Spring 2021

Essays on Geography and Firm Dynamics

Marcos Ribeiro Frazao
Yale University Graduate School of Arts and Sciences, mrfrazao@gmail.com

Abstract

Essays on Geography and Firm Dynamics

Marcos Ribeiro Frazao

2021

In this dissertation, I study the effects of geography on firm dynamics. I do so from three different perspectives. The first two chapters explore how geography affects sales of branded retail products in the United States. **Chapter 1** focuses on the spread of sales over time and space. It describes each component of sales and discusses the potential role of word-of-mouth in the growth of brands. **Chapter 2** studies the cross-section relation of brand sales and distance. The chapter analyses how physical trade frictions and information frictions interact and generate the observed reduction in customer base as distance increases. **Chapter 3** uses a reduced-form approach to analyze the border-effect using Brazilian exporters' data. I show that serving a neighboring country increases the probability of entry, the expected growth of sales and reduces the probability of exit.

Chapter 1 studies the spread of sales of branded products in the United States. Distance affects the cost of moving goods and people across space. Recent evidence suggests that geography also affects the flow of information. To investigate this hypothesis, I study the causes of brand sales growth over time and space. I analyze data from a large set of branded retail products sold in different regions in the United States and document a series of stylized facts about their life-cycle. I find that brands typically sell to a small number of locations that tend to be geographically close. Growth usually happens around previously successful markets. Furthermore, I decompose sales into three components: customer base, prices, and quantities per customer. Almost all of the variation in brand sales, both across locations and over

time, comes from the first term. The evidence suggests that geography plays a vital role in customer acquisition, but not due to differences in prices. Motivated by these findings, I propose a model in which information about brands' existence spreads geographically, similarly to how contagious diseases spread. Consumers aware of a brand might 'infect' others with that knowledge, and the probability of contagion depends on their location. Additionally, brands have different costs to deliver their goods to different markets. I use the predictions for the correlation of brand sales and customer base across regions to estimate the model using Simulated Methods of Moments and find that information frictions are more severe between distant locations. Furthermore, eliminating the role of distance in contagion increases consumer welfare by 32.5%. These results highlight the importance of geography for the spread of information about brands. This relationship allows for the description of brand dynamics across space and has significant welfare implications.

Chapter 2 studies the cross-section relations of sales and distance. How does distance affect the sales of brands in the United States? To answer that question, I use brand data across 44 different regions in the US. At the brand level, most of the effect of distance on sales is associated with a reduced number of customers in distant locations rather than sales per customer. At the aggregate level, most sales reduction comes from having fewer brands serving other areas. To understand what drives these patterns, I introduce a trade model between regions with shipping costs and search frictions between brands and final consumers. The estimates suggest that the shipping costs are a log-linear function of distance. Information frictions, however, are lower at the origin and are not affected by distance otherwise.

Chapter 3 describes the post-entry dynamics of Brazilian exporters. This article documents post-entry sales dynamics of Brazilian exporters and how they can depend on the set of countries that they exported to in previous years. Controlling for

marginal costs and selection on idiosyncratic demand, the results on firm's sales dynamics are similar to the ones in Fitzgerald et al. (2016) for Irish firms, and are robust to the inclusion of destination-year controls. The main contribution of this article is to investigate how these dynamics are affected by the set of destinations served by the firm in the previous periods. The evidence collected here suggests that sales to destinations that are close to the ones that were previously served by the firm tend to be higher, to grow more, and that the firm is less likely to exit from those locations. These *geographic spillovers* seem to contribute to more successful endeavors, and to be economically relevant when describing the dynamics of exporters. Finally, the evidence suggests that these spillovers are correlated with higher idiosyncratic demand in those destinations, but are not associated with lower fixed and sunk costs individually faced by firms.

Essays on Geography and Firm Dynamics

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Marcos Ribeiro Frazao

Dissertation Director: Samuel Kortum

June 2021

To Vitoria, the love of my life.

Contents

Acknowledgments

I would like to express my deep and sincere gratitude to my dissertation committee: Sam Kortum, Costas Arkolakis, and Michael Peters. Without their dedication, patience, and invaluable mentorship, this wouldn't be possible. Throughout the Ph.D., I have learned a great amount from them, and I will be forever thankful for their time and encouragement.

I thank the participants of the Yale International Trade Workshop and the University of Minnesota Workshop in Trade and Development, especially Lorenzo Caliendo, Ana Cecilia Fieler, Guillermo Noguera, John Eric Humphries, Tim Kehoe, and Manuel Amador. Besides their valuable comments, they taught me the best scholastic environments cannot abstain from the sense of camaraderie and great food.

I also thank the professors that I had the pleasure to assist with their lectures: Tony Smith, Zhen Huo, Giuseppe Moscarini, William Nordhaus, Zvika Neeman, and Aleh Tsyvinski. The Yale students are as lucky as I am to hear their teachings.

My sincere thanks also go to the staff of the Department of Economics. I am especially thankful to Nicole D'Aria, whose diligence saved me from missing a handful of deadlines, and Pam O'Donnell, the cornerstone of the department.

I was fortunate to be surrounded by friends throughout the program. I am grateful to Eduardo, Marcelo, Akshay, Paolo, Ian, Nic, Conor, Brian, Soonwoo, Ro'ee, Oren, Lucas, Martin, Luis, Marianne, Joao Paulo, Amy, Pedro, Natalia, Marina, Estella, and Analu. Their laughs made it all easier.

Finally, I am incredibly grateful to my family. I thank my mom, dad, and brother for their unconditional support, my wife Vitoria, who likes to sign contracts with me, and Alice, who I do not know yet, but I already miss dearly. They taught me the true meaning of the word *saudade*.

List of Figures

List of Tables

Chapter 1

Brand Contagion: The Popularity of New Products in the United States

1.1 Introduction

Spatial economics and international trade have blossomed in recent decades by exploring the economic implications of geography. Traditionally, these frameworks consider frictions that make it costly to move goods and people. Recent evidence, however, suggests that distance might also impact economic outcomes by making it harder for information to flow between places. I investigate this hypothesis using data on branded retail products sold in different regions in the United States. Brands allow consumers to identify products that they know. Therefore, analyzing brand data is a natural starting point to evaluate the role of information frictions between

Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

producers and final consumers, what I refer to as *product awareness*.

I use the Nielsen Homescan data to construct a panel of over 150,000 brands in 44 major regions in the US over a decade. The data reveal new stylized facts about the growth of brands in the US and the key factors behind that growth. I show that sales are usually concentrated in a small number of locations that tend to be geographically close. Older brands serve more markets and have higher sales, especially in areas close to their previous top markets.

To explore what drives these patterns, I break sales down into the customer base and sales per customer. Differences in the number of customers explain almost all of the variation in brand sales. Sales per customer are stable for a given brand, both in the cross-section and time-series dimension. Furthermore, I use data on bar-coded products to break sales per customer into price levels and quantities per customer. Neither component changes systematically over space or as the brand gets older.

These results suggest that trade costs affecting prices cannot explain the dynamics of brand sales and customer base in the US. Hence, I take an alternative approach and consider that the flow of information about products between different regions is the primary driver of customer acquisition.

Information about products can be important at different stages of the consumer decision process. The literature on product adoption often considers it to be a process with several steps. Consumers might gather data about which products exist and try to infer their quality and other features. However, this process invariably starts with product awareness, *i.e.* knowledge about their existence. Here, as in Kalish (1985), I reduce the adoption process to two stages: awareness and the decision to buy the product.

I propose a parsimonious model in which the main feature is the evolution of the number of consumers aware of each brand. The key aspect of this stochastic

process is contagion: consumers aware of a brand might 'infect' others with that knowledge, both locally and in other regions. Some areas might be more connected than others, which makes contagion more likely. Geography may be critical if the distance between locations decreases the probability of contagion. The estimated model confirms it is. With this feature the model is able to replicate the stylized facts discussed above.

In the model, consumers' decision to buy a product also relies on its price. They only buy from brands they know and that have the lowest price among their category. For this reason, the choice of the stochastic process for brand awareness is critical. The assumptions made here allow for a simple characterization of the price distribution. This object is necessary to compute the probability that a brand offers the cheapest product within a category and that the consumer actually buys it. Therefore, a brand's success depends on reaching consumers and offering lower prices than competitors.

I estimate the model for the year 2016, using a Simulated Method of moments. The model replicates the previously mentioned stylized facts about brand dynamics. The estimates show that distance is vital for contagion. A consumer is 50 times more likely to spread awareness of a product to neighboring regions than to the other side of the country. Furthermore, I use the model to conduct counterfactuals about the welfare gains of reducing frictions. Eliminating the role of distance on contagion increases consumers' welfare by 32.5% in the baseline estimation. It is important to note that this result might be sensitive to changes in key parameters, such as the demand elasticity and the distribution of brands productivities. Nevertheless, these results show the importance of geography for information flows. The role of geography for contagion is essential for the description of brand dynamics, and might have significant welfare implications.

Literature Review

The prominent role of informational frictions found here is in line with the recent literature on international trade. Chaney (2014) provides a framework in which exporters search for connections in other countries, and those connections subsequently help them find others they can sell to. By considering these information frictions, he can replicate the geographic patterns of entry by French exporters. The evidence found here suggests that the same forces that affect trade internationally are also present domestically.

This paper also contributes to the long literature of product and technology adoption following the initial works of Rogers (1962) and Bass (1969) in which consumers' adoption of new products rely on their interactions with other consumers, generating S-shaped adoption curves. The works that followed focused on two aspects of product adoption: the behavioral reasons behind consumer decisions to adopt new products and the mathematical characterization and estimation of these processes. Recent developments in the literature also use data on the particular social network of consumers and find evidence of the importance of their connections in the adoption of certain products.[1]

These advances are important for understanding how particular businesses grow, but there are a couple of considerations to be made when evaluating the aggregate implications of product awareness and contagion. First, we want to consider a broad set of products instead of focusing only on a few, since we want our results to be representative of a large part of the economy. My data allows me to do so, as it covers most items that final consumers buy. The other concern has to do with the economic environment. In order to evaluate counterfactuals we need the model to

[1] Hauser et al. (2006) provide a literature review on product innovation and adoption in Marketing Science, including a discussion on the current topics that are being researched in the field.

describe how brands compete in different locations in equilibrium. For this reason, we develop a general equilibrium model, where many brands exist but consumers are not aware of all of them. Brands might have different costs to serve different locations, and for a given product the consumer will choose the least expensive one that they are aware of.

In economics, there is a vast literature that assesses the aggregate effects of product creation and adoption. Romer (1987) provided the initial framework that links product creation and economic growth. Much has been done in this field and, more recently, Perla (2019) evaluates the effects of incorporating a slow diffusion process for new products that can explain patterns observed in the life-cycle of industries.

This paper also contributes to the long literature on the life-cycle of products. Argente et al. (2018a) also use Nielsen data, to emphasize the short life cycle of products. There are fundamental differences between their data approach and mine, however. First, they use retail scanner data, while I use the Nielsen homescan panel data. While they can compute each retailer's overall quantity, my choice of data allows me to construct measures of how many people consume each product. This choice is important since the customer base is the central object of my paper. Second, we differ in the product definition. They consider the UPC, which uniquely identifies the bar-code of a product. In my model, information frictions affect the set of products that consumers know. For this reason, I use the brand code constructed by the University of Chicago Booth. Each brand consists of several UPCs that have a similar visual identification. Argente et al. (2018a) point out that firms introduce new UPCs to replace old ones. Therefore, I do not observe short product life cycles as they do.

The fact that firms spend more than $200 billion annually in advertisements

just in the US[2], suggests that reaching potential customers is very valuable. By recognizing this, Gourio and Rudanko (2014) introduce search frictions for firms to reach consumers, which explains long-term customer relationships and can rationalize patterns in investment volatility by firms. In the context of International Trade Arkolakis (2010) formulates a model in which reaching consumers is a costly activity, allowing him to reconcile the positive correlation between firm entry and market size and the existence of small exporters in different countries.

This paper also explores how information frictions shape the growth of brands over time. As time passes, the brand accumulates more potential consumers because of the continuing search and contagion. This is related to Drozd and Nosal (2012), where firms continuously invest in increasing their customer base in different markets. Their model generates persistent pricing-to-market and quantitatively accounts for several puzzles in International Macroeconomics.

The stochastic structure of the model presented here draws heavily from Eaton et al. (2019). There, random search between firms generates contacts for potential suppliers. Distance affects search because firms are more likely to establish a relationship with the ones that are closeby. In that framework, firms buy intermediate inputs from the lowest-cost supplier among their contacts. Here, as in Lenoir et al. (2018) the search friction affects the relationships with final consumers. In my model, consumers buy from the lowest cost brand among the ones they know. Furthermore, I include the possibility of contagion among consumers. This deviation allows me to address the stylized facts for brands in the United States.

This paper pushes the understanding of how information frictions in the form of imperfect product awareness by final consumers can explain the dynamics of brands and economic aggregates over a geographic space. In the next section, I describe the

[2]Industry figures from GroupM/Statista.

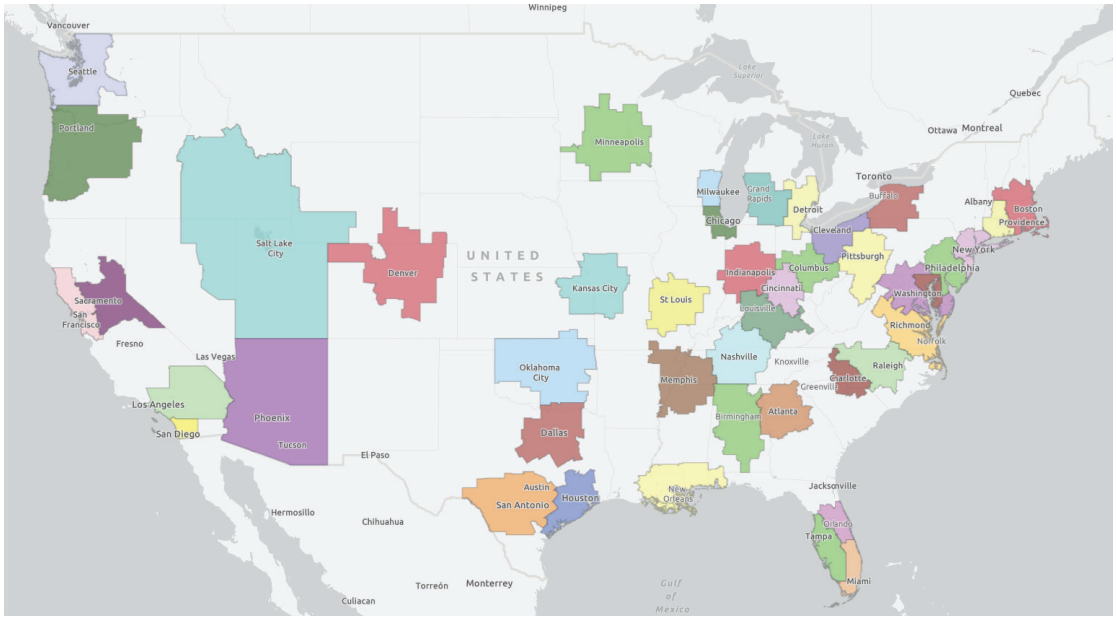

Figure 1.1: 44 representative Scantrack® Markets.

data and proceed to introduce the model.

1.2 Data and Preliminary Evidence

In this work, I use the Nielsen Homescan panel data between 2007 and 2016. It tracks the purchases of about 50,000 panelists per year in the US, with detailed information about households and products. The households' information we are interested in is where they live, what products they buy in a given year, and how much they spend on each product. The geographic unit that is used throughout the paper is the Scantrack® Markets, defined by Nielsen. There are 52 of these regions that cover all major metropolitan areas in the US. During the period evaluated, only the weights of 44 of these regions are designed to represent the whole region's demographics. I restrict my attention to these locations, which account for roughly 70% of the US population.

The Nielsen Homescan data presents two possibilities for defining a product: its

UPC bar code and its assigned brand code. According to the American Marketing Association dictionary: "A brand is a name, term, design, symbol or any other feature that identifies one seller's good or service as distinct from those of other sellers.". Since I am interested in the effects of product awareness, I naturally focus on the brand code as the product definition, as it presents the necessary aspects for consumers to identify and recall the product. Another reason to avoid using the UPC for this purpose is that many similar products might have different UPCs. For example, a 12 fl oz bottle of soda has a different UPC than the same soda in a 24 fl oz bottle.

I define the customer base of a brand as the number of households that bought the product at least once during a year. For each year, I compute the brands' customer base in each location and their total sales. The following are summary statistics for all brands and locations in 2016.

Table 1.1: Summary Statistics for Brands in 2016

# of Brands	Avg Number of Locations Served	Avg Total Sales	Avg Customer Base
73,118	11.48	$ 5,372,218	331,077

We can see that most brands are present in a few locations, as the average number of locations served is 11 out of a total of 44. It is also the case that their sales and number of customers are fairly concentrated. To evaluate that, I first compute the sales and number of customers in each brands' top market. Then, I calculate their sales and customers' share in other locations relative to their top market. Finally, I take the average of these measures over all brands, for their first, second, third locations and so on. Zeros are not accounted for in this calculation. If a brand sells to 7 locations during that year, the 0 in their 8th market is not included in the average. Figure 1.2 displays the results.

The figures suggest that as we move away from each brand's top markets, their

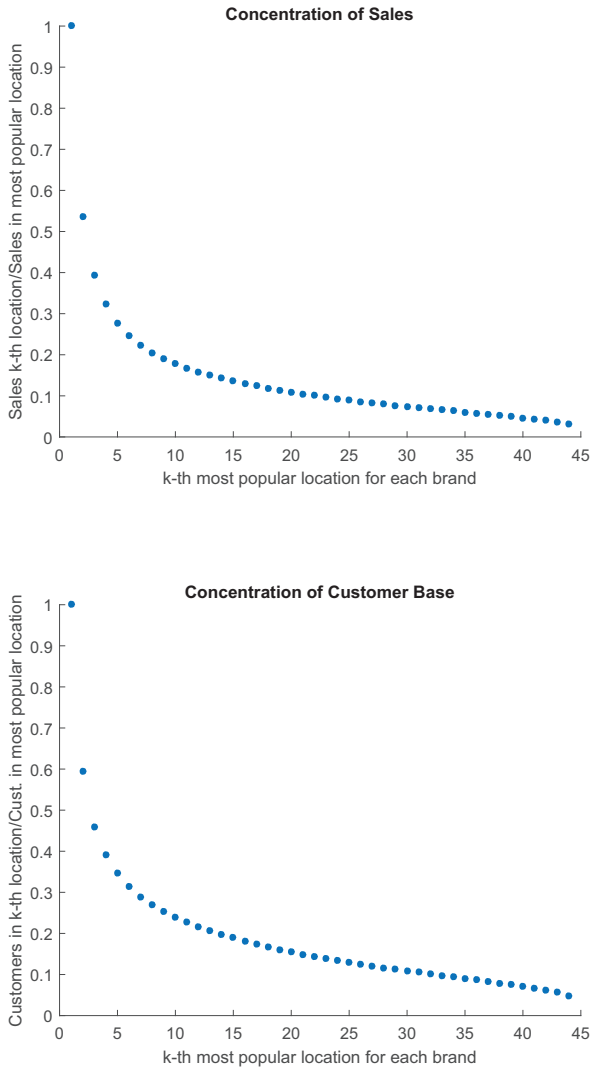

Figure 1.2: Concentration of sales and customers

Note: Each dot represents the brand's ratio of sales (customers) in their k-th market with respect to their top market, averaged across all brands that serve at least k markets.

number of customers and sales fall sharply and by comparable rates. The decline is considerably more accentuated than the reduction in total sales and population as we move away from the largest markets. One possible reason for this is that the same product might have higher prices in different markets, which would reduce their number of customers and possibly their sales. But there are other possible explanations for this pattern that do not rely on differences in prices, such as the geographic contagion mechanism presented in the model section.

To learn why brands grow, I investigate what happens to sales and customer base as they age. I start by constructing age variables for brands. For those that entered after 2007, I use the first year they appear in the sample to compute their age. After that, I build an index for their sales, customer base, and sales per customer by normalizing the variables by their initial values. Finally, I take the average over all brands in each age group, weighted by their initial sales. This measure tells me, for example, that brands in their second year had 2.5 times more sales than when they entered. In the next figure, we can see that older brands sell significantly more and to more people than they used to. However, the change in sales per customer is more modest, with an increase of just 20% after eight years.

Although sales per customer are relatively stable, it can still be the case that prices are decreasing. The price decline could explain the vast increase in the customer base that we observe, as consumers would shift their consumption towards the cheaper older brands. To investigate this hypothesis, I break sales per customer down into two components: prices and quantities per customer.

To have a good measure of prices by brand, I construct a price index for each brand in each location that they sell to during a particular year, as well as a nationwide price index for the brand. I start by constructing the average price for each UPC by computing their sales divided by quantity in each location. After that, I divide

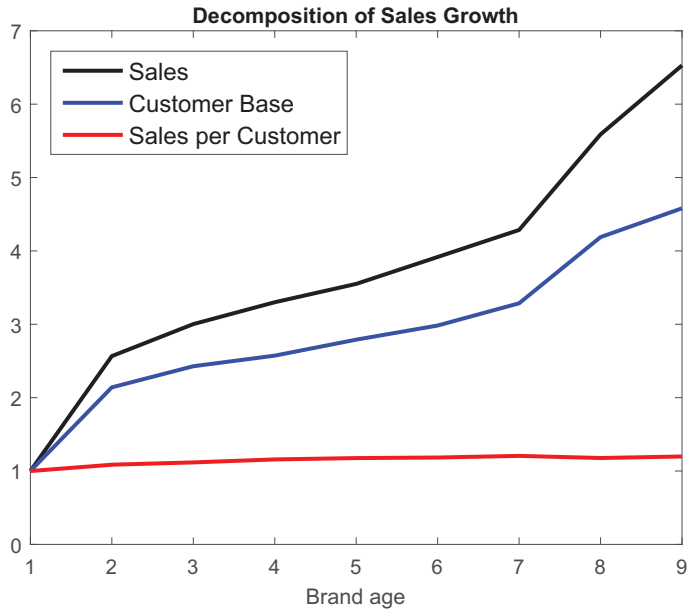

Figure 1.3: Evolution of sales

Note: Each brand's sales, customer base, and sales per customer are normalized to 1 for their initial year. The lines plot the weighted average of these indices for brands of different ages.

their average prices by their nation-wide average price in the first year that they appear in the data set. The brand-level price index in a location is then constructed by averaging all the brand's UPC indices sold there, weighted by their particular sales. However, some corrections need to be made, especially for the inclusion of new UPCs in the data. If a new UPC enters the data and there were already UPCs being sold beforehand, the new UPC's initial index is re-scaled using the price index for that brand in the entry period, excluding the entrant UPC. If no other UPC was being sold at that time, the entrant UPC's initial price index is re-scaled based on the most recent price index of the brand.

To highlight these corrections' importance, suppose that a brand sells a 12 oz can of soda, and its price doubles every year. By the 3rd year, the price index for the can would be 4. If in that year they launch a 24 oz bottle of the exact same soda with

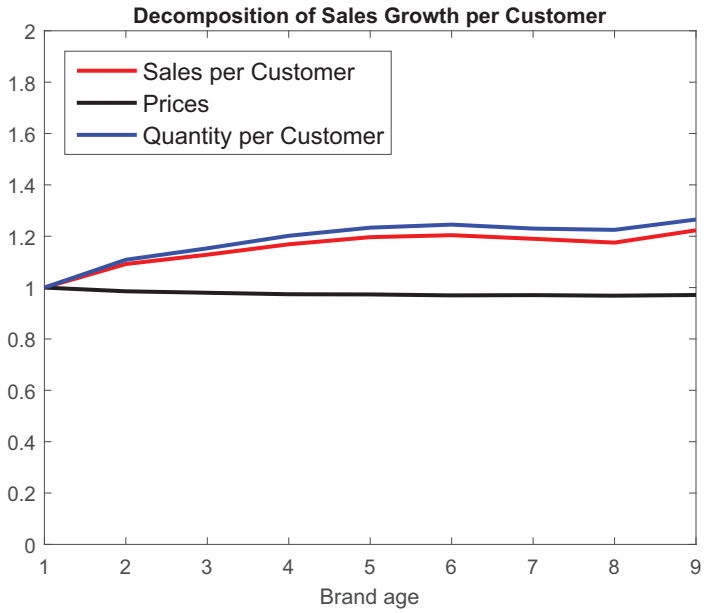

Figure 1.4: Evolution of sales per customer

Note: Each brand's sales per customer, prices, and quantity per customer are normalized to 1 for their initial year. The lines plot the weighted average of these indices for brands of different ages.

the same price-per-ounce as the can, the UPC price index of the bottle would be 1, and its inclusion would bias the brand-level price-index down. The correction makes the bottle's initial price index to be 4 and avoids this type of bias. Unfortunately, the correction cannot account for potential changes in the quality of new UPCs or unit-price changes associated with different packages.

After computing prices, I construct a quantity per customer index by dividing the brand sales by their customer base times the price index. Again, these measures are normalized by their initial values to see how age affects them. The next figure shows how these components affect sales per consumer. We can see that, on average, prices do not change much as brands get older. The small movement we have observed in sales per customer is accounted for by changes in quantities purchased. But again, these variations are dwarfed by the magnitude of the shift in the customer base.

Geography plays a significant role in the dynamics of brands' sales and customers. It is helpful to restrict our attention to a subset of the data to visualize these effects. Ideally, one would take brands from a particular location and track their sales and customers in different places over time. However, in our case, this is not feasible. The UPCs that appear in the data set are re-coded versions of the products real UPCs, so we cannot match them directly with firm data and find what their origin would be. Even if that would be possible, the headquarters of the firm might be different from where production happens. Also, some of these firms have several plants, so they can engage in production in different places across the country. These features make the definition of an origin questionable for some brands. To circumvent this issue, I assign origin by looking at the first location that sells a product from a particular brand. This definition is not used in the model, but by selecting brands with a particular origin, one can observe their growth patterns on a map, which is a good starting point.

With this in mind, I have selected the 168 brands that entered the data in 2008, and whose first sale was in Cleveland. The reason for the location choice is that Cleveland is relatively central, surrounded by other representative markets, and it is not small. In the next table, we can see the decline in the number of brands that are operating at a given year, and also that their average number of UPCs follows a hump-shaped pattern similar to the brand sales and customers.

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	All
Number of Brands	168	100	82	70	62	59	61	53	45	168
UPCs/Brand	1.44	2.18	3.22	4.00	3.92	3.56	3.26	3.13	2.86	4.07

Table 1.2: Summary statistics for brands that entered in 2008 and only served Cleveland during their first year.

Figures 1.14 to 1.21 in the Appendix show color-coded maps of the US displaying the total sales and customer base of these selected brands in each location for their first 4 years (2008, 2009, 2010, and 2011). There are a few things to note. First, the patterns observed in sales are very similar to the ones observed in customers. Sales are smaller at first and grow over time before they fall. In the beginning, most sales happen in Cleveland and surrounding areas, as well as in the two largest markets New York and Los Angeles. As time passes, not only do sales spread over other regions, but it seems like areas that are close to those two large markets also experience an unusually large increase in sales and customer acquisition.

To quantify the effects of geography in sales, I construct a measure of similarity between brands' sales in any two places. I start by computing the normalized sales of a brand in each location by dividing the total amount that a brand sells there in a given period by their average sales across all locations that they serve. This procedure generates a vector for each brand that indicates if the amount they sell in a location is above or below the brand's national average, making the observations for brands comparable. Next, I calculate the correlation between the normalized sales of brands in those two places for each pair of locations. If this correlation is high for pairs that are close to each other and low for pairs that are further away, it means that places where a brand has high sales tend to be closer to other places where sales are high, and conversely that places with low sales are also close to other low sales places. This is precisely what we observe by plotting these pairwise correlations and the distance between two locations, as in Figure 1.5.

I do the same procedure for all the components of brand sales that we observe: customer base, the price index, and the remainder, representing an index for the average quantity individual households buy in a given location. Figure 1.6 displays the results. We can see that the customer base behaves in the same way as sales.

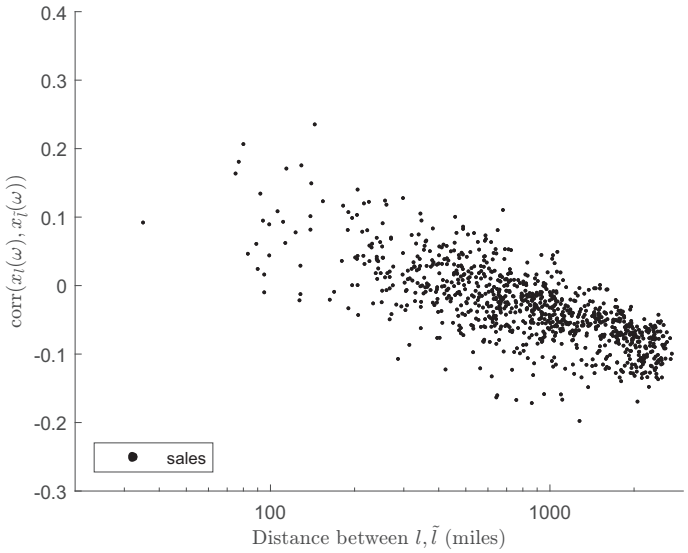

Figure 1.5: Correlation of brands relative sales in two locations

Interestingly, prices and individual quantities bought do not share the same patterns. This suggests that the geographic pattern that we observe for sales must be explained by something that affects customer acquisition in places that are close to each other, while not being explained by differences in prices.

After observing the geographic concentration of sales and customer base, I can compute similar correlations between pairs of locations in different periods. The lagged correlations measure the similarity of the variables in one region today and another tomorrow. The figure suggests that sales and customer base also display geographic persistence, in the sense that a higher level of sales in one location today is associated with higher sales in nearby areas tomorrow.

One potential explanation for that is what I call contagion, in which consumers that are aware of a brand can inform unaware consumers that are nearby. This mechanism is similar to the one which Chaney (2014) uses to model French exporter's entry into markets. However, the decisions that the consumer face in that setting are

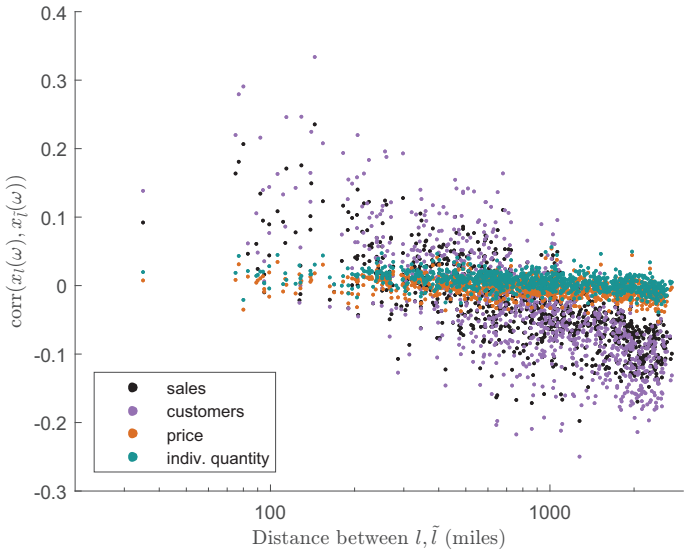

Figure 1.6: Correlation of brands relative sales in two locations

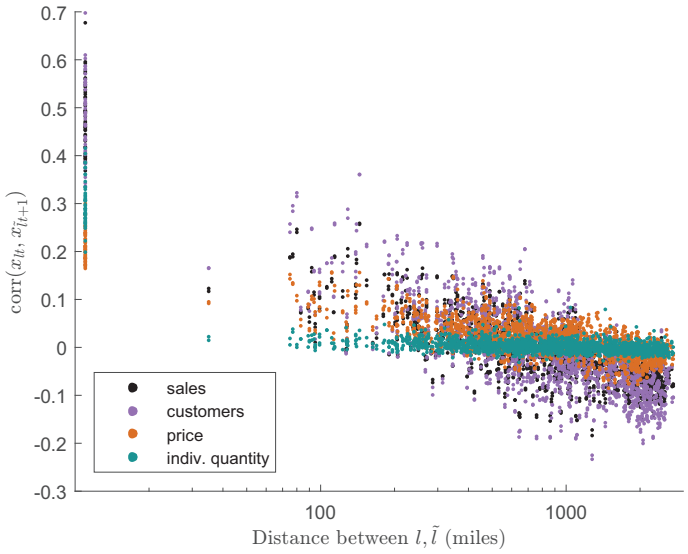

Figure 1.7: Correlation of brands relative sales one period ahead

very simplistic, in the sense that if a consumer is aware of a good they demand one unit of it. Here, I provide a framework in the spirit of Eaton et al. (2019), in which

information frictions prevent buyers from having access to all sellers, and they choose to purchase the least expensive variety that they are aware of. The model generates predictions about the distribution of brands' customers in different locations that can be brought to the data, and there is effective competition in the sense that not only must a brand be known, but it must also be cheaper than its competitors so that it makes a sale. This competitive setting might, therefore, be more suited for counterfactuals and welfare analysis.

1.3 Model

The model describes an economy that is composed of a discrete number of locations populated by consumers. They buy different varieties of products, and many brands can produce each variety. The key feature of the model is the description of a brand's customer base dynamics. Consumers buy a product if they are aware of the brand and if it is the cheapest among the known brands for that variety.

The evolution of the number of consumers aware of a brand - the *potential* customer base - is governed by a stochastic process. This process describes the brand actively searching for consumers in each region and the possibility of consumers spreading this information to others. If this contagion is more likely among consumers close to each other, the model predicts the patterns of geographic concentration and persistence observed in Figures 1.6 and 1.7.

The nature of the stochastic process comes from the literature in epidemiology. The particular modeling choices made here generate a closed-form solution for the distribution of brands' potential customers. This way, it is easy to compute the probability that the product is the cheapest one found by the consumer. I start by describing the evolution of the potential customer base. Then, I move to production

technology, consumers' preferences, and equilibrium outcomes.

Evolution of brand awareness

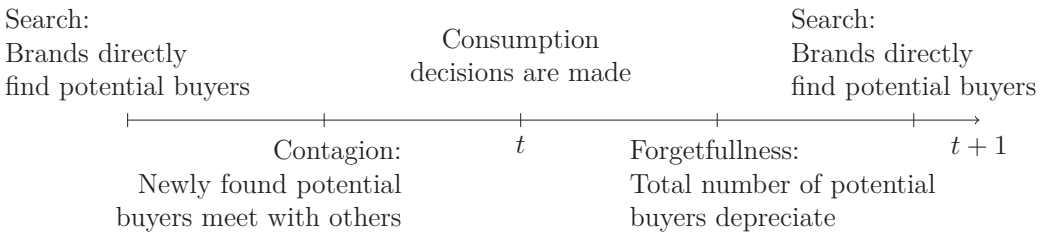

Figure 1.8: Timing of the model.

When a brand is created, it searches *directly* for consumers in all markets. The number of consumers it reaches this way in location l follows a Poisson distribution, $N_{0,l}^D \sim \text{Poisson}(\lambda_l)$, independent across all regions.

After that, the consumers who have been contacted can inform others who were previously unaware and spread the information about the brand's existence, which is akin to how contagious diseases spread. Let $N_{0,l}$ denote the number of consumers that the brand has matched in its first period in location l. This random variable is

$$N_{0,l} = N_{0,l}^D + \sum_{i=1}^{N_{0,l}^D} n_{0,l}(i) + \sum_{m \neq l} \sum_{i=1}^{N_{0,m}^D} n_{0,ml}(i).$$

The $n(i)$ random variables denote the number of potential consumers that each aware consumer can "infect" locally and in other regions. A high draw for $N_{0,m}^D$ also impacts the expected number of consumers in market l due to contagion. Note that $N_{0,l}$ is a compound random variable since the number of elements in the summations is random.

My goal is to have a simple characterization for the potential customer base of

brands with different ages, and dealing with compound random variables can be challenging. For this reason, I use the Generalized Poisson distribution, henceforth GP, which has properties that make the compounding due to contagion straightforward. This distribution was introduced by Consul and Jain (1973) as an extension of the Poisson distribution. It includes an extra parameter that allows for the variance to exceed the mean. If $X \sim GP(\lambda, \phi)$, then $\mathbb{E}(X) = \frac{\lambda}{1-\phi}$ and $\mathbb{V}(X) = \frac{\lambda}{(1-\phi)^3}$. The Poisson distribution is the special case in which $\phi = 0$. Notice that increasing ϕ above 0 increases the mean, but increases the variance by even more. Shoukri and Consul (1987) describe the use of the Generalized Poisson to characterize the total number of people infected by a contagious disease in a setting that is reasonably similar to our framework here. The properties that simplify the characterization of the potential number of buyers are the following[3]:

If $X \sim GP(\lambda_X, \phi), Y \sim GP(\lambda_Y, \phi)$ are independent, then $X + Y = Z \sim GP(\lambda_X + \lambda_Y, \phi)$.

If $X \sim GP(\lambda_X, \phi_X)$ and $\{Y_i\} \sim GP(\phi_Y, \phi_Y)$ is a sequence of *iid* random variables that are also independent from X, then $X + \sum_{i=1}^{X} Y_i \sim GP(\lambda_X, \phi_X + \phi_Y)$.

Property 1 is similar to the convolution property of the Poisson distribution, and it allows to add up independent draws of GP. Property 2 allows for compounding, and it is extremely helpful in the context of contagion. Notice that the number of Y_i random variables in the summation is X, which is a random variable itself.

I assume that the number of consumers that each aware consumer is able to reach in their own location is $n_{0,l}(i) \sim GP(\phi, \phi)$, with $\phi \in [0, 1)$. Recall that $N_{0,l}^D$ is simply Poisson distributed, which means that $N_{0,l}^D \sim GP(\lambda_l, 0)$. The second property implies that $N_{0,l}^D + \sum_{i=1}^{N_{0,l}^D} n_{0,l}(i) \sim GP(\lambda_l, \phi)$. Consumers also contribute to the spread

[3]The first property is derived in Consul (1989). The proof of the second property consists of an induction argument on the p.m.f of the resulting distribution to show that it is equal to the p.m.f. of $GP(\lambda_X, \phi_X + \phi_Y)$.

of awareness to other locations. I assume that the total number of consumers that each aware consumer in m finds in location l, is $\sum_{i=1}^{N_{0,m}^D} n_{0,ml}(i) \sim GP(\lambda_m \lambda_{ml}, \phi)$. Noting that the total number of consumers that became aware with information coming from different locations is independent, we can use the first property of the GP to write that

$$N_{0,l} \sim GP\left(\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}, \phi\right).$$

After contagion happens, the brand sells its product to consumers. The average number of potential consumers that a newly created brand has in location l is

$$\mathbb{E}(N_{0,l}) = \frac{[\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}]}{1 - \phi}.$$

We can see the influence of the other locations in the expansion of the customers in l, through the contagion parameters λ_{ml}. This is also evident when we compute the covariance of the number of aware consumers in two locations $\text{Cov}(N_{0,l}, N_{0,m}) = [\lambda_l \lambda_{lm} + \lambda_m \lambda_{ml}](1-\phi)^{-2}$. The higher λ_{ml} and λ_{lm} the higher the correlation between location l, m.

As we have seen in the data, the correlation between customers in close regions is higher. This suggests that λ_{lm} is higher if locations l and m are geographically close. Since I want to study the effects of distance on the information frictions, I model the contagion parameters as a function of distance: $\lambda_{lm} = C_0 \exp(-C_1 \text{distance}_{l,m})$. This choice allows having a simple description of how geography affects the contagion friction and greatly reduces the number of contagion parameters to be estimated.

Before the next period, some of the consumers aware of the product forget about

its existence. I consider a survival process similar to a binomial survival process for a random variable with a Poisson distribution. Say that $R \sim \text{Poisson}(\lambda)$ describes the number of people that know about a brand. If each of these individuals is independent and equally likely to forget about the brand existence, and the probability of it happening at the individual level is δ_b, then the distribution of the individuals that remember the brand conditional on X is $R|X \sim \text{Bin}(X;(1-\delta_b))$. The unconditional distribution of the consumers that remember is $R \sim \text{Poisson}((1-\delta_b)\lambda)$. Unfortunately, this survival process does not preserve the form of General Poisson distribution. However, the Quasi-Binomial distribution of type-II introduced by Consul and Mittal (1975) does. If $X \sim GP(\lambda,\phi)$ and $R|X \sim QBDII(X;(1-\delta_b),\phi/\lambda)$, then $R \sim GP((1-\delta_b)\lambda,\phi)$. Since I use this operation quite frequently, I define the following operator

$$\text{If } X \sim GP(\lambda,\phi), \text{ then } \chi(X)|X := QBDII(X;(1-\delta_b),\phi/\lambda).$$

Evaluating $\chi(X)$ before knowing the value of X implies that $\chi(X) \sim GP((1-\delta_b)\lambda,\phi)$. Therefore, the total number of potential customers that did not forget about the brand is

$$\chi(N_{0,l}) \sim GP\left((1-\delta_b)\left[\lambda_l + \sum_{m\neq l}\lambda_m\lambda_{ml}\right],\phi\right).$$

After that, the brand directly searches for consumers as before. The number of consumers that they find is independently distributed $N^D_{1,l} \sim \text{Poisson}(\lambda_l)$, and contagion happens as in the previous period. Since the number of newly found potential consumers is independent from the ones that remember the brand, we can

write

$$N_{1,l} = \chi(N_{0,l}) + N_{1,l}^D + \sum_{i=1}^{N_{1,l}^D} n_{1,l}(i) + \sum_{m \neq l} \sum_{i=1}^{N_{1,m}^D} n_{1,ml}(i)$$
$$N_{1,l} \sim GP\left(\left[\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}\right] + (1-\delta_b)\left[\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}\right], \phi\right).$$

In general, for a brand with age a, we have that

$$N_{a,l} \sim GP\left(\sum_{i=0}^{a} (1-\delta_b)^i \left[\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}\right], \phi\right) := GP(\Lambda_{l,a}, \phi).$$

where $\Lambda_{l,a} := \sum_{i=0}^{a} (1-\delta_b)^i \left[\lambda_l + \sum_{m \neq l} \lambda_m \lambda_{ml}\right]$. Given the properties of the GP, the mean and variance of the potential customer base is

$$\mathbb{E}(N_{la}) = \frac{\Lambda_{la}}{1-\phi}, \quad \mathbb{V}(N_{la}) = \frac{\Lambda_{la}}{(1-\phi)^3}.$$

The average number of potential customers grows with age, and so does the variance. The assumption that there is a positive probability that consumers forget about the brand guarantees that the process is stationary.

This process is related to epidemiologic metapopulation models[4]. There, infected people in one location might spread the disease to others in different places. The main difference here is that consumers are only contagious when the brand directly found them. This assumption limits the intertemporal effects that contagion might have on the evolution of the customer base. However, it allows for a closed-form

[4]See Sattenspiel (2009) for examples.

characterization of the distribution of all brands' potential customer base given their age. This helps keep track of how many brands each *consumer* is aware of, which is a key market condition that governs the probability that a brand actually sells its product.

To evaluate the *effective* number of customers for each brand, we move to the description of the costs that the brand faces and how consumers choose which products they buy.

Technology

The data doesn't assign an origin to a brand. Consequently, it is harder to consider the effects of geography on the costs of delivering goods to different locations, as it is not possible to model the production cost plus the shipping costs explicitly. However, the model circumvents this problem by considering that brands are assigned a vector of productivities to deliver their good in different locations. This setting allows for the interpretation that regions with low costs are close to the production site, and the ones with high costs are hard for the brand to reach. This interpretation is also consistent with several plants producing a single brand's goods, which is the case for many well established retail products. The data also suggests that differences in brand prices across regions are not essential to describing the geographic patterns of their sales. Therefore, I choose to simplify the cost structure of brands and allow for a richer characterization of awareness evolution.

A brand can deliver the good that they produce in every location $l = 1, \ldots, \mathcal{L}$, but at different costs. Each brand has a linear productivity in each location z_l. The measure of brands of all ages that have their productivity vector greater or equal to $\mathbf{z}$, element by element, is[5]

[5]One way to interpret the description of the measure of brands is to consider that a measure

$$M(\mathbf{z}) = \left(\sum_{l=1}^{\mathcal{L}} \frac{z_l}{T_l^{1/\theta}} \right)^{-\theta}.$$

This implies that for $z_l > 0$, the measure of brands of all cohorts that have productivity higher than z_l is $M_l(z_l) = T_l z_l^{-\theta}$, and the correlation between the productivities of a brand in any two locations is $1/\theta$. The measure of brands that have delivery cost below c is then $\mu_l(c) = T_l c^{\theta}$. After every period a fraction of δ_e of brands of all ages cease to exist and are replaced by entrant brands. This implies that the measure of brands of age a that have unit costs below c in location l is $\mu_l^a(c) = \delta_e(1-\delta_e)^a \mu_l(c) = \delta_e(1-\delta_e)^a T_l c^{\theta}$.

The parameters T_l govern how costly it is, on average, to deliver a good to a particular location. In the data, regions might have different price distributions for various reasons, such as rents or distance from where production happens. In the model, T_l can capture these cost differences. In equilibrium, however, the price index of a location also depends on how well information about brands circulates.

Consumers

All consumers have the same utility function. They value consumption in a single period according to a CES aggregator over the quantity that they consume of each variety j.

$$U(\mathbf{q}) = \sum_{t=0}^{\infty} \beta^t \left[\int_0^1 q_t(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}.$$

ϵ^{-1} of brands draw $\mathbf{z}$ from a multivariate Pareto distribution as in Mardia (1962) with CDF $1 - \left(\sum_{l=1}^{\mathcal{L}} \frac{z_l}{(\epsilon T_l)^{1/\theta}} - \mathcal{L} + 1 \right)^{-\theta}$, and to take the limiting case as $\epsilon \to 0$.

There is perfect substitution within a given variety j. Consumers in l are endowed with $\bar{X}_l/L_l$ units of cash every period. Their budget constraint is $\int_0^1 p_t(j)q_t(j)dj = \bar{X}_l/L_l$, where j indexes the varieties. They are aware of many brands that produce each variety. They are also indifferent about which brand produces the good. Therefore, their decision process consists of looking at all the costs from brands they know as quotes and buying from the cheapest one. Now, I describe the distribution of the number of brands consumers know.

Consumers are equally likely to be hit by any type of shock that either makes them aware or unaware of a brand. Therefore, the intensity that consumers in l are matched with a brand with age a is simply the average number of consumers reached by brands of that cohort, divided by the local population: $\frac{\Lambda_{l,a}}{L_l(1-\phi)}$. Hence, the distribution of the number of quotes that consumers have with cost below c is Poisson distributed, with the following parameter

$$\begin{aligned}\rho_l(c) &= \sum_{a=0}^{\infty}\int_0^c \frac{\Lambda_{l,a}}{L_l(1-\phi)}d\mu_l^a(c) = \frac{\delta_e T_l}{L_l}\left(\sum_{a=0}^{\infty}\frac{(1-\delta_e)^a\Lambda_{l,a}}{1-\phi}\right)c^\theta \\ &= \nu_l c^\theta.\end{aligned}$$

This implies that the probability that a buyer encounters no quotes below c for a given variety is $\exp(-\rho_l(c))$. The effective price paid by the consumer is the lowest quote they can find. The price distribution in l is given by integrating over all varieties in the unit interval [0,1]:

$$G_l(p) = 1 - \exp\left(-\nu_l p^\theta\right).$$

This is the distribution of quotes from brands of all ages. To study brand sales evolution, I derive the quote distribution of brands with a particular age from evaluating the probability that a consumer buys from a brand from that cohort. The distribution of quotes that a buyer in l receives from brands with age a also follows a Poisson distribution, this time with parameter

$$\rho_l^a(c) = \int_0^c \lambda_l^a d\mu_l^a(c) = \frac{\delta_e T_l}{L_l} \left(\frac{(1-\delta_e)^a \Lambda_{l,a}}{1-\phi} \right) c^\theta := \nu_{l,a} c^\theta.$$

Hence, the distribution of the lowest quotes coming from brands with age a in location l is

$$G_{l,a}(c) = 1 - \exp\left(-\nu_{la} c^\theta\right).$$

The probability that a variety is bought from a brand with age a can be found by solving the following integral $\pi_{la} = \int_0^\infty \Pi_{a' \neq a}[1 - G_{la'}(c)] dG_{la}(c)$, which implies that

$$\pi_{la} = \frac{\nu_{la}}{\nu_l} = \frac{(1-\delta_e)^a \Lambda_{l,a}}{\sum_{a'=0}^\infty (1-\delta_e)^{a'} \Lambda_{l,a'}}.$$

This equation shows two forces at play that affect the probability that consumers buy from a given cohort. On the one hand, as brands get older, they become more well-known, as shown by the Λ_{la} term that increases with age. However, as time passes, a brand is also more likely to exit, as we can see on the $(1-\delta_e)^a$ term. Those two forces generate a hump-shaped curve for the total customer base and sales of particular cohorts, similar to what we observe in the data.

Price Index and equilibrium

To characterize the equilibrium, I start by constructing the price index, which imposes a necessary parameter restriction for its existence. Second, I describe each brand's effective customer base, which is needed to compute its sales, and discuss the market clearing conditions in detail.

For the CES utility, the price index is given by $P_l = [\int_0^\infty p^{1-\sigma} dG_l(p)]^{\frac{1}{1-\sigma}}$, where p is the effective price paid by consumers. As mentioned above, the price paid is the lowest quote that consumers can find, which implies that $G_l(p) = 1 - \exp(-\nu_l p^\theta)$. Solving the integral, we find that

$$P_l = \nu_l^{-\frac{1}{\theta}} \left[\Gamma\left(1 - \frac{\sigma - 1}{\theta}\right)\right]^{\frac{1}{1-\sigma}}$$

This equation implies that we need to impose the parameter restriction that $\theta > \sigma - 1$, for the equilibrium to be well-defined.

The CES preferences imply that a consumer spends $c_l^{1-\sigma} P_l^{\sigma-1} \frac{\bar{X}_l}{L_l}$ in a variety with cost c_l. To find the total sales of a particular brand, we need to characterize the total number of consumers that effectively buy their product. Let $B_{la}(c_l)$ denote the random variable that describes the customer base of a brand with age a and that has cost c_l. As previously discussed, age matters for the distribution of their potential customer base N_{la}, but each potential consumer has a probability $\exp(-\nu_l c_l^\theta)$ of actually buying the product. This implies that, given a realization of N_{la}, the number of buyers follows a binomial distribution with parameters N_{la} and $\exp(-\nu_l c_l^\theta)$. We can use the total law of expectations to show that

$$\mathbb{E}[B_{la}(c_l)] = \mathbb{E}[\mathbb{E}[B_{la}(c_l)|N_{la}]] = \mathbb{E}[\exp(-\nu_l c^\theta)N_{la}] = \frac{\exp(-\nu_l c_l^\theta)\Lambda_{la}}{1-\phi}.$$

We can see that on average older and efficient brands have more customers. However, randomness still plays an essential role in the outcomes of specific brands. First, even a brand with high productivity might be unlucky to match consumers who have found cheaper alternatives. As time passes, this effect is mitigated by the expected increase in their potential customer base. But the trajectory of consumers who are aware of the brand is also random, and some brands will be more successful in finding consumers. In particular, if a brand directly finds consumers in one location, they might 'infect' others randomly. If the contagion parameters are larger for close regions, this luck can also spread to neighboring places. This process generates the pattern we observe in the data: the customer base of a brand is geographically concentrated and persistent.

In the data, we observe the same geographic pattern for sales. We can see this in the model by writing the brand sales as $x_{la}(c_l) = c_l^{1-\sigma} P_l^{\sigma-1} \frac{\bar{X}_l}{L_l} B_{la}(c_l)$. Therefore, the expected sales of a brand is

$$\mathbb{E}[x_{la}(c_l)] = c_l^{1-\sigma} P_l^{\sigma-1} \frac{\bar{X}_l}{L_l} \frac{\exp(-\nu_l c_l^\theta)\Lambda_{la}}{1-\Phi_{la}}.$$

In equilibrium, the fact that the demand of each brand is met by supply also implies that

$$\sum_{a=0}^{\infty} \int_0^{\infty} \mathbb{E}[x_{la}(c)] d\mu_{al}(c) = \bar{X}_l.$$

Model Predictions

The model generates other predictions that can be brought to the data. Going back to the definition of the effective customer base of a brand $B_{la}(c)$, we know that $B_{la}(c)|N_{la} \sim Bin(N_{la}, \exp(-\nu_l c^{\theta}))$. This implies that a brand that has N_{la} potential customers has a probability $(1 - \exp(-\nu_l c^{\theta}))^{N_{la}}$ of not selling to any of them. Since $N_{la} \sim GP(\Lambda_{la}, \Phi_a)$, we can use the p.m.f. of the Generalized Poisson distribution to numerically compute the probability that a brand of age a and cost c finds at least one customer in l as

$$\begin{aligned} \mathbb{P}(B_{la}(c) > 0) &= 1 - \mathbb{P}(B_{la}(c) = 0) \\ &= 1 - \sum_{k=0}^{\infty} (1 - \exp(-\nu_l c^{\theta}))^k \frac{\Lambda_{la}(\Lambda_{la} + k\phi)^{k-1}}{k!} e^{-\Lambda_{la} - k\phi}. \end{aligned}$$

We can then integrate over all brands costs to find the mass of brands that operate in l with age a

$$F_{la} = \int_0^{\infty} \mathbb{P}(B_{la}(c) > 0) d\mu_{la}(c).$$

Furthermore, we can add these cohorts to find the total number of brands that serve each location as $F_l = \sum_a F_{la}$.

One way to compute the average number of buyers among brands that operate

in a market is to add up all the customers that brands find in a given location and divide by the number of brands that are selling in that market:

$$R_{la} = \int_0^\infty \mathbb{E}[B_{la}(c)] d\mu_{la}(c) = \int_0^\infty \frac{\exp(-\nu_l c^\theta)\Lambda_{la}}{1-\Phi_{la}} d\mu_{la}(c) = \frac{\Lambda_{la}}{1-\Phi_l} \frac{\delta_e(1-\delta_e)^a T_l w^{-\theta}}{\nu_l}.$$

The average number of buyers among brands that actively sell to location l is then

$$\bar{b}_{la} = \frac{R_{la}}{\tilde{F}_{la}}, \text{ and } \bar{b}_l = \frac{\sum_a R_{la}}{\sum_a F_{la}}.$$

These moments are targeted in the estimation algorithm and are helpful to identify the probability that a consumer forgets about the brand δ_b.

Furthermore, there are closed-form solutions for the variance of customer base, but the solution for the covariances is quite involved. These moments are crucial for the identification of the contagion parameters. To proceed with the estimation, I rely on the Simulated Method of Moments. For thousands of brands, I simulate draws for productivities, the evolution of the potential customer base, and whether consumers effectively buy the brands' products. I then use the model-generated data to compute the correlation between the customer base and brands' sales in different locations. I describe the algorithm in more detail in the next section.

1.4 Estimation

In this section I describe the algorithm to estimate the model. The first thing to notice is that the parameter β does not affect the equilibrium allocations, and in our

context is only relevant for welfare calculations. When performing those exercises we choose a range for β that is compatible with the literature on intertemporal discounting.

I start by recovering an estimate for the elasticity of substitution, σ. If I want to identify the true elasticity of demand, I should control for supply-side endogeneity of prices. However, the model has the strong assumption that sales happen at cost, so *under the model assumptions* this is not an issue. Therefore, even if the estimate for the demand elasticity might not be generalized to other contexts, it allows the model to generate predictions about quantities sold that are aligned with the data.

Under the model assumptions, if a consumer buys a good with price p, the quantity that they acquire is:

$$q_l(p) = p^{-\sigma} P_l^{\sigma-1} \frac{\bar{X}_l}{L_l}.$$

The equation states that the quantity demanded depends on a location-specific factor, including the local price index, individual spending, and the price. Since the measures that I have for prices and quantities are indices computed for each brand, I regress the quantities sold by a brand in a location on the price paid there, including location and brand fixed effects.

$$\ln(q_l(\omega)) = -\hat{\sigma} \ln(p_l(\omega)) + I_l + I_\omega + \epsilon_{\omega,l},$$

where ω denotes an individual brand, $q_l(\omega)$, their quantity index and $p_l(\omega)$ their price index. So $\hat{\sigma}$ is the estimate for the elasticity of substitution that is consistent with the patterns of local demand observed in the data.

The estimate found here is $\hat{\sigma} = 0.1257$, which is a low number in comparison to the literature. There are a few reasons why this might be the case. One of them is relying on variation for a brand's quantities per customer in different locations. As shown before, this variable does not change much. Therefore, this could be capturing the inelastic demand for quantities for a particular good once the consumer decides to buy it. For example, one might still buy a single bag of chips for lunch, even when there are significant discounts. This might be problematic when computing the welfare measures, as this parameter also governs the elasticity across different varieties. For this reason, it is important to consider alternative values for σ when conducting these analyses.

After that, I estimate the remaining parameters of the model, using a Simulated Method of moments. Let $\Theta = (\theta, \delta_e, \delta_b, \boldsymbol{\lambda}, \boldsymbol{\phi})$. For every Θ I pick a vector of T_l that rationalizes the difference in prices observed in the data.

$$T_l(\Theta) = \left(\bar{P}_l \Gamma \left(\theta - \frac{\sigma - 1}{\theta} \right)^{\frac{1}{1-\sigma}} \right)^{-\theta} \bar{\nu}_l^{-1},$$

where $\bar{\nu}_l := \nu_l / T_l$ and $\bar{P}_l$ is the average of the normalized prices of brands in location l. The parameter T_l defines the average draw of costs in each location and directly affects the price level in a region. A higher T_l increases the brand probability of having low costs in that location ex-ante. But it also influences the competition landscape since it affects the distribution of costs and the probability that the brand effectively sells its product to a potential consumer.

I separate the moments that are targeted by the algorithm into two groups: $m_1(\Theta)$ and $m_2(\Theta)$. The first group represents moments that do not require simulation and can be computed directly. They are the number of brands that sell in each location,

their average customer base, and average sales $F_l, \bar{b}_l, \bar{x}_l$. For the year of 2016, I can assign age for brands with $a = 0, \ldots, 8$, so I also target the number of brands, average customer base and average sales by cohort $(F_{la}, \bar{b}_{la}, \bar{x}_{la})_{a=0}^{8}$. The evolution of the number of brands is important to identify the parameter δ_e, which determines the exogenous exit rate of brands. The evolution of the average customer base and sales provide information about the depreciation of customer base δ_b.

The second set of parameters requires simulation. I choose a large number of brands to simulate $\bar{K}$. Then I draw the productivity vectors from the following CDF: $1 - \left(\sum_{l=1}^{\mathcal{L}} \frac{z_l}{(K^{-1}T_l)^{1/\theta}} - \mathcal{L} + 1\right)^{-\theta}$. This distribution is convenient as it has closed-form solutions for the marginal distributions $F(z_l)$, and the conditional distributions $F_l(z_l|(\bar{\mathbf{z}}))$, where $\bar{\mathbf{z}}$ is any subset of $\mathbf{z}$. This allows me to draw $\mathbf{z}$ sequentially: $\bar{z}_1 \sim F_1(z_1)$, $\bar{z}_2 \sim F_2(z_2|\bar{z}_1), \ldots, \bar{z}_{\mathcal{L}} \sim F_{\mathcal{L}}(z_{\mathcal{L}}|\bar{z}_1, \ldots, \bar{z}_{\mathcal{L}-1})$. After that, I simulate the evolution of brand awareness by sequentially following the steps of customer acquisition described in the model: direct search, contagion based on the previous draws and then the number of the customers that forget. I divide the $\bar{K}$ number of brands in cohorts with size proportional to the model implications for brand survival, that is $1 - \delta_e$ to the power of the cohort, and compute their trajectories until the current year. This way, the distribution of brands age is similar to the model predictions.

With the values of productivities and potential customer base in hand, I compute the probability that a consumer buys a product from each brand as $\exp(-\nu_l c_l^{\theta})$. I use the brand-location specific probability of actually selling the good and the number of potential customers to draw the number of actual customers. After that, I compute brand sales by multiplying the customer base by the amount spent by consumer $c_l^{1-\sigma} P_l^{\sigma-1} \frac{\bar{X}_l}{L_l}$. This way, I can compute the same normalized correlations of brand sales and brand customer base as shown in Figure 1.6. I also compute age-specific correlations both in the model and in their data. Those moments are the key elements

that guide the choice of the contagion parameters, since the greater the contagion parameters between two regions, the larger their current correlation of sales and customers.

Finally, the algorithm searches for the set $\hat{\Theta}$ that solves the following problem

$$\hat{\Theta} = \text{argmin}_{\Theta}||m(\Theta) - \bar{m}||, \quad \text{s.t. } \Theta > 0, \quad \phi < 1, \quad \theta > \hat{\sigma} - 1,$$

where the model implied moments are $m(\Theta) = [m_1(\Theta), m_2(\Theta)]$, and $\bar{m}$ are their data counterparts.

I have intentionally not included the lagged correlations of Figure 1.7 as a moment to be targeted. This way, I can evaluate how well the model performs in describing the dynamics of the geographic spread of brands by contrasting the predictions of the estimated model with these moments.

Now I discuss the results of the estimation and the counterfactuals.

1.5 Results and Counterfactuals

The following summarizes the estimated parameters, and the moments that are associated with their identification.

Figure 1.9 plots the values of the estimated $\hat{\lambda}_l$ against the total sales for all brands in each market l in the year of 2016.

The estimated model predicts a crucial role for distance in the strength of contagion. To see that, compare the two closest regions, D.C. and Baltimore, with the ones that are furthest away, Seattle and Miami. For the cities on the opposite side of the country, contagion is very unlikely. The contagion parameter between them is

Table 1.3: Estimated parameters and associated moments

Parameter	Value	Target
σ	0.1257	regression
θ	4.1753	$corr(x_{la}, x_{ma})$
δ_e	0.1560	F_{la}
δ_b	0.2925	$\mathbb{E}(b_{la})$
ϕ	0.2072	$\mathbb{V}(b_{la})$
C_0	0.5316	$corr(b_{la}, b_{ma})$
C_1	0.0015	$corr(b_{la}, b_{ma})$
λ_l	See Figure 1.9	$\mathbb{E}(b_{l0})$

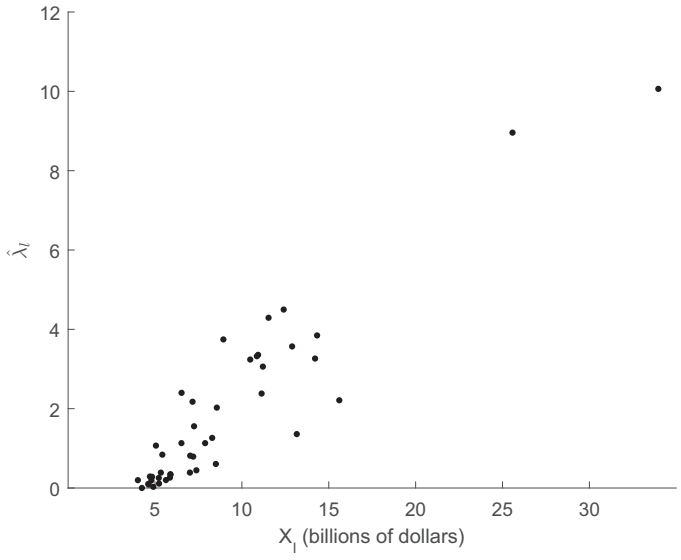

Figure 1.9: Estimated values of λ_l and total sales in location l

$C_0 * exp(-C_1 2730) = 0.0089$. The contagion parameter between the two cities in the middle of the east coast, on the other hand, is quite large $C_0 * exp(-C_1 35) = 0.5044$. On average, an informed consumer in D.C. spreads information to more than 0.5 consumers in Baltimore. This is more than 50 times the contagion probability between the two distant regions.

To further investigate how geography affects the outcome of brands in the model, I return to Figures 1.5 and 1.6. They describe how close regions tend to be more similar concerning sales and number of customers. I use the model simulated data to compute the same correlations and plot both the data and model variables for sales and customer base.

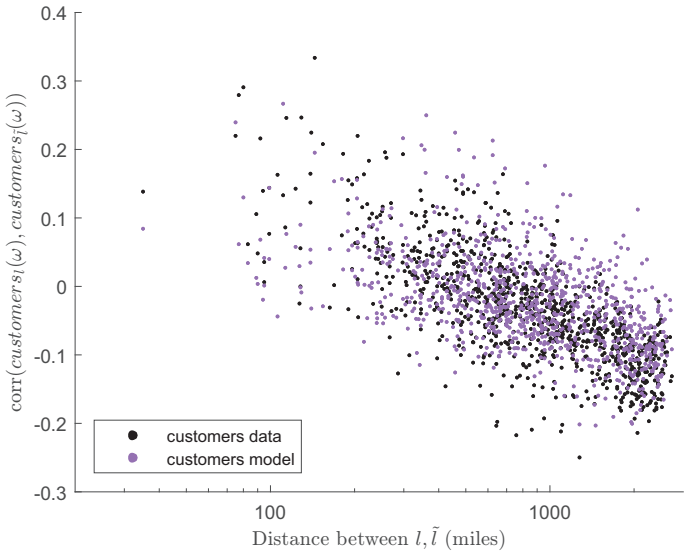

Figure 1.10: Correlation of brands relative customer base: data x model

We can see that the model performs quantitatively well in replicating the geographic concentration patterns observed in the data for sales and customer base. Furthermore, the model predictions for prices and individual quantities are not geographically correlated. The model also replicates the increasing trajectory for average

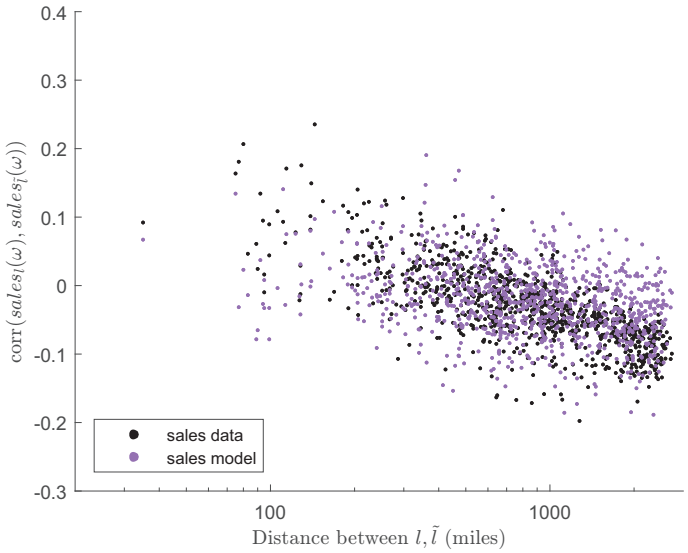

Figure 1.11: Correlation of brands relative sales: data x model

sales and customer base, although with higher levels.

I also evaluate the model predictions with respect to the geographic persistence, by computing the model-predicted correlations for sales and customer base as in Figure 1.7.

We can see that the model generates geographic persistence, but to a lesser degree than observed in the data. One reason behind that is the strong assumption imposed on the contagion process that only newly found potential customers might spread the information to others. As discussed before, this assumption reduces the influence that the current customer base has in acquiring customers in the future. Since contagion is more likely to happen in closer regions, this assumption reduces the degree of geographic persistence. Now, I discuss the counterfactuals, using the estimated model as the benchmark.

First, I evaluate the implications of reducing information frictions. I use the model to address the welfare gains of eliminating the role of geography in contagion.

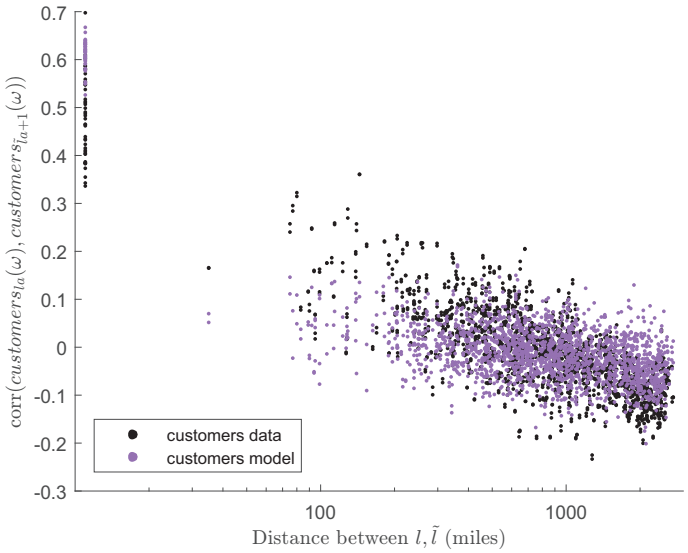

Figure 1.12: Correlation of brands relative customers one period ahead: data x model

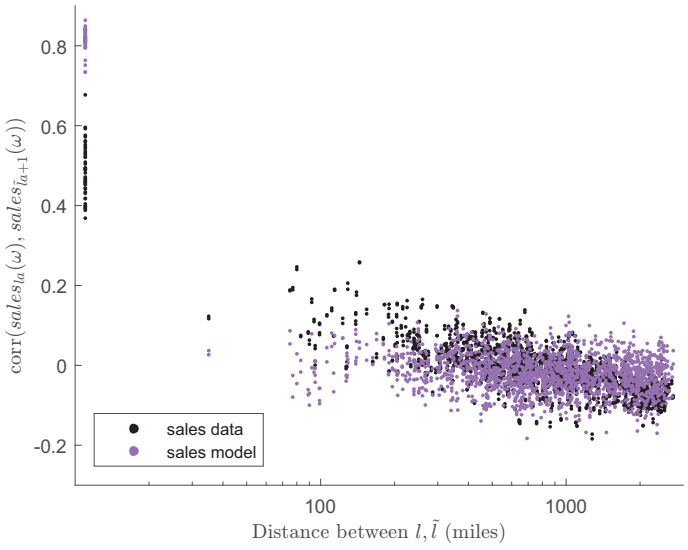

Figure 1.13: Correlation of brands relative sales one period ahead: data x model

In the benchmark, contagion parameters are defined as $\lambda_{lm} = C_0 exp(C_1 \text{distance}_{l,m})$. I assume that in the counterfactual economy, distance does not matter. There,

the contagion parameter between any two locations is $\tilde{\lambda}_{lm} = C_0 = 0.5316$. Notice that I compare two stationary economies, meaning that the counterfactual economy *always* had $\tilde{\lambda}_{lm}$ governing the evolution of brand awareness. The computation of new outcomes is straightforward since all equations derived before still hold for the new economy.

The welfare for consumers with CES utility function can expressed by their real income, which is $W_l = \bar{X}_l / P_l L_l$. Since the income and populations remain the same, our measure of welfare gain is $\Delta\tilde{W}_l = \tilde{W}_l / W_l - 1 = P_l / \tilde{P}_l - 1$. The price level in the new economy for all locations is

$$\tilde{P}_l = \tilde{\nu}_l^{-1/\theta} \left[\Gamma \left(1 - \frac{\sigma - 1}{\theta} \right) \right]^{\frac{1}{1-\sigma}},$$

where $\tilde{\nu}_l = \frac{\delta_e T_l}{L_l} \left(\sum_{a=0}^{\infty} \frac{(1-\delta_e)^a \tilde{\Lambda}_{l,a}}{1-\phi} \right)$.

The welfare gains are simply $\Delta\tilde{W}_l = \left(\frac{\sum_{a=0}^{\infty} (1-\delta_e)^a \tilde{\Lambda}_{la}}{\sum_{a=0}^{\infty} (1-\delta_e)^a \Lambda_{la}} \right)^{1/\theta} - 1$. The average welfare gains across all 44 locations is 32.5%, which is a large number and very likely to be dependent on the value of particular parameters. Moreover, some places are more affected by this change than others. While New York has the smallest increase in welfare, 18%, Portland's gains are the highest: 40%. In the counterfactual economy, information about the existence of brands circulates more. Therefore, consumers are more likely to find brands with lower prices, and consequentially increase consumption. Next, I evaluate the welfare gains of eliminating the effects of geography on costs.

In models where shipping costs among locations are explicit, this exercise can be conducted by reducing them to zero. Consider the case of a firm that has a cost c to produce and sell their good locally but face an iceberg cost of τ to deliver it to

another location. In this case, the effect of geography is eliminated by setting $\tau = 1$. This makes selling to other sites as expensive as selling domestically. The same logic is applied in this exercise.

The productivity vector $\mathbf{z}$ accounts for differences in the cost of selling to different locations. We can consider the situation where all brands have their costs reduced to their lowest z_l^{-1} in all areas. This way, each brand might sell its good to all locations at its lowest cost, similar to what eliminating shipping costs would do.

Again, I consider that the counterfactual economy is stationary, meaning that all brands faced the alternative distribution of costs since their beginning. As in the previous exercise, to calculate the welfare, I need to compute the new price indices P_l'. However, this is not straightforward anymore. Consider the productivity vector of a brand $z = (z_1, \ldots, z_{\mathcal{L}})$. In the counterfactual economy, its productivity vector would be $z' = (\overline{z}, \ldots, \overline{z})$, where $\overline{z} = \max(z_1, \ldots, z_{\mathcal{L}})$. The distribution of the maximum entry of the productivity vector does not have the same properties of the original marginal distribution. I briefly describe how I compute the new price distribution numerically.

First, I follow the same steps as the estimation procedure to draw the productivity vector for millions of brands. After that, I find the highest productivity entry for each brand. I assign a brand's cost in every location as the inverse of its highest value $\overline{z}_l$. Then, I estimate the density of the new cost distribution using a standard kernel function. After that, I multiply the estimated density function by the measure of brands simulated $\bar{K}$. The result is $d\mu'(c)$ analogous to $d\mu_l(c)$ in the original model. Here, $\mu'(c)$ is the measure of brands with the *lowest* cost below c. Notice that the brand's cost now does not depend on the location they sell to, but their ability to reach customers is still location-specific.

To compute the price distribution, I calculate the intensity that consumers in l

find quotes below c under the new cost distribution

$$\rho_l'(c) = \left[\sum_{a=0}^{\infty} \frac{\Lambda_{la}\delta_e(1-\delta_e)^a}{L_l(1-\phi)}\right] \int_0^c d\mu'(c).$$

The new price distribution is $G_l'(c) = 1 - \exp(-\rho_l'(c))$, and I can compute the new price indices as

$$P_l' = \left[\int_0^{\infty} p^{1-\sigma} dG_l'(p)\right]^{\frac{1}{1-\sigma}}.$$

Now, I can compute the welfare gains in each location $\Delta W_l'$. The average of the welfare gains across all locations is 57%. This is also a considerable improvement in consumption associated with reducing the costs of brands that consumers are already aware of. However, these estimates are potentially sensitive to the estimation of the demand elasticity and other parameters of the model.

The interpretation of the last counterfactual as the reduction of shipping costs requires another important caveat. Suppose that the model had a considerably larger number of locations, *i.e.* take $\mathcal{L}$ to be 1000. When the largest productivity is selected, the odds that any brand achieves a high number is substantial. Therefore, the cost reduction associated with geography is also conflated with increased expected productivity associated with more draws.

Nevertheless, the counterfactuals show that significant welfare gains are associated with reducing the impacts geography has on contagion and reducing the costs of brands that consumers already know.

1.6 Conclusion

This paper studies how brand sales evolve over time and space in the United States. I show evidence of a significant role for geography in the growth of brands. Unlike traditional spatial economics and trade models, these differences are not associated with changes in costs and prices.

To reconcile the data with economic modeling, I posit that geography can affect a brand's customer base through informational frictions that are not directly associated with changes in prices. I provide a parsimonious model of the geographic spread of brand awareness that relies on contagion: customers aware of a brand might infect unaware consumers with their knowledge. The model can replicate the stylized facts about how brand sales and customer base evolve. Furthermore, the estimates show that the proposed information frictions are more severe between distant locations. The model can also contrast the welfare costs associated with reducing the role of geography in prices and the flow of information. Finally, it can also evaluate how much expected revenue a single aware customer generates by considering the spread of this knowledge and the probability that consumers reached will effectively buy the product.

1.7 Appendix - Tables and Figures

Table 1.4: Summary Statistics for Locations in 2016

Location	# of Brands	Sales ($MM)	Location	# of Brands	Sales ($MM)
New York	28,771	33,967	Portland	17,664	7,161
Los Angeles	25,569	25,574	Orlando	17,952	7,024
Philadelphia	24,922	15,605	Richmond	17,879	7,009
Boston	22,731	14,328	Salt Lake City	16,317	6,528
Chicago	25,238	14,212	Sacramento	19,482	6,521
Washington, D.C.	23,094	13,158	Hartford	16,195	5,889
Dallas	22,151	12,890	Birmingham	16,317	5,886
San Francisco	20,017	12,406	Cincinnati	17,613	5,857
Miami	20,279	11,534	Indianapolis	16,398	5,630
Houston	21,592	11,207	Oklahoma City	15,025	5,421
Tampa	21,541	11,134	Nashville	15,933	5,336
Phoenix	21,615	10,933	Baltimore	16,893	5,230
Atlanta	21,050	10,862	St. Louis	17,696	5,217
Detroit	21,687	10,480	San Diego	15,591	5,054
Seattle	20,254	8,935	Grand Rapids	16,043	4,901
Denver	19,429	8,557	Columbus	17,492	4,827
Cleveland	20,047	8,502	Charlotte	17,453	4,793
Raleigh	18,840	8,290	Kansas City	16,236	4,707
San Antonio	18,908	7,884	Louisville	16,303	4,637
Pittsburgh	17,976	7,380	Buffalo	17,407	4,611
Minneapolis	19,695	7,247	Milwaukee	15,535	4,246
New Orleans	16,928	7,198	Memphis	13,514	4,014

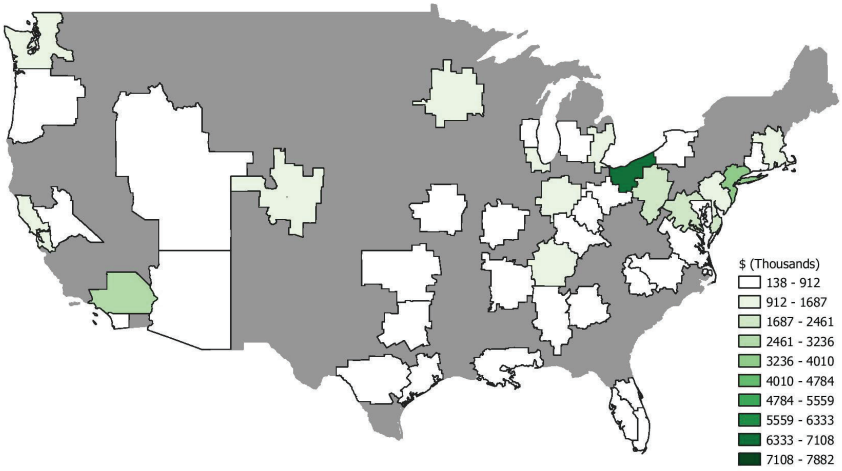

Figure 1.14: Sales of Brands that started in Cleveland - 2008

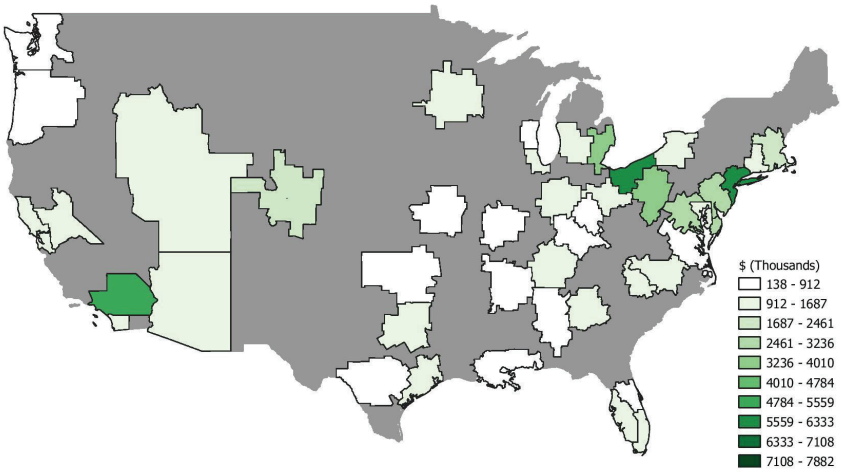

Figure 1.15: Sales of Brands that started in Cleveland - 2009

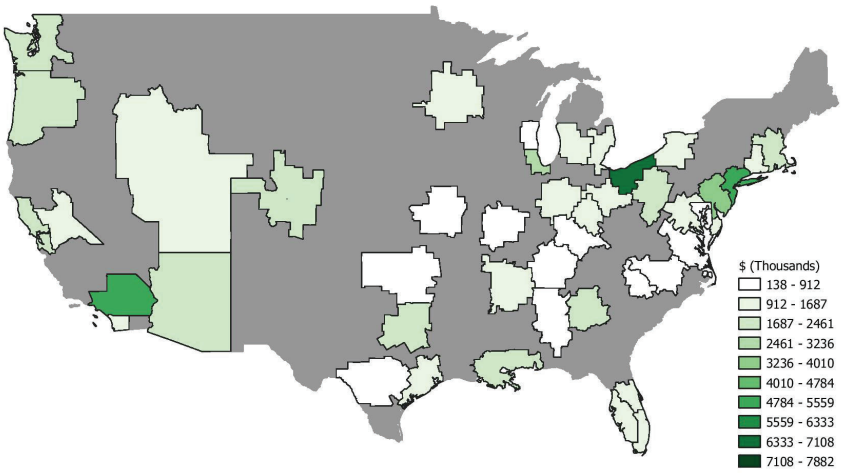

Figure 1.16: Sales of Brands that started in Cleveland - 2010

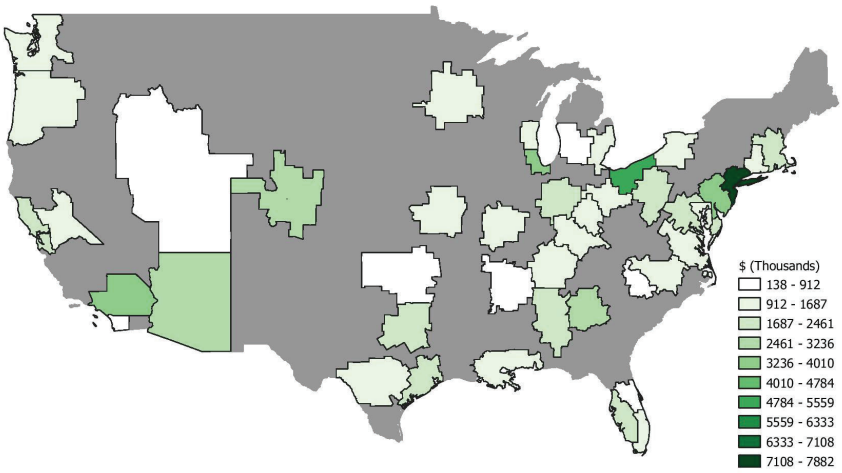

Figure 1.17: Sales of Brands that started in Cleveland - 2011

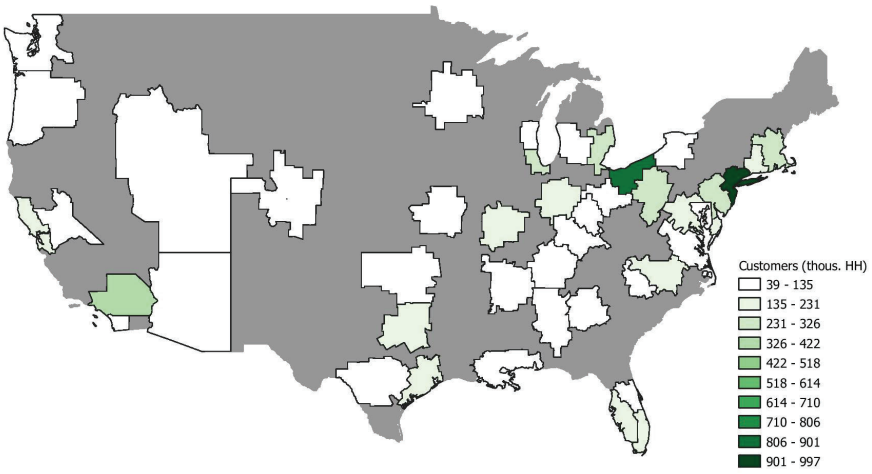

Figure 1.18: Customers of Brands that started in Cleveland - 2008

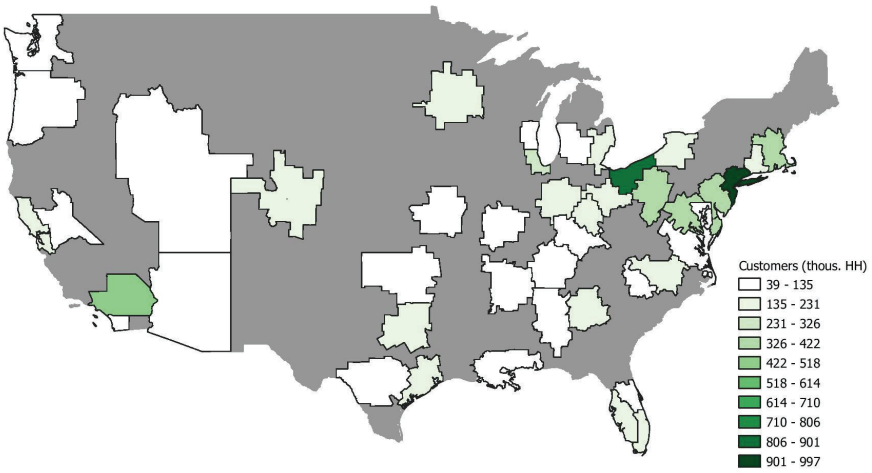

Figure 1.19: Customers of Brands that started in Cleveland - 2009

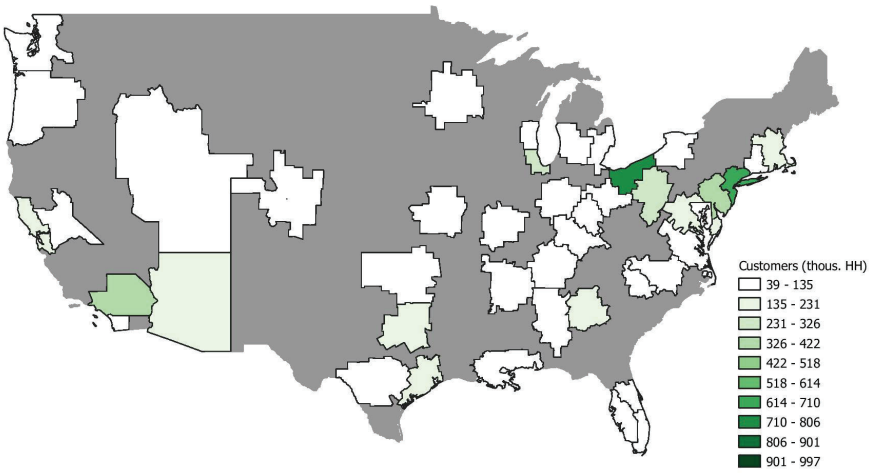

Figure 1.20: Customers of Brands that started in Cleveland - 2010

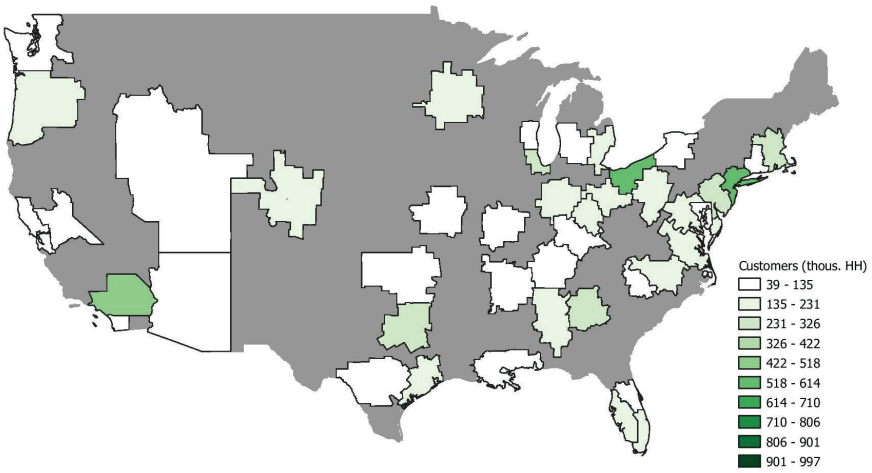

Figure 1.21: Customers of Brands that started in Cleveland - 2011

Chapter 2

Internal Gravity and Customer Base

2.1 Introduction

The gravity relation introduced by Tinbergen (1963) states that the bilateral trade flows between two countries are directly proportional to their size and inversely proportional to their distance. This empirical regularity was found extensively in International Trade data and became an important tool to understand how distance affects the flow of goods between different locations. In recent works, such as Allen and Arkolakis (2014) and Caliendo et al. (2019), this relation is also successful in describing the domestic trade and migration patterns within a country.

Traditionally, spatial economic models that produce this relation rely on the assumption that trade costs increase with distance (Anderson, 1979). As two locations are further away, it is more costly to ship goods between them. Consequently, it is

Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

less expensive to buy a particular good close to where it is produced than in distant locations, which generates an incentive for people in one place to consume more goods from nearby regions.

Recent evidence, however, points toward another way in which distance reduces trade. Eaton et al. (2019) and Chaney (2014) assume that information frictions make it harder for sellers to reach consumers in distant locations. This setting allows the authors to reproduce empirical regularities found on French exporters' data. The information frictions differ from physical trade costs since they generate other predictions for prices and entry. Therefore, measuring how much of the effect of distance on the gravity relation comes from physical trade costs instead of information frictions is significant for quantitative models of trade.

In this paper, I use brand-level data in different regions in the United States to separate the effects of both types of frictions. Using brand-level data is appropriate for studying information frictions between final consumers and firms because brands summarize information about products that allows consumers to identify them[1]. Moreover, studying the market for coffee, Draganska and Klapper (2011) shows that brand awareness varies across consumers and directly affects the probability that they buy particular products.

In this chapter, the relevant measure of distance is between the brand origin and their potential customers. This definition allows us to compare the effects of different types of costs that brands face to reach their consumers directly. It is important to highlight that this approach is fundamentally different from the previous chapter. The main element of the previous analysis was the possibility that contagion happened *among* consumers. Therefore, the relevant measures of distance there are

[1]According to the American Marketing Association dictionary "A brand is a name, term, design, symbol or any other feature that identifies one seller's good or service as distinct from those of other sellers."

among potential consumers in different regions.

I use the Nielsen Homescan panel data to construct a data set of more than 150,000 unique brands, their sales, and customer base in 44 different regions in the United States. The data provides novel insights into the effects of distance on sales of retail products. At the brand level, distance greatly affects brands' customer base, but it doesn't reduce individual sales by the same amount.

Furthermore, I also investigate the aggregate sales of all brands from a particular region. I can decompose the effect of distance on aggregate sales into the number of brands, average customer base, and average individual sales. For the aggregate bilateral sales, the number of brands that enter each market is the most important component that explains the decline in sales as distance increases. Furthermore, due to selection, the average customer base of brands that sell to distant locations is actually higher than the average brand that sells locally. Finally, these effects are more drastic when distances are small, which contrasts with the original gravity relation.

To separate the effects of distance due to physical trade costs and information frictions, I use a spatial model of trade between regions. Brands face shipping costs to deliver their goods to other areas and are also subject to information frictions. I assume that final consumers only buy products from brands that they know. Consumers and brands are matched using a standard matching function, where the intensity of these matches depends on the brand's and the consumer's location.

I calibrate the model to match the data for the year of 2016. The results show that the physical trade costs are increasing with distance and the information friction parameters are lower at the origin, but are otherwise not affected by distance.

2.2 Empirical Evidence

The Nielsen Homescan panel data surveys around 50,000 households every year and contains detailed information about the retail products they buy. The survey also assigns weights to households based on their demographics to make the sample representative. Between the years 2007-2016, there are 44 regions in the USA with a representative sample. I measure a brand's total sales in each of these markets by computing the weighted sum of the total value of products from that brand sold to households in that location. Furthermore, I also construct a measure of customer base by counting the number of households that bought products from each brand, again using the weights provided by Nielsen.

There are 150,206 different brands in the data. However, they do not have an origin assigned. To circumvent this issue, I look at all the brands that I observe their entry period, which could be anytime between 2008-2016, and assign their location as the first market that they serve. This assumption is in line with most spatial economic models that predict that firms start selling locally before reaching out to other regions.

Out of all the brands in the sample, I can assign an origin to 43% of them. However, due to the selection of younger and potentially smaller brands, they correspond to only 4% of the total sales in all years. On average, brands with origin assigned serve 7.38 markets, as opposed to 11.24 in the full sample.

The inclusion of new cohorts mitigates this problem. For 2016, I observe brands with ages ranging from 0 to 7 years. In this year, the brands with origin account for 7% of total sales. For this reason, I choose to calibrate the model targeting the data moments of this year. But before that, I present some regressions that describe the effects of distance on brand-level sales and aggregate flows considering all the years

in the sample.

Let $x_{ij}(\omega)$ denote the total sales of brand ω, with origin i in market j in a given year. I compute the total bilateral sales between i and j in that year (X_{ij}), by adding all the sales in j of brands located in i:

$$\underbrace{X_{ij}}_{\text{Total Sales}} = \sum_{\omega \in \Omega_{ij}} \underbrace{x_{ij}(\omega)}_{\text{Sales of Brand } \omega} \quad .$$

The first assessment evaluates how distance affects the aggregate bilateral flows. To do so, I regress the log of the bilateral flows on the log of distance and dummies that cover each combination of origin and destination in different periods. This is the same setting as the traditional gravity relation, and if the magnitudes are similar the coefficient $\hat{\beta}$ should be equal to -1.

Table 2.1: Gravity relation
$\log(X_{ijt}) = \alpha + \beta \log(\text{dist}_{ij}) + \gamma_{it} + \delta_{jt} + \epsilon_{ijt}$

	log(total_sales)	
	Full Sample	2016
log(distance)	-0.196^{***}	-0.315^{***}
	(0.002)	(0.011)
Observations	17,423	1,936
R^2	0.938	0.804
Adjusted R^2	0.935	0.795
Residual Std. Error	0.314 (df = 16640)	0.548 (df = 1848)
Note:		*p<0.1; **p<0.05; ***p<0.01

We can see that distance is negatively associated with the bilateral flows between two locations. However, the coefficients' magnitude is lower than the traditional gravity relation, which would be -1. This could be because the sample selects mostly

young and small brands. To mitigate this, I also run the same regression considering only the year 2016. The point estimate goes from -0.196 to -0.315, which is still lower than the evidence from the international trade literature.

For the next analyses, I consider a set of equally spaced dummies between locations instead of the log of distance as an explanatory variable. In this setting, I use the local sales as reference, where distance would be zero otherwise. The reason for this choice is to consider the possibility that the effect of distance on the flow of retail goods varies. For example, it can be the case that it doesn't matter whether two locations are 1500 or 2000 miles apart, but it could still make a difference if they are separated by 15 or 20 miles.

I also use the measure of each brand's customer base to decompose their sales in a given location by the number of customers and the individual sales. The sales of brand ω from j in market i can be decomposed as

$$x_{ij}(\omega) = \underbrace{b_{ij}(\omega)}_{\text{Customers}} \quad \underbrace{v_{ij}(\omega)}_{\text{Indiv. Sales}} .$$

I regress the brand-level sales and their components on the set of distance dummies, brand-specific dummies and destination specific dummies. The next table summarizes the result, and the next figure plots the coefficients for sales and customer base for each distance bin.

There are two critical takeaways from this evaluation. First, the drop between selling locally and at other locations is quite sharp, and the effects of distance further fade away as it increases. Second, the decomposition shows that 80%of the effect of distance on sales is due to the brand having a smaller customer base in distant locations, not by reducing individual sales.

Table 2.2: Effect of distance on brand-level variables
$\log(x_{ijt}(\omega)) = \alpha + \sum_{k=1}^{10} \beta_k D_k(dist_{ij}) + \gamma_t(\omega) + \delta_{jt} + \epsilon_{ijt}(\omega)$

	log(sales)	log(indiv_sales)	log(customers)
distance bin 1	-0.297^{***}	-0.054^{***}	-0.242^{***}
	(0.006)	(0.003)	(0.005)
distance bin 2	-0.367^{***}	-0.066^{***}	-0.301^{***}
	(0.006)	(0.003)	(0.005)
distance bin 3	-0.409^{***}	-0.072^{***}	-0.336^{***}
	(0.006)	(0.003)	(0.005)
distance bin 4	-0.443^{***}	-0.079^{***}	-0.364^{***}
	(0.006)	(0.003)	(0.005)
distance bin 5	-0.479^{***}	-0.080^{***}	-0.399^{***}
	(0.007)	(0.003)	(0.006)
distance bin 6	-0.493^{***}	-0.082^{***}	-0.411^{***}
	(0.007)	(0.003)	(0.006)
distance bin 7	-0.510^{***}	-0.078^{***}	-0.432^{***}
	(0.007)	(0.003)	(0.006)
distance bin 8	-0.540^{***}	-0.081^{***}	-0.459^{***}
	(0.007)	(0.003)	(0.006)
distance bin 9	-0.523^{***}	-0.085^{***}	-0.438^{***}
	(0.007)	(0.003)	(0.007)
distance bin 10	-0.543^{***}	-0.082^{***}	-0.461^{***}
	(0.009)	(0.004)	(0.008)
Observations	1,389,225	1,389,225	1,389,225
R^2	0.567	0.724	0.533
Adjusted R^2	0.525	0.697	0.488

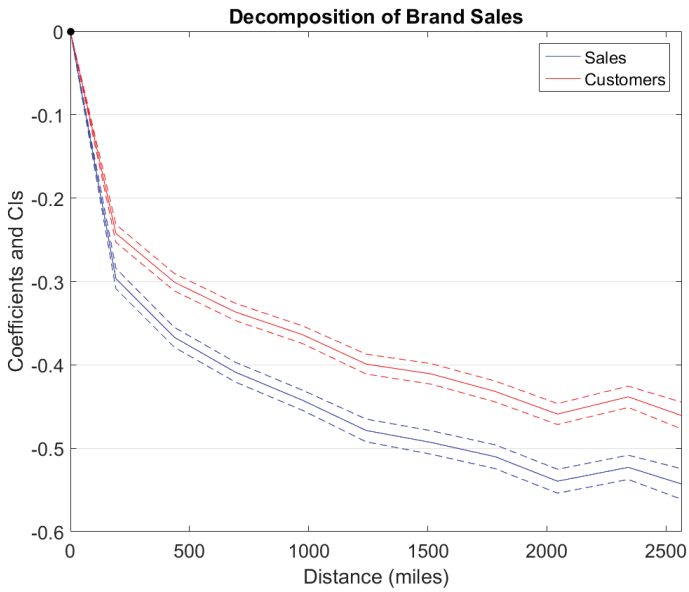

Figure 2.1: Effect of distance on brand-level variables

The total sales can also be decomposed as:

$$X_{ij} = \underbrace{N_{ij}}_{\text{Numb. brands}} \underbrace{\left(\frac{\sum_\omega b_{ij}(\omega)}{N_{ij}}\right)}_{\text{Avg. Cust.}} \underbrace{\left(\frac{\sum_\omega b_{ij} v_{ij}(\omega)}{\sum_\omega b_{ij}(\omega)}\right)}_{\text{Avg. Indiv. Sales}}.$$

Notice that now there are three elements in the decomposition: the number of brands, their average customer base, and average individual sales. I regress all those variables in a set of dummies for distance and destination and origin fixed effects. The next table summarizes the results.

As expected there is not much action in the individual sales. It is interesting though, that now the effect comes from the number of brands that serve each market. Curiously, the average number of customers is larger for destinations that are not the origin. The reason for this is because brands that only sell locally tend to be quite smaller than the ones that reach to other regions. Since the aggregate regressions

Table 2.3: Effect of distance on aggregate variables
$\log(X_{ijt}) = \alpha + \sum_{k=1}^{8} \beta_k D_k(dist_{ij}) + \gamma_{it} + \delta_{jt} + \epsilon_{ijt}$

	log(total_sales)	log(numb_brands)	log(avg_customers)	log(avg_indiv_value)
distance bin 1	-0.984^{***}	-1.226^{***}	0.384^{***}	-0.143^{***}
	(0.017)	(0.008)	(0.010)	(0.013)
distance bin 2	-1.132^{***}	-1.399^{***}	0.404^{***}	-0.138^{***}
	(0.017)	(0.008)	(0.010)	(0.013)
distance bin 3	-1.241^{***}	-1.475^{***}	0.372^{***}	-0.138^{***}
	(0.017)	(0.008)	(0.010)	(0.013)
distance bin 4	-1.311^{***}	-1.534^{***}	0.363^{***}	-0.140^{***}
	(0.018)	(0.009)	(0.011)	(0.013)
distance bin 5	-1.408^{***}	-1.593^{***}	0.346^{***}	-0.161^{***}
	(0.018)	(0.009)	(0.011)	(0.014)
distance bin 6	-1.468^{***}	-1.647^{***}	0.336^{***}	-0.157^{***}
	(0.018)	(0.009)	(0.011)	(0.014)
distance bin 7	-1.515^{***}	-1.676^{***}	0.321^{***}	-0.159^{***}
	(0.019)	(0.009)	(0.011)	(0.014)
distance bin 8	-1.561^{***}	-1.665^{***}	0.289^{***}	-0.185^{***}
	(0.021)	(0.010)	(0.013)	(0.016)
Observations	17,423	17,423	17,423	17,423
R^2	0.938	0.972	0.842	0.540
Adjusted R^2	0.935	0.971	0.835	0.518

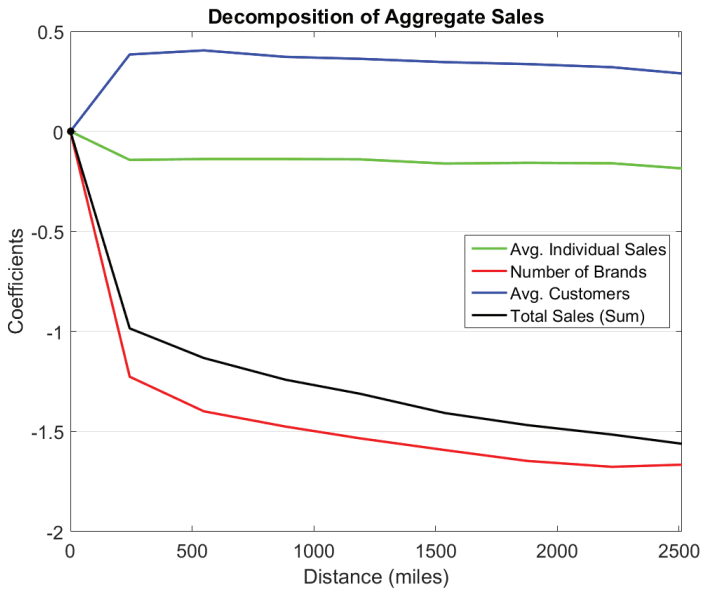

Figure 2.2: Effect of distance on aggregate variables

lump all those brands together by the distance of the market that they serve, we find that, on average, the customer base of brands increase as we consider distant locations.

These findings point towards the customer base of brands as the main driver of their sales in different locations, and decision to enter in different markets. In the next section, I propose a model that rationalize these findings, and is able to separate between the physical trade frictions that affect the cost of delivering goods and information frictions that might affect how many people actually know the brand in different locations.

2.3 Model and Results

Consider an economy with $i = 1, 2, \ldots, N$ locations. There is a continuum of varieties $j \in [0, 1]$ of final goods. Those goods are produced using only labor as a production factor, and each location i is endowed with L_i workers that supply one unit of labor

inelastically at the equilibrium wage w_i.

Technology

For a given variety j, location i is able to produce it using the linear production function $q = z_i(j)l_i$. The measure of potential producers from i that have efficiency $z_i(j) \geq z$ is $\mu_i^Z(z) = T_i z^{-\theta}$. There is an iceberg cost $\tau_{ni} \geq 1$ of delivering one unit of a good from i to n. For a given productivity level $z_i(j)$, the unit cost of delivering a good to destination n is $\frac{w_i \tau_{ni}}{z_i(j)}$. Therefore, the measure of firms in i that are able to deliver a good to n at a unit cost below c is $\mu_{ni}(c) = T_i(w_i \tau_{ni})^{-\theta} c^{\theta}$.

Consumers

Each consumer in location n has a CES utility function over all the varieties j

$$U(q) = \left[\int_0^1 q(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}.$$

There is perfect substitution within a given variety j, which implies that the consumer will choose the supplier that offers him the lowest price. However, the consumer is not aware of all the suppliers of a particular variety. The knowledge about the suppliers is subject to a matching friction between firms and consumers.

Matching

The number of matches that happen between consumers from n and firms from i with cost to deliver there below c is

$$M_{ni}(c) = \frac{\lambda_{ni} L_n \mu_{ni}(c)}{1-\gamma} L_n^{1-\varphi} S_n(c)^{-\gamma}, \quad \varphi, \gamma \geq 0,$$

where $S_n(c) = \sum_i \lambda_{ni}\mu_{ni}(c)$ are all the potential sellers. Summing across all sources, we have that the number of matches in location n is:

$$M_n(c) = \sum_{i=1}^{N} M_{ni}(c) = \frac{1}{1-\gamma} L_n^{1-\varphi} S_n(c)^{1-\gamma}.$$

Consistent with this matching function, we have that the intensity with which a buyer in location n encounters a seller from i with unit cost equal to c is

$$\lambda_{ni}(c) = \lambda_{ni} L_n^{-\varphi} S_n(c)^{-\gamma}$$

Aggregating over all potential suppliers from each source and cost below c implies that the number of quotes that a buyer gets below that cost level follows a Poisson distribution with parameter

$$\begin{aligned} \rho_n(c) &= \sum_{i=1}^{N} \int_0^c \lambda_{ni}(c) d\mu_{ni}(c) = L_n^{-\varphi} \int_0^c S_n(c')^{-\gamma} \sum_{i=1}^{N} \lambda_{ni} d\mu_{ni}(c') \\ &= L_n^{-\varphi} \int_0^c S_n(c')^{-\gamma} dS_n(c') = \frac{L_n^{-\varphi} S_n(c)^{1-\gamma}}{1-\gamma}. \end{aligned}$$

Notice that we can write

$$\begin{aligned}\rho_n(c) &= \frac{L_n^{-\varphi}}{1-\gamma}\left[\sum_{i=1}^{N}\lambda_{ni}\mu_{ni}(c)\right]^{1-\gamma} \\ &= \frac{L_n^{-\varphi}}{1-\gamma}\left[\sum_{i=1}^{N}\lambda_{ni}T_i(w_i\tau_{ni})^{-\theta}c^{\theta}\right]^{1-\gamma} = \frac{L_n^{-\varphi}}{1-\gamma}\left[\Upsilon_n c^{\theta}\right]^{1-\gamma} \\ &= \nu_n c^{\theta(1-\gamma)},\end{aligned}$$

where $\Upsilon_n = \sum_{i=1}^{N}\lambda_{ni}T_i(w_i\tau_{ni})^{-\theta}$, and $\nu_n = \frac{L_n^{-\varphi}}{1-\gamma}\Upsilon_n^{1-\gamma}$.

Therefore, the probability that a buyer encounters no quotes below c is given by $\exp(-\rho_n(c)) = \exp(-\nu_n c^{\theta(1-\gamma)})$. Therefore, the distribution of the lowest costs in location n has the following C.D.F.

$$G_n(c) = 1 - \exp\left(-\nu_n c^{\theta(1-\gamma)}\right).$$

which is the the equilibrium distribution of prices in location n.

Bilateral Trade Shares and Price distributions

The distribution of quotes below c that a buyer in location n receives from a particular location i also follows a Poisson distribution, with parameter

$$\rho_{ni}(c) = \frac{M_{ni}(c)}{L_n} = \frac{\lambda_{ni}\mu_{ni}(c)L_n^{-\varphi}S_n(c)^{-\gamma}}{1-\gamma} = \nu_{ni}c^{\theta(1-\gamma)}.$$

Hence, the best distribution of the lowest quotes that location i offers to n is

$$G_{ni}(c) = 1 - \exp\left(-\nu_{ni} c^{\theta(1-\gamma)}\right),$$

where $\nu_{ni} = \lambda_{ni} T_i (w_i \tau_{ni})^{-\theta} \frac{L_n^{-\varphi} \Upsilon_n^{-\gamma}}{1-\gamma}$. The probability that a supplier in n comes from i is $\pi_{ni} = \int_0^\infty \prod_{s \neq i} [1 - G_{ns}(c)] dG_{ni}(c)$. Solving this integral generates the following result

$$\pi_{ni} = \frac{\nu_{ni}}{\nu_n} = \frac{\lambda_{ni} T_i (w_i \tau_{ni})^{-\theta}}{\Upsilon_n}.$$

Moreover, the distribution of prices in n coming from i of goods that are actually bought there is the same as the price distribution in that location, since we can show that $G_n(c) = \frac{1}{\pi_{ni}} \int_0^c \prod_{s \neq i} [1 - G_{ns}(q)] dG_{ni}(q)$. This imply that the bilateral trade share is also given by $\frac{X_{ni}}{X_n} = \pi_{ni}$.

Furthermore, we can use the price distribution in n to compute its price index as

$$P_n = \left[\int_0^1 p(j)^{1-\sigma} di\right]^{\frac{1}{1-\sigma}} = \left[\int_0^\infty p^{1-\sigma} dG_n(p)\right]^{\frac{1}{1-\sigma}},$$

using the Moment Generating Function for G_n we have that

$$P_n = \nu_n^{-\frac{1}{\theta(1-\gamma)}} \Gamma\left(1 - \frac{\sigma - 1}{\theta(1-\gamma)}\right)^{\frac{1}{1-\sigma}}.$$

Here, we see that in order to have a well defined price index we need to impose the parameter restriction that $\theta(1-\gamma) > \sigma - 1$.

A balanced Equilibrium can be found by finding the wages and bilateral trade shares that satisfy

$$w_i L_i = \sum_{n=1} \pi_{ni} w_n L_n$$

and $\pi_{ni} = \frac{\lambda_{ni} T_i (w_i \tau_{ni})^{-\theta}}{\sum_{k=1}^{N} \lambda_{nk} T_k (w_k \tau_{nk})^{-\theta}}$

This imply that the equilibrium wages and bilateral flows are only functions of the trade frictions τ_{ni}, λ_{ni}, population L_i, Technology T_i and shape parameter θ. Other important equilibrium objects (such as the Price Indices) that are important for welfare analysis will depend on σ, γ, φ]

Implications for Individual Producers

A firm operates as long as it has at least one buyer. The number of buyers that firms have do not depend only on their productivity, but also on their luck in finding consumers. For a firm in i that has cost c of delivering to n, the number of encounters it has with consumers from that location is Poisson distributed, with parameter

$$e_{ni}(c) = L_n \lambda_{ni}(c), \quad \lambda_{ni}(c) = \lambda_{ni} L_n^{-\varphi} S_n(c)^{-\gamma}.$$

Having met a buyer, they make a sale with probability $\exp(-\nu_n c^{\theta(1-\gamma)})$. Combining the two results, the number of customers in n buying a variety from supplier in i with unit cost of delivering to n equal to c is distributed following a Poisson distribution with parameter $\eta_{ni}(c)$ given by

$$\eta_{ni}(c) = \lambda_{ni} L_n^{1-\varphi} [S_n(c)]^{-\gamma} \exp(-\nu_n c^{\theta(1-\gamma)}).$$

this is also the expected number of buyers on that location.

Now consider a firm from i that can produce a good and deliver it domestically at cost c. Their cost of delivering the good to location n is $c\tau_{ni}$ and consequentially their total customer base is also Poisson distributed with parameter

$$\eta_i^W(c) = \sum_{n=1}^{N} \eta_{ni}(c\tau_{ni}).$$

Therefore, the probability that a potential producer from with domestic unit cost c has at least 1 buyer anywhere is $1 - \exp(-\eta_i^W(c))$. Furthermore, the measure of firms that operate from i is

$$F_i = \int_0^\infty \left[1 - \exp\left(-\eta_i^W(c)\right)\right] d\mu_{ii}(c)$$

Measure of sellers

The measure of firms from i selling to n can be found by

$$N_{ni} = \int_0^\infty \left[1 - \exp\left(-\eta_{ni}(c\tau_{ni})\right)\right] d\mu_{ii}(c).$$

The total measure of suppliers that serve n is therefore $N_n = \sum_{i=1}^{N} N_{ni}$.

Measure of Relationships

Let R_{ni} be the total number of seller-buyer connections from i to n, we have that

$$\begin{aligned} R_{ni} &= \int_0^\infty \eta_{ni}(c) d\mu_{ni}(c) = \\ &= \frac{T_i (w_i \tau_{ni})^{-\theta} \lambda_{ni} L_n^{1-\varphi} \Upsilon_n^{-\gamma}}{(1-\gamma)\nu_n} = \frac{\nu_{ni}}{\nu_n} L_n = \pi_{ni} L_n. \end{aligned}$$

Also, the average number of customers in n from firms in i is

$$\bar{b}_{ni} = \frac{R_{ni}}{N_{ni}}.$$

Individual Sales

Consider a firm from i that has a cost of delivering domestically equal to c. If they are matched with a consumer in n that buys their product, the expected value of their sales is going to be

$$x_{ni}(c) = X_n P_n^{\sigma-1} p_{ni}(c)^{1-\sigma} = X_n P_n^{\sigma-1} (c\tau_{ni})^{1-\sigma}$$

Notice how this is decreasing in the iceberg trade cost, but it is not affected by the informational friction.

Calibration

In order to recover the parameters of the model, I use only the data from 2016, since the selection problems are less severe for the last year in the sample.

I directly observe the population of each location, and take the shape parameter of the productivity distribution to be $\theta = 4.87$. This is the same value for the shape parameter in Eaton et al. (2011), for the distribution of productivities in the French manufacturing sector.

With that in mind, the set of parameters we need to calibrate are $(\sigma, \gamma, \varphi, \boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{T})$. The algorithm that identifies each parameter is divided in three parts. The first part pins down $(\varphi, \gamma, \boldsymbol{\lambda})$, using data on average customers from each pair of origin and destination. We find $\tilde{\lambda}_{ni} := \lambda_{ni} L_n^{\frac{(1-\varphi)}{(1-\gamma)}}$ that is able to precisely match the model's prediction for the average number of buyers that brands from i have in destination n, jointly with γ that matches the variance of the number of buyers. After that, if we assume that the population size is orthogonal to the bilateral informatoin frictions λ_{ni}, we can regress $\log(\tilde{\lambda}_{ni})$ on $\log(L_n)$, so that the regression coefficient is an estimate of $\frac{1-\varphi}{1-\gamma}$, and the error is an estimate of λ_{ni}.

The second part of the calibration exploits the data on the aggregate bilateral flows and expenditure X_n. We normalize $\tau_{ii} = 1$ for every location, and using the system of equations that pins-down the bilateral flows and equilibrium wages I can find the τ_{ni} and T_i that match the bilateral flows and the total expenditure.

The parameter σ is the one that remains to be estimated. I use the sales per costumer of brands that sell to more than one location to pin it down. Noticing that $\frac{x_{ni}(c)}{x_{ii}(c)} = \left(\frac{X_n}{X_i} \frac{P_n^{\sigma-1}}{P_i^{\sigma-1}}\right) \tau_{ni}^{1-\sigma}$, I regress $\log(x_{ni}(c)/x_{ii}(c))$ on $\log(\tau_{ni})$, including origin and destination dummies. The estimated coefficient provides an estimate of $1 - \sigma$.

Algorithm

Module 1

Let $\tilde{\lambda}_{ni} := \lambda_{ni} L_n^{1-\frac{(\varphi)}{(1-\gamma)}}$. Our first goal is to write the average number of buyers from a brand that is from i and n as a function of γ, $\tilde{\lambda}_{ni}$ and observables. The average number of buyers in$n$ of brands in i is $\bar{b}_{ni} = \frac{R_{ni}}{N_{ni}}$, where we know that $R_{ni} = \pi_{ni} L_n$ and $N_{ni} = \int_0^\infty [1 - \exp(-\eta_{ni}(c\tau_{ni}))]\, d\mu_{ii}(c)$. We can define $y = T_i w_i^{-\theta} c^\theta$ and change the variables of integration, so we get that

$$\tilde{\eta}_{ni}(y) := \eta_{ni}(c\tau_{ni}) = \tilde{\lambda}_{ni} \left[\frac{\tilde{\lambda}_{ni} y}{\pi_{ni} L_n}\right]^{-\gamma} \exp\left[\frac{-1}{(1-\gamma)} \left[\frac{\tilde{\lambda}_{ni} y}{\pi_{ni} L_n}\right]^{1-\gamma}\right]$$

and

$$\bar{b}_{ni} = \frac{\pi_{ni} L_n}{\int_0^\infty [1 - \exp(-\tilde{\eta}_{ni}(y))]\, dy}.$$

Given γ, I can find the matrix of $\tilde{\lambda}_{ni}$ that match what we observe in the data for b_{ni}, π_{ni} and L_n.

After that, I compute the variance of b_{ni}. I choose the set $\{\gamma, \tilde{\lambda}(\gamma)\}$ that makes the model generated variance of b_{ni} for all pairs of locations equal to their data counterparts.

In order to recover φ, I make an identification assumption that λ_{ni} is orthogonal to L_n. I regress $\ln \lambda_{ni}$ on a full set of origin and destination fixed effects. Then I regress the destination effects on $\ln L_n$ to recover $\hat{\beta} = 1 - \varphi/(1-\gamma)$. Hence, $\varphi = (1 - \hat{\beta})(1 - \gamma)$ and I can finally recover the information friction parameters as

$$\lambda_{ni} = \frac{\tilde{\lambda}_{ni}}{L_n^{\hat{\beta}}}.$$

Module 2

I assume that $\tau_{ii} = 1$ for all locations. This implies that there are no shipping costs associated with selling locally. I then choose the remaining iceberg costs and productivity shifters to match

$$\frac{\pi_{ni}}{\pi_{nn}} = \frac{\lambda_{ni} T_i (w_i \tau_{ni})^{-\theta}}{\lambda_{nn} T_n w_n^{-\theta}}, \quad w_i = \frac{\bar{X}_i}{L_i}.$$

Module 3

The last step is to recover the parameter σ. To do so, I use the model predictions for the individual sales of the same brand in different destinations. I regress $\ln(v_{ni}(\omega)/v_{ii}(\omega))$ on a set of origin and destination factors, and the estimated $\ln \tau_{ni}$. The OLS coefficients represent an estimate for $\sigma - 1$.

Results

The following table summarizes the estimated coefficients.

The congestion parameters found in the estimation suggest that congestion on the side of the seller is not very important, but it is significant on the buyers side. In the model, this means that the number of matches that a brand finds in a location is almost independent of that location population. However, competition matters a lot. The more sellers available in a location, the harder it is for a brand to be

Table 2.4: Estimated coefficients

Description	Parameter	Value
Efficiency Het.	θ	4.87
Congestion Seller	γ	0.0119745
Congestion Buyer	φ	1
Matching Friction	λ_{ni}	Various
Iceberg Costs	τ_{ni}	Various
Technology Shifter	T_n	Various
Elast. Subst.	σ	1.165

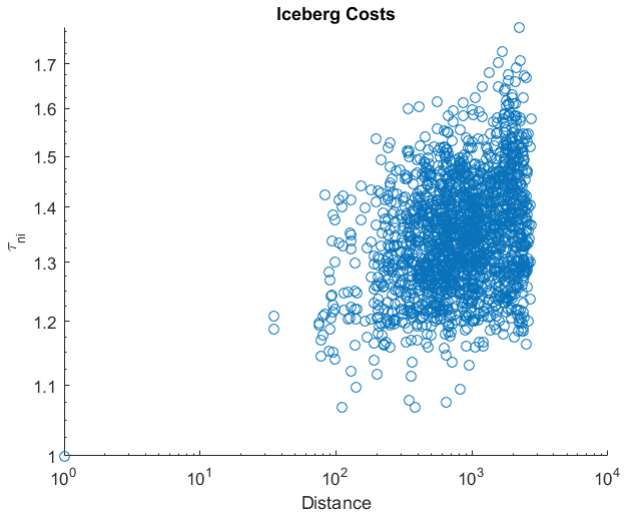

Figure 2.3: Physical frictions by distance

matched there. This is an interesting result and, since the parameters are close to their boundaries, they indicate that another matching function might have a better fit.

Furthermore, the elasticity of substitution is also quite low. This is expected, because these parameters are estimated using regressions on the individual values for the same brands in different locations. As we have seen in the empirical analysis that these quantities do not change much for different locations.

In the next figures I plot the estimated physical and information frictions between all two pairs of locations, by the distance between them.

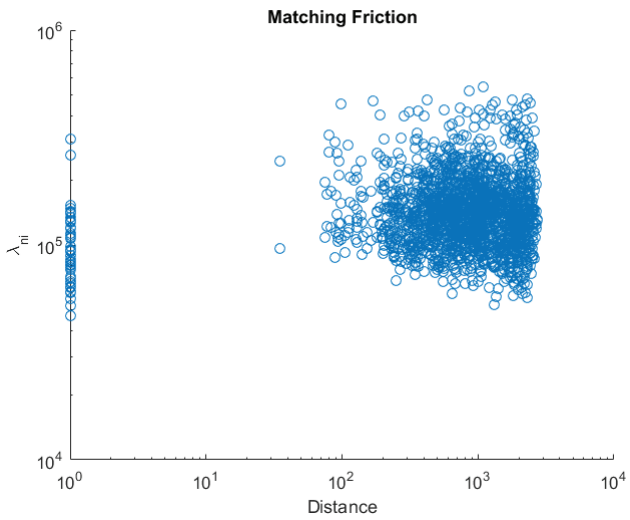

Figure 2.4: Information frictions by distance

Both plots show that there is a lot of variation with respect to those estimates, and distance cannot fully predict the iceberg costs and information frictions. Nevertheless, the results for the iceberg costs are expected: two locations that are further away tend to have higher transportation costs.

The result for the information frictions, on the other hand, are quite puzzling. They suggest that, on average, brands find it easier to connect with consumers in locations that are different than their origin.

2.4 Conclusion

This paper introduces new findings concerning the role of distance for brand sales in the United States. Here, I show that the reduction in brand sales associated with increasing distances is primarily due to the decrease in their number of customers at the brand-level. At the aggregate level, however, this reduction comes from entry: there are fewer brands that serve distant locations.

To understand what drives these patterns, I propose a model of trade among dif-

ferent locations that includes two types of frictions. The first one is the traditional iceberg cost. If a brand is selling to an area that differs from the origin, they incur a physical cost. The second friction is informational. Buyers are not aware of all the brands, and they only buy from the ones they know. The set of brands that a consumer is aware of is determined by random matching with brands from all locations. Furthermore, the intensity in which buyers and brands are matched depends on their location.

The estimation shows that, as expected, iceberg costs are increasing with distance. However, no clear relationship between distance and the intensity of the information frictions is found. This might suggest that the direct flow of information between brands and consumers is less dependent on geography than the cost of shipping goods.

Chapter 3

Geographic Spillovers and Exporter's Growth

3.1 Introduction

There is a vast literature documenting the dynamic of firms in the context of both closed and open economies. The subject is relevant since many decisions that ultimately shape aggregate economic variables are taken inside the firm, such as buying inputs, hiring, and investing in capital, innovation, and marketing.

In particular, firms that are part of a global economy choose which foreign countries that they serve, and their pricing and quantity strategy for their products in each destination. These decisions are not trivial since different countries are subject to different juridic systems, use their own currencies, have potential language barriers, and trade costs associated with distances that often surpass the ones of selling domestically.

Since the previously described features of selling to a particular foreign market should be reasonably similar across all the firms from the domestic country, most

of the research on firm dynamics in open economies focuses on the evolution of the productivity of firms as being the only driving force at the firm level[1]. This is a natural approach, but it misses a few empirical regularities, that could only be captured by assuming other types of heterogeneity across firms.

The fact that firms can have different paths in terms of the countries that they serve has been addressed by a few recent papers. In Morales et al. (2019), the authors use Chilean customs data and observe a path dependency in the countries that a firm serves: they are more likely to enter markets that are similar to the ones that they are already serving. In that context, similar means that they share common characteristics, such as official language, per-capita GDP, border, or lay in the same continent. They estimate a structural model that attributes these differences to lower sunk costs in entering destinations that are similar to the ones that a firm already supplies to.

The role of distance in establishing particular paths for firms' destinations has also been scrutinized by Chaney (2014). In that paper, he proposes a network model in which firms search for costumers in locations that are distributed in a geographical space. Firms are then more likely to find customers that are located close to the ones that they already sell to. Using French firm-level data, the model is simulated and it can rationalize several features of the distribution of the distance of exports across French firms that are hard to fit under more standard assumptions.

Both papers are focused on this type of proximity effect on the extensive margins of exporting. To the best of my knowledge, there is no work available on the consequences of these forces on the intensive margins, and the dynamic of sales in a broader sense. Hence, this paper tries to document and quantify these effects using

[1]Examples that are a part of this literature include Arkolakis (2016) and Costantini and Melitz (2007).

Brazilian customs firm-level data.

Besides the novelty in documenting these features, the quantification exercise is important for a few of reasons. If those quantities are considered to be large, by incorporating these effects in their analysis economists will be able to have better predictions on the dynamics of exporters. This collaborates to more precise estimates of relevant economic variables, such as short-run and long-run trade elasticities, and international trade effects on TFP.

Also, the empirical exercises presented here for post-entry dynamics can shed light on what are the potential mechanisms driving the aforementioned proximity effects on entry. For example, we document that proximity between the set of previous destinations served by a firm and a new destination leads to higher sales in the first period there. This suggests that it is not likely that entry was facilitated due to smaller fixed or sunk costs, but that the higher profitability came from larger demand faced by the firm in that location.

This article is directly related to the empirical literature of post-entry export sales dynamics, such as Eslava et al. (2015), and Ruhl and Willis (2016). These papers use transaction- and plant-level data on Colombian firms, and describe the growth path of exporters, specially in their early years. The closest article from this literature is Fitzgerald et al. (2016), since the methodology here draws heavily from their work. The reason for that choice is twofold. First, we want to be sure to control for year-specific shocks at the firm level. The reason for that is to avoid documenting sales dynamics that arrive due to shocks that affect the firm as a whole, such as the evolution of the firm's productivity. The second reason is that we want to check how the results depend on controlling for selection. As we will see, since these proximity effects reduce the probability of exit at any given time, controlling for selection using ex-post spell length will have major implications for the estimates of their effect on

sales.

Since we only observe the value of sales, we are silent with respect to disentangling price and quantity dynamics as in Fitzgerald et al. (2016) themselves, Vicard et al. (2016), Piveteau (2015), and Timoshenko et al. (2016). This is unfortunate, because separating the two gives important information about the underlying driving mechanisms for the sales dynamics, as it is done in Foster et al. (2016), for example. Ultimately, the objective of this research project is to fully describe the mechanisms behind these geographical effects on exporter dynamics. Hence, the characterization of quantity and price dynamics, coupled with a model that can rationalize them, are the natural follow up steps of this project.

The remaining sections of this paper are organized as follows: section II describes the Brazilian customs data, section III discuss the empirical methods and results, and section IV concludes. Supporting material can be found in the Appendix.

3.2 Data

We use yearly data from the universe of customs declarations for merchandise exports from Brazil. The data is collected by SECEX, a government office which is a part of the MDIC (Ministry of Industry, Foreign Trade and Services). The data includes all declared exports, by firms and individuals, covering the years of 1990-2001. In order to make the variables comparable, we convert each transaction to 1992 dollars. Finally, we compute the yearly sales for each firm to each country we aggregate all their exports in each year.

The bilateral data on geography comes from CEPII GeoDist database. Here we focus on the border effects, so we only use the data on which country shares borders with another.

Finally, it is worth detailing how the variables tenure and spell are constructed. The tenure variable keeps track of the amount of years a firm has continuously served a particular destination, at the current period. In order to assign a value to it, we need to observe their first year serving that market. For example, if we observe that a firm only exported to France in the years 1991, 1992 and 1993, the value assigned to the tenure variable would be 1 in 1991, 2 in 1992 and 3 in 1993. Notice that if we observe that this firm was also selling to France in 1990 we would not be able to pin-down the value for the tenure variable, because our data set starts in 1990 so we do not observe the year in which the firm entered in that destination.

The spell variable assigns the ex-post number of years that the firm has continuously spent in a particular destination. In our previous example of a firm that served France only for the years of 1991, 1992, and 1993, the value of the spell variable would be 3 in 1991, 3 in 1992, and 3 in 1993. It is important to notice that even tough this variable is the same for a continuous exporting experience, a firm can have more than one continuous experience in a given destination if they exit and enter there again. In this case, the firm would potentially have two different values for their spell in the destination, depending on the year of the observation, which is why we also need to index this variable by date.

Furthermore, notice that we do not treat for *churning* in any way, meaning that whenever a firm goes from not selling to selling to a country it counts as an entry, regardless of their previous experiences, and the same goes for exiting, so that we do not distinguish between temporary and permanent exits.

Since many successful experiences are going to have missing values if we consider the precise spell length we top-code them at 6 years, which is half the length of the our data span. By doing so, some values for ex-post spell that were unknown but are definitely larger than 6 years are assigned the value of 6, the maximum level of

the top-coded spell. We later denote these group as 6+, to remind the reader that it encompasses all spells greater or equal than 6 years.

For our estimations we will not consider the sales of firms that have censured tenure, and censured spell below the top-code of 6 years. However, we still use these dropped observations to compute the border variables that are going to be used as regressors later on because these variables describe the firm's export activities in other countries and potentially other years, so we want to be sure that all the information is being considered. The construction of these variables is going to be explained in the next section, when they are introduced.

3.3 Empirical Approach

3.3.1 Sales

Table 3.1 in the appendix presents the results of all the exercises performed in this section for sales, that are controlling for the spell length.

The first empirical exercise we do is to replicate the specification in Fitzgerald et al. (2016), for the dynamics of sales. Using Irish customs data, they want to observe how sales change due to how long the firm has been selling in a given country. In order to control for aspects that can affect the firm as a whole, such as changes in their idiosyncratic productivity, they include firm-year fixed effects. Notice that these coefficients are properly identified, since the same firm can sell to multiple destinations and the coefficient is common across destinations in a given period. Also, in order to control for selection within the firm-year on unobserved heterogeneity in idiosyncratic demand, they separate the exporting experiences of firms according to their ex-post duration. Hence, we run the following specification

$$x_t^{ik} = \delta_t^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \right) + \epsilon_t^{ik}. \tag{3.1}$$

The variable x_t^{ik} is the log of sales of firm i in country k, f_t^i is the firm-year fixed effect, $\boldsymbol{a}_t^{ik}$ is a vector of indicator variables for the tenure ("age") of firm i in market k, in the same way $\boldsymbol{s}_t^{ik}$ is a vector of indicator variables for the ex-post spell length of firm i in country k at time t. The Kronecker product between the tenure and spell dummy variables ($\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik}$) generates a set of dummy variables for the tenure-spell pair. Since we choose the single-period spells to be our reference, each element of $\boldsymbol{\beta}$ represents difference between the sales of given pair of tenure-spell with respect to the reference group, all else equal.

The sole difference between this specification and the one that they use in their paper is the that they use destination controls that are constant over time (δ^k), where we allow these controls to change over time (δ_t^k). We believe this to be more adequate, since there could be changes in the aggregate demand from country k due to macroeconomic fluctuations, and to changes in bilateral trade costs. This inclusion is potentially relevant, since the time frame covered by our data includes major changes in trade costs, such as the creation of the Mercosur. As we can see in Table 3.1 in the appendix, the estimates barely change regardless of the choice, which serve as a robustness check of this specification. All the figures that we display in the body of the text include destination-year controls. The results and figures for the specifications with the destination controls restricted to be constant over time are reported in the appendix.

In order to compare the predicted ratios of export sales for different tenures we can use the properly exponentiated estimates of $\boldsymbol{\beta}$. We report the values of the ratio

of predicted sales, for every pair of tenure-spell, with respect to predicted sales of a single year spell. The dashed lines represent the 95% confidence interval for those values. The results are similar to the ones in Fitzgerald et al. (2016): longer spells are associated with a larger initial size and faster growth, and the initial and final sales are similar for the same spell.

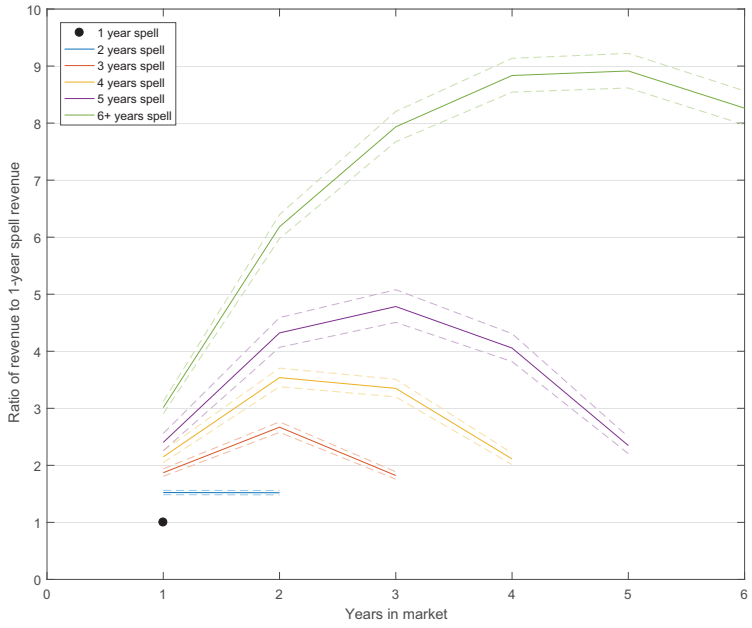

Figure 3.1: Ratio of predicted sales - specification (1)

Moving forward, we want to observe if selling to neighboring countries affect the sales dynamics of firms. The first step is to construct a measure for that, so we define the dummy variable b_t^{ik}. This variable takes the value of 1 if the firm was active in a country that shares a border with country k during period t. Next, we include those variables in the previous specification by separating the tenure-spell cases, between the ones in which the firm exported to a bordering country in the previous periods and the ones it did not. The regression is:

$$x_t^{ik} = \delta_t^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \otimes b_{t-1}^{ik} \right) + \epsilon_t^{ik}. \tag{3.2}$$

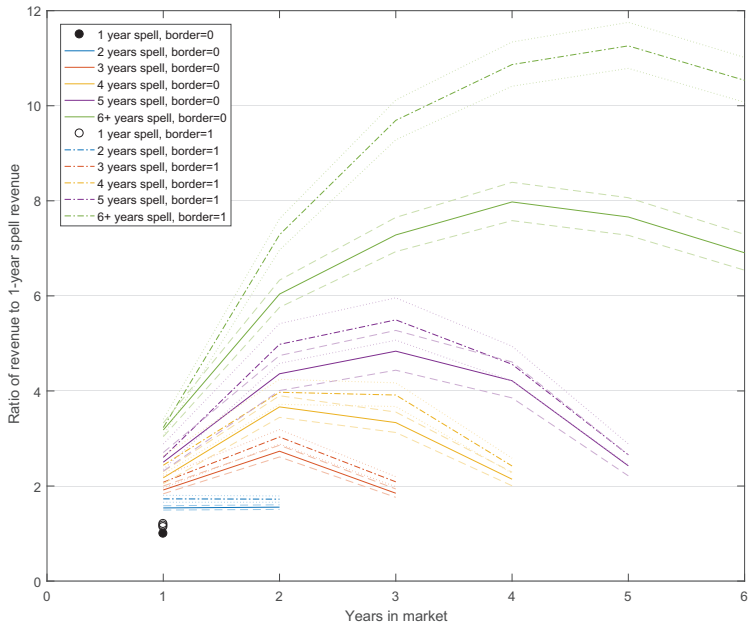

Figure 3.2: Ratio of predicted sales - specification (2)

In Figure 3.2, again by exponentiating the estimates on $\boldsymbol{\beta}$, we can compare the sales predictions among different tenure-spell-border tuples. There, the reference level is sales for one period spells for countries in which the border variable is zero. We can see that for a given tenure-spell pair it does not seem to make big difference in the estimates if the firm was serving a neighboring country in the previous period or not. The notable exception happens when the spell group is the one of 6 or more years. As we are going to see more clearly in the regressions on probability of exit, this seems to happen because the border effects, which we will denote by *geographic spillovers*, seem to positively affect the length of spells. However, by fully controlling for ex-post spell length, there is no difference in the dynamics. In the case of 6+ year

spells, we have top-coded it so spell lengths more years also fall into this category. Hence, within this group, we see that the geographic spillovers are indeed quite large. For example, in their fourth year of tenure of the 6+ spell group the predicted sales if $b_{t-1}^{ik} = 1$ are about 37.5% larger in the same tenure-spell for a firm that does not served any neighboring market in the previous period.

One problem with this approach is that our analysis might be polluted from firms that are switching between 0 and 1 for the border variable. This wouldn't be a problem if we were regressing it on a relative variable, such as the growth rate from one period to the other, but since x_t^{ik} is measured in levels, the interpretation of the $\boldsymbol{\beta}$ coefficients is not clear. Trying to mitigate this inconvenience, we define $\tilde{b}_t^{ik}$ as a dummy variable that takes the value of 1 if the firm i operates in a country that shares a border with k *throughout the whole spell*, and is 0 if the firm does not sell to any bordering country *throughout the whole spell*. Hence, many observations such that these variables changed will have a missing value to that variable. They will not help to estimate the $\boldsymbol{\beta}$, but we include them in the regressions to have better estimates of the firm and destination effects. We regress

$$x_t^{ik} = \delta_t^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \otimes \tilde{b}_t^{ik} \right) + \epsilon_t^{ik}. \tag{3.3}$$

The qualitative results seem to be unchanged from the previous specifications. The predicted sales for when $\tilde{b}_t^{ik} = 1$ are still quite similar to the ones when $\tilde{b}_t^k$, with the exception of the 6+ spell group, which seems to have an even larger gap now.

Now, we focus on the effects of the geographic spillovers unconditioned to the ex-post spell length. We regress the following specification

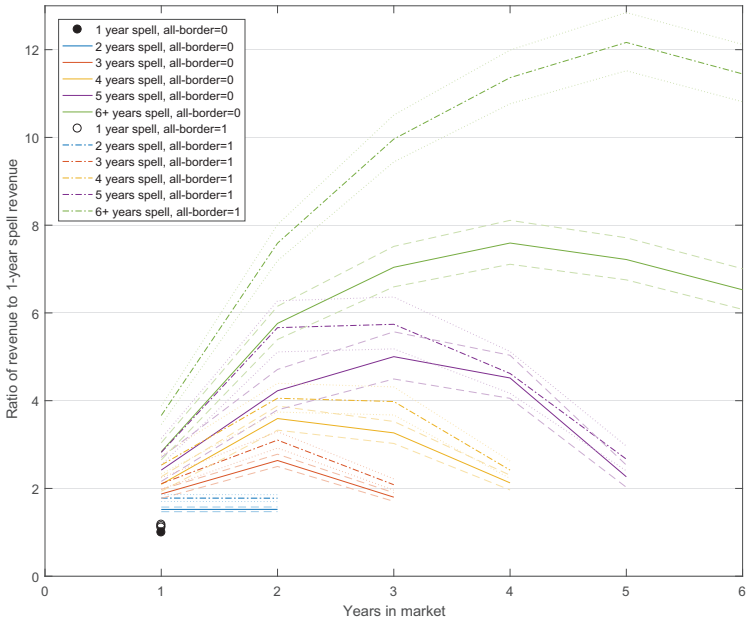

Figure 3.3: Ratio of predicted sales - specification (3)

$$x_t^{ik} = \delta_t^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \tilde{b}_t^{ik} \right) + \epsilon_t^{ik}. \tag{3.4}$$

The results of these regressions are presented in Table 3.2 in the appendix. The reference group is now the entrants (1 year tenure) with the border variable $\tilde{b}_t^{ik} = 0$.

The effects of the geographic spillovers is indeed much larger in this context, and it increases with tenure. At one year of tenure, the predicted sales for the firm with $\tilde{b}_t^{ik} = 1$ is about 22% larger than the firm with $\tilde{b}_t^{ik} = 1 = 0$. This gap increases to about 87% in their eighth year of tenure, which suggests that these effects are indeed large.

The effects are larger here for a few reasons. First of all, we are going to see in the next section that the border variable is associated with a lower probability of exit at any tenure level. As we have seen in the previous results, the longer spells are

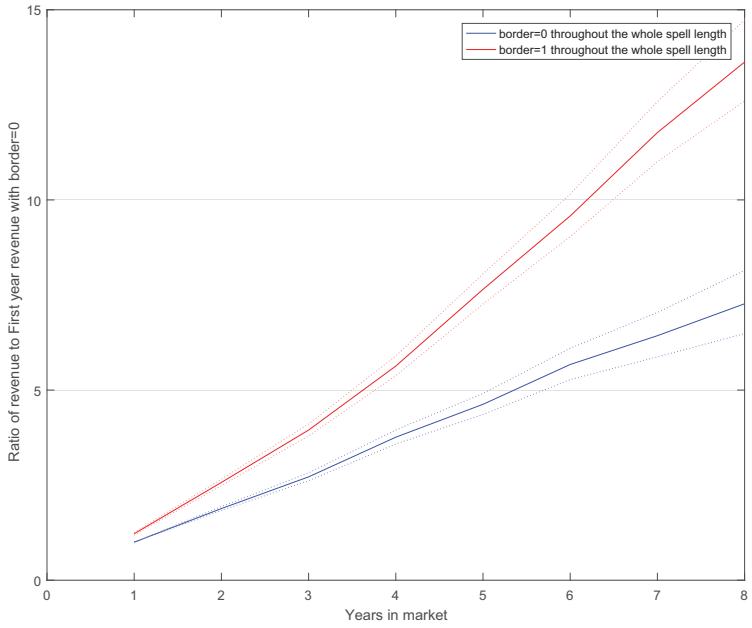

Figure 3.4: Ratio of predicted sales - specification (4)

associated with higher growth. Those two features of the data explains why when we stop controlling for spell we see that the group of observations with $\tilde{b}_t^{ik} = 1$, which has a longer expected spell-length, has a higher expected growth. Another possible explanation for these figures arrives if the fact that the firm was selling to a neighboring country in the previous year has a positive effect on growth. In that case, the two groups that we are comparing here are the extreme ones in the sense that one of them was not exposed to this effect in any period, and the other was exposed in every period. Hence, these positive effects on growth would accumulate over time for the latter, which would increase the difference between the level of sales of these two groups over time.

3.3.2 Probability of Exit

Now, we focus on the Probability of a firm exiting a market, given their tenure. Again, we control for firm-year fixed effects, and destination fixed effects, as in Fitzgerald et al. (2016), and destination-year controls, in order to check for the robustness of the results. All the regression results from this section can be found in Table 3.3, at the end of the text. The linear probability model including the destination-year controls is:

$$Pr\left[X_{t+1}^{ik} = 1 | X_t^{ik} = 0\right] = \delta_t^k + f_t^i + \boldsymbol{\beta}' \boldsymbol{a}_t^{ik} + \epsilon_t^{ik}. \tag{3.5}$$

The variable X_t^{ik} is an indicator variable that takes the value of one if firm i sells to destination k during the year t. Again, $\boldsymbol{a}_t^{ik}$ is a set of dummies that represent the year of tenure, with the reference level being the first year of tenure.

Figure 3.5 displays the results, which are similar to the ones in Fitzgerald et al. (2016). As we can see, the probability of exit decreases with tenure, and the fall is sharper in the initial years.

Now, we want to check if selling to a neighboring country has any effect on the probability of firms exiting a given market. We run the following regression:

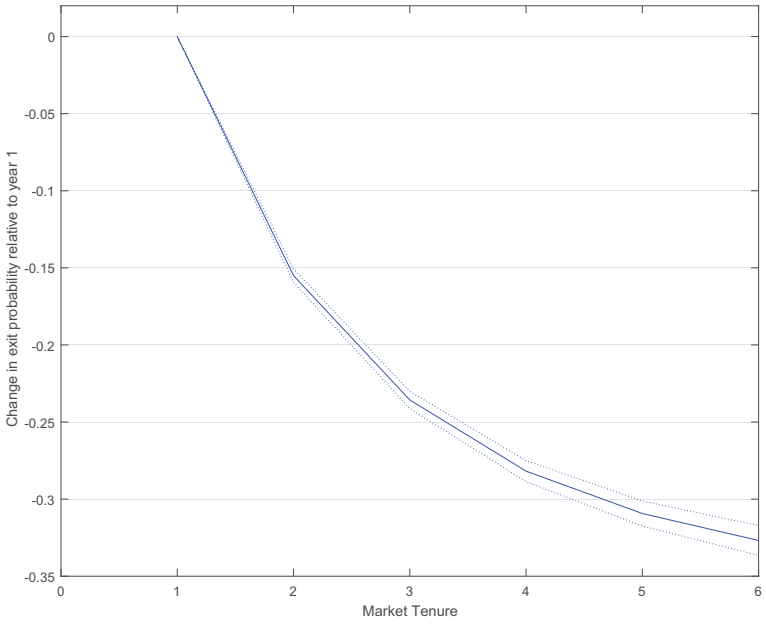

Figure 3.5: Exit Probability and Market Tenure - specification (5)

$$Pr\left[X_{t+1}^{ik}=1|X_{t}^{ik}=0\right]=\delta_{t}^{k}+f_{t}^{i}+\boldsymbol{\beta}'\left(\boldsymbol{a}_{t}^{ik}\otimes b_{t}^{ik}\right)+\epsilon_{t}^{ik}. \tag{3.6}$$

Figure 3.6 displays the results and it is clear that the geographic spillovers have a large negative impact on the probability of a firm exiting a market. The probability of a firm exiting a market right after its first year of tenure is about 9pp lower if the firm was selling to a neighbor of that country during that year. This gap is slowly reduced over tenured years, but even in the 6th year of tenure it is about 3pp.

This result could just be driven by the observations with $b_t^{ik}=1$ having larger sales than their peers with the same tenure. In order to check this conjecture, we include the log of the current sales (w_t^{ik}) on the probability of exit regressions. Hence, we run:

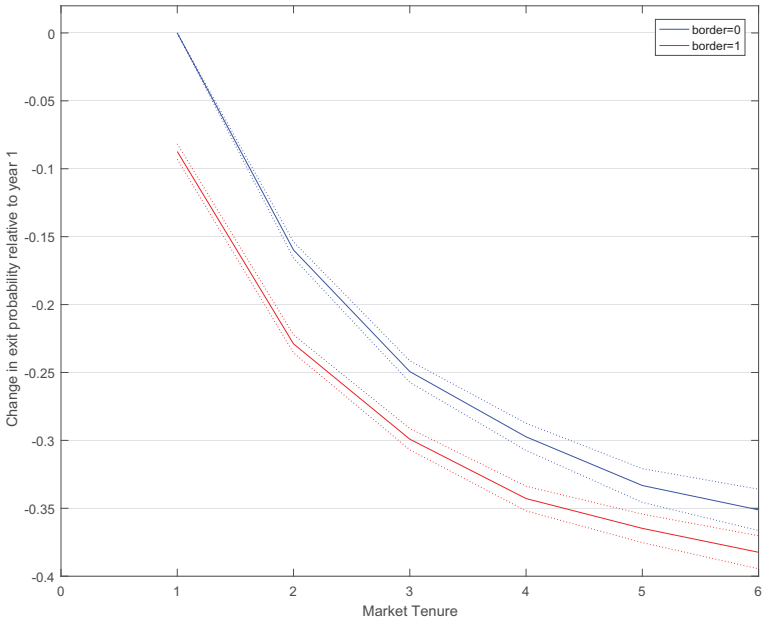

Figure 3.6: Exit Probability and Market Tenure - specification (6)

$$Pr\left[X_{t+1}^{ik}=1|X_{t}^{ik}=0\right]=\delta_{t}^{k}+c_{t}^{i}+\boldsymbol{\beta}'\left(\boldsymbol{a}_{t}^{ik}\otimes b_{t}^{ik}\right)+w_{t}^{ik}+\epsilon_{t}^{ik}. \tag{3.7}$$

Figure 3.7 displays the results of regression (7). We can see that including the log of sales affects our estimates. It reduces the difference for the probability of exit in the first year of tenure to about 7pp, and around the 5th year of tenure the gap seem closed. Hence, part of the negative effect of the geographic spillovers on the probability of exit seems to be explained by a positive correlation between selling to a neighboring country and the size of the sales. But even controlling for that we find that the geographic spillovers have a large effect in reducing the probability of exit, specially in the first years.

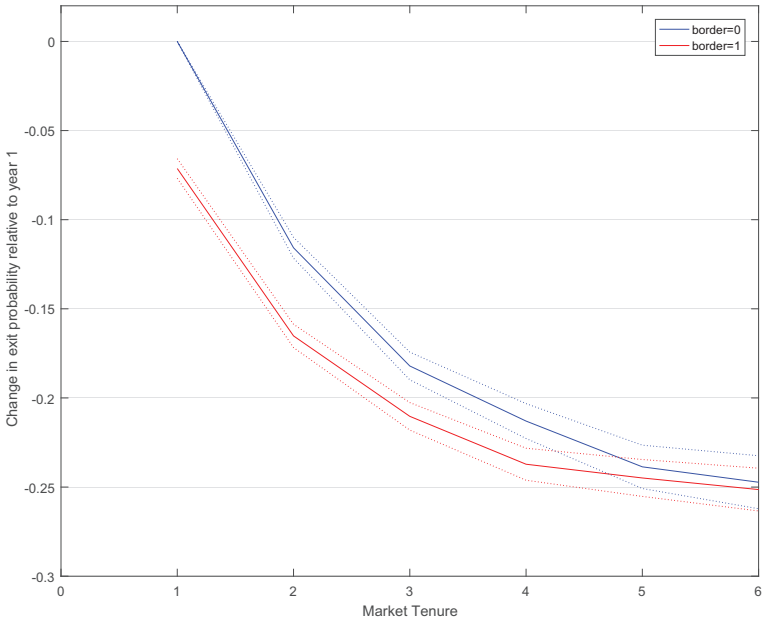

Figure 3.7: Exit Probability and Market Tenure - specification (14)

3.4 Conclusion

This paper contributes to the literature of ex-post exporter dynamics. We start by replicating the results on sales dynamics, and probability of exit in found in Fitzgerald et al. (2016) for Irish data in our Brazilian customs data. As in that paper we find that, controlling for destination and firm-year fixed effects, longer spell-lengths are associated with a larger initial sales levels and higher initial growth. Also, for a given spell-length, the predicted level of sales in the entry period is similar to the one of the period that the firm exits a market. Furthermore, we also find that the probability of exit is decreasing with tenure. These results are robust to the inclusion of time varying destination controls, which endorses their findings.

After that, we introduce variables that can describe the effect of geographic spillovers in our estimations. We understand geographic spillovers as being the possibility that previous years activity in neighboring countries have positive effects in

a given location.

We conclude that controlling for tenure, firm-year and destination-year fixed effects, the geographic spillovers are associated with larger revenues, lower probabilities of exit, which contribute to longer spells, and higher growth. Also, since the effect on lower probabilities of exit persists after controlling for the level of sales, it seems that the positive effects of geographic spillovers do not only rely on having larger levels of sales in a given period.

Furthermore, when we use controls for the selection on idiosyncratic demand faced by the firm, by including ex-post spell length, the effects of the geographic spillovers vanish. This is not surprising, since we are effectively controlling for how successful a particular endeavor of a firm has been. This is also a good check that these controls are effective in characterizing selection on idiosyncratic demand, and that they are good "sufficient statistics" on the dynamic of firms.

These results generate some interesting insights. Morales et al. (2019) and Chaney (2014) provide two different stories that could explain why firms are more likely to enter in destinations close to the ones that they are already selling to. In the first paper, the idea that fixed and sunk costs would be lower is the driving mechanism for their observed increase in the entry probability, whereas in the second paper this happens because there is a higher probability of finding new customers close to your current ones. The findings here that the initial sales are higher due to the border effect endorses the second explanation, since we would expect that the initial sales would be lower otherwise.

Furthermore, we have also observed that these geographical effects are not only related to entry, but that they are relevant during the whole exporting period, and they are associated with lower probabilities of exit. Those patterns are helpful in providing guidance to future research about the mechanisms driving these geographic

spillovers.

3.5 Appendix

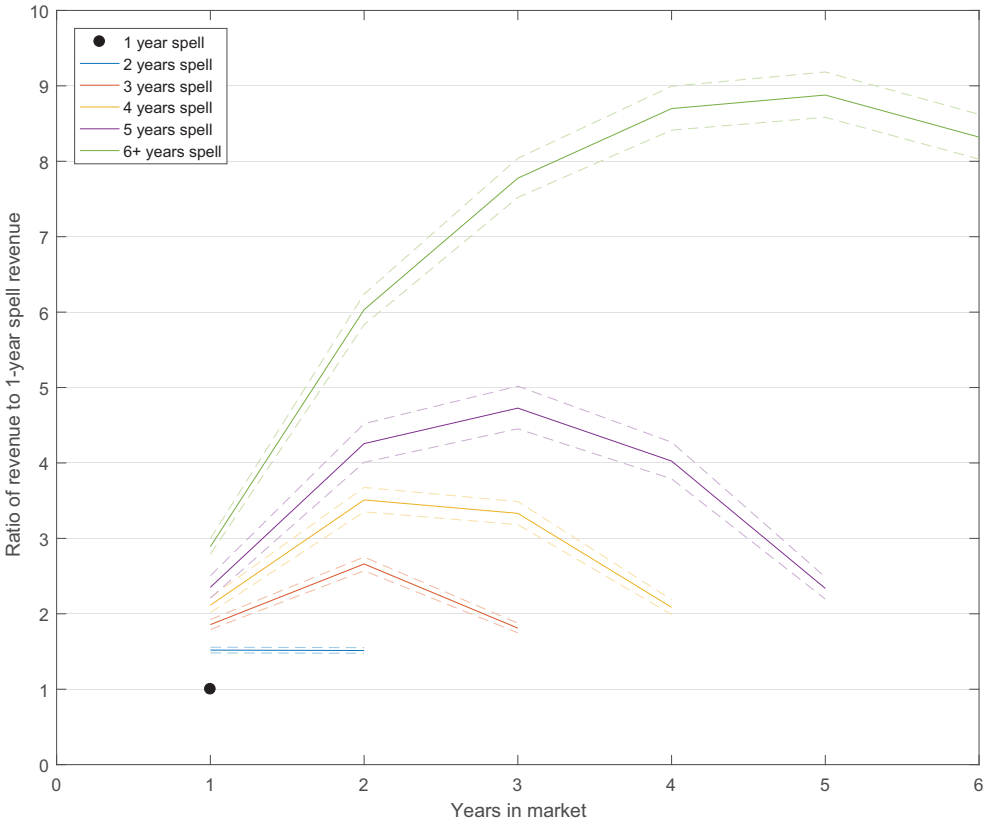

Figure 3.8: Ratio of predicted sales - specification (1 rest)
$x_t^{ik} = \delta^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \right) + \epsilon_t^{ik}$

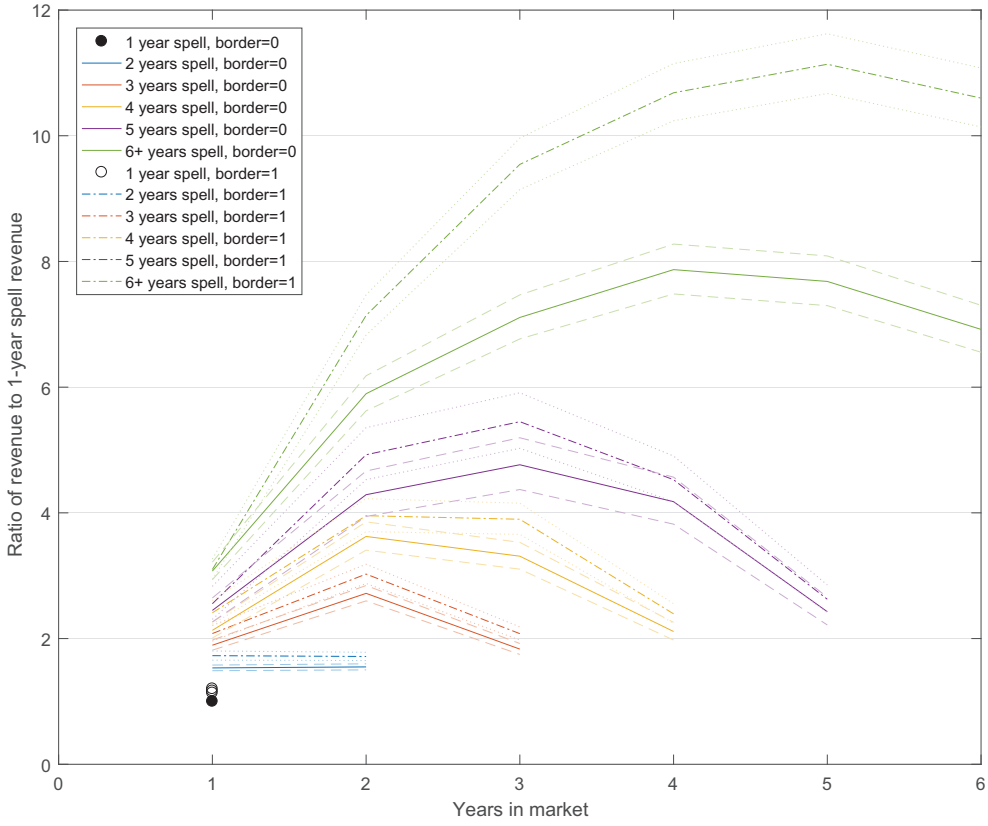

Figure 3.9: Ratio of predicted sales - specification (2 rest)
$x_t^{ik} = \delta^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \otimes b_{t-1}^{ik} \right) + \epsilon_t^{ik}$

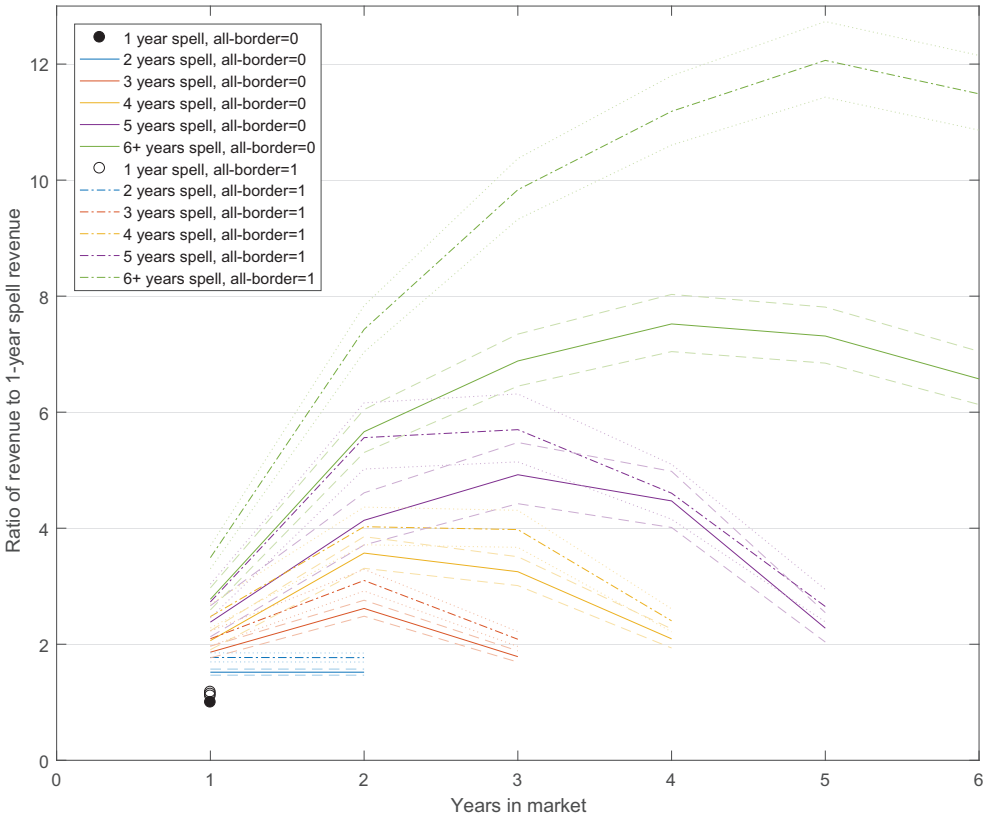

Figure 3.10: Ratio of predicted sales - specification (3 rest)
$x_t^{ik} = \delta^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \boldsymbol{s}_t^{ik} \otimes \tilde{b}_t^{ik} \right) + \epsilon_t^{ik}$

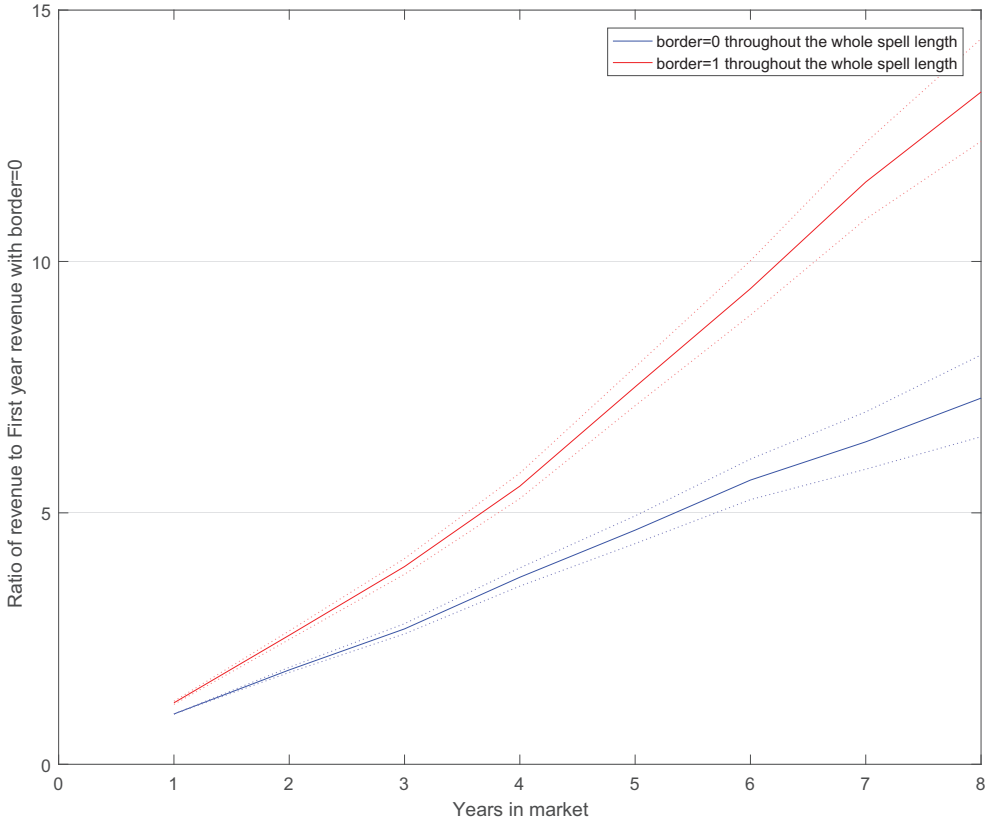

Figure 3.11: Ratio of predicted sales - specification (4 rest)
$x_t^{ik} = \delta^k + f_t^i + \boldsymbol{\beta}' \left(\boldsymbol{a}_t^{ik} \otimes \tilde{b}_t^{ik} \right) + \epsilon_t^{ik}$

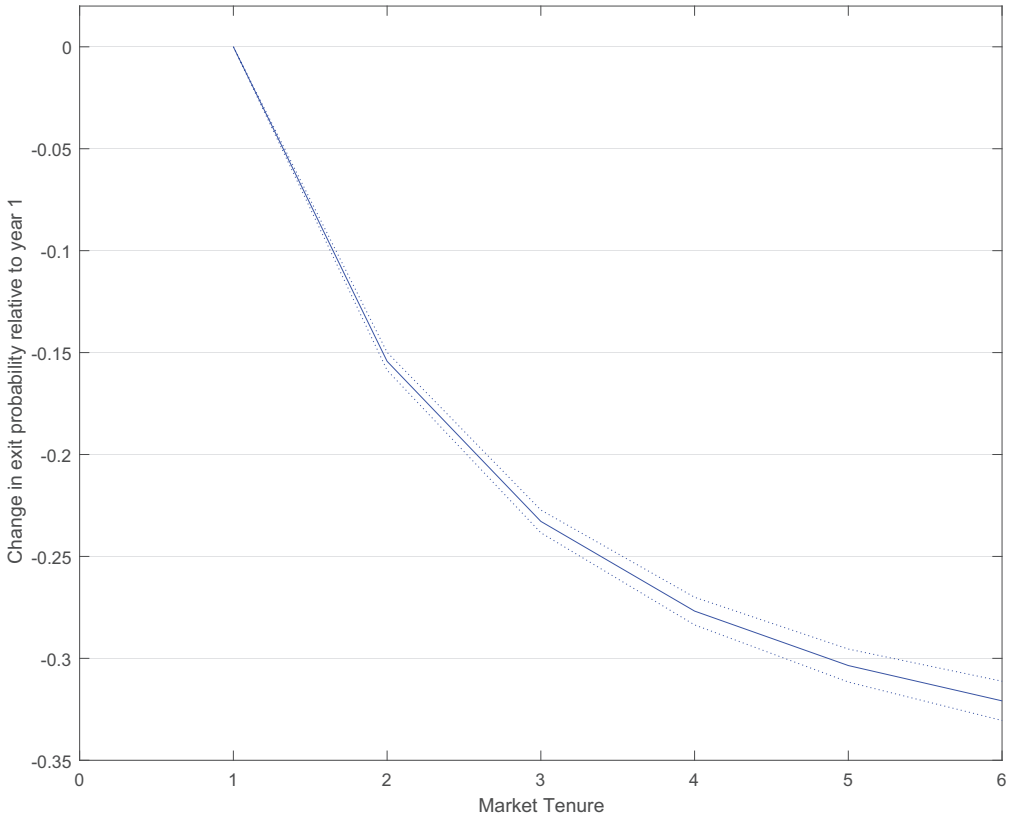

Figure 3.12: Difference in Exit Probability - specification (5 rest)
$Pr\left[X_{t+1}^{ik} = 1 | X_t^{ik} = 0\right] = \delta^k + f_t^i + \boldsymbol{\beta}' \boldsymbol{a}_t^{ik} + \epsilon_t^{ik}$

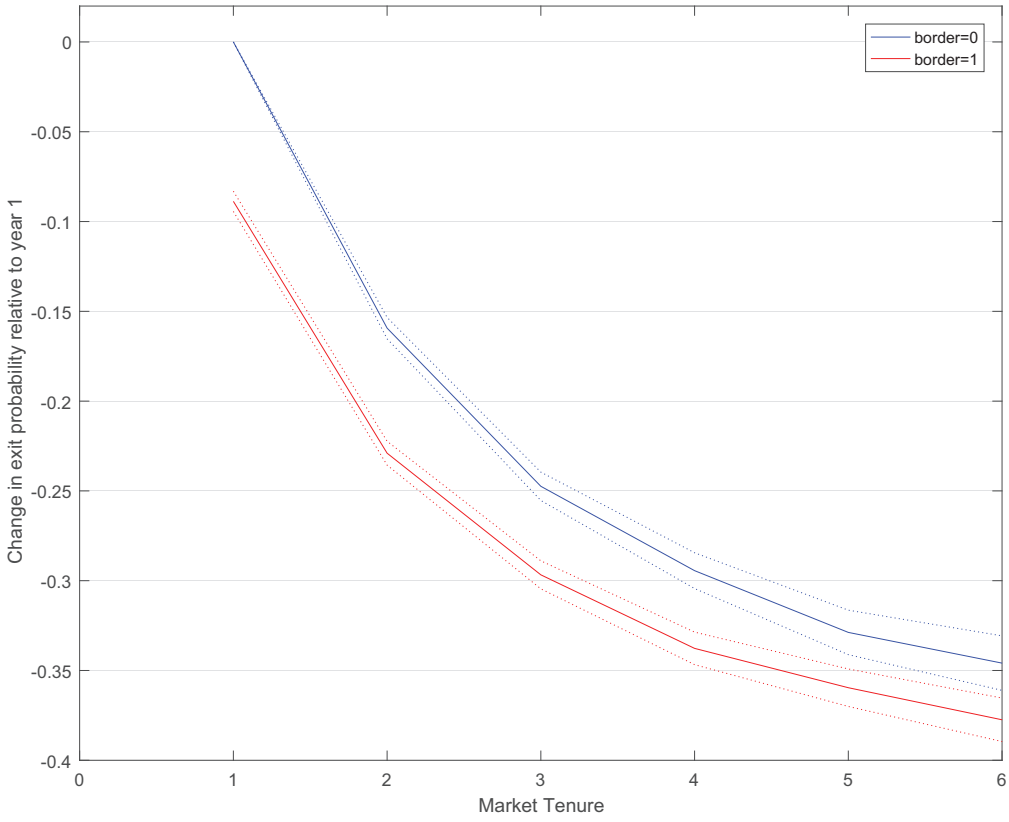

Figure 3.13: Difference in Exit Probability - specification (6 rest)
$Pr\left[X_{t+1}^{ik}=1|X_t^{ik}=0\right]=\delta^k+c_t^i+\boldsymbol{\beta}'\left(\boldsymbol{a}_t^{ik}\otimes b_t^{ik}\right)+\epsilon_t^{ik}$

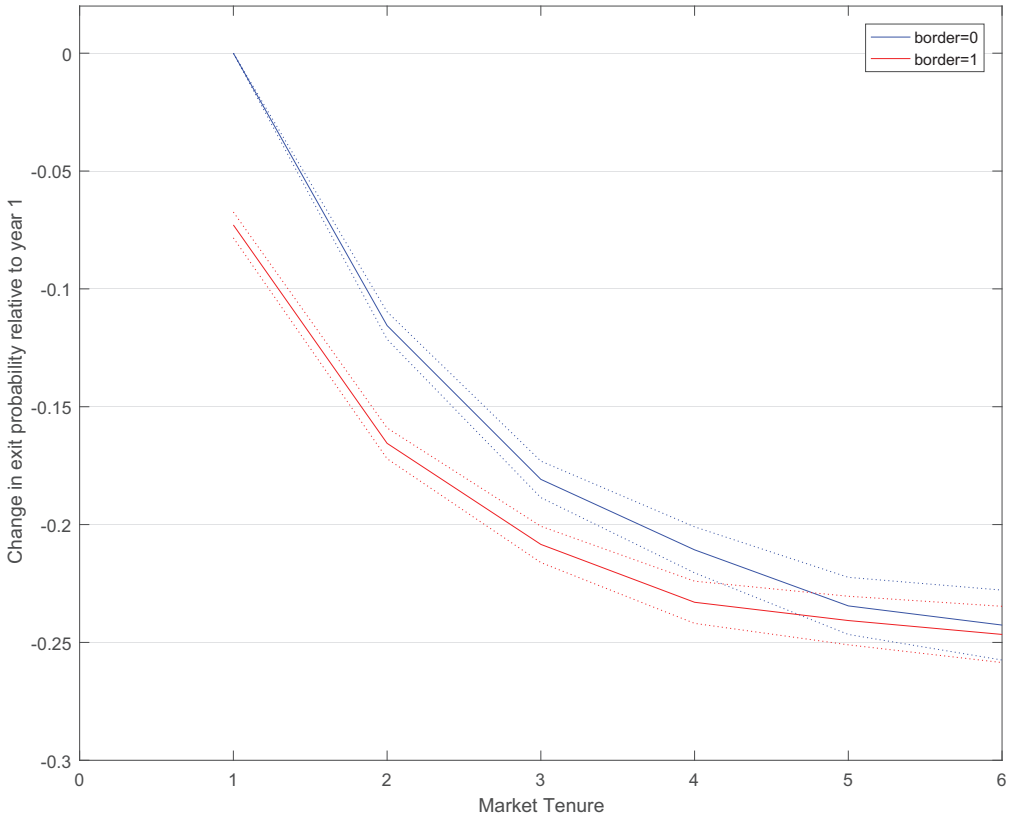

Figure 3.14: Difference in Exit Probability - specification (7 rest)
$Pr\left[X_{t+1}^{ik}=1|X_t^{ik}=0\right]=\delta^k+c_t^i+\boldsymbol{\beta}'\left(\boldsymbol{a}_t^{ik}\otimes b_t^{ik}\right)+w_t^{ik}+\epsilon_t^{ik}$

Table 3.1: Regressions on sales controlling for selection

Spec.	(1 rest)	(1)	(2 rest)		(2)		(3 rest)		(3)	
Border	NA	NA	0	1	0	1	0	1	0	1
Ten.Spell										
1.1	NA	NA	NA	0.16***	NA	0.16***	NA	0.13***	NA	0.13***
				(0.01)		(0.01)		(0.01)		(0.01)
1.2	0.42***	0.42***	0.43***	0.55***	0.43***	0.55***	0.42***	0.57***	0.42***	0.58***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
2.2	0.41***	0.42***	0.44***	0.54***	0.44***	0.54***	0.42***	0.57***	0.42***	0.58***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
1.3	0.62***	0.63***	0.64***	0.73***	0.65***	0.73***	0.62***	0.74***	0.63***	.75***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
2.3	0.98***	0.98***	1.00***	1.11***	1.00***	1.11	0.96***	1.13***	0.97***	1.13***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
3.3	0.59***	0.60***	0.61***	0.73***	0.61***	0.74***	0.58***	0.74***	0.59***	0.73
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
1.4	0.75***	0.76***	0.76***	0.88***	0.78***	0.89***	0.72***	0.91***	0.74***	0.93***
	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
2.4	1.26***	1.26***	1.29***	1.37***	1.30***	1.38***	1.27***	1.40***	1.28***	1.40***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
3.4	1.20***	1.21***	1.20***	1.36***	1.20***	1.37***	1.18***	1.38***	1.18***	1.38***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
4.4	0.74***	0.75***	0.75***	0.87***	0.76***	0.89***	0.74***	0.88***	0.76***	0.88***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
1.5	0.86***	0.88***	0.90***	0.94***	0.92***	0.96***	0.87***	1.00***	0.88	1.04***
	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)
2.5	1.45***	1.46***	1.46***	1.59***	1.47***	1.61***	1.42***	1.72***	1.44***	1.73***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)	(0.05)
3.5	1.55***	1.56***	1.56***	1.70***	1.58***	1.70***	1.59***	1.74***	1.61***	1.75***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)
4.5	1.39***	1.40***	1.43***	1.51***	1.44***	1.52***	1.50***	1.53***	1.51***	1.53***
	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)	(0.04)	(0.06)	(0.05)
5.5	0.85***	0.85***	0.89***	0.97***	0.89***	0.98***	0.82***	0.97***	0.82***	0.98***
	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)	(0.04)	(0.05)	(0.05)
1.6+	1.06***	1.10***	1.12***	1.13***	1.16***	1.17***	1.02***	1.25***	1.04	1.30***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
2.6+	1.80***	1.82***	1.77***	1.97***	1.80***	1.99***	1.73***	2.00***	1.75	2.03***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
3.6+	2.05***	2.07***	1.96***	2.26***	1.99***	2.27***	1.93***	2.29***	1.95***	2.30***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
4.6+	2.16***	2.18***	2.06***	2.37***	2.08***	2.36***	2.02***	2.41***	2.03***	2.43***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
5.6+	2.18***	2.19***	2.04***	2.41	2.04***	2.42***	1.99***	2.49***	1.98***	2.50***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
6.6+	2.12***	2.11***	1.93***	2.36***	1.93***	2.35***	1.88***	2.44***	1.88***	2.44***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)
N	364,422	364,422	364,422		364,422		291,960		291,960	
R^2	0.68	0.69	0.68		0.69		0.70		0.71	
Adj-R^2	0.52	0.53	0.52		0.53		0.52		0.52	

Table 3.2: Regressions on sales not controlling for selection

Spec.	(4 rest)		(4)	
Border	0	1	0	1
Tenure				
1	NA	0.20***	NA	0.20***
		(0.01)		(0.01)
2	0.63***	0.94***	0.63***	0.94***
	(0.01)	(0.02)	(0.01)	(0.02)
3	0.99***	1.37***	1.00***	1.37***
	(0.02)	(0.02)	(0.02)	(0.02)
4	1.31***	1.71***	1.32***	1.73***
	(0.02)	(0.02)	(0.03)	(0.02)
5	1.54***	2.02***	1.53***	2.04***
	(0.03)	(0.03)	(0.03)	(0.03)
6	1.73***	2.25***	1.74***	2.26***
	(0.04)	(0.03)	(0.04)	(0.03)
7	1.86***	2.45***	1.86***	2.47***
	(0.05)	(0.03)	(0.05)	(0.03)
8	1.98***	2.60***	1.98***	2.61***
	(0.06)	(0.04)	(0.06)	(0.04)
N	261,960		261,960	
R^2	0.69		0.69	
Adj-R^2	0.49		0.49	

Table 3.3: Probability of Exit

Spec.	(5 rest)	(5)	(6 rest)		(6)		(7 rest)		(7)	
Border	NA	NA	0	1	0	1	0	1	0	1
Tenure										
1	NA	NA	NA	-0.09***	NA	-0.09***	NA	-0.07***	NA	-0.07***
				(0.003)		(0.003)		(0.003)		(0.003)
2	-0.15***	-0.16***	-0.16***	-0.23***	-0.16***	-0.23***	-0.12***	-0.17***	-0.12***	-0.17***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
3	-0.23***	-0.24***	-0.25***	-0.30***	-0.25***	-0.30***	-0.18***	-0.21***	-0.18***	-0.21***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
4	-0.27***	-0.28***	-0.29***	-0.34***	-0.30***	-0.34***	-0.21***	-0.23***	-0.21***	-0.24***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
5	-0.30***	-0.31***	-0.33***	-0.36***	-0.33***	-0.36***	-0.23***	-0.24***	-0.24***	-0.25***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)
6	-0.32***	-0.33***	-0.35***	-0.38***	-0.35***	-0.38***	-0.24***	-0.25***	-0.25***	-0.25***
	(0.005)	(0.005)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)
log(sales)	NA	NA	NA		NA		-0.06***		-0.06***	
							(0.0005)		(0.0005)	
N	402,582	402,582	402,582		402,582		402,582		402,582	
R^2	0.51	0.51	0.51		0.52		0.54		0.54	
Adj-R^2	0.29	0.29	0.29		0.30		0.33		0.34	

Bibliography

Albornoz, F., Calvo Pardo, H., Corcos, G., and Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, 88(1):17–31.

Albornoz, F., Fanelli, S., and Hallak, J. C. (2016). Survival in export markets. *Journal of International Economics*, 102:262 – 281.

Allen, T. and Arkolakis, C. (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics*, 129(3):1085–1140. 48

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1):106–116. 48

Argente, D., Lee, M., and Moreira, S. (2018a). How do firms grow? the life cycle of products matters. 2018 Meeting Papers 1174, Society for Economic Dynamics. 5

Argente, D., Lee, M., and Moreira, S. (2018b). How do firms grow? the life cycle of products matters. 2018 Meeting Papers 1174, Society for Economic Dynamics.

Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade. *Journal of Political Economy*, 118(6):1151–1199. 6

Arkolakis, C. (2016). A unified theory of firm selection and growth. *The Quarterly Journal of Economics*, 131(1):89–155. 72

Bagwell, K. (2005). The economic analysis of advertising. *forthcoming, Handbook of Industrial Organization*, 3.

Bailey, M., Cao, R. R., Kuchler, T., Stroebel, J., and Wong, A. (2017). Measuring social connectedness. Working Paper 23608, National Bureau of Economic Research.

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5):215–227. 4

Bronnenberg, B. and Albuquerque, P. (2003). Geography and marketing strategy in consumer packaged goods. *Advances in Strategic Management*, 20.

Bronnenberg, B., Dube, J.-P., Gentzkow, M., and Shapiro, J. (2014). Do pharmacists buy bayer? informed shoppers and the brand premium.

Bronnenberg, B. J., Dhar, S. K., and Dubé, J. H. (2009). Brand history, geography, and the persistence of brand shares. *Journal of Political Economy*, 117(1):87–115.

Bronnenberg, B. J., Dubé, J.-P. H., and Gentzkow, M. (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6):2472–2508.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, 87(3):741–835. 48

Campa, J. and Goldberg, L. (2005). Exchange rate pass-through into import prices. *The Review of Economics and Statistics*, 87(4):679–690.

Caplin, A., Dean, M., and Leahy, J. (2017). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. NBER Working Papers 23652, National Bureau of Economic Research, Inc.

Caplin, A., Leahy, J., and Matějka, F. (2015). Social Learning and Selective Attention. NBER Working Papers 21001, National Bureau of Economic Research, Inc.

Chandrasekaran, D. and Tellis, G. (2007). A critical review of marketing research on diffusion of new products. *Review of Marketing Research*, 3.

Chaney, T. (2013). The Gravity Equation in International Trade: An Explanation. NBER Working Papers 19285, National Bureau of Economic Research, Inc.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634. 4, 15, 49, 72, 87

Chevalier, J. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *J. Marketing Res*, pages 345–354.

Consul, P. C. (1989). *Generalized Poisson Distributions: Applications and Properties*. Marcel Dekker, Inc., New York. 19

Consul, P. C. and Jain, G. C. (1973). A generalization of the poisson distribution. *Technometrics*, 15(4):791–799. 19

Consul, P. C. and Mittal, S. P. (1975). A new urn model with predetermined strategy. *Biometrische Zeitschrift*, 17(2):67–75. 21

Costantini, J. and Melitz, M. (2007). The dynamics of firm-level adjustment to trade liberalization. *The Organization of Firms in a Global Economy, edited by E Helpman, D Marin, and T Verdier. Cambridge: Harvard University Press.* 72

Dinlersoz, E. and Yorukoglu, M. (2008). Informative advertising by heterogeneous firms. *Information Economics and Policy*, 20(2):168–191.

Draganska, M. and Klapper, D. (2011). Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research*, 48(4):653–669. 49

Drozd, L. A. and Nosal, J. B. (2012). Understanding International Prices: Customers as Capital. *American Economic Review*, 102(1):364–395. 6

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498. 65

Eaton, J., Kramarz, F., and Kortum, S. (2019). Firm-to-Firm Trade: Exports, Imports, and the Labor Market. 2019 Meeting Papers 702, Society for Economic Dynamics. 6, 16, 49

Eslava, M., Tybout, J., Jinkins, D., Krizan, C., and Eaton, J. (2015). A Search and Learning Model of Export Dynamics. 2015 Meeting Papers 1535, Society for Economic Dynamics. 73

Evenett, S. and Venables, A. (2002). Export growth in developing countries: Market entry and bilateral trade flows.

Fitzgerald, D., Haller, S., and Yedid-Levi, Y. (2016). How exporters grow. 2016 Meeting Papers 499, Society for Economic Dynamics. 3, 73, 74, 76, 78, 83, 86

Foster, L., Haltiwanger, J., and Syverson, C. (2016). The Slow Growth of New Plants: Learning about Demand? *Economica*, 83(329):91–129. 74

Gourio, F. and Rudanko, L. (2014). Customer Capital. *Review of Economic Studies*, 81(3):1102–1136. 6

Hauser, J., Tellis, G. J., and Griffin, A. (2006). Research on innovation: A review and agenda for "marketing science". *Marketing Science*, 25(6):687–717. 4

Hillberry, R. and Hummels, D. (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review*, 52(3):527–550.

Kalish, S. (1985). A new product adoption model with price, advertising, and uncertainty. *Management Science*, 31(12):1569–1585. 2

Lenoir, C., Mejean, I., and Martin, J. (2018). Search Frictions in International Good Markets. 2018 Meeting Papers 878, Society for Economic Dynamics. 6

Lim, K. (2017). Firm-to-firm Trade in Sticky Production Networks. 2017 Meeting Papers 280, Society for Economic Dynamics.

Luttmer, E. G. J. (2007). Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics*, 122(3):1103–1144.

Luttmer, E. G. J. (2011). On the Mechanics of Firm Growth. *Review of Economic Studies*, 78(3):1042–1068.

Mahajan, V. and Peterson, R. A. (1979). Integrating time and space in technological substitution models. *Technological Forecasting and Social Change*, 14(3):231 – 241.

Mardia, K. V. (1962). Multivariate pareto distributions. *Ann. Math. Statist.*, 33(3):1008–1015. 24

Mayer, T. and Zignago, S. (2011). Notes on cepii's distances measures: The geodist database. Working Papers 2011-25, CEPII.

Morales, E., Sheu, G., and Zahler, A. (2019). Extended Gravity. *Review of Economic Studies*, 86(6):2668–2712. 72, 87

Oakes, J., Andrade, K., Biyoow, I., and Cowan, L. (2015). Twenty years of neighborhood effect research: An assessment. *Current Epidemiology Reports*, 2.

Perla, J. (2019). A model of product awareness and industry life cycles. Working paper. 5

Piveteau, P. (2015). An empirical dynamic model of trade with consumer accumulation. Technical report, Mimeo. 74

Rasouli, S. and Timmermans, H. (2013). Influence of social networks on latent choice of electric cars: A mixed logit specification using experimental design data. *Networks and Spatial Economics*, 16.

Rogers, E. M. (1962). Diffusion of Innovations. New York: The Free Press of Glencoe, 1962. *Social Forces*, 41(4):415–416. 4

Romer, P. M. (1987). Growth Based on Increasing Returns Due to Specialization. *American Economic Review*, 77(2):56–62. 5

Ruhl, K. J. and Willis, J. L. (2016). New exporter dynamics. *International Economic Review* (forthcomming). 73

Sattenspiel, L. (2009). *The Geographic Spread of Infectious Diseases: Models and Applications*. Princeton University Press, Princeton. 22

Shoukri, M. M. and Consul, P. C. (1987). *Some Chance Mechanisms Generating the Generalized Poisson Probability Models*, pages 259–268. Springer Netherlands, Dordrecht. 19

Sovinsky, M. (2008). Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074.

Timoshenko, O., Bastos, P., and Dias, D. (2016). Learning, prices, and firm dynamics. Working papers, The George Washington University, Institute for International Economic Policy. 74

Tinbergen, J. (1963). Shaping the world economy. *The International Executive*, 5(1):27–30. 48

Vicard, V., Rebeyrol, V., and Berman, N. (2016). Demand learning and firm dynamics:evidence from exporters. 2016 Meeting Papers 517, Society for Economic Dynamics. 74