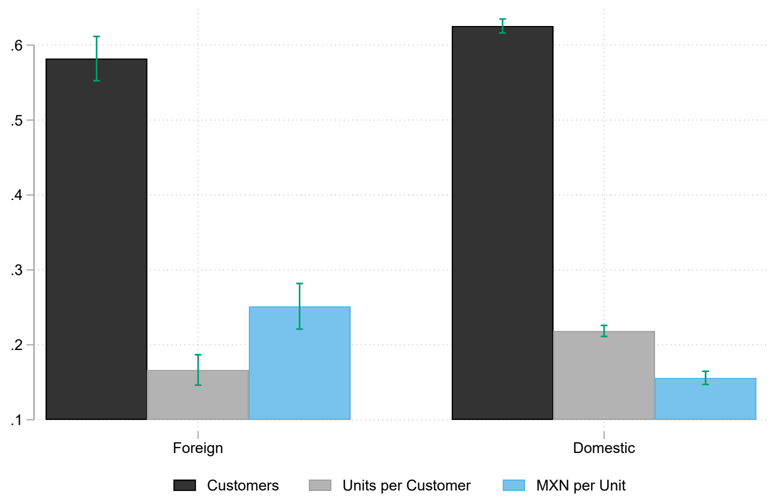


Figure 5: The extensive customer margin is by far the most important

Notes: Figure represents the coefficients obtained on log sales from firm-year-level regressions of log customers, log units per customer and log monetary value per unit, respectively on log sales, product category and year fixed effects. I run each of the three regressions separately for foreign and domestic firms. 95% confidence intervals are represented using the bars. Table B.8 shows the numerical values of these coefficients and an alternative specification.

within firms over time. This is true for both domestic and foreign firms. It appears that for domestic firms, the customer margin seems slightly more important, while the value margin seems relatively less important than for foreign firms. Relating this finding to the last fact about the product life-cycle, I interpret the rapid decay of sales of products introduced by foreign firms as they age as customers being attracted to new products in the initial quarters of the life of a product, and customers being increasingly attracted away from these products and towards newer products from the same firm, or other or newer products from other firms. By contrast, the sustained sales of domestic products even after two years could suggest that even if early customers grow tired of these products and stop buying them, more customers arrive to replace them as the products age.

This fact may vary a lot by industry. Because the regressions control for firm fixed effects, effectively comparing years in which firms grow faster than their own average, I am not excessively worried about industry variation. Nevertheless, I study this by taking the analysis to the firm-product category level. Appendix Figure A.8 shows that across the top 20 categories, the customer margin is the most important one in 19 categories, the only exception being perfume for which it makes sense that the price obtained per purchase would matter a lot. This finding confirms that this analysis is capturing an important dimension of how firms grow in this industry.

This fact is less amenable than others to concerns about firm size, partly because here the point is not that there is not a large difference between domestic and foreign firms. Nevertheless, we show in Appendix Figure A.9 that the result is robust when looking only at firms in the top quarter of the firm size distribution or only at firms in the bottom quarter of the firm size distribution. The results are also robust to firm age heterogeneity (not shown).

5.4 The Intensive Customer per Product Margin is Key to Customer Growth

Because the extensive customer margin appears to be so important according to the previous analysis, I turn to study the question of how to acquire more customers. To double their number of customers, firms may do one of two things, *ceteris paribus*

(i) double the number of markets they operate in - the extensive margin

(ii) double the number of customers they reach in each market - the intensive margin

Markets can be understood in several different ways: in terms of geography, distribution channels, or product. The firm upgrading literature has mostly focused on the extensive margin. The intuition behind is that firms face a barrier in accessing new markets, such as exports (Atkin et al. (2017a)) or large, tender-based markets (Hjort et al. (2020)). A few recent papers have shown that access to better inputs through trade with high-income countries leads the most performing firms to expand their product scope (Bas and Paunov (2021), Kee (2015)). Intuitively, it means that if there are customers who do not like the current products that the firm has in its portfolio, but may like slightly different products, the firm may increase sales by iterating on its current varieties. Product scope expansion might be easier for firms than quality upgrading which often means adopting new technologies or techniques. However, product scope expansion still imposes some fixed cost on the firm.

I study this by expanding on Einav et al. (2021)'s analysis and looking at the following exact decomposition:

$$\text{Customers} \equiv \text{Products} \times \frac{\text{Customers}}{\text{Products}}$$

which yields, taking logs

$$\log(\text{Customers}) = \log(\text{Products}) + \log(\text{Customers per product})$$

I perform two regressions, regressing each element in turn on the log of the number of customers:

$$\log(\text{Products})_{kt} = \alpha + \beta_M \log(\text{Customers})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{6a}$$

$$\log(\text{Customers per product})_{kt} = \alpha + \beta_C \log(\text{Customers})_{kt} + \eta_k + \vartheta_t + \epsilon_{kt} \tag{6b}$$

where η_k are firm fixed effects and ϑ_t are year fixed effects. By construction

$$\beta_M + \beta_C \equiv 1$$

Figure 6 shows the coefficients obtained for domestic and foreign firms. For the latter, the two margins do not appear to be extremely different. By contrast, for domestic firms the intensive margin accounts for over 70% of the variation in the number of customers, while the

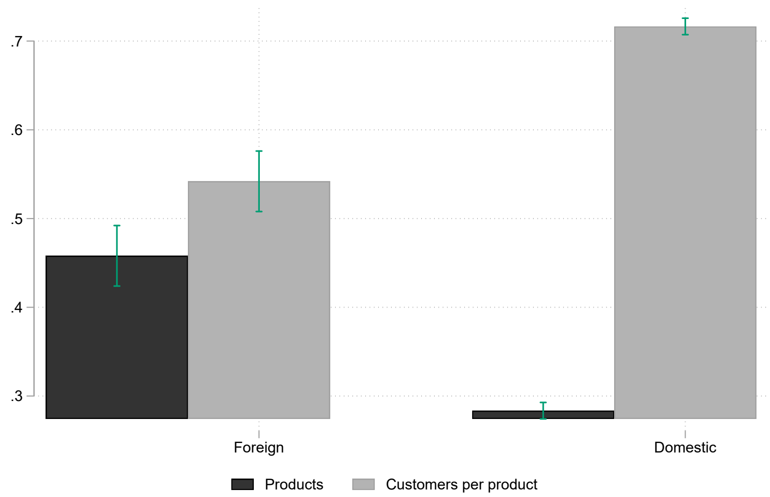


Figure 6: The intensive customer per product margin matters more for domestic firms

Notes: Figure represents the coefficients obtained on log customers from firm-year-level regressions of log products and log customers per product, respectively on log customers, product category and year fixed effects. I run each regression separately for foreign and domestic firms. 95% confidence intervals are represented using the bars. See Appendix Table B.9.

extensive margin accounts for less than 30%. This suggests that it is easier for domestic firms to grow by convincing more customers to buy their existing products, than to grow by adding additional products to their portfolio. This is coherent with the first two facts showing that for domestic firms, the growth of the sales of surviving goods is a strong determinant of firm-level sales growth, and that the products introduced by domestic firms acquire customers over time. Moreover, I showed in Section 3 that when I exclude small domestic firms that only carry one good, domestic firms have a lower rate of introduction of new products than foreign firms on average. This may be a strategic response to the difficulty they face in introducing new products.

A potential concern here is that the difference between domestic firms and foreign firms be entirely driven by size heterogeneity. We cannot control for firm size in the above regressions as we are controlling for firm fixed effects. The concern would then translate into thinking that when a firm is already large and has many customers, faster growth over time cannot be achieved by selling a product to more consumers faster, but by releasing more products faster, while when a firm is small and has few customers (result obtained for domestic firms), it is possible to accelerate customer acquisition. When looking separately at four categories of firms, binned based on their sales, the difference between foreign and domestic firms holds out qualitatively for all bins, alleviating this concern. In Appendix Figure A.10, I show similar figures as above for firms in the bottom 25% and the top 25% of the size distribution. The results are also robust to firm age heterogeneity (not shown).

5.5 The New Customers of Older Domestic Products are Poorer

In this last subsection, I ask who are the customers of aging products. In particular, are the consumers who start buying an older domestic product different from the consumers who start buying an older foreign product? I estimate the following equation:

$$\log y_{i,u,t,g} = \alpha + \sum_{a=1} \beta_a D_a + \lambda_{jt} + \theta_c + \delta_g + u_{i,u,t} \tag{7}$$

where i is a good observed in a certain quarter t in a city g, u is the individual who purchased it, a is a potential age of the good in quarters (between 0 and 14). j is a product category and c is a cohort-quarter. I regress the annual expenditure of new customers u of good i in quarter t on dummies for the age of this product, product category interacted with quarter fixed effects, and cohort fixed effects (corrected following Deaton (1997)'s suggestion to avoid collinearity). I restrict the analysis to products born between 2011Q1 and 2012Q2 for the same reasons as in Subsection 5.2. I perform this regression separately for domestic and foreign products. I control for city fixed effects δ_g based on where customer u lives.

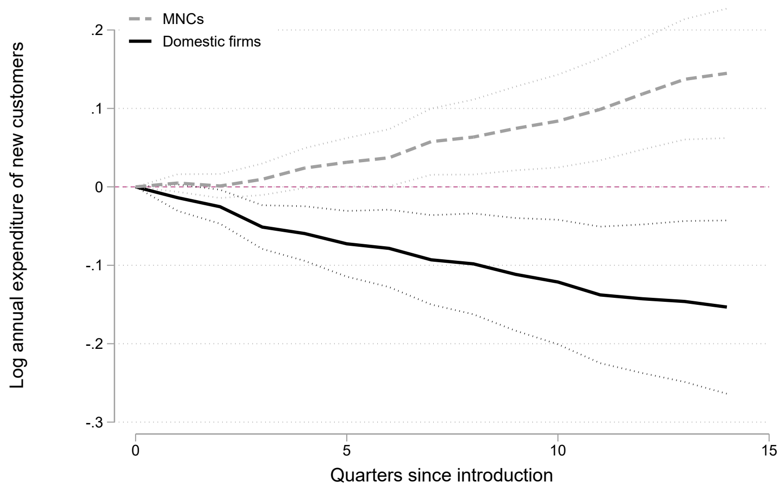


Figure 7: The new customers of older domestic products are poorer

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from an individual-product-quarter level regression described in Equation (7) where I regress the average annual expenditure of the new customers of a product in a given quarter on these dummies, quarter by product category fixed effects, cohort fixed effects and city fixed-effects. I run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines. Table B.10 shows the numerical values.

Figure 7 shows the coefficients obtained on each of the 14 quarter-age dummies. The new customers who start buying a foreign product as it ages are richer than the customers who started buying the foreign product in the very first quarters of its existence. By contrast, the new customers who start buying a domestic product as it ages are significantly poorer than the customers who started consuming the same product in its initial quarters of existence. This fact suggests that the process by which domestic firms acquire demand is very different from the one foreign firms go through. This is confirmed by an alternative description of consumers: in Figure A.11, we show the results of a similar regression where the outcome is the average socio-economic status of new customers, measured on a scale from 1 to 6. We find that customers who buy older foreign products hail from a higher socio-economic status that customers who buy new foreign products, while this is not true for domestic products.

A potential concern is again that the difference between domestic firms and foreign firms be entirely driven by size heterogeneity. The findings are robust to controlling for firm size as measured by firm sales, as shown in Appendix Table B.10, Columns (3) and (4). Binning firms into four different categories based on their sales does not work here. This is because the products released by large firms typically acquire a lot of consumers within the first few months and then do not acquire many new customers, making the regression impossible to estimate just for very large firms. Similarly the products released by very small firms get few customers in total, so the regression is poorly estimated.[21] The results are robust to firm age heterogeneity (not shown).

One may think that there is a supply-side explanation to this pattern. For example, if domestic products are distributed through certain stores and foreign products are distributed through other stores, it may be that there the distribution dynamics specific to each type of firm account for the divergence the new customers profile over time. Specifically, if domestic products are better able to go to more remote, and poorer areas, then the new customers they are getting are not arriving later because of learning but because the products were not available to them before. I am not too concerned about this because the regression includes city fixed effects, which suggests that this finding holds within cities.

I address this concern by asking whether the new cities that are reached by products as they age are different depending on whether the product is sold by a domestic firm (I would expect the cities to be increasingly small and rural) or by a foreign firm (I would expect the products to only diffuse to big cities). To answer this question, I implement a similar estimation equation as shown in Equation (7) but at the barcode-city-quarter level, where I only keep the new cities where the barcode makes sales in each quarter after it is born. The results are shown in Appendix Table B.11. I find that the new cities that products reach when they are older are smaller in population (Columns (1) and (2)) and less dense (Columns (3) and (4)) than the cities reached when the products are new. However, this trend is common to both domestic and foreign goods. I find that the new cities that domestic product reach when they are older are further away from Mexico City, the capital, than the cities by the same products when they were new. This trend does not exist for foreign products: in fact, it appears to be almost the opposite. However, the coefficients for both set of products are very noisy, so it does not appear to be a solid threat to out interpretation of fact # 5.

6 Model of Demand for Foreign Goods over Time

Motivated by these reduced-form results, I estimate a model of demand for consumption goods where consumers choose between foreign and domestic products. I allow consumers' relative valuation for foreign goods to vary over time. The main output of interest from this estimation is to quantify the relative preference for foreign goods and the role of time in disentangling an

[21] The results are robust to estimating the regressions just for firms in the second to last quarter of the firm size distribution, or just for firms in the second quarter of the firm size distribution (not shown).

a priori taste for foreign goods from a real one.

6.1 Model and Identification

I follow Deaton and Muellbauer (1980) and estimate demand for each product within a product category k, relative to a "baseline product" defined as the most popular domestic product available in the market at the beginning of the period. For each product j, the demand specification is:

$$\ln s^k_{jmt} - \ln s^k_{j^0_{mk}mt} = -\alpha(p_{jt} - p_{j^0_{mk}mt}) + \phi_j + \delta_t \tag{8}$$

$$+ \sum_{n=0}^{t-2010} \mathbb{I}(\text{Periods Since Entry}_{jt} = n)\mu^F_n + \sum_{n=0}^{t-2010} \mathbb{I}(\text{Periods Since Entry}_{jt} = n)\mu_n + \epsilon_{jt}$$

$$j = 1, ..., J_{kmt},\ \ m = 1, ..., M,\ \ t = 1, ...T$$

where s^k_{jmt} is the share of product j within the category k in market m at period t, defined as

$$s_{jmt}(y) = \frac{q_{jmt}(y)}{q_{Jmt}(y)}$$

where q_{jmt} is the volume consumed (mostly, mL for liquids and grams for solids) of good j and q_{Jmt} is the volume consumed on the product category as a whole, in the market m at time t. J_{kmt} represents all the possible products one could choose from that product category. I define j^0_{km} the product which has the largest quantity within J_{km0}, the set of domestic products available at the beginning of times in market m. The rest of the analysis will be in relation to this product.[22] In what follows, j is a barcode, t is a quarter (there are 26 distinct quarters in the data) and m is a state (there are 30 states in the data). Alternative choices are discussed below.

Identification in this estimation comes from the variation of the products available across time and across markets and the variation in the price of these products. A concern is that price is set by firms who take demand into account. I therefore propose to use Hausman et al. (1994) instruments. For each product observed in market m at time t, I compute the average price for that product in the rest of the country at time t, and use this $\bar{p}_{jmt} = p_{j,-m,t}$ as an instrument for price. The condition for this instrument to work is that it is relevant: indeed $\bar{p}_{jmt}$ is a strong predictor of the price p_{jmt} and that it is exogenous. The idea here is that the only way by which a change in the price of good j in markets $-m$ can influence the price of j in market m is through a change in supply, such as an input shock rising prices across a region, and not a change in demand, which seems reasonable in this context.

6.2 Model Estimates

The estimates from the model estimated using OLS and IV are presented in Table B.12. I report the estimates of the price coefficient, the age dummies and the age interacted with foreign dummies. The first observation is that the coefficients are fairly consistent across the

[22]Note that when this baseline product is not consumed in a given market m at period t, I lose the observation. This will be fixed in future work.

two specifications. It is perhaps easier to interpret the results of this regression as shown in the two graphs below. First, I show the coefficients obtained on the age dummies. The age dummies exhibit a bell-shaped pattern, which reminds of the data shown in Subsection 5.2: the utility consumers get from a product increases over time, then declines. It is highest relatively close to the introduction of the product. The most important finding, however, is that although foreign goods at first enjoy an even higher effect of time, this premium declines over time and becomes negative after two years. Note that the initial positive effect is small, because the foreign premium is already incorporated in the product fixed effects, shown below. However, these coefficients show that the premium decreases over time.

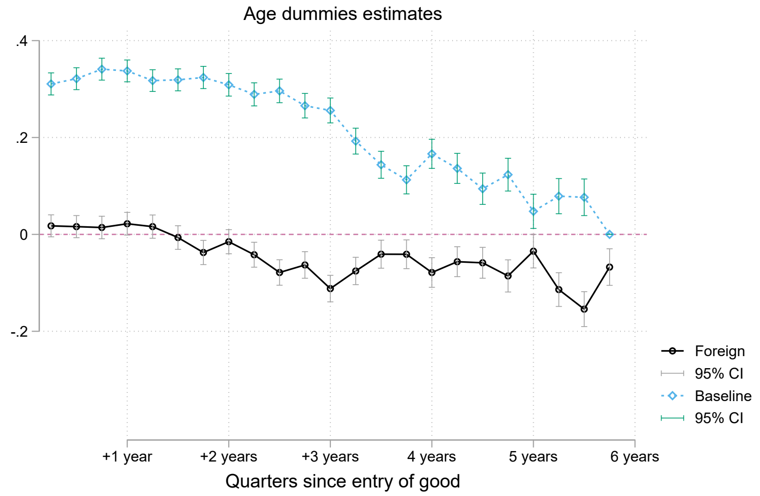


Figure 8: Estimates of the age and age × foreign dummies

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, from the estimation of Equation (8) where I regress log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects and product fixed effects. Appendix Table B.12 shows the numerical values.

In this estimation, I control for product fixed effects. The second result of interest is therefore the shape of the distribution of the product fixed effects. The distribution of foreign products ϕ_j seems to first-order dominate that of domestic products, indicating that consumers, conditional on price, value foreign products more.

A relevant dimension of heterogeneity is income. I divide households into two groups, one belonging to the top socioeconomic status categories used by Kantar (AB, C or C+) and one in the rest (D, D+ or E), and estimation the demand equation 8 again separately for these two groups. The shape of the age dummy parameters is very similar across the two groups are similar, as shown in Appendix Figure A.15. The distributions of product fixed effects are also similar across the two groups, as shown in Figure 10, although there seems to be a more accentuated effect of age on product valuation for households in the higher SES categories (right-hand side panel). Perhaps more strikingly, households in the lower SES categories appear to start off with a higher relative valuation for foreign products, and this valuation does not decrease much over time. The decrease observed in Figure 9 seems to be driven instead by households in the higher socio-economic categories.

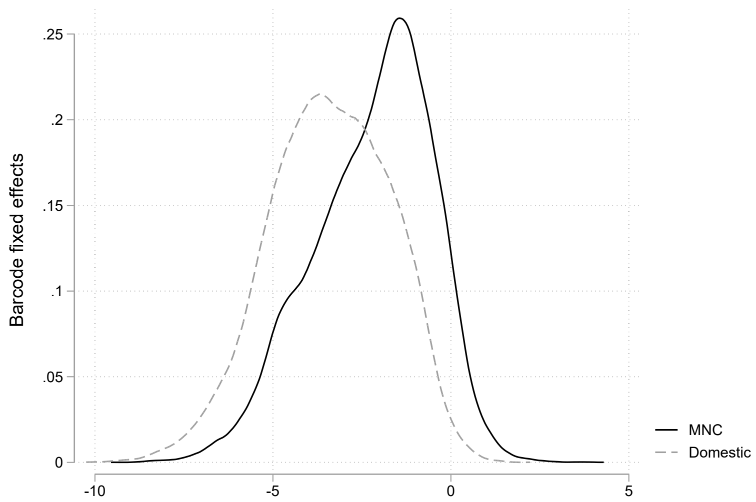
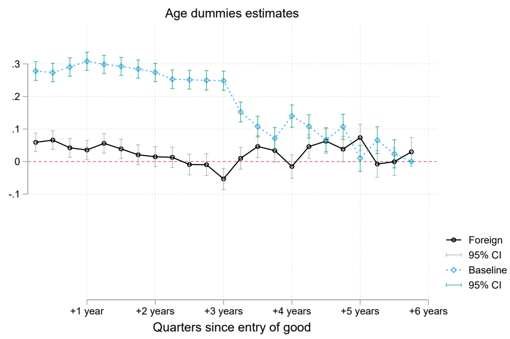
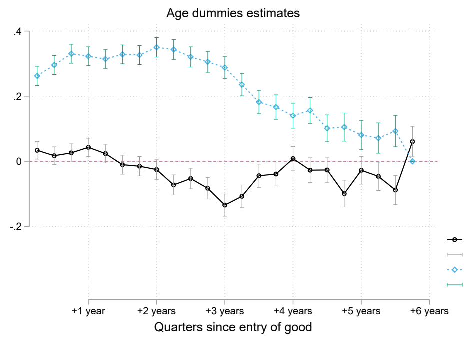


Figure 9: Estimates of the product fixed effects

Notes: Figure represents the coefficients obtained on the product dummies from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, dummies for the age of the product in quarters, and the age interacted with foreign, quarter fixed effects and product fixed effects.



(a) Households in D, D+ or E SES



(b) Households in AB, C or C+ SES

Figure 10: Estimates of the age and age × foreign dummies

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects and product fixed effects; separately for households of low socio-economic status (Panel a) and high socio-economic status (Panel b). Appendix Table B.13 shows the numerical values. Appendix Figure A.15 shows the estimates of the product fixed effects.

This is important because it suggests that these households are changing their opinion faster over time and are therefore driving the increase in demand for domestic products over time. This is consistent with a literature in economics and beyond which shows that the consumption of the rich dictates the preferences of the rest (Colson-Sihra and Bellet (2021)).

6.3 Robustness Checks

There are many other ways to estimate this equation. For example, it would be interesting to estimate how valuation for products map to valuation for specific characteristics. I propose to estimate the following equation:

$$\ln s^k_{jmt} - \ln s^k_{j^0_{mk}mt} = -\alpha(p_{jt} - p_{j^0_{mk}mt}) + \beta(X^k_j - X_{j^0_{mk}m}) + \delta_t \tag{9}$$

$$+ \sum_{n=0}^{t-2010} \mathbb{I}(\text{Periods Since Entry}_{jt} = n)\mu^F_n + \sum_{n=0}^{t-2010} \mathbb{I}(\text{Periods Since Entry}_{jt} = n)\mu_n + \epsilon_{jt}$$

where X_j are the characteristics of product j, and $X_{j^0_{mk}m}$ are the characteristics of the baseline product. For practical reasons, I only control for the difference in volume between the two goods.[23] However, I have shown in Section 4 that it seems to be the characteristic that matters the most within product categories. I show the results of this estimation in Appendix Figure A.14. The age dummies exhibit a very similar pattern. It is now possible to identify the foreign premium, which is worth about 1.25 for new products. This premium decreases over time below 1 on average, confirming the finding using the product fixed effects.

The main results are also robust to estimating the equation at the month-state level and at the quarter-city level (not shown).

6.4 Interpretation

Since the foreign premium declines with consumers' experience of the goods, it indicates that it is not entirely due to an underlying intrinsic quality, which by definition does not change over time. Based on this, and the qualitative evidence about uncertainty presented in the introduction, my main hypothesis is that the decline of the foreign premium is due to uncertainty that resolves over time, as consumers learn about other goods. I formalize this interpretation in Appendix C. While poor consumers are attracted to low prices, they cannot afford to waste money on bad products. Instead, richer consumers are attracted to novelty, they can afford to experiment and are therefore able to discover high-quality domestic products.

An alternative interpretation is that domestic firms improve over time, perhaps introducing new goods that imitate foreign products, and therefore stealing away their business. The last section attempts to provide evidence of uncertainty.

7 Mechanisms

In the last section of the paper, I start by presenting evidence of state dependence, which is stronger for domestic goods than for foreign goods. I then attempt to show that this differential state dependence seems to be more attributable to learning than to loyalty or search (Dubé et al. (2010)). I present evidence that at least part of this state dependence is due to learning,

[23] The other characteristics are strings (for example, diet or not diet, cherry or stawberry, paper or plastic, etc.) so it is difficult to control for the difference between two products. For now, I control for whether each characteristic in the list is different from the characteristic the baseline product has in its own list. In future work, I could use text analysis to measure the difference between the ensemble of characteristics.

by showing that conditional on purchasing a good again, consumers increase the quantity they buy, but this effects fades over time (second result). I show that state dependence is higher in products categories for which quality is more salient (third result), and for which we don't expect imitation to be more frequent. Last, I show that there seems to be less state dependence, and therefore less learning, for lower-income households, as shown in the demand model.

7.1 Evidence of State Dependence in Consumer Choice

I establish state dependence by studying consumers' repeat behavior. I therefore transform the data into a consumer-year-barcode matrix. Conserving only the products that are available to consumers from one year to the next, I look at the probability of buying a good in a given year depending on whether the consumer has purchased the good the year before. I estimate the following equation:

$$y_{ikt} = \alpha + \beta y_{ik,t-1} + \gamma D_k + \delta D_k \times y_{ik,t-1} + \delta_t + \xi_i + \mu_k + \epsilon_{ikt} \tag{10}$$

where i is the consumer, k is the barcode, and t is the year. The parameter of interest is β, which measures the effect of the consumer's past individual experience with the good k at $t-1$, $y_{ik,t-1}$ on current consumption (at t) of the barcode. The second parameter of interest is the interaction coefficient δ, which measures the differential impact of experience with a brand for a domestic brand compared to a foreign brand. I control for year fixed effects. One may be worried about individual selection. If some individuals just have strong tastes for certain products, we might find spurious state dependence. Controlling for individual fixed effects ξ_i takes care of this. One may also be worried about product selection: if some goods are simply a lot better than others, we might also find state dependence. In my preferred specification, I therefore also include product fixed effects.

The results are described in Table 2. In Column (1) I find that previous experience with a barcode is predicted to raise the probability of consumption of that barcode in the following year by 0.42, *ceteris paribus*, a very high effect compared to the control mean of less than 1%. The data therefore exhibit state dependence. Moreover, although a domestic good is less likely to be consumed if the consumer has not tried it before, conditional on being consumed it is more likely to be consumed again, an effect of 0.036, almost 10% of the size effect of the previous consumption for all goods. This is a key result, as it shows that although consumers might be reluctant to buy domestic goods in general, therefore alleviating the product selection concern, once they have tried it they are more likely to buy it again than once they have tried a foreign good.

In Column (2) I include product fixed effects to eliminate production selection concern. I find that the repeat coefficient is lower, suggesting some initial selection, although still very high at 0.33. However the interaction coefficient with the dummy for domestic goods is higher at 0.045, well above 10% of the effect of previous consumption.

A question is therefore what can domestic firms do in order to stimulate the purchase of their new products, which we have shown struggle more to sell than the new products of foreign firms? I propose to look at how the probability of consuming a new barcode is influenced by the

	Current consumption					
	All barcodes		Not previously consumed barcodes			
	(1)	(2)	(3)	(4)	(5)	(6)
Domestic	-0.002	0.000	-0.001	0.000	-0.001	0.000
	(0.000)	(.)	(0.000)	(.)	(0.000)	(.)
Previous consumption of barcode	0.416	0.329				
	(0.000)	(0.000)				
Previous consumption of barcode X Domestic	0.036	0.045				
	(0.000)	(0.000)				
Previous consumption of brand			0.012	0.009		
			(0.000)	(0.000)		
Previous consumption of brand X Domestic			0.005	0.006		
			(0.000)	(0.000)		
Previous consumption of firm					0.007	0.005
					(0.000)	(0.000)
Previous consumption of firm X Domestic					0.006	0.006
					(0.000)	(0.000)
Hhd FEs	Yes	Yes	Yes	Yes	Yes	Yes
Barcode FEs	No	Yes	No	Yes	No	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.008	0.008	0.004	0.004	0.002	0.002
N hhd	6888	6888	6888	6888	6888	6888
N	178991568	178991568	176952267	176952267	176952267	176952267
R2	0.19	0.24	0.01	0.06	0.01	0.06

Table 2: Evidence of state dependence, larger for domestic goods

Notes: Table shows the results of a barcode-household-year level regression of a dummy that turns on if the individual has consumed the barcode in the current year on a dummy for consumption in the previous year, a dummy that turns on if the product if Mexican, and the interaction of these two dummies. In Columns 1 and 2, previous-year consumption is defined as previous consumption of the barcode. In Columns 3 to 6, I only look at current consumption of barcodes that were not consumed by the household in the previous year. Past consumption is defined at the brand level (3 and 4) or firm level (5 and 6), excluding the barcode itself. I always control for household and year fixed effects, while sometimes including barcode fixed effects. Standard errors are reported in parenthesis.

previous consumption of a brand. I therefore estimate Equation 10 using only products that the consumer has not purchased before, and regress the probability of consuming them on a dummy for whether the consumer has consumed any other product from that brand (excluding the barcode of interest) in the previous year. Column (3) shows that having consumed a brand is predicted to raise the probability of consumption of a barcode non previously consumed by 0.01, *ceteris paribus*, a huge effect compared to the baseline probability of 0.004. This suggests that consumers are inferring from their previous exposure to brands that they might like other products from that brand. The first coefficient shows that consumers are less likely to buy a new domestic good than a new foreign good, which is not surprising given the result in Column (1). However, the last coefficient in Column (3) shows that having experience with a domestic brand has a 42% larger effect on the probability to consume a previous not consumed barcode from that brand, than experience with a foreign brand. This effect is confirmed, and larger in proportion, when including product fixed effects in Column (4).

In Columns (5) and (6), I repeat the analysis with experience at the firm level. The effects are smaller but comparable. This makes sense because the name of a firm is less salient to consumers than the name of the brand they purchase goods from, so they learn less from the signal. Similarly, while national reputations matter as underlined by Cagé and Rouzet (2015) and Bai et al. (2019), I expect their signal to be less strong than that of firms or brands[24]. Importantly, the domestic firm experience effect is almost 100% larger than the foreign firm experience effect.

7.2 Evidence of learning

The previous subsection provides evidence of state dependence. This does not necessarily mean that consumers are learning: instead, Dubé et al. (2010) propose loyalty and search as explanations. In order to investigate the learning hypothesis, I exploit data corresponding to consumers' repeat purchases within. The strength of the scanner data, as underlined by Dubois et al. (2022), is that I observe the intensive margin of the purchase of consumers over time. If I see that consumers who buy a good for the nth time buy a larger quantity than those buying a good for the n-1th time, it suggests that consumers are using what they learned from their experience the n-1th time. I estimate the following equation:

$$y_{ikt} = \alpha + \sum_n \beta_n \{N_{ikt} = n\} + \sum_n \delta_n D_k \times \{N_{ikt} = n\} + p_{ikt} + d_k + \delta_t + \xi_i + \epsilon_{ikt} \tag{11}$$

Where again i is the consumer and k the barcode but t is the month. y_{ikt} is the log quantity of the product k purchased by consumer i at time t. I regress this on a series of dummies that represent the order in which the barcode k was purchased by the household. I am interested in the coefficients $\{\beta_n\}_{n=1}^{N}$ which describe the order in which that particular purchase took place, relative to the barcode and the household who purchased it. It captures the learning effect associated with the repeat purchase of the same product. I interact a dummy D_k that turns on if the product is domestic with the dummies for the order of the purchase. Because of the

[24] I do not show the results for countries here as 100% of consumers have tried at least one barcode from Mexico so I cannot estimate the interaction coefficient on country of origin and Domestic. Results are available upon request.

importance of having a precise estimation of the order in which households purchased the goods, I only keep new products in this estimation, where new products are defined as described in Section 3.2. I control for the price paid by consumer i for product k at period t. Last, I include household fixed effects ξ_i, month fixed effects δ_t and barcode fixed effects d_k.

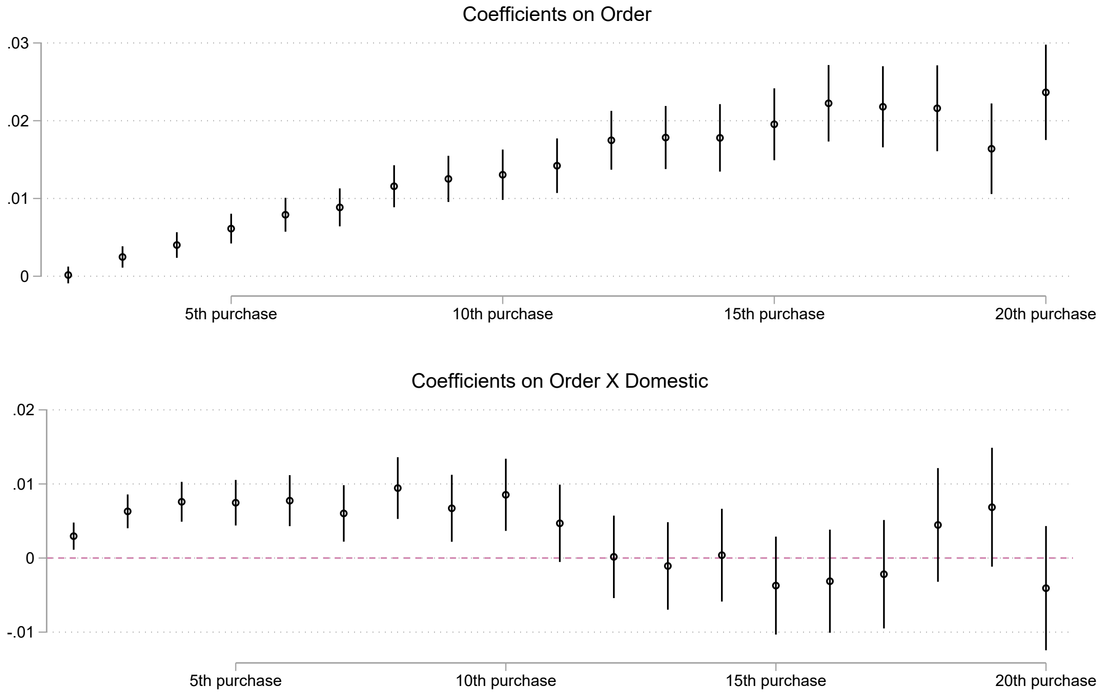


Figure 11: Evidence of learning behavior

Notes: Figures shows the results of a purchase-household level regression, for barcodes purchased up to 20 times, of the quantity of the barcode chosen during the purchase, on dummies describing the order in which the barcode was purchased by the household (coefficients shown in the top panel) and on dummies interacting the order and the fact that the good is produced by a domestic firms (coefficients shown in the bottom panel). I always control for household fixed effects, barcode fixed effects and year fixed effects. Standard errors are reported in parenthesis.

Figure 11 shows the results estimated for up to twenty purchases of the same barcode. The first panel of the figure shows that for all barcodes, being purchased for the 10th time by the same household is predicted to raise the quantity purchased by 1.5 to 2%. This effect is small but statistically significant, and must be assessed against the fact that 75% of purchases have 1 as a quantity. This strongly supports the learning hypothesis. Importantly, the coefficients stagnate starting at the 12th purchase, which also makes sense given that there is a limited amount of information that any household might be able to learn about a particular good. This element of convergence over time is underlined by Dubé et al. (2010) as differentiating learning from other explanations for state-dependence. The bottom panel shows the coefficients obtained on the interaction of the order of the purchase and the dummy representing whether the goods are produced by a domestic firm. The first 11 coefficients are positive, statistically significant and economically important as their size is comparable to the size of the coefficients on the order (up to 1 percentage point additional for the 6th purchase for example). This suggests that there is more learning going on for domestic goods, as consumers change their behavior more as they repeat their purchase of a domestic good than when they repeat the purchase of a foreign good.

One can ask whether these patterns hold for the size of barcode purchases, within products: if consumers are uncertain about the ability of a certain brand to deliver high-quality milk for example, they might start by buying a small bottle of milk, and only if they are satisfied and buy it again will they buy a bottle of a larger quantity. This is indeed the case and it is confirmed in Appendix Figure A.13, where I present this analysis based on two alternative different definitions of goods. It is more difficult to identify new products following the method described in Section 3.2 when using larger product definitions, which may make the analysis less precise, but the results are consistent with the above.

7.3 Importance of quality

The previous subsection suggests that there is some consumer learning. What are consumers learning about? The marketing literature has studied how consumers may learn about the social value of products for example. In this context, this is unlikely because these consumption goods are not prone to social value. Therefore, the hypothesis is that consumers learn about intrinsic product quality. I study this by exploiting heterogeneity across product categories, relying on the fact that quality may matter more for some categories than others.

For example, if food products are of low quality, it may imply that household members get sick. On the contrary, if paper products are of low quality, it has less dramatic implications. Therefore, the model implies that consumers would take more precaution in buying new food products than in buying new paper products. Further, they would learn more from consuming good brands of food products than from consuming good brands of paper products.

I test this by estimating the following equation:

$$y_{u,ik,t} = \alpha + \beta y_{uk,t-1} + \gamma D_i + \delta D_i \times y_{uk,t-1} + \xi_u + \epsilon_{u,i,t} \tag{12}$$

which is similar to Equation 10: $y_{u,ik,t}$ is current consumption of product i (brand k) by household u and time t, $y_{uk,t-1}$ is previous consumption by household u of brand k excluding barcode i, but importantly the dummy D_i represents whether the quality of that product is salient or not.

I start by comparing food to non-food products. Column (1) of Table 3 shows the results. Previous exposure to a brand is predicted to increase the consumption of any product by 0.018 *ceteris paribus*, similar to what I showed in Table 2. This effect increases by 0.005 or almost a third when the product is a food product as opposed to a non-food product, confirming that when quality is more important, there is more state dependence and therefore potentially more learning.

However, food and non-food products are different not only with respect to the importance of quality, but also in terms of the structure of demand, which may influence learning. For example, households typically buy food every week or even several times a week, while they probably buy toilet paper once a month or even less frequently. This would make learning slower for the latter products, regardless of the importance of quality. Moreover, food and non-food

	Current consumption		
	All	Milk	Pads
	(1)	(2)	(3)
Previous consumption	0.018	0.021	0.012
	(0.000)	(0.000)	(0.000)
Salient category	0.002	-0.005	-0.005
	(0.000)	(0.000)	(0.000)
Previous consumption X Salient	0.005	0.010	0.005
	(0.000)	(0.000)	(0.000)
Hhd FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Control mean	0.009	0.009	0.009
N hhd	11966	11966	11966
N	448683433	15059005	15365529
R2	0.01	0.01	0.01

Table 3: Learning by quality salience

Notes: Table shows the household-barcode-year-level regression of a dummy indicating whether the household has consumed the barcode in the observation year on a dummy indicating whether the household has consumed products for a similar brand in the previous year, whether the product belongs to a category for which quality is salient, and an interaction of these two dummies. In the first column, I look at all products and the salient category is food. In the second column, I keep only milk-products and the salient category is infant formula. In the third category, I keep baby diapers and sanitary pads and the salient category is baby diapers. I always control for household and year fixed effects. Standard errors are reported in parenthesis.

products are different in terms of supply chains and marketing strategies, which may influence learning as well. So finding different learning speeds between food and non-food products might be attributable to these factors and not to quality salience.

Therefore, I exploit the heterogeneity of quality salience among more narrow product categories. Most people would agree that the safety of food being fed to babies and young children is relatively more important than the safety of other foods. In Column (2), I therefore compare the effect of brand exposure on the probability of buying an infant formula products to the effect of brand exposure on the probability of buying another milk product, a comparable product in terms of both supply and frequency of purchase. I find that the effect of exposure to the brand is much larger for infant formula products: it increases the probability of buying a product from the same brand by 0.01, an effect that is about 50% that of the baseline previous state dependence effect in that category. Similarly, in Column (3) I compare the effect of brand exposure on the probability of buying baby diapers to the effect of brand exposure on the probability of buying adult sanitary pads. I again find that the brand effect is larger for baby products. The fact that products for which intrinsic quality is more salient exhibit more state dependence confirms the hypothesis that the state dependence we see is about learning, and learning is about intrinsic quality.

7.4 Price sensitivity

I now turn to the last element of the learning hypothesis: that income is a binding constraint in choosing to experiment with unknown goods. I propose to estimate the following equation:

$$y_{u,ik,t} = \alpha + \beta y_{uk,t-1} + \gamma D_u + \delta D_u \times y_{uk,t-1} + \mu_i + \epsilon_{u,i,t} \tag{13}$$

Which again estimates the impact of $y_{uk,t-1}$, previous consumption of brand k, on $y_{u,ik,t}$, the current consumption of product i (brand k) by household i and time t, but importantly the dummy D_u represents whether the households' expenditure in year $t-1$ puts them in the bottom quarter of the expenditure distribution in the sample in the year t, which I will henceforth call the "low-expenditure" group. I cannot control for household fixed effects in this context, and therefore control for barcode fixed effects μ_i instead.

	Current consumption			
	Brand	Barcode	Firm	Country
	(1)	(2)	(3)	(4)
Previous consumption	0.024	0.344	0.020	0.014
	(0.000)	(0.000)	(0.000)	(0.000)
Low expenditure	-0.003	-0.007	-0.002	-0.003
	(0.000)	(0.000)	(0.000)	(0.000)
Previous consumption X Low expenditure	-0.008	-0.066	-0.007	-0.003
	(0.000)	(0.000)	(0.000)	(0.000)
Barcode FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Control mean	0.009	0.016	0.006	0.009
N hhd	11966	11966	11966	11966
N	448683433	459375213	448683433	448683433
R2	0.05	0.18	0.05	0.05

Table 4: There is relatively less learning among the poor

Notes: Table shows regression of a dummy that turns on if the individual has consumed the barcode in the current year on a dummy for consumption in the previous year, a dummy that turns on if the household is in the bottom quarter of the expenditure distribution, and the interaction of these two dummies. In Column 1, previous-year consumption is defined as previous consumption of the brand that the product belongs to, while I only look at current consumption of products that were not consumed before. In Columns 3 and 4 I do the same but for the firm and country that the product belongs to, respectively. Column 2 looks at previous consumption of the product itself and therefore includes all products, whether consumed or not in the current year. I always control for barcode and year fixed effects. Standard errors are reported in parenthesis.

Table 4 shows the results of the estimation in the same order as in Table 2. I find that having consumed a brand is predicted to raise the probability to purchase a previously not-purchased product by a large, positive and significant amount, *ceteris paribus*. This effect is considerably smaller for low-expenditure households. Focusing on Column (1), I find that the effect is a third smaller for low-expenditure households. As a benchmark, I show in Column (2) the effect of having consumed the exact same good, which is again about 15 times larger than the brand effect, is much smaller for low-expenditure households. Last, Columns (3) and (4) show robustness checks looking at larger groups: firm and country, and the results are similar.

8 Conclusion

The presence of foreign firms in developing countries generates welfare gains in many ways, including because they provide consumers with goods that they like. However, if these foreign firms get so large that domestic firms cannot get started, it is detrimental to many governments' industrial policy objectives and ultimately to product variety.

In this context: the consumer packaged goods sector in Mexico, I show that foreign goods command a 16 to 20% price premium, controlling for observables. This premium speaks to the high-quality reputation enjoyed by foreign firms, in a context of quality uncertainty. In parallel, I show that domestic firms have a harder time introducing new goods in the market than foreign firms: for the former, growth mostly comes from older goods, and new goods take years until they reach the sales enjoyed by new foreign goods from the start. I show that in order to grow their customer base, it seems easier for domestic firms to try and attract customers to their existing products than to try to expand their product scope.

I then estimate a demand model and recover the valuation for foreign goods, which declines over time. I provide evidence that this decline can be attributed to consumers learning about the quality of domestic substitutes over time. This means that there are domestic products with a high enough quality, which could do as well as foreign products. There is potentially room for policy intervention, in the form of a subsidy for marketing efforts for example. More generally, demand-side interventions can be important complements to supply-side interventions aiming to help firms produce higher-quality goods.

References

Ackerberg, Daniel A, "Advertising, learning, and consumer choice in experience good markets: an empirical examination," *International Economic Review*, 2003, *44* (3), 1007–1040.

Afrouzi, Hassan, Andres Drenik, and Ryan Kim, "Growing by the Masses: Revisiting the Link between Firm Size and Market Power," *Available at SSRN 3703244*, 2020.

Aguilar, Arturo, Emilio Gutierrez, and Enrique Seira, "The effectiveness of sin food taxes: Evidence from Mexico," *Journal of Health Economics*, 2021, *77*, 102455.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *Available at SSRN 3376129*, 2019.

Argente, David, Munseob Lee, and Sara Moreira, "The Life Cycle of Products: Evidence and Implications," *Available at SSRN 3163195*, 2019.

Arkolakis, Costas, "Market penetration costs and the new consumers margin in international trade," *Journal of political economy*, 2010, *118* (6), 1151–1199.

Armington, Paul S, "A theory of demand for products distinguished by place of production," *Staff Papers*, 1969, *16* (1), 159–178.

Atkin, David, "Trade, tastes, and nutrition in India," *American economic review*, 2013, *103* (5), 1629–63.

_ , Amit K Khandelwal, and Adam Osman, "Exporting and firm performance: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 2017, *132* (2), 551–615.

_ and Dave Donaldson, "Who's getting globalized? The size and implications of intra-national trade costs," Technical Report, National Bureau of Economic Research 2015.

— , **Azam Chaudhry, Shamyla Chaudry, Amit K Khandelwal, and Eric Verhoogen**, "Organizational barriers to technology adoption: Evidence from soccer-ball producers in Pakistan," *The Quarterly Journal of Economics*, 2017, *132* (3), 1101–1164.

— , **Benjamin Faber, and Marco Gonzalez-Navarro**, "Retail globalization and household welfare: Evidence from mexico," *Journal of Political Economy*, 2018, *126* (1), 1–73.

Bagwell, Kyle, "The economic analysis of advertising," *Handbook of industrial organization*, 2007, *3*, 1701–1844.

Bai, Jie, "Melons as lemons: Asymmetric information, consumer learning and quality provision," Technical Report, Working paper 2018.

— , **Ludovica Gazze, and Yukun Wang**, "Collective reputation in trade: Evidence from the Chinese dairy industry," Technical Report, National Bureau of Economic Research 2019.

Bas, Maria and Caroline Paunov, "Disentangling trade reform impacts on firm market and production decisions," *European Economic Review*, 2021, *135*, 103726.

Batra, Rajeev, Venkatram Ramaswamy, Dana L Alden, Jan-Benedict EM Steenkamp, and Shivaramaiah Ramachander, "Effects of brand local and nonlocal origin on consumer attitudes in developing countries," *Journal of consumer psychology*, 2000, *9* (2), 83–95.

Bems, Rudolfs and Julian Di Giovanni, "Income-induced expenditure switching," *American Economic Review*, 2016, *106* (12), 3898–3931.

Bennett, Daniel and Wesley Yin, "The market for high-quality medicine: Retail chain entry and drug quality in India," *Review of Economics and Statistics*, 2019, *101* (1), 76–90.

Bronnenberg, Bart J and Jean-Pierre Dubé, "The formation of consumer brand preferences," *Annual Review of Economics*, 2017, *9*, 353–382.

— , — , **Matthew Gentzkow, and Jesse M Shapiro**, "Do pharmacists buy Bayer? Informed shoppers and the brand premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Cagé, Julia and Dorothée Rouzet, "Improving "national brands": Reputation for quality and export promotion strategies," *Journal of International Economics*, 2015, *95* (2), 274–290.

Chen, Maggie X and Min Wu, "The value of reputation in trade: Evidence from alibaba," *Review of Economics and Statistics*, 2020, pp. 1–45.

Colson-Sihra, Eve and Clement Bellet, "Does Inequality Affect the Perception of Needs?," *Available at SSRN 3270814*, 2021.

Crawford, Gregory S and Matthew Shum, "Uncertainty and learning in pharmaceutical demand," *econometrica*, 2005, *73* (4), 1137–1173.

Deaton, Angus, *The analysis of household surveys: a microeconometric approach to development policy*, World Bank Publications, 1997.

— **and John Muellbauer**, "An almost ideal demand system," *The American economic review*, 1980, *70* (3), 312–326.

Dranove, David and Ginger Zhe Jin, "Quality disclosure and certification: Theory and practice," *Journal of Economic Literature*, 2010, *48* (4), 935–63.

Dubé, Jean-Pierre, Günter J Hitsch, and Peter E Rossi, "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 2010, *41* (3), 417–445.

Dubois, Pierre, Rachel Griffith, Martin O'Connell et al., "The Use of Scanner Data for Economics Research," Technical Report, Toulouse School of Economics (TSE) 2022.

Duflo, Esther and Emmanuel Saez, "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly journal of economics*, 2003, *118* (3), 815–842.

Einav, Liran, Peter Klenow, Jonathan D Levin, and Raviv Murciano-Goroff, "Customers and retail growth," Technical Report, Working Paper 2021.

Faber, Benjamin, "Trade liberalization, the price of quality, and inequality: Evidence from mexican store prices," 2014.

Fajgelbaum, Pablo D and Amit K Khandelwal, "Measuring the unequal gains from trade," *The Quarterly Journal of Economics*, 2016, *131* (3), 1113–1180.

Foster, Andrew D and Mark R Rosenzweig, "Learning by doing and learning from others: Human capital and technical change in agriculture," *Journal of political Economy*, 1995, *103* (6), 1176–1209.

Ge, Ying, Huiwen Lai, and Susan Chun Zhu, "Multinational price premium," *Journal of Development Economics*, 2015, *115*, 181–199.

Goldberg, Pinelopi K and Tristan Reed, "Income Distribution, International Integration, and Sustained Poverty Reduction," Technical Report, National Bureau of Economic Research 2020.

Hausman, Jerry, Gregory Leonard, and J Douglas Zona, "Competitive analysis with differenciated products," *Annales d'Economie et de Statistique*, 1994, pp. 159–180.

Hjort, Jonas, Vinayak Iyer, and Golvine De Rochambeau, "Informational Barriers to Market Access: Experimental Evidence from Liberian Firms," Technical Report, National Bureau of Economic Research 2020.

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.

Hulland, John, Honorio S Todiño Jr, and Donald J Lecraw, "Country-of-origin effects on sellers' price premiums in competitive Philippine markets," *Journal of International Marketing*, 1996, *4* (1), 57–79.

Hummels, David and Peter J Klenow, "The variety and quality of a nation's exports," *American Economic Review*, 2005, *95* (3), 704–723.

Jiménez-Hernández, Diego and Enrique Seira, "Should the government sell you goods? Evidence from the milk market in Mexico," Technical Report, Stanford University Working Paper 2021.

Kee, Hiau Looi, "Local intermediate inputs and the shared supplier spillovers of foreign direct investment," *Journal of Development Economics*, 2015, *112*, 56–71.

Khandelwal, Amit, "The long and short (of) quality ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450–1476.

Macchiavello, Rocco, Atonu Rabbani, and Christopher Woodruff, "The market for training services: a demand experiment with Bangladeshi garment factories," *American Economic Review*, 2015, *105* (5), 300–304.

Mobius, Markus and Tanya Rosenblat, "Social learning in economics," *Annu. Rev. Econ.*, 2014, *6* (1), 827–847.

Orlove, Benjamin S, *The allure of the foreign: Imported goods in postcolonial Latin America*, Ann Arbor: University of Michigan Press, 1997.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Papadopoulos, Nicolas and Louise A Heslop, *Product-country images: Impact and role in international marketing*, Routledge, 1993.

Perla, Jesse, "A model of product awareness and industry life cycles," *Unpublished Paper, University of British Columbia*, 2019.

Ragasa, Catherine, Hiroyuki Takeshima, Antony Chapoto, and Shashi Kolavalli, *Substituting for rice imports in Ghana*, Vol. 6, Intl Food Policy Res Inst, 2014.

Schmalensee, Richard, "Product differentiation advantages of pioneering brands," *The American Economic Review*, 1982, *72* (3), 349–365.

Schott, Peter K, "Across-product versus within-product specialization in international trade," *The Quarterly Journal of Economics*, 2004, *119* (2), 647–678.

Shapiro, Carl, "Premiums for high quality products as returns to reputations," *The quarterly journal of economics*, 1983, *98* (4), 659–679.

Shimp, Terence A and William O Bearden, "Warranty and other extrinsic cue effects on consumers' risk perceptions," *Journal of Consumer research*, 1982, *9* (1), 38–46.

Shin, Sangwoo, Sanjog Misra, and Dan Horsky, "Disentangling preferences and learning in brand choice models," *Marketing Science*, 2012, *31* (1), 115–137.

Tirole, Jean, "A theory of collective reputations (with applications to the persistence of corruption and to firm quality)," *The Review of Economic Studies*, 1996, *63* (1), 1–22.

Verhoogen, Eric, "Firm-Level Upgrading in Developing Countries," 2021.

A Appendix Figures

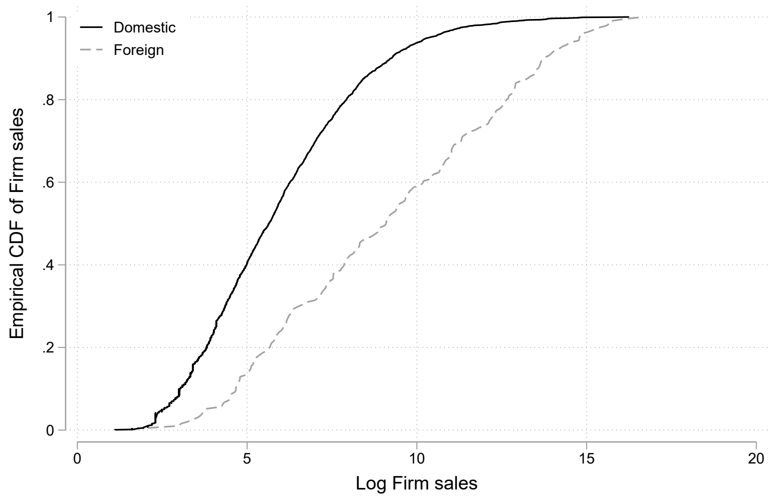


Figure A.1: Firm size distribution

Notes: Figure plots the distribution of the log sales of the firms observed in the dataset, where sales are defined as the total yearly expenditure of the households in that dataset on that firm. I separate out the distribution between foreign and domestic firms.

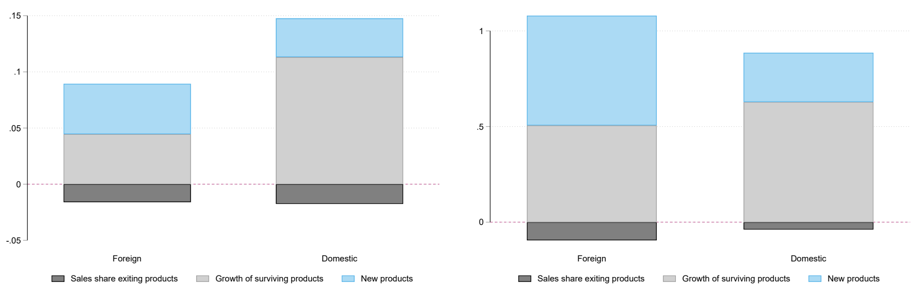


Figure A.2: Growth decomposition, by firm size

Notes: Figure represents the average firm-level year-to-year growth components, separated out by the origin of the firms. The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products. The left panel shows the decomposition for firms belonging to the top quarter of the firm sales distribution of each year. The right panel shows the decomposition for firms belonging to the bottom quarter of the firm sales distribution of each year.

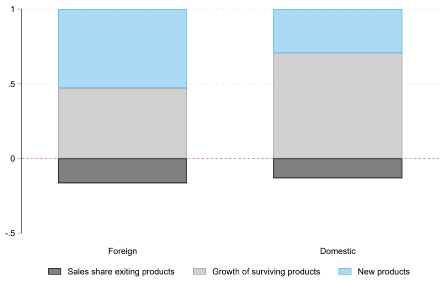


Figure A.3: Growth decomposition, relative terms

Notes: Figure represents the relative firm-level year-to-year growth components, separated out by the origin of the firms. The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products. Each component is then measured relative to the average growth rate of firm sales. The left panel shows the decomposition for firms belonging to the top quarter of the firm sales distribution of each year. The right panel shows the decomposition for firms belonging to the bottom quarter of the firm sales distribution of each year.

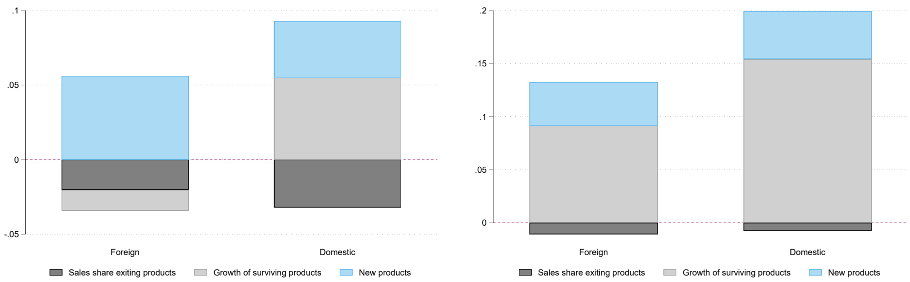


Figure A.4: Growth decomposition, by firm age

Notes: Figure represents the average firm-level year-to-year growth components, separated out by the origin of the firms. The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products. The left panel shows the decomposition for firms that are four years or older. The right panel shows the decomposition for firms that are less than four years old.

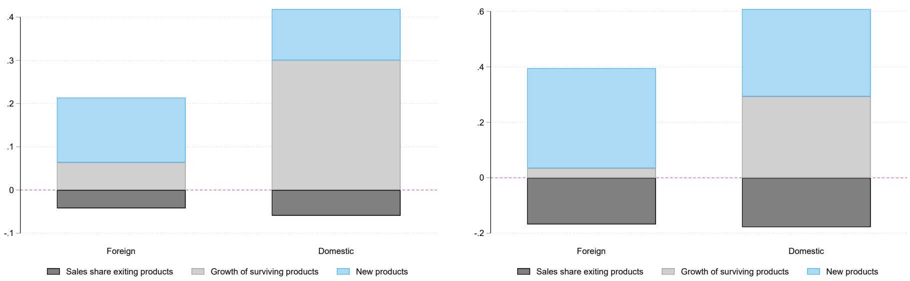


Figure A.5: Growth decomposition, by time horizon

Notes: Figure represents the average firm-level growth components, separated out by the origin of the firms, studying growth over two years (top panel) and over five years (bottom panel). The sales of exiting products is the average share of sales that the product that exited between the previous year and the current year represented in the previous year. The growth of surviving products is the average growth rate of the total sales of products that survived between the previous year and the current year. The new products components is the product of the average rate of introduction of new products and the average sales of new products compared to older products. The left panel shows the decomposition for firms belonging to the top quarter of the firm sales distribution of each year. The right panel shows the decomposition for firms belonging to the bottom quarter of the firm sales distribution of each year.

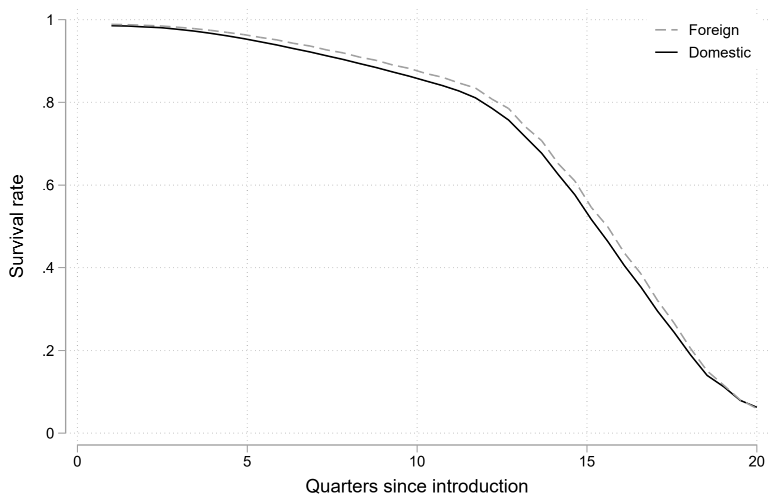


Figure A.6: Survival rate of foreign and domestic products, by quarter

Notes: Figure plots the survival rate of new products in the dataset over time, measured in quarters since the product was introduced. The death of a product is defined at the last quarter it appears in the dataset. I separate the survival rate by whether the product belongs to a foreign or a domestic firms.

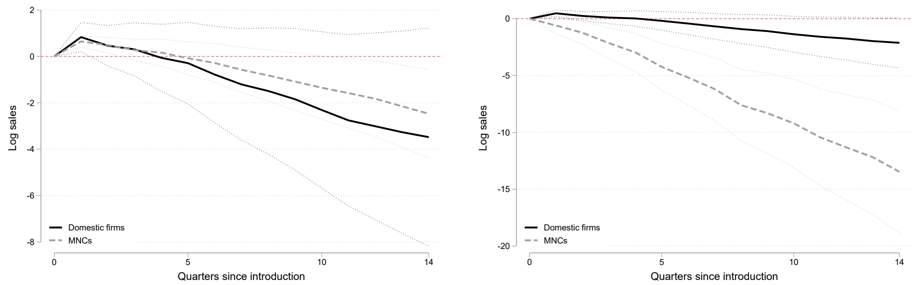


Figure A.7: Product life-cycle, by firm size

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression where I regress the sales of a product in a given quarter on these dummies, quarter by product category fixed effects, and cohort fixed effects. I run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines. The left panel shows the results for products released by firms in the top quarter of the size distribution. The right panel shows the results for products released by firms in the bottom quarter of the size distribution.

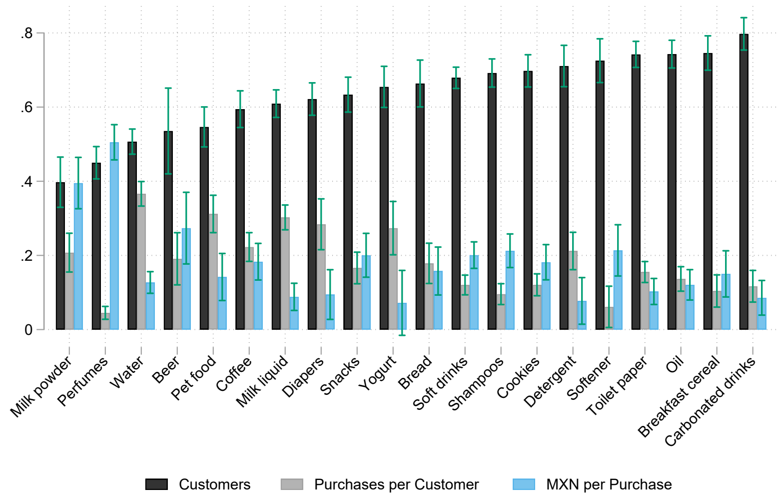


Figure A.8: Customer/Quantity/Value decomposition, by product category

Notes: Figure represents the coefficients obtained on log sales from firm-year-level regressions of log customers, log purchases per customer and log monetary value per purchase, respectively on log sales and year fixed effects. I run each of the three regressions separately for each of the top 20 product categories (one observation is the sales of a firm in a given a year in the category). 95% confidence intervals are represented using the bars.

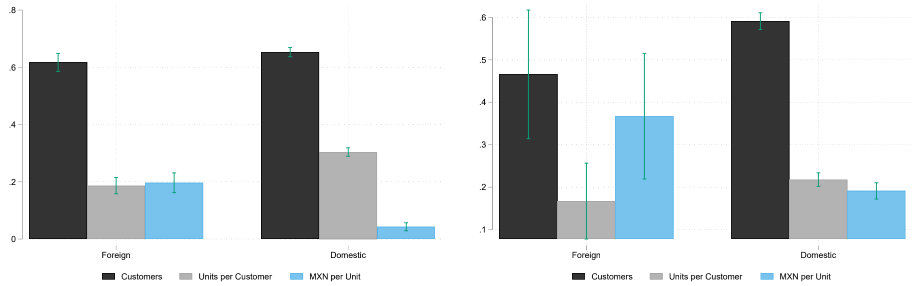


Figure A.9: Customer/Quantity/Value decomposition, by firm size

Notes: Figure represents the coefficients obtained on log sales from firm-year-level regressions of log customers, log units per customer and log monetary value per unit, respectively on log sales, product category and year fixed effects. I run each of the three regressions separately for foreign and domestic firms. 95% confidence intervals are represented using the bars. The left panel of this figure shows the results of the decomposition for firms in the top quarter of the firm size distribution while the right panel shows the results for firms in the bottom quarter of the firm size distribution.

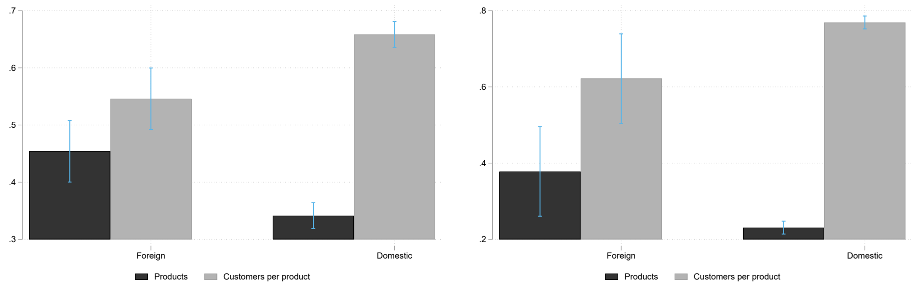


Figure A.10: Customer growth decomposition, by firm size

Notes: Figure represents the coefficients obtained on log customers from firm-year-level regressions of log products and log customers per product, respectively on log customers, product category and year fixed effects. I run each regression separately for foreign and domestic firms. 95% confidence intervals are represented using the bars. The left panel shows the decomposition for firms in the top quarter of the firm size distribution, and the right panel shows the decomposition for firms in the bottom quarter of the firm size distribution.

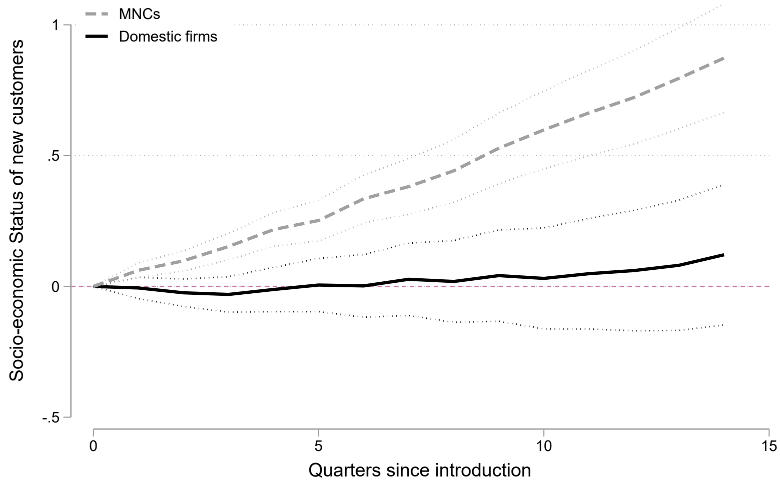


Figure A.11: Socio-economic status of new customers of older products (1-6)

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from an individual-product-quarter level regression described in Equation (7) where I regress the average socio-economic status, measured on a scale from 1 to 6, of the new customers of a product in a given quarter on these dummies, quarter by product category fixed effects, cohort fixed effects and city fixed-effects. I run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines.

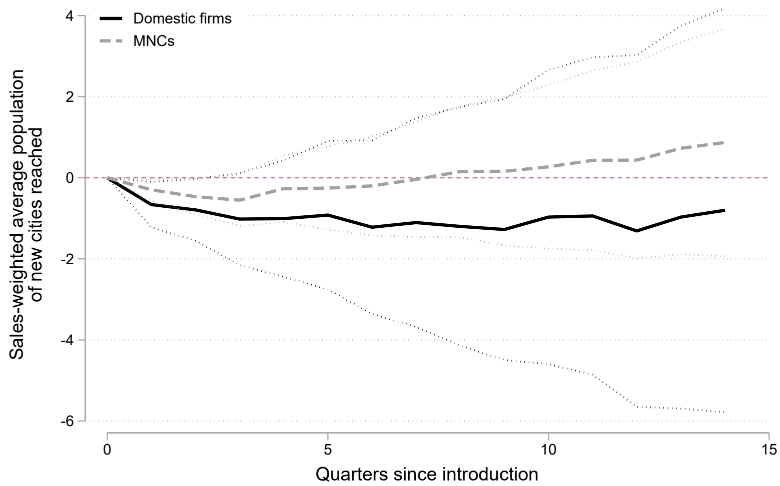
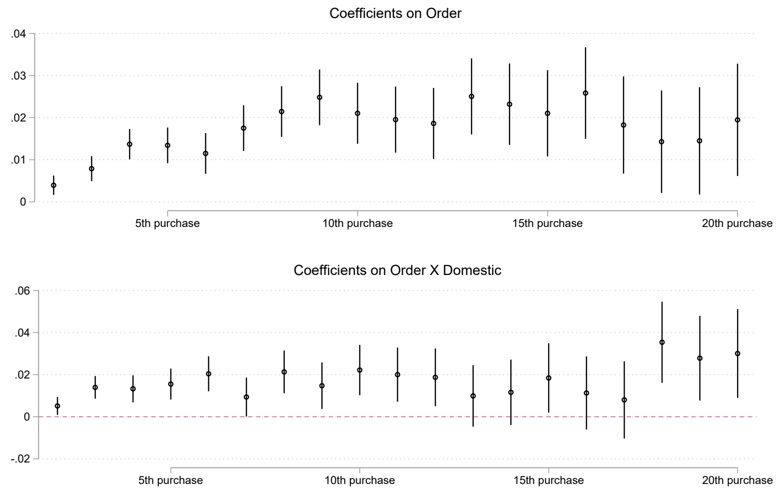
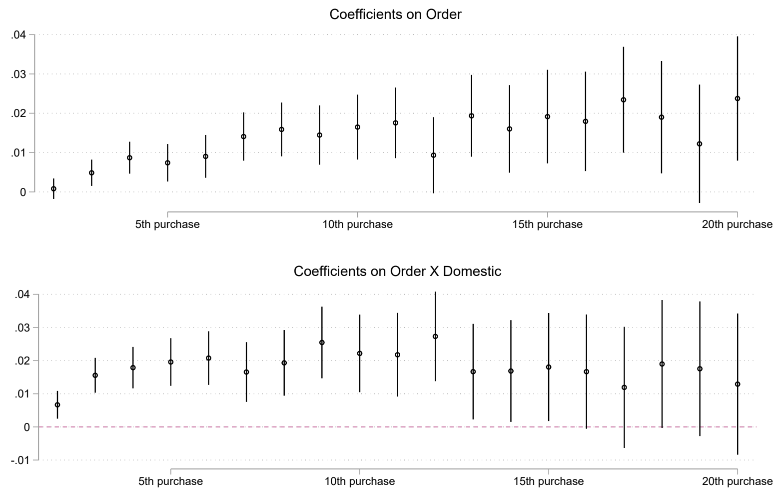


Figure A.12: New cities characteristics, by quarter and product type

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression where I regress the sales-weighted average of the population of the new cities in which a product sells in a given quarter on these dummies, quarter by product category fixed effects, and cohort fixed effects. I run two separate regressions for domestic and foreign products, respectively. 95% confidence intervals are represented by the lighter lines.



(a) Products defined as firm-brand-category-4 characteristics



(b) Products defined as firm-brand-category-2 characteristics

Figure A.13: Evidence of learning behavior, robustness by size

Notes: Figures show the results of a purchase-household level regression, for products purchased up to 20 times, of the size of the product (defined as firm-brand- category-4 characteristics in Panel (a) and 2 characteristics in Panel (b)) chosen during the purchase, on dummies describing the order in which the product was purchased by the household (coefficients shown in the top panel) and on dummies interacting the order and the fact that the good is produced by a domestic firms (coefficients shown in the bottom panel). I only keep products that are new relative to the dataset. I always control for household fixed effects, product fixed effects and year fixed effects. Standard errors are reported in parenthesis.

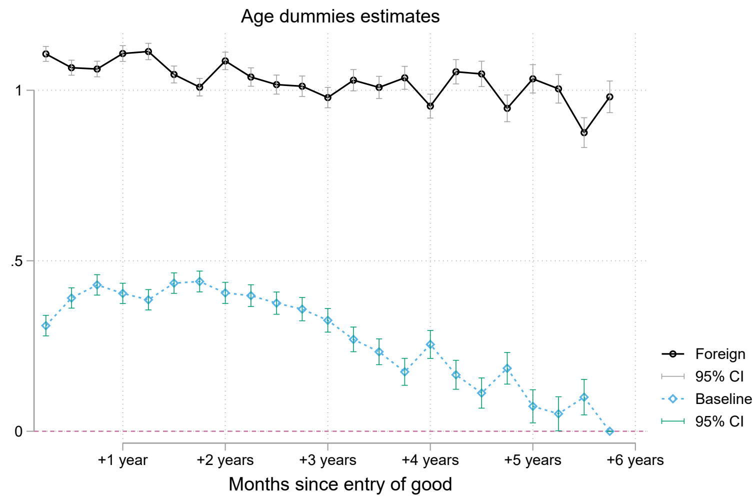
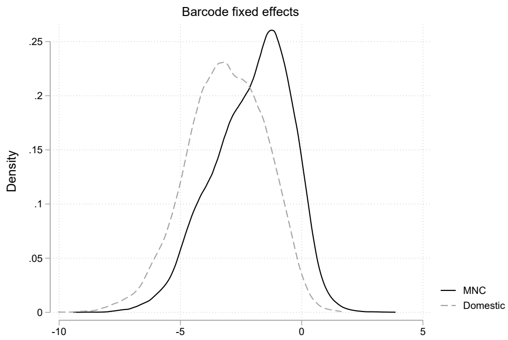
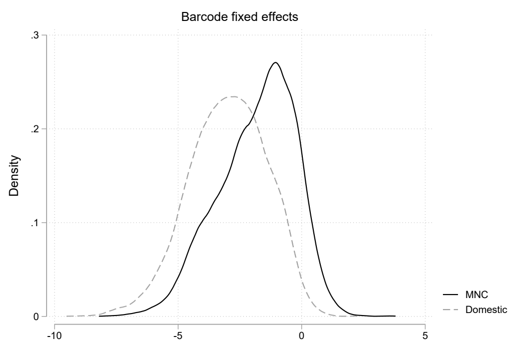


Figure A.14: Structural estimation without product FEs

Notes: Figure represents the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects, cohort fixed effects and product characteristics. Appendix Table B.14 shows the numerical values.



(a) Households in D, D+ or E SES



(b) Households in AB, C or C+ SES

Figure A.15: Estimates of the age and age × foreign dummies

Notes: Figure represents the coefficients obtained on the dummies for product, from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, dummies the age of the product in quarters, and the age interacted with foreign quarter fixed effects; separately for households of low socio-economic status (Panel a) and high socio-economic status (Panel b). Appendix Table B.13 shows the numerical values.

B Appendix Tables

	ENIGH			KWP			Difference	
	mean	sd	N	mean	sd	N	diff	p
Number of household members	3.94	1.99	26238	4.37	1.83	8414	0.425	0.00
Number of women in household	2.03	1.27	26238	2.29	1.22	8414	0.263	0.00
Age head of household	48.32	15.59	26238	45.61	14.02	8412	-2.707	0.00
Finished primary	0.84	0.37	26238	0.96	0.20	8414	0.120	0.00
Finished secondary	0.71	0.45	26238	0.86	0.34	8414	0.148	0.00
Finished Post-secondary	0.26	0.44	26238	0.26	0.44	8414	-0.006	0.25
Works full time	0.75	0.43	26238	0.75	0.43	8414	0.006	0.30
Number of cars	0.52	0.79	26238	0.56	0.66	8414	0.040	0.00
Number of PCs	0.31	0.61	26238	0.33	0.47	8414	0.020	0.01
Access to Internet (0/1)	0.19	0.39	26238	0.24	0.42	8414	0.043	0.00
Number of color TVs	1.44	0.92	26238	1.87	0.98	8413	0.432	0.00
Number of fridges	0.83	0.43	26238	0.96	0.19	8412	0.137	0.00
Number of microwaves	0.41	0.51	26238	0.70	0.46	8414	0.289	0.00
Number of bedrooms	2.01	0.97	25696	2.20	0.97	8412	0.191	0.00
Debit or credit card (0/1)	0.21	0.41	26238	0.27	0.44	8414	0.057	0.00
Monthly expenditure (MXN)	2391	1562	26238	1095	647	8414	-1,296	0.00

Table B.1: Household-level summary statistics, KWP vs ENIGH in 2010

Notes: Table compares summary statistics of the main dataset used in the analysis (Kantar World Panel or KWP) in 2010 in Columns (4)-(6) against the official expenditure survey (ENIGH 2010) in Columns (1)-(3). ENIGH provides the national reference values for household characteristics, income and expenditures. When relevant, the variable described is measured for the head of household (adult man if two working-age adults are present). Ownership of a debit or credit card is a variable in the ENIGH survey and it is coded in the KWP to 1 if the household is ever reported to use a card as a mode of payment. Expenditure in the ENIGH survey cannot be compared exactly to expenditure in the KWP but is constructed based on similar categories (spending on personal care, household care, and food for at-home consumption). On average KWP goods represent 46% of expenditure on at-home consumption measured in ENIGH.

	Log unit price of rice in				
	Ghana	Nigeria	Niger	Timor Leste	Yemen
	(1)	(2)	(3)	(4)	(5)
Imported	0.356	0.280	0.367	0.037	0.199
	(0.010)	(0.016)	(0.102)	(0.011)	(0.012)
Geographical FEs	Yes	Yes	No	Yes	No
Indiv FEs	No	No	Yes	No	Yes
Years covered	2005-2017	2015-2018	2011, 2014	2001, 2008	2005
Mean of Dep. Var.	3.25	5.35	4.87	-0.15	5.37
R2	0.81	0.26	0.23	0.09	0.53
N	7983	25760	2002	4489	3044

Table B.2: Imported goods cost a lot more in developing countries

Note: data comes from LSMS surveys financed by the World Bank and subsequent rounds of surveys run by the national statistical agencies of each country. I kept the items which were described by their country of origin and for which both a "foreign" item and a "domestic" item was available according to the survey. I then computed the unit price by dividing the total price reported by the total quantity reported. Then for each country and each item I regressed the log unit price on a dummy for whether the good was imported, a time dummy and regional dummies.

	Log price					
	(1)	(2)	(3)	(4)	(5)	(6)
MNC HQ low-income	0.237	0.096	0.082	0.092	0.105	0.125
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MNC HQ high-income	0.162	0.167	0.163	0.175	0.198	0.216
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
City by month FEs	6011	6011	6011	6011	6011	6011
Product FEs	180	180	10912	16647	23635	26397
Volume control	No	Yes	Yes	Yes	Yes	Yes
R2	0.53	0.87	0.91	0.92	0.93	0.94
N	43759653	43759369	43758021	43697214	42445758	37548230

Table B.3: MNC price premium by origin of FDI

Notes: Table reports regressions of the log price on two dummies for whether the firm is foreign and its headquarter is in a high-income country, whether the firm is foreign and its headquarter is in a low-income country as defined by the World Bank[25]. The default category is a domestic firm. The regressions include city by month fixed effects, and various level of characteristics. Column (1) only controls for the product category by subcategory fixed effects. Columns (2) through (6) also control for the volume of the barcode. Columns (3) through (6) include narrower fixed effects defined by product characteristics. The row "product FEs" describes how many fixed effects are absorbed in the regression.

	Share new products					
	(1)	(2)	(3)	(4)	(5)	(6)
Foreign	-0.046	-0.035	-0.041	0.046	0.047	0.037
	(0.012)	(0.013)	(0.013)	(0.007)	(0.007)	(0.007)
Firm sales, mMXN		-0.025	0.017		-0.004	-0.003
		(0.006)	(0.009)		(0.003)	(0.005)
Number of old varieties			-0.000			0.000
			(0.000)			(0.000)
Firm leader in category			-0.062			-0.013
			(0.034)			(0.017)
Category FEs	No	No	Yes	No	No	Yes
Baseline share (domestic)	0.27	0.27	0.27	0.11	0.11	0.11
N	12127	12127	12126	10008	10008	10008
R2	0.06	0.06	0.12	0.01	0.01	0.06

Table B.4: Firm-level new goods introduction rate

Notes: Table reports the firm-year level regression of the share of new products among the total number of products the firm has in a given year on a dummy for whether the firm is Mexican, the firm sales in million MXN, the number of not new varieties, firm fixed effects and other firm-level characteristics described in the table. The first three columns show all firms I have data for. Columns (4), (5) and (6) only show firms that sell at least one not new good in the year of observation. Standard errors are reported in parenthesis.

	All	Mexican	Foreign
Growth sales	0.11	0.14	0.08
	(0.35)	(0.48)	(0.17)
Product Life Cycle Component	0.06	0.10	0.03
Growth of Surviving	0.08	0.12	0.04
	(0.32)	(0.45)	(0.17)
Sales Share of Exit	-0.02	-0.02	-0.02
	(0.04)	(0.05)	(0.04)
New Products Component	0.04	0.04	0.04
Entry Rate	0.16	0.16	0.16
	(0.12)	(0.15)	(0.09)
Entrants Relative Sales	0.26	0.24	0.27
	(0.28)	(0.31)	(0.26)
Obs	8885	8010	875

Table B.5: Firm growth decomposition

Notes: Table shows the results from the decomposition of annual growth of sales at the firm-year level, as defined in Equation (3) and show in Figure 5.1. For each firm and year starting in 2011, I compute the contribution of new products the number of new products and their sales in their first year of activity. Table shows the sales-weighted average across all firms and years. The first column groups all firms, while the second and third column separate firms by whether they have received foreign investment ("Foreign") or not ("Mexican"). Standard errors are shown in parenthesis.

	Products observed every quarter of their life-cycle				
	All	Complete	Right-censored	Left-censored	Both R and L
Number	34322	4019	18593	22808	11098
Mean quarters	15	9	20	16	23
Median quarters	17	9	23	21	23
	All products ever observed				
	All	Complete	Right-censored	Left-censored	Both R and L
Number	66089	18558	35821	22808	11098
Mean quarters	10	5	13	16	23
Median quarters	9	4	14	21	23

Table B.6: Length of observation of products in quarters

Notes: Right-censored products are products that are still observed in the last 6 months of 2015. Left-censored Products are products that are already present in the first 6 months of 2010.

	Log sales				Log quantities		Log price	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age=2 quarters	0.659	1.051	0.660	1.067	0.541	0.969	0.158	0.055
	(0.079)	(0.092)	(0.079)	(0.089)	(0.080)	(0.094)	(0.037)	(0.038)
Age=3 quarters	0.558	1.319	0.561	1.356	0.322	1.162	0.308	0.085
	(0.111)	(0.125)	(0.111)	(0.121)	(0.111)	(0.127)	(0.052)	(0.051)
Age=4 quarters	0.422	1.543	0.426	1.601	0.063	1.278	0.446	0.121
	(0.149)	(0.166)	(0.149)	(0.161)	(0.149)	(0.168)	(0.070)	(0.068)
Age=5 quarters	0.276	1.796	0.282	1.872	-0.194	1.437	0.599	0.157
	(0.190)	(0.209)	(0.190)	(0.203)	(0.190)	(0.212)	(0.089)	(0.086)
Age=6 quarters	0.055	1.967	0.062	2.061	-0.528	1.498	0.732	0.203
	(0.232)	(0.255)	(0.232)	(0.247)	(0.232)	(0.257)	(0.109)	(0.104)
Age=7 quarters	-0.142	2.088	-0.134	2.201	-0.848	1.514	0.877	0.238
	(0.274)	(0.301)	(0.274)	(0.292)	(0.275)	(0.304)	(0.129)	(0.123)
Age=8 quarters	-0.393	2.236	-0.383	2.367	-1.209	1.546	1.016	0.283
	(0.317)	(0.348)	(0.317)	(0.337)	(0.318)	(0.351)	(0.149)	(0.142)
Age=9 quarters	-0.617	2.396	-0.606	2.546	-1.554	1.605	1.171	0.316
	(0.360)	(0.395)	(0.360)	(0.383)	(0.361)	(0.398)	(0.169)	(0.161)
Age=10 quarters	-0.825	2.542	-0.813	2.710	-1.888	1.625	1.305	0.364
	(0.404)	(0.442)	(0.404)	(0.429)	(0.404)	(0.446)	(0.189)	(0.180)
Age=11 quarters	-1.061	2.630	-1.047	2.817	-2.227	1.630	1.447	0.404
	(0.446)	(0.488)	(0.446)	(0.474)	(0.447)	(0.493)	(0.210)	(0.199)
Age=12 quarters	-1.267	2.745	-1.252	2.949	-2.544	1.621	1.566	0.461
	(0.490)	(0.536)	(0.490)	(0.520)	(0.491)	(0.541)	(0.230)	(0.219)
Age=13 quarters	-1.479	2.876	-1.461	3.098	-2.858	1.613	1.723	0.500
	(0.534)	(0.584)	(0.534)	(0.566)	(0.535)	(0.589)	(0.251)	(0.238)
Age=14 quarters	-1.706	2.999	-1.688	3.243	-3.228	1.646	1.843	0.569
	(0.578)	(0.632)	(0.578)	(0.614)	(0.579)	(0.638)	(0.271)	(0.258)
Age=15 quarters	-1.972	3.135	-1.952	3.401	-3.583	1.677	2.012	0.597
	(0.623)	(0.681)	(0.623)	(0.661)	(0.623)	(0.687)	(0.292)	(0.278)
Firm sales, mMXN			0.007	0.164				
			(0.004)	(0.006)				
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	No	No	No	No	Yes	Yes	Yes	Yes
Initial quarter mean	5.47	4.71	5.47	4.71	5.47	4.71	5.47	4.71
N	16836	12424	16836	12424	16678	12173	16678	12173
R2	0.23	0.28	0.23	0.32	0.33	0.30	0.56	0.59

Table B.7: Barcode life-cycle: log sales

Notes: Table shows the results from the product-quarter-level regression shown in Equation (4) of log sales (Columns (1)-(4)), log quantities (Columns (5) and (6)) or log price (Columns (7) and (8)) on dummies for the age of the product in quarters, product interacted with quarter fixed effects, cohort fixed effects and sometimes firm-level sales. Results from Colums (1) and (2) are also shown in Figure 5.2. I only keep products that survived at least 14 quarters and compute their total sales or quantities, or sales-weighted price in the dataset in each quarter. Standard errors are shown in parenthesis.

	Within firms across time			Across firms within time		
	Customers	Items per C	MXN per item	Customers	Items per C	MXN per item
Mexican firms	0.626	0.176	0.198	0.736	0.200	0.064
	(0.005)	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)
N	10934	10934	10934	12081	12081	12081
R2	0.97	0.90	0.86	0.81	0.43	0.03
	Within firms across time			Across firms within time		
	Customers	Items per C	MXN per item	Customers	Items per C	MXN per item
Foreign firms	0.582	0.141	0.277	0.795	0.205	0.000
	(0.015)	(0.008)	(0.015)	(0.007)	(0.005)	(0.008)
N	1126	1126	1126	1157	1157	1157
R2	0.99	0.97	0.90	0.91	0.61	0.00

Table B.8: Sales decomposition

Notes: Table shows the results of the firm-year-level decomposition of sales as explained in Equation (5a). On the left-hand side, I show the coefficients obtained on each firm-level regressions of the log number of customers, log number of items sold to each customer, and log monetary value of each item sold; respectively on the log of sales and firm and year fixed effects. The regressions are computed separately for Mexican and foreign firms, respectively. I show these results graphically in Figure 5. On the right-hand side, I show the coefficients obtained from three similar regressions without firm fixed effects, which amounts to considering each observation as a separate firm and interpreting the coefficient as heterogeneity across firms. Standard errors are shown in parenthesis.

	Within firms across time		Across firms within time	
	Products	Customers per P	Products	Customers per P
Mexican firms	0.284	0.716	0.474	0.526
	(0.005)	(0.005)	(0.003)	(0.003)
N	11273	11273	12599	12599
R2	0.95	0.95	0.67	0.72
	Within firms across time		Across firms within time	
	Products	Customers per P	Products	Customers per P
Foreign firms	0.458	0.542	0.626	0.374
	(0.017)	(0.017)	(0.008)	(0.008)
N	1150	1150	1186	1186
R2	0.98	0.96	0.82	0.63

Table B.9: Customers decomposition

Notes: Table shows the results of the firm-year-level decomposition of the number of customers as explained in Equation (6a). On the left-hand side, I show the coefficients obtained on each firm-level regressions of the log number of products and log customers buying each product, respectively on the log of number of customers and firm and year fixed effects. The regressions are computed separately for Mexican and foreign firms, respectively. I show these results graphically in Figure 6. On the right-hand side, I show the coefficients obtained from two similar regressions without firm fixed effects, which amounts to considering each observation as a separate firm and interpreting the coefficient as heterogeneity across firms. Standard errors are shown in parenthesis.

			Log expenditure				SES (1-6)	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age=2 quarters	0.011	0.002	0.014	-0.005	0.014	-0.005	0.093	0.008
	(0.008)	(0.012)	(0.007)	(0.011)	(0.007)	(0.011)	(0.018)	(0.027)
Age=3 quarters	0.012	-0.005	0.015	-0.017	0.016	-0.017	0.140	-0.020
	(0.010)	(0.015)	(0.009)	(0.014)	(0.009)	(0.014)	(0.023)	(0.033)
Age=4 quarters	0.024	-0.030	0.029	-0.051	0.030	-0.052	0.209	-0.031
	(0.013)	(0.018)	(0.012)	(0.018)	(0.012)	(0.018)	(0.030)	(0.042)
Age=5 quarters	0.045	-0.025	0.053	-0.055	0.054	-0.055	0.285	-0.027
	(0.016)	(0.023)	(0.015)	(0.022)	(0.015)	(0.022)	(0.038)	(0.052)
Age=6 quarters	0.060	-0.051	0.067	-0.077	0.069	-0.077	0.340	-0.033
	(0.020)	(0.027)	(0.018)	(0.026)	(0.018)	(0.026)	(0.046)	(0.062)
Age=7 quarters	0.066	-0.057	0.080	-0.087	0.082	-0.087	0.439	-0.059
	(0.023)	(0.032)	(0.022)	(0.030)	(0.022)	(0.030)	(0.054)	(0.073)
Age=8 quarters	0.092	-0.075	0.110	-0.107	0.112	-0.107	0.495	-0.039
	(0.027)	(0.037)	(0.025)	(0.035)	(0.025)	(0.035)	(0.063)	(0.084)
Age=9 quarters	0.105	-0.090	0.123	-0.122	0.125	-0.122	0.568	-0.052
	(0.030)	(0.041)	(0.028)	(0.039)	(0.028)	(0.039)	(0.071)	(0.094)
Age=10 quarters	0.119	-0.111	0.139	-0.141	0.142	-0.141	0.678	-0.041
	(0.034)	(0.046)	(0.032)	(0.044)	(0.032)	(0.044)	(0.079)	(0.105)
Age=11 quarters	0.131	-0.123	0.158	-0.154	0.161	-0.154	0.755	-0.076
	(0.037)	(0.051)	(0.035)	(0.049)	(0.035)	(0.049)	(0.087)	(0.116)
Age=12 quarters	0.148	-0.147	0.179	-0.182	0.182	-0.182	0.832	-0.069
	(0.041)	(0.056)	(0.039)	(0.053)	(0.039)	(0.053)	(0.096)	(0.128)
Age=13 quarters	0.172	-0.155	0.203	-0.187	0.207	-0.187	0.909	-0.079
	(0.045)	(0.061)	(0.042)	(0.058)	(0.042)	(0.058)	(0.104)	(0.139)
Age=14 quarters	0.192	-0.166	0.226	-0.197	0.229	-0.198	0.999	-0.078
	(0.048)	(0.066)	(0.045)	(0.063)	(0.046)	(0.063)	(0.113)	(0.150)
Age=15 quarters	0.204	-0.165	0.240	-0.205	0.244	-0.207	1.082	-0.033
	(0.052)	(0.071)	(0.049)	(0.067)	(0.049)	(0.067)	(0.122)	(0.162)
Firm sales, mMXN					0.001	0.013		
					(0.001)	(0.003)		
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Initial quarter mean	9.76	9.72	9.76	9.72	9.76	9.72	2.43	2.38
N	447113	219106	447113	219106	447113	219106	447113	219106
R2	0.02	0.03	0.13	0.13	0.13	0.13	0.05	0.05

Table B.10: Barcode life-cycle: characteristics of new customers

Notes: Tables represents the coefficients obtained on the dummies for the age of the product in quarters from a individual-product-quarter level regression described in Equation (7) where I regress the annual expenditure of the new customers of a product in a given quarter on these dummies, quarter by product category fixed effects and cohort fixed effects. In Columns (1) and (2) I don't have city fixed effects. In Columns (3) and (4) I add city fixed effects, and these are the coefficients which are represented in Figure 7. In Columns (5) and (6) I look at consumers' socio-economic status, which is an integer between 1 (lowest) and 6 (highest). For each outcome, I run two separate regressions for foreign and domestic products, respectively. Standard errors are shown in parenthesis.

	Log population		Log density		Log distance to CDMX	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)	(5)	(6)
Age=2 quarters	-0.610	-0.324	-0.534	-0.467	-0.048	-0.180
	(0.211)	(0.424)	(0.192)	(0.379)	(0.127)	(0.260)
Age=3 quarters	-0.786	-0.369	-0.856	-0.595	-0.068	0.013
	(0.306)	(0.633)	(0.278)	(0.569)	(0.182)	(0.381)
Age=4 quarters	-0.855	-0.562	-0.776	-0.998	-0.153	0.619
	(0.420)	(0.900)	(0.379)	(0.804)	(0.246)	(0.535)
Age=5 quarters	-0.944	-0.470	-0.895	-1.053	-0.165	0.507
	(0.548)	(1.194)	(0.493)	(1.057)	(0.320)	(0.702)
Age=6 quarters	-1.023	-0.444	-0.660	-1.238	-0.249	0.787
	(0.668)	(1.474)	(0.598)	(1.302)	(0.388)	(0.864)
Age=7 quarters	-1.066	-0.406	-0.712	-1.652	-0.370	1.514
	(0.787)	(1.778)	(0.705)	(1.561)	(0.457)	(1.035)
Age=8 quarters	-1.222	-1.299	-0.659	-1.631	-0.448	1.072
	(0.918)	(2.065)	(0.821)	(1.821)	(0.532)	(1.208)
Age=9 quarters	-1.263	-0.754	-0.873	-1.949	-0.292	2.068
	(1.036)	(2.326)	(0.928)	(2.059)	(0.600)	(1.364)
Age=10 quarters	-1.230	-0.384	-0.839	-1.938	-0.434	2.144
	(1.162)	(2.618)	(1.039)	(2.312)	(0.673)	(1.533)
Age=11 quarters	-1.646	-1.442	-0.862	-2.082	-0.467	2.006
	(1.285)	(2.925)	(1.151)	(2.574)	(0.745)	(1.706)
Age=12 quarters	-1.614	-0.828	-0.740	-2.621	-0.458	2.533
	(1.418)	(3.229)	(1.268)	(2.835)	(0.820)	(1.879)
Age=13 quarters	-1.732	-1.199	-0.945	-2.550	-0.351	2.516
	(1.550)	(3.501)	(1.384)	(3.088)	(0.896)	(2.047)
Age=14 quarters	-2.143	-0.575	-1.173	-3.419	-0.453	3.174
	(1.681)	(3.792)	(1.502)	(3.352)	(0.972)	(2.221)
Age=15 quarters	-1.911	0.677	-1.092	-2.128	-0.437	3.016
	(1.812)	(4.087)	(1.616)	(3.607)	(1.046)	(2.391)
Product X Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Initial quarter mean	14.14	14.22	5.26	5.26	6.01	5.97
N	4673	1802	5502	2151	5370	2093
R2	0.16	0.25	0.11	0.18	0.11	0.17

Table B.11: Barcode life-cycle: log sales-weighted average characteristics of new cities

Notes: Tables represents the coefficients obtained on the dummies for the age of the product in quarters from a product-quarter level regression described in Equation (7) where I regress the average population of the new cities reached by a product in a given quarter on these dummies, quarter by product category fixed effects, and cohort fixed effects in Columns (1) and (2). For each outcome, I run two separate regressions for foreign and domestic products, respectively. In Columns (3) and (4), I reproduce the regressions ran in Columns (1) and (2) with the average income per capita of the new cities and in Columns (5) and (6) with the poverty rate of the new cities. Standard errors are shown in parenthesis.

	OLS				IV			
	Price	s.e.			Price	s.e.	F statistic	
	-0.37	0.003			-0.54	0.004	3192885	
	N obs	N markets	N barcodes		N obs	N markets	N barcodes	
	2236295	720	47526		2038998	720	38642	
				Age Dummies				
Age $= n$	μ_n	s.e.	μ_n^F	s.e.	μ_n	s.e.	μ_n^F	s.e.
	0.26	0.01	0.03	0.01	0.31	0.01	0.02	0.01
2	0.27	0.01	0.02	0.01	0.32	0.01	0.02	0.01
3	0.29	0.01	0.02	0.01	0.34	0.01	0.01	0.01
4	0.29	0.01	0.03	0.01	0.34	0.01	0.02	0.01
5	0.27	0.01	0.02	0.01	0.32	0.01	0.02	0.01
6	0.28	0.01	0.00	0.01	0.32	0.01	-0.01	0.01
7	0.29	0.01	-0.04	0.01	0.32	0.01	-0.04	0.01
8	0.27	0.01	-0.01	0.01	0.31	0.01	-0.02	0.01
9	0.25	0.01	-0.03	0.01	0.29	0.01	-0.04	0.01
10	0.27	0.01	-0.08	0.01	0.30	0.01	-0.08	0.01
11	0.24	0.01	-0.06	0.01	0.27	0.01	-0.06	0.01
12	0.23	0.01	-0.10	0.01	0.26	0.01	-0.11	0.01
13	0.17	0.01	-0.07	0.01	0.19	0.01	-0.08	0.01
14	0.13	0.01	-0.03	0.01	0.14	0.01	-0.04	0.01
15	0.11	0.01	-0.05	0.01	0.11	0.01	-0.04	0.02
16	0.15	0.01	-0.07	0.01	0.17	0.02	-0.08	0.02
17	0.13	0.02	-0.06	0.01	0.14	0.02	-0.06	0.02
18	0.09	0.02	-0.06	0.02	0.09	0.02	-0.06	0.02
19	0.13	0.02	-0.11	0.02	0.12	0.02	-0.09	0.02
20	0.05	0.02	-0.06	0.02	0.05	0.02	-0.03	0.02
21	0.09	0.02	-0.14	0.02	0.08	0.02	-0.11	0.02
22	0.06	0.02	-0.15	0.02	0.08	0.02	-0.15	0.02
23	0.00	0.00	-0.08	0.02	0.00	0.00	-0.07	0.02

Table B.12: Results: estimation of the demand model with barcode fixed effects

Notes: Tables shows the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, and other relevant statistics from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects and product fixed effects. Figures 8 and 9 show the age dummies and the product dummies estimates, respectively.

	SES E, D+, D				SES C+,C,AB			
	Price	s.e.			Price	s.e.	F statistic	
	-0.60	0.005	.	.	-0.65	0.006	.	.
	N obs	N markets	N barcodes		N obs	N markets	N barcodes	
	1419210	720	30807		1231184	720	29158	
				Age Dummies				
Age $= n$	μ_n	s.e.	μ_n^F	s.e.	μ_n	s.e.	μ_n^F	s.e.
	0.28	0.01	0.06	0.01	0.26	0.02	0.03	0.01
2	0.27	0.01	0.07	0.01	0.30	0.01	0.02	0.01
3	0.29	0.01	0.04	0.01	0.33	0.01	0.03	0.01
4	0.31	0.01	0.04	0.01	0.32	0.01	0.04	0.01
5	0.30	0.01	0.06	0.02	0.31	0.01	0.02	0.01
6	0.29	0.01	0.04	0.02	0.33	0.01	-0.01	0.01
7	0.28	0.01	0.02	0.02	0.33	0.01	-0.02	0.02
8	0.27	0.01	0.01	0.02	0.35	0.02	-0.03	0.02
9	0.25	0.01	0.01	0.02	0.34	0.02	-0.07	0.02
10	0.25	0.01	-0.01	0.02	0.32	0.02	-0.05	0.02
11	0.25	0.02	-0.01	0.02	0.31	0.02	-0.08	0.02
12	0.25	0.02	-0.05	0.02	0.29	0.02	-0.13	0.02
13	0.15	0.02	0.01	0.02	0.24	0.02	-0.11	0.02
14	0.11	0.02	0.05	0.02	0.18	0.02	-0.04	0.02
15	0.07	0.02	0.03	0.02	0.17	0.02	-0.04	0.02
16	0.14	0.02	-0.02	0.02	0.14	0.02	0.01	0.02
17	0.11	0.02	0.05	0.02	0.16	0.02	-0.03	0.02
18	0.07	0.02	0.06	0.02	0.10	0.02	-0.03	0.02
19	0.11	0.02	0.04	0.02	0.11	0.02	-0.10	0.02
20	0.01	0.02	0.07	0.02	0.08	0.02	-0.03	0.02
21	0.07	0.02	-0.01	0.02	0.07	0.02	-0.05	0.02
22	0.02	0.02	-0.00	0.02	0.09	0.02	-0.09	0.02
23	0.00	0.00	0.03	0.02	0.00	0.00	0.06	0.02

Table B.13: Results: estimation of the demand model separately for low and high SES

Notes: Tables shows the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, and other relevant statistics from the estimation of Equation (9) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects and product fixed effects, estimated separately for households of low socio-eocnomic status and of high socio-economic status. Figure 10 shows the age dummies and the age interacted with foreign dummies, while Appendix Figure A.15 shows the product fixed effects.

	OLS				IV			
	Price	s.e.			Price	s.e.	F statistic	
	-0.62	0.002	.		-0.69	0.003	17047958	
	N obs	N markets	N barcodes		N obs	N markets	N barcodes	
	2244859	720	56090		2042406	720	42050	
			Age Dummies					
Age $= n$	μ_n	s.e.	μ_n^F	s.e.	μ_n	s.e.	μ_n^F	s.e.
Baseline	.	.	1.18	0.01	.	.	1.19	0.01
1	0.31	0.01	1.10	0.01	0.31	0.02	1.11	0.01
2	0.39	0.01	1.06	0.01	0.39	0.02	1.07	0.01
3	0.42	0.01	1.06	0.01	0.43	0.02	1.06	0.01
4	0.39	0.01	1.12	0.01	0.40	0.02	1.11	0.01
5	0.38	0.01	1.10	0.01	0.39	0.02	1.11	0.01
6	0.43	0.01	1.05	0.01	0.43	0.02	1.05	0.01
7	0.43	0.01	1.01	0.01	0.44	0.02	1.01	0.01
8	0.39	0.01	1.08	0.01	0.41	0.02	1.09	0.01
9	0.39	0.02	1.03	0.01	0.40	0.02	1.04	0.01
10	0.38	0.02	1.00	0.01	0.38	0.02	1.02	0.01
11	0.36	0.02	1.00	0.01	0.36	0.02	1.01	0.02
12	0.33	0.02	0.97	0.01	0.33	0.02	0.98	0.02
13	0.27	0.02	1.02	0.02	0.27	0.02	1.03	0.02
14	0.23	0.02	1.02	0.02	0.23	0.02	1.01	0.02
15	0.19	0.02	1.01	0.02	0.17	0.02	1.04	0.02
16	0.25	0.02	0.96	0.02	0.25	0.02	0.95	0.02
17	0.18	0.02	1.03	0.02	0.17	0.02	1.05	0.02
18	0.12	0.02	1.04	0.02	0.11	0.02	1.05	0.02
19	0.20	0.02	0.91	0.02	0.18	0.02	0.95	0.02
20	0.09	0.02	1.00	0.02	0.07	0.02	1.03	0.02
21	0.09	0.02	0.95	0.02	0.05	0.03	1.00	0.02
22	0.10	0.03	0.87	0.02	0.10	0.03	0.88	0.02
23	0.00	0.00	0.98	0.02	0.00	0.00	0.98	0.02

Table B.14: Results: estimation of the demand model without fixed effects

Notes: Tables shows the coefficients obtained on the dummies for the age of the product in quarters, and the age interacted with foreign, and other relevant statistics from the estimation of Equation (8) where I regress the log difference between share of the product and share of the baseline domestic product in its category on these dummies, quarter fixed effects, cohort fixed effects and product characteristics. Appendix Figure A.14 shows the age dummies graphically.

C Model

I introduce a conceptual framework in which I highlight the hypothesis that uncertainty about product quality generates this premium and how it can lead to a barrier for domestic firms.

C.1 Setup

In this model, agents are consumers faced with a new good of unknown quality x. x is a random variable from a known distribution with prior mean μ_0. Agents choose whether to purchase the good or not, maximizing

$$u(\mu) = \max\{\mu - \beta_i p, 0\}$$

where β_i represents price-sensitivity. For simplicity, I suppose there are only two possible types of agents: $\beta_i \in \{\beta_L, \beta_H\}$ where $\beta_H > \beta_L$. γ_i represent the market share of each type, and $\gamma_H > \gamma_L$. This game is dynamic. In each period, agents decide whether to purchase the good or not. If they purchase the good, they immediately learn the true quality x. They can then use this information to decide whether they will purchase the good in the next period. I call this learning through individual experimentation "individual learning". Agents buy at most three times in a row (three times if they like the good, one time if they don't). Agents may also learn through observation: by looking at whether people who have tried the good continue to purchase it or not, they can update their belief about the quality of the good. I call this "social learning".

C.2 Individual learning

Suppose that at period t no agent has purchased the good yet. Everyone has the same prior μ_t. Each agent decides whether

$$\mu_t - \beta_i p \lesseqgtr 0$$

If one agent decides to purchase the good and not the other, it must mean that it is the agent with the lesser price-sensitivity β_L. She immediately learns the true quality x. However, the other agents don't learn anything from observing this initial purchase. $\mu_{t+1} = \mu_t$.

C.3 Social learning

In the following period $t + 1$, the leader assesses whether

$$x - \beta_L p \lesseqgtr 0$$

Upon observing this decision, the other agents learn whether

$$x \lesseqgtr \beta_L p$$

and can update their belief accordingly:

$$\mu_{t+2} = \begin{cases} < \mu_{t+1} & \text{if } x < \beta_L p \\ \geq \mu_{t+1} & \text{if } x \geq \beta_L p \end{cases}$$

Importantly, this is anonymous: consumers do not need to know who has purchased the good, only that the good has been purchased several times, as an agent who has decided against purchasing the good in the initial period rationally expects that people who have decided to try are less price-sensitive than himself.

C.4 Sales trajectories

This model generates the following sales trajectory for a "successful" product $x > \beta_H p$, when $\mu_0 < x$ ("uncertainty") and for when $\mu = x$ ("no uncertainty"), where each point represents the fixed effects of the product aging one quarter on its sales compared to the initial quarter. Figure C.1 strongly resembles Figure 4 which plots the coefficient on the dummies for the age

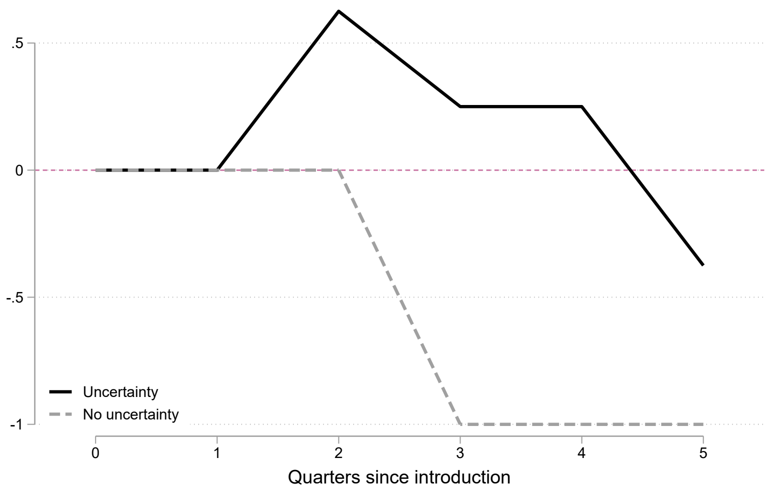


Figure C.1: Sales trajectories of a successful product given uncertainty

Notes: Figure represents the simulated coefficient on the dummies for age of the product in quarters one would get by running a product-quarter level regression of these dummies on sales of a successful product in the model.

in quarters of the products in the data.