NORTHWESTERN UNIVERSITY

Essays on Firms and Labor Markets in Developing Countries

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Managerial Economics and Strategy

By
Miguel Ángel Talamas Marcos

EVANSTON, ILLINOIS

June 2022

Abstract

This dissertation contains two chapters. The first one is on microenterprises in developing countries and how they face competition from large corporations. The second one is on estimating the causal effect of childcare availability on the formation and persistence of gender gaps in the Mexican labor market.

The first chapter is motivated by hundreds of millions of microenterprises in emerging economies facing increased competition from the entry and expansion of large firms that offer similar or identical products. This chapter studies how one of the world's most prevalent microenterprises, neighborhood shops, confront competition from convenience chains (e.g., 7-Eleven) in Mexico. To address the endogeneity in time and location of chains' store openings, I construct an instrument that captures a cost reduction and profitability increase for chains but not for shops. The instrument exploits two key differences between them: i) chains build economies of scale from cost-sharing across stores in adjacent cities, and ii) within cities, chains are more than three times as likely to open stores on wide streets. The results show that an increase from zero to the average number of chain stores in a neighborhood reduces the number of shops by 17%. This reduction is driven by a decrease in shop entry and not by an increase in shop exit. Shops retain 95% of their customers and their sales of fresh products, but customers visit shops less often and spend less on non-fresh and packed goods. I find evidence consistent with shops having comparative advantages stemming from being small and owner-operated, such as lower agency costs, building relationships with the community, and offering informal credit.

The second chapter estimates the effect of childcare availability on parents' employment using the timing of grandmothers' death– the primary childcare provider in Mexico– as identifying variation. I use a triple-difference to disentangle the effect of coinhabiting grandmothers' deaths due to their impact on childcare from their effects due to alternative

mechanisms. Through their impact on childcare availability, grandmothers' deaths reduce mothers' employment rate by 12 percentage points (27 percent) and do not affect fathers' employment. The negative effect on mothers' employment is smaller where public daycare is more available, or private daycare or schools are more affordable.

Acknowledgements

I am grateful to my committee for their ongoing feedback and advice throughout the years. Nancy Qian, my committee chair, thank you for pushing me to give the best and telling me what I sometimes did not want but needed to hear to improve my research. Seema Jayachandran, thank you for always giving me constructive advice on improving the empirics and suggesting additional analysis to expand the depth of my research. I am also thankful for your invaluable support and guidance through the publication process of the second chapter of this thesis. Luis Rayo, I will always cherish those long discussions about the economic framework underlying the empirical findings. Thank you for asking the precise questions that made me take a step back and think about the broader meaning of my research. Ameet Morjaria, thank you for helping me frame and contextualize my research in the economics literature and being a genuine colleague in the projects we work on together. To all of you, I am not only thankful for all of what you did as my committee members but for having you as role models and a source of inspiration.

Many other professors and colleagues from the Strategy and MEDS Departments at Kellogg School of Management and the Economics Department at Northwestern University have played a significant role in developing ideas and analysis. In particular, I want to thank Chris Udry for leading the Development Advising Group, where we had an opportunity to present and discuss our research every week.

I thank my mom, who taught me the value of hard work and perseverance with her example. Your entrepreneurial spirit and effort opened the doors for my undergraduate and graduate studies. I also thank my sisters, who have supported, motivated, and encouraged me. I am specially grateful to my wife, Andrea, for all the questions and discussions that enriched my papers and presentations. Moreover, you and our children, Miguel and Isabella, have filled my life with love.

Table of Contents

List of Tables

List of Figures

Chapter 1

Surviving Competition:

Neighborhood Shops vs Convenience Chains

1.1 Introduction

In developing countries, a stunning 214 million microenterprises account for 84% of all firms, 40% of employment, and 21% of value-added.[1] As economies develop, these small firms face increased competition, even an existential threat, from larger rivals with similar product offerings. Yet microenterprises continue to exist in overwhelming numbers (Hsieh and Olken (2014); Atkin et al. (2019)), raising the question of to what extent competition from large firms affects microenterprises and how the surviving ones compensate for disadvantages in scale?

This paper makes progress in answering this question by studying one of the most prevalent microenterprises, sole-proprietor neighborhood shops (henceforth shops). Shops are commonplace in developing countries; for example, in Mexico, there are 600 thousand *tienditas*, in the Philippines, there are 1.3 million *sari-sari*, and in India, there are 12 million *kiranas*.[2] Across the globe, shops share some common characteristics, such as offering a wide variety of food and drinks and their neighbors being their primary customers. I study the context of Mexico, where shops are vital to the economy. They represent one out of every eight firms, 4% of total employment, and they have the largest market share in the food and beverages industry, 31%.[3]

[1]Source: SME Finance Forum (2019)

[2]Sources: Economic Times (2019); Philstar (2017)

[3]Source: Economic Census 2019 and ENIGH 2018. Supermarkets have 17%, food markets 20%, specialized

In the last two decades, shops in Mexico faced increased competition from the entry and expansion of convenience chains (e.g., 7-Eleven) that expanded from fewer than 2,000 stores to more than 23,000. These convenience chains (henceforth chains) are a direct competitor to shops because they have a significant overlap in their product offering, are also small relative to other retail outlets, and mainly capture incidental purchases. Chains have advantages over shops due to their economies of scale, which allow them to share costs across their stores, better bargain with suppliers, have a lower cost of capital, and more productivity-enhancing investments. Chains may also represent lower search costs to consumers because of their store uniformity within each chain and their location on wide streets with big signs. However, shops may have advantages over chains, such as their relationships with the neighbors, offering informal credit, having lower agency costs (few or no employees), and not paying taxes. This paper studies the effect of the expansion of chains on shops.

To measure chains' impact on shops, I assembled a rich collection of data, including each of the shops covered by the economic censuses between 1999 and 2019.[4] For these close to two million shops, I also gained access to their confidential micro-level performance measures, such as revenue and profits. I link this detailed firm data with household income and expenditure surveys spanning from 2006 to 2018, which include information on what households buy, where they buy it, and how much they pay for it.

Even with a rich collection of data, estimating chains' impact on shops is challenging because chains' entry time and location are endogenous. Neighborhoods with higher demand for the products offered by both shops and chains will have better outcomes for shops and will also be more attractive for chains to enter. For example, a new park in a neighborhood may increase foot traffic and demand for drinks and snacks for both shops and chains. This type of unobservable shocks that increase or decrease demand for both chains and shops will

stores (e.g., bakeries, *tortillerias*, rotisseries) 30%, and convenience chains 2%.

[4]Economic censuses cover all establishments in the country. Establishments that are not covered are those that open and close in between census waves.

upward bias ordinary least squares estimates, even after controlling for city-wide trends and neighborhood time-invariant characteristics.

To address the endogeneity in chains' entry, I use an instrumental variable that varies across neighborhoods and time, and it leverages the complementarity between two key differences between chains and shops. First, chains exploit advantages from opening stores in nearby cities, such as same-chain stores sharing distribution center, trucks, inspectors, and regional offices. This cost-sharing in distribution, transportation, marketing, overhead, and other costs reduces chains' average cost and makes each store more profitable, leading to regional economies of scale for each chain. I measure these economies of scale using a Herfindahl-Hirschman index without normalization, which increases with the number of chain stores in adjacent cities and their concentration. And second, different from shops that are commonly located next to the owners' house, chains open on wide streets to target driving and bus-riding customers. I measure the neighborhood's suitability for chains using the prevalence of wide streets.

The instrument is the interaction between the measure of economies of scale, which varies across time and cities, and the prevalence of wide streets in the neighborhood, which varies within city and across neighborhoods. The identification assumption is that when a chain has more stores in cities near city A, neighborhoods suitable for chains in city A become more attractive in the eyes of the chain because of economies of scale, but do not change from the incumbent shop perspective, except for the increased probability of a chain opening in the neighborhood.[5] City-year and neighborhood fixed effects absorb the uninteracted terms and control for city-wide trends and neighborhood time-invariant characteristics.

I organize the main results into three categories. First, I find that each additional chain

[5]Placebo tests in figure A.8 show that the instrument is not related to customers (neighbors) characteristics likely correlated with demand such as income, number of cars, expenses, and demographics. Moreover, Appendix A.2 documents that regional economies of scale are firm-specific. The empirical section also discusses other threats to validity, such as potential spillovers or violations to the exclusion restriction.

store in a neighborhood reduces the number of shops by 4.6, implying that an increase from zero to the average number of chain stores in a neighborhood (6.4) reduces the number of shops by 17%. The number of exits of shops does not increase, making the 18% reduction in number of shop entries the main driver of the decrease in number of shops. Second, the negative effects on shops' performance concentrate along the extensive margin. At the neighborhood level, shops' total profits, revenue, value-added, inventories, total employed, and total hours worked decline between 20 and 30%. However, these adverse effects are less than a third in magnitude at the shop level (intensive margin), between 0 and 7%.

Third, I find that customers continue to purchase in shops, but they do so less and less often. An increase in chain stores from zero to their average number in a neighborhood decreases the probability of neighbors purchasing in shops by 4.5%. Those who continue to purchase in shops do so 7% less often and buy 10% less. The effect on neighbors' purchases differs across product categories. Chains do not affect household expenditure in shops on fresh products such as fresh sweet bread, fruit, and vegetables, which are often sourced daily by shop owners from central markets. Still, chains decrease household purchases in shops of packed and standardized products like sodas, milk, and bottled juices.

Why do shops survive? I find evidence consistent with shops leveraging comparative advantages and adjusting in response to competition. Shops' productivity, measured by the output-input ratio, is unaffected by chains, because shops respond to the decrease in revenue with a reduction in purchases and inventories. Shops also specialize in products that are harder to source and ensure quality, such as fresh products. Moreover, I find that shops less affected by chains are smaller and owner-operated. Anecdotal evidence and heterogeneity in the effects of chains on shops are consistent with these shops having comparative advantages in building relationships with their customers, facing lower agency costs, and screening their neighbors to provide them with informal credit to buy in the shop. For example, shops supply 16% of all the credit in the food and beverages industry and 78% of all credit not supplied through credit cards. In a context where consumers are both credit and cash constrained, this becomes a critical advantage.

The results are consistent with a standard competitive model with differentiated competition between chains and shops presented in Section III. The model highlights economic mechanisms driving shop entry and exit as the shop industry transitions from a steady-state without chains to a steady-state with them. The model presents this transition at the shop level and at the industry level to illustrate why the adverse effects on shops' performance concentrate along the extensive margin.

This paper contributes to two strands of literature. First, it contributes to the literature on competition in developing countries.[6] Busso and Galiani (2019) and Bergquist and Dinerstein (2020) use randomization to obtain variation in the number of entrants across markets and study the effect of competition on prices and markups.[7] Different from these two papers, where competing firms are fairly similar, in Jensen and Miller (2018) boat builders are heterogeneous in quality, and market integration leads to better outcomes for high-quality builders and exits for low-quality ones. In Bao and Chen (2018) and Atkin et al. (2018) domestic firms compete against multinationals. This paper adds to this literature by providing evidence on how microenterprises compete against large firms. I show how one of the smallest firms in Mexico, the shop, responds to competition from some of the largest companies in the country that operate hundreds, even thousands, of small convenience stores using a novel instrument and micro-level data of close to two million shops across twenty years.

The second literature this paper relates to is barriers to small-firm growth in developing countries. Karlan et al. (2015), Bruhn et al. (2018), Bloom et al. (2012), and McKenzie and

[6]In developed countries, the literature on entry and competition initiated by Bresnahan and Reiss (1991, 1990) is more extensive. This paper is most closely related to prior work that has studied the effects of increased competition in retail markets and the expansion of Walmart in the United States (Jia (2008); Matsa (2011); Basker (2005); Basker and Noel (2009); Hausman and Leibtag (2007); Holmes (2011); Haltiwanger et al. (2010)).

[7]Papers in this literature also include, for example, Macchiavello and Morjaria (2020) that study the effect of competition on relational contracts.

Woodruff (2017) study the role of consulting services and management practices, Alfaro-Urena et al. (2019) and Atkin et al. (2017) study access to international buyers, Atkin et al. (2017) study technology adoption, and De Mel et al. (2008) and Fafchamps et al. (2014) study access to finance and capital. The heterogeneous effects of competition from chains on shops highlight a potential understudied trade-off for small firm growth. On the one hand, when shops grow, they may access more customers and exploit economies of scale. On the other hand, they might lose comparative advantages from being small and owner-operated that differentiate them from large chains and allow them to survive.

The following section provides background information on shops, chains, and competition between them. In Section IV, I describe the data sources and document that chains have regional economies of scale and that within cities, chains are more than three times as likely to open on wide streets. Section V presents the empirical strategy and discusses potential concerns to the validity of the instrument. In Section VI, I present the main results. Section VII includes ample robustness checks for alternative instrument specifications, for including different sets of controls, and alternative standard errors. In Section VIII, I discuss the heterogeneity of the effects on shops and potential consumer welfare implications across the income distribution. The last section concludes.

1.2 Background

One out of ten firms and one out of four retail firms in Mexico is a neighborhood shop. They are, on average, 28 square meters, employ 1.7 people (mostly owner and family), and are frequently located next to the owners' house.[8] Shops use different sourcing channels to offer a wide variety of products. Large producers deliver packed, branded, and standardized products such as bread, dairy, cold meats, sodas, beer, and snacks directly to shops. Dif-

[8]Sources for this and next paragraph: Economic Census (2019); COFECE (2020)

ferent than for these products, shops source fresh fruit and vegetables from central markets known as *centrales de abasto* and also offer products that they make themselves (e.g., bread, sandwiches, pastries) or that they source from nearby bakeries and *tortillerias*. Shops in Mexico are often perceived as more than just a store:

> For most Mexicans, shops are much more than a purchase location ... they are places where it's possible to find what we need because the owner knows us to perfection. The owners' relationships with the people make them a central link of the community ... [the shops] have also been, since always, the meeting place of neighborhoods. ... In them we learn about solidarity, personal finance, and trust in one's word.
>
> Coca-Cola Mexico (2020)

Oxxo, 7-Eleven, Circle K, 3B, Dunosusa, and Tiendas Neto are the most prominent chains that have rapidly expanded in the last two decades, reaching more than 22,000 stores in 2019. Chain stores are between 20 and 50 square meters (plus parking), employ between 6 and 10 people, and stock between 300 and 800 SKUs. Even though chains and shops are often within a couple of meters, chains are in high-traffic locations next to wide streets.[9] Chains are perceived as very successful. In particular, the largest of them, OXXO, is perceived as the most successful food and beverages retailer in the country (above Walmart Mexico).[10] OXXO's annual sales per squared meter are 89,500 MXN, above Walmart Mexico's 88,800 MXN,[11] and OXXO's EBITDA margin, a measure of operating profitability, is 11.2% (larger than that of all large supermarkets in Mexico, including Walmart with an EBITDA margin of 9.3%). Even though large producers deliver their products directly to both chains and shops, one of the keys to chains' success is their logistics operations with distribution centers around the country that facilitate the distribution to each store in the chain of fruits, vegetables, liquor, medicines, cellphones, party supplies, among other products.[10]

[9]Source: Milenio (2016); El Universal (2015)
[10]Source: El Financiero (2014)
[11]Source: El Financiero (2020)

Chains are a direct competitor to shops because they are similar in size, capture incidental purchases of consumers, and have a significant overlap in their product offering. Additionally, chains may represent an existential threat to shops because they have advantages in scale that allow them to share costs across stores, have more bargaining power with suppliers, have a lower cost of capital, and invest in productivity-enhancing technologies. Chains also represent lower search costs to consumers because they have uniformity across same-chain stores, are located on wide streets, have big signs, and operate 24/7. According to government officials, chambers of commerce, and market research companies, between 5 and 35 shops close for each additional chain-store.[12] However, chains could have a limited effect on shops because shops may have comparative advantages in building relationships with consumers, tailoring their product offering, offering informal credit to the neighbors, and providing similar prices.[13]

1.3 Model

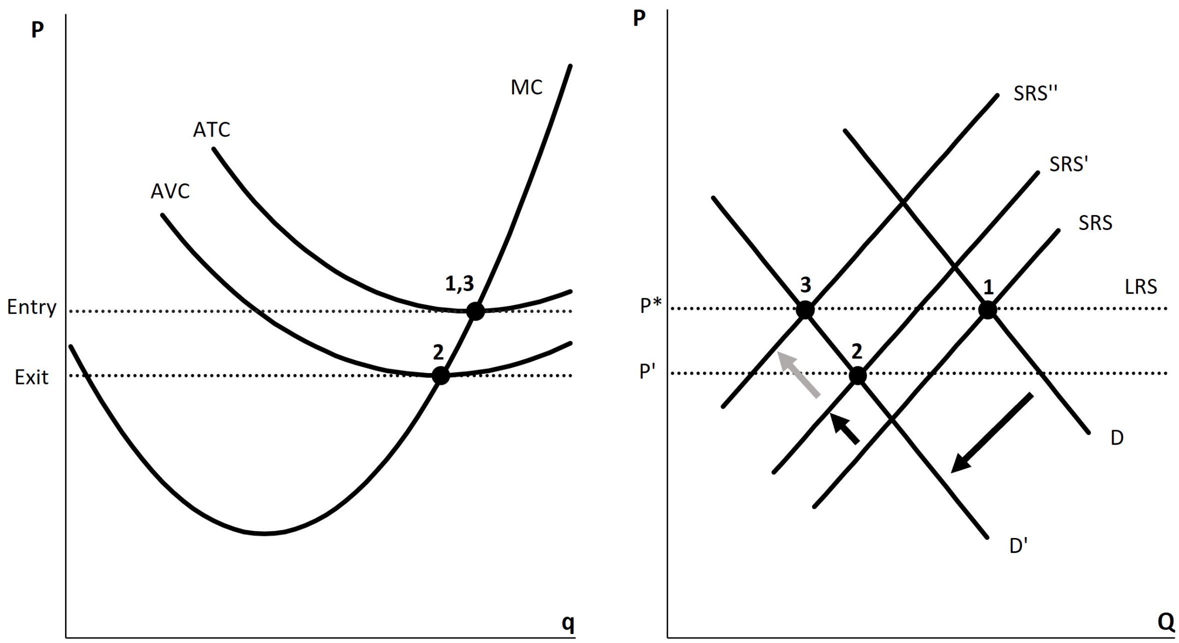
This section presents a standard model of differentiated competition consistent with the adverse impact on shops occurring mainly along the extensive margin and the reduction in number of shops being driven by a reduction in shop entry.

Consider a competitive industry with many homogeneous firms (i.e., all shops in a given neighborhood), each facing sunk entry costs and standard u-shaped marginal and average costs. Assume free entry of firms and a high exogenous exit rate due to a fraction of them facing a sizable idiosyncratic shock, e.g., the owner's death.[14] I model the arrival of chains, an imperfect substitute, as a downward shift in industry-level demand (i.e., demand from all consumers in the neighborhood). Figure 1.1 depicts the cost curves of a representative shop

[12]Source: El Universal (2017); El Universal (2015); El Financiero (2018)
[13]Source: Gonzalez Sanchez and Gaytán (2015); Milenio (2016)
[14]This assumption is consistent with a 10% yearly exit rate.

on the left side, and the neighborhood level supply and demand curves on the right side.

Before chains' entry (point 1), the equilibrium price is given by the intersection of the short-run supply (SRS) and demand curves, which is also equal to the minimum average total cost (ATC), inclusive of the sunk entry cost. At this price, potential entrants are indifferent between entering or not. Because price is above average variable cost (AVC), incumbents have short-term economic profits. Assume this equilibrium is a steady-state with new firms replacing those that exit due to their idiosyncratic shocks.[15]



Firm (Shop) Level | Industry (Neighborhood) Level

Figure 1.1: Differentiated Competition with Entry Costs

Note: The figure on the left contains the marginal cost (MC), average variable cost (AVC) and average total cost (ATC) curves. The sunk entry cost drives the difference between the ATC and AVC. The figure on the right plots the transition from the long term equilibrium (1) to a short term equilibrium (2) caused by the entry of a differentiated competitor shifting the demand curve from D to D'. At (2), firms that face the idiosyncratic shocks exit, but new firms do not enter. This exit without replacement leads to a shift upward of the supply curve from SRS to SRS' and a new long run equilibrium in (3).

Now suppose chains enter. Provided the resulting downward shift in demand is large relative to the sunk entry cost, the intersection of the new demand curve and the SRS curve

[15]Firms that face the idiosyncratic shock exit, shifting the short-run industry supply curve up, increasing the equilibrium price, and making entry profitable. Entry shifts the short-run supply curve back down until the potential entrants are indifferent between entering or not, and the price returns to its long-run equilibrium (ATC = MC).

will occur below the minimum AVC.[16] In this case, shops face short-term losses and begin to exit. This process shifts up the SRS curve until it intersects the demand at a price equal to the minimum AVC, point 2, where incumbents are indifferent between exiting or not. This new short-run equilibrium has lower price, profits, and revenue.

As time progresses, some shops exit due to their idiosyncratic shocks, but new firms do not replace them because the price is below the minimum ATC. These exits without replacement gradually shift up the short-run supply curve until the price equals the minimum ATC (point 3). This new steady-state differs from the first (point 1) at the neighborhood level because it has lower profits and revenue. Provided the fraction of shops facing idiosyncratic shocks has not changed, there will be fewer exits and entries in the new steady state because fewer shops exist.

The empirical section quantifies in detail the predictions of the model. As a preview:

1. The decrease in number of shops, represented by the distance between the steady-state before chains (point 1) and the steady-state after chains (point 3), is 17%.
2. The total effect on shop entry has two components with the same sign. The first is the lack of entries between the short-run equilibrium (point 2) and the new steady-state (point 3), and the second is the effect of moving from the steady-state without chains (point 1) to the steady-state with chains (point 3), where there are fewer shops and less entry and exit. The net of these two effects is a decline in shop entries by 18%.
3. The total effect on shop exit also has two components but with opposite signs. The first is the exits caused by chains' entry making some shops unprofitable (moving from point 1 to 2), and the second is the effect of moving from the steady-state without chains (point 1) to the steady-state with chains (point 3), where there are fewer shops and less entry and exit. The latter effect dominates the former, and the net effect of

[16] Alternatively, suppose the intersection of the new demand curve and the SRS curve occurs above the minimum AVC (not depicted). In that case, incumbents' profits decrease but not enough to incur short-term losses and exit.

chains is a decrease in number of shop exits by 8%.

4. The reduction in profits at the neighborhood level is 27%, represented by the distance between points 1 and 3 at the industry level times the vertical distance between point 3 and the AVC at the shop level.
5. The negative effect on profits at the shop level is less than a third of that at the industry level, 7%, because the reduction in number of shops compensates for the adverse impact of chains.[17]

1.4 Data

1.4.1 Sources

The three most important data sources for the paper are: i) Economic Censuses (1999, 2004, 2009, 2014, 2019) collected by the Mexican Statistics Institute (INEGI), ii) Income and Expenditure Surveys (2006, 2008, 2010, 2012, 2014, 2016, 2018) collected by INEGI, and iii) Open Street Maps.

The Mexican Economic Censuses cover all the firms in the country without any restriction,[18] and the confidential part includes microdata on, among other variables, revenue, profits, employment, investment, operations, and location. The Economic Censuses classify the establishments according to the North America Industrial Classification System for Mexico (SCIAN), which has subtle differences that represent a significant advantage relative to the North America Industrial Classification System for the United States (NAICS). Unlike the NAICS, with a code for supermarkets (445110) and one for both convenience stores and shops (445120), in the SCIAN classification, shops, chains, and supermarkets have different

[17] With current assumptions, surviving shops are as well-off after the entry of chains. The model can be extended to allow heterogeneity in shops, for example, in their entry cost. This heterogeneity would lead to the long-run supply having a positive slope and chains' entry having a negative impact at the shop level.

[18] Include both formal and informal firms without minimum size requirements.

codes, 461110, 462112, and 462111, respectively. I further classify establishments with the 462112 code, composed of convenience stores (minimarkets), into two categories based on ownership: firms with more than 100 establishments as chains and those with only one store as hybrid stores. In number, hybrid stores are equivalent to 3% of neighborhood shops and convenience chains. I do not include hybrid stores in the analysis, except when comparing the effect of chains on hybrid stores and neighborhood shops.

Starting in 2009, INEGI added an establishment identifier to the Economic Censuses. To track establishments before 2009, I use the establishment identifiers created by Busso et al. (2018). The result is an establishment-level panel from 1999 to 2019.

The biyearly Income and Expenditure Surveys (ENIGH) of 2006-2018 contains data on what households buy, where they buy it, and how they pay. The sample of the ENIGH has grown throughout the years. For 2006, it contains responses for little more than twenty thousand households, and by 2018 it included responses for more than seventy thousand.

INEGI's geostatistical framework for urban Mexico divides the country into states, then into municipalities, then into localities, then into urban census tracts (AGEBs). The data has between 37,000 and 47,000 AGEBs (depending on census year) with an average size between 25 and 50 blocks, 650 households, and 2,000 people. AGEBs are perfectly delimited by streets, avenues, or any other trait easily identifiable in the field. INEGI designed the AGEBs to facilitate the data recollection process by enumerators in the field. Figure A.3 displays the frequency distribution of AGEBs by the number of chain stores and shops.

There are two additional data sources. INEGI's geostatistical framework and the Population Censuses (2000, 2010). I use INEGI's geostatistical framework to overlay the map of AGEBs and streets. Then, I compute the length of each street type within each AGEB. I use the Population Censuses of 2000 and 2010 for i) alternative specifications that predict the number of chain stores using machine learning instead of the IV, and ii) for specifications that use population census data as controls. I present the results for these specifications in the robustness section.

After merging the different data sources, the sample includes the most populous 655 municipalities in Mexico, which have an average population of 115,000. The distribution of these cities by size is: i) small: 508 towns with an average population of 37,000, ii) medium: 120 towns with an average population of 262,000, and iii) large: 29 towns with an average population of 880,000.

I use AGEBs to construct neighborhoods. I draw a buffer of 1km from the center of each AGEB. I define a neighborhood as the union of AGEBs that overlap with each buffer.[19] On average, there are 12 AGEBs in each neighborhood, 370 blocks, 30,000 people, 68 shops, and 5 chain stores. The robustness section shows results using alternative definitions of neighborhoods created with different buffer sizes (0km, 0.25km, 0.5km, 0.75km, 1.25km, 1.5km, and 2km). There are two reasons for using a buffer larger than 0km (running the analysis at the AGEB level). The first one is statistical power, which is only an issue when using ENIGH data. Using neighborhood fixed effects limits the sample to neighborhoods where households were interviewed by the ENIGH at least two times. The larger the neighborhood size, the more neighborhoods meet this condition. The second reason is to ensure that the neighborhood is large enough to capture the effect on all the shops affected by the entry of a chain store so that spillovers to other neighborhoods do not bias the estimates.

Using Open Street Maps, I classify trunk, primary, secondary and tertiary streets as wide and the remaining categories not wide. Of the resulting classification, 21% of the total street length is wide, and of the remaining 79% of non-wide, 95% is residential streets. I construct a measure of the prevalence of wide streets by adding the lengths of all wide streets in the neighborhood and normalizing it by its size, specifically dividing by the square root of its area. This measure ranges from 0 to 63 (less than 1% of neighborhoods have 0), the average is 10, and the standard deviation is 7.6.

[19]Figure A.5 has a visual representation of how neighborhoods are constructed.

1.4.2 Summary Statistics

Table 1.1: Summary Statistics Shops and Chains

	Shops	Chain Stores
Number	1,787,952	42,101
Annual Profits (000's MXN)	67	1,689
Annual Revenue (000's MXN)	251	10,098
Expenses (000's MXN)	183	8468
Value Added (000's MXN)	69	2,027
Total Employed	1.8	6.4
Profits per Worker (000's MXN)	40	329
Revenue per Worker (000's MXN)	145	2,111
Initial Resale Inventory (000's MXN)	11	514
Final Resale Inventory (000's MXN)	12	679
Fixed Assets (000's MXN)	66	1,928
Publicity (000's MXN)	0.1	32
HH Purchase Probability (Week)	0.85	0.16
HH Purchase Probability | Purchasing in Chain (Week)	0.68	1
HH Number of Days Visited per Week	4.00	0.30

Source: Economic Censuses and Income and Expenditure Surveys

There are stark differences between shops and chains. On average, chain stores, relative to shops, have 30 times the revenue, 25 times the profits, four times the employees, six times the profits per employee, and seven times the revenue per employee. Chains are 2-5 times larger in squared meters than shops (4-10 times with parking), yet, this difference in physical size is not enough to explain the differences in profits and revenue.[20] Eighty-five percent of households purchase at least once a week in shops. The probability of buying in a chain is significantly lower, 15%. However, for households who purchase in chains, the probability of purchasing in shops is 17 percentage points lower than for the average household, consistent with these two types of establishments being substitutes. The shops' exit rate from one census to the next is 40%, implying a 10% annual exit rate. For chains, the yearly exit rate

[20]Chain stores are, on average, 187 squared meters or 420 squared meters, including parking. shops are between 40 and 100 square meters.

is below 3%.

Shops and chains have a significant overlap in their product offering. As Figure 1.2 displays, the five most popular products for shops (sodas, milk, eggs, tortillas, and bread), which represent almost half of their revenue, are also available and among the 12 most popular food products in chains.

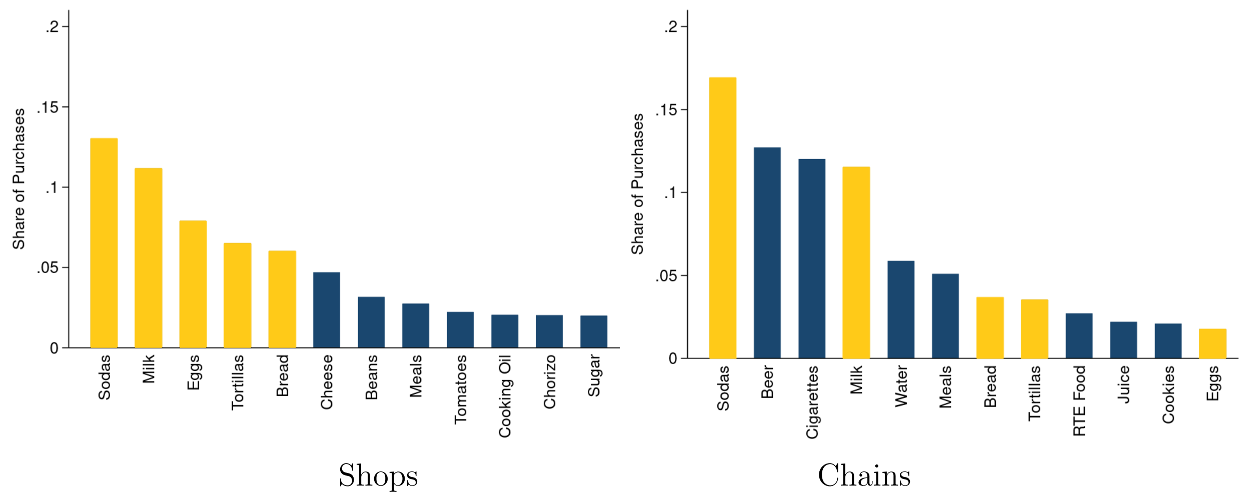


Figure 1.2: Share of Store Sales for Top 12 Products

Source: Income and Expenditure Survey (ENIGH 2018)

A limitation of the consumption data is that it does not include information on quality or brand, which makes price comparisons between chains and shops inconclusive. To partially address unobservable differences in quality, I include household fixed effects in the price paid in chains and shops comparison. Figure A.4 presents the estimates of the difference of price paid and size purchased in chains and shops. For most goods, the differences in prices and sizes are not statistically significant. For other goods, a volume discount might explain the price difference. For example, sodas are cheaper per liter in shops, but households also purchase larger sizes in shops. Similarly, rice is more expensive per kg in shops, but households purchase smaller rice bags in shops.

1.4.3 Importance of Economies of Scale and Wide Streets for Chains

I use two key differences between chains and shops to construct the instrument. The first one is that chains have advantages in opening stores in nearby cities. Some of these advantages are cost-sharing in transportation, marketing, distribution, and overhead costs that generate regional economies of scale. Other potential benefits include specialization and brand building. If these advantages of opening in nearby cities are significant, chains will open stores in cities close to each other to exploit them. The map in Figure 1.3 shows that chains' store openings between 2016 and 2020 present a spatial correlation consistent with advantages from opening stores in nearby cities.

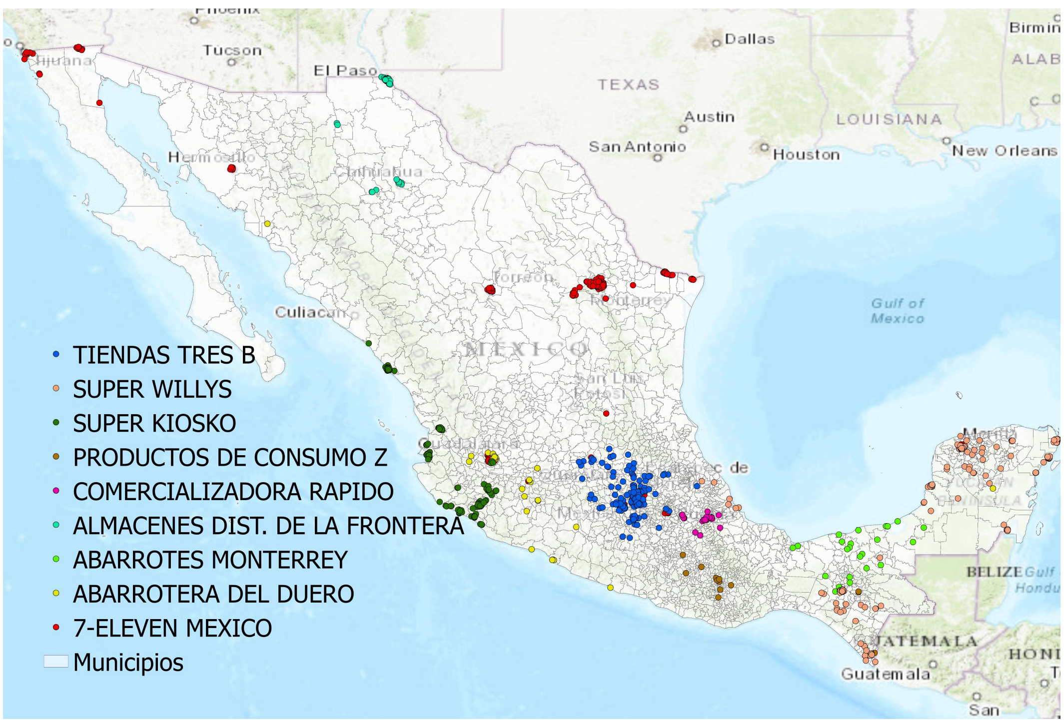
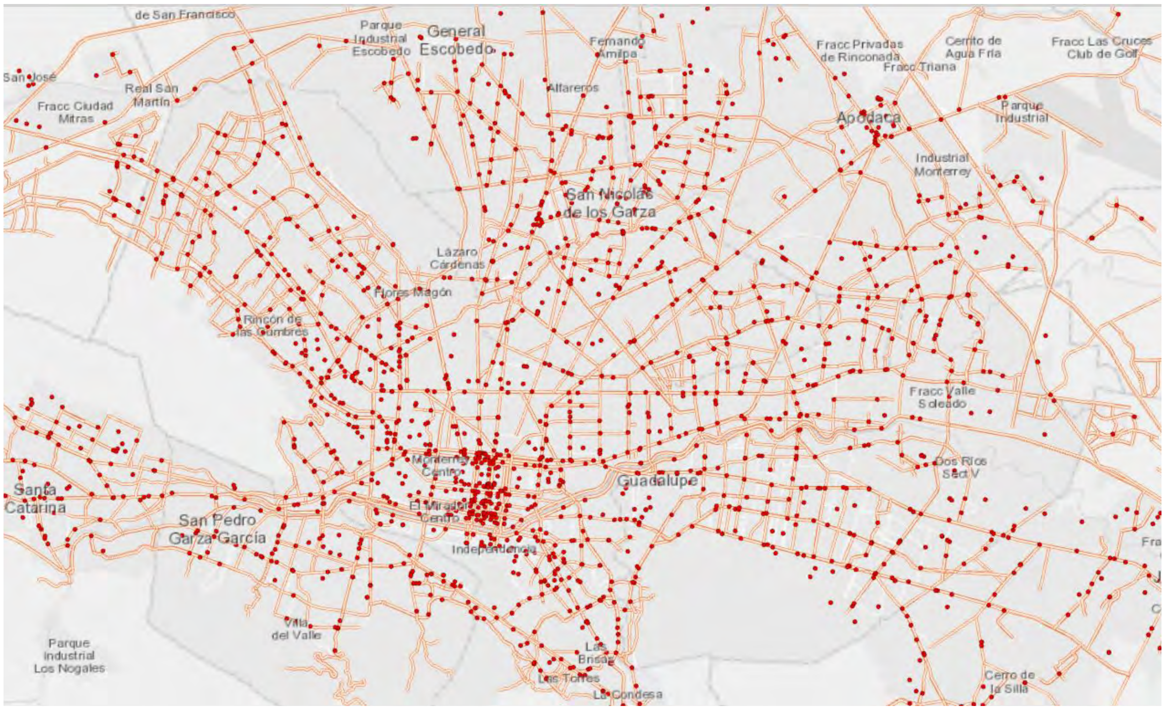
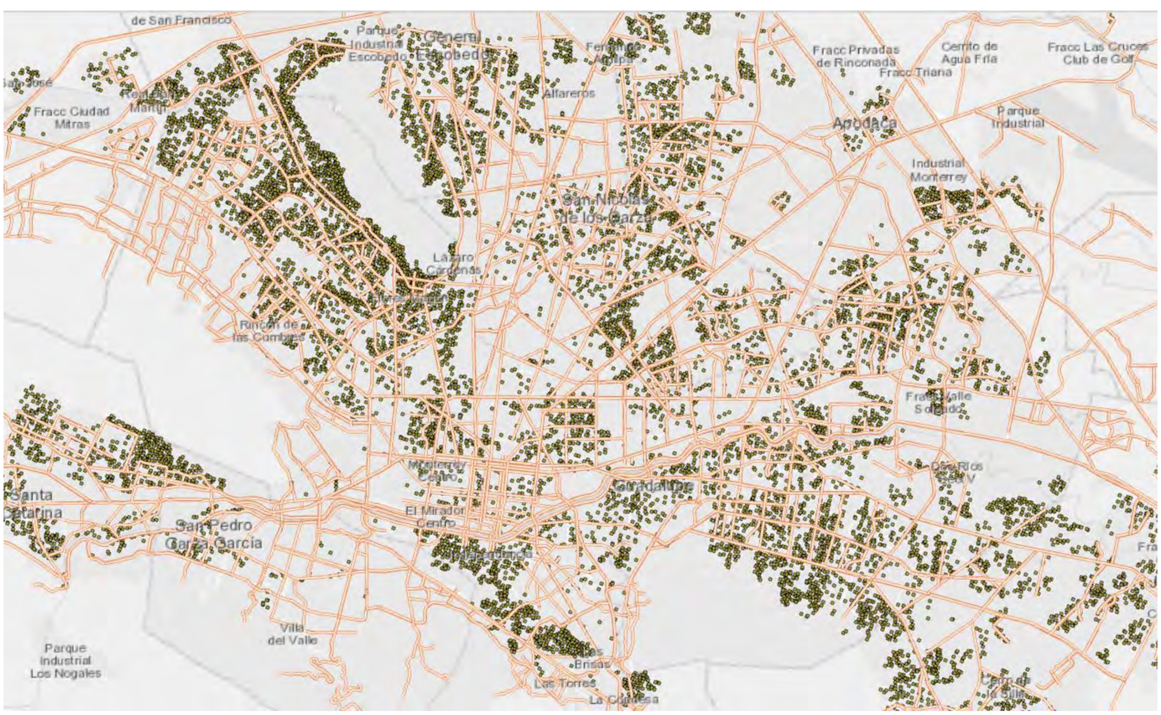


Figure 1.3: Spatial Correlation in Store Openings by Chain (2016-2020)

Note: The map plots the chains' openings between 2016 and 2020 by chain using data from DENUE 2020. Only nine of the largest twenty chains by number of stores in the country are used in the map for exposition purposes.



Wide Streets and Chains



Wide Streets and Shops

Figure 1.4: Wide Streets, Chains, Shops of Monterrey

Source: Open Street Maps, DENUE 2020
Note: The maps plot wide streets and chain stores or shops. Wide streets are those classified as Trunk, Primary, Secondary, and Tertiary on Open Street Maps.

The analysis in Appendix A.2 measures the relevance of these advantages on determining the number of stores each chain has in a city. In particular, after controlling with firm-city, year-city, and year-firm fixed effects, 19 additional same-chain stores in nearby cities are associated with one more store in the city. Stores in nearby cities account for 9% of the total variation in the number of stores each chain has in a city.[21] These advantages of opening stores in nearby cities are firm-specific: the positive correlation dissipates when using the number of different-chain stores (competitors) in cities nearby. Moreover, the number of competitors in nearby cities accounts for less than 0.0001% of the variation in the number of stores each chain has in a city.

The second key difference is that even though both stores coexist a couple of meters away, shops are usually located next to the owners' houses, and chains are located next to wide streets to target car and bus traffic customers. Chains also place big signs, offer parking spots, and provide a speedy process to enter, purchase, and leave. If traffic customers are essential for chains, they will mainly locate on wider streets to target them. The top map of Figure 1.4 shows that chains are almost exclusively located on wide streets, and the bottom map shows this is not the case for shops. More formally, Figure A.6 displays the distribution of distance from each store to the closest wide street. Almost 80% of chain stores are within 25 meters of a wide street, while only 20% of shops are this close to a wide street.

1.5 Empirical Strategy

Chains' entry time and location across cities and within cities (across neighborhoods) are endogenous to shops' outcomes. The endogeneity arises from joint determination: neighborhoods with higher demand for products offered by both store types have better outcomes for shops and are more attractive for chains. This positive correlation in demand leads to

[21] The 9% is the R-squared of the model after demeaning by all the fixed effects.

an upward bias of the effects of chains on shops if estimated using OLS.

I control for time-invariant neighborhood characteristics and city-wide trends using city-year and neighborhood fixed effects. However, fixed effects do not control for neighborhood-level unobservable shocks. For example, a new park may increase foot traffic and demand for drinks and snacks for both store formats. If I compare neighborhoods where chains enter to those where they do not, I would implicitly compare neighborhoods that received a positive demand shock for shops to those that did not. To address this issue, I use an instrument that reduces the costs and increases the profitability of chains, but not of shops.

As shown in Equation 1.1, the instrument exploits two key differences between chains and shops: i) chains have regional economies of scale,[22] and ii) chains locate on wide streets. The left of the interaction is a Herfindahl–Hirschman Index without normalization that measures regional economies of scale and increases in both the number of chain stores in nearby towns and their concentration. Specifically, it is the square root of the sum of the squared number of stores per chain in nearby cities.[23] Nearby cities are those adjacent to the city and cities adjacent to those (1^{st} and 2^{nd} degree neighbors).[24] Instead of a measure of regional advantages that aggregates across firms, it is possible to use one measure and one instrument for each chain. The robustness section shows that the results are almost identical when using one measure and one instrument per chain, but the main specification has the advantage of a stronger first stage.

[22] Jia (2008) and Holmes (2011) have used this intuition to model the expansion of Walmart in the US.

[23] In the Robustness section, I repeat the estimation, but i) without squaring the number of chain stores and without taking the square root, and ii) without taking the square root. The main specification has the advantage of having a conceptual link to the Herfindahl–Hirschman index and a stronger first stage.

[24] The robustness section presents results using only 1^{st} degree neighbors and also using 3^{rd} degree neighbors. The results are similar and consistent.

$$Z_{n,c,t} = \underbrace{\left(\sum_f (\#StoresNearbyTowns_{f,c,t})^2\right)^{1/2}}_{\text{Economies of Scale}_{c,t}} \times \underbrace{\frac{Total\ wide\ streets\ length_{n,c}}{Area_{n,c}^{1/2}}}_{\text{Prevalence of Wide Streets}_{n,c}} \quad (1.1)$$

The advantages of opening stores in nearby cities provides variation at the city and year level, but they do not predict where, within cities, new chain stores will locate. To predict the location within cities, I construct a measure of suitability for chains based on the prevalence of wide streets in the neighborhood, which is the right side of the interaction in Equation 1.1. I add up the total length of wide streets in each neighborhood and normalize it by dividing it by the squared root of the neighborhood area.[25] The instrument is the interaction of the regional advantages measure and the suitability measure, and it captures that when chains open stores in nearby cities, suitable locations this city become more attractive for chains. The instrument only uses variation from the interaction of the measures; two-way fixed effects absorb the individual components. Figure A.7 displays the increasing relationship between the instrument and the number of chain stores in the neighborhood.

Equations 1.2 and 1.3 are the first and second stages of the 2SLS estimation.

$$CS_{n,c,t} = \gamma_1 Z_{n,c,t} + \zeta_{n,c} + \eta_{c,t} + \mu_{n,c,t} \quad (1.2)$$

$$Y_{n,c,t} = \beta_1 \widehat{CS}_{n,c,t} + \zeta_{n,c} + \eta_{c,t} + \epsilon_{n,c,t} \quad (1.3)$$

where n denotes neighborhood, c denotes city, t denotes census year, and f denotes firm.

[25] I divide by the squared root of the neighborhood area so that the prevalence of wide streets measure captures the density of wide streets. Neighborhoods have different sizes because they are unions of census tracts. I use the square root of the area so that the numerator and denominator units are in meters.

Equations 1.2 and 1.3 include neighborhood fixed effects, $\zeta_{n,c}$, and city-year fixed effects, $\eta_{c,t}$. CS stands for the number of chain stores and $Y_{n,c,t}$ is the outcome of interest; for example number of shops, revenues, profits, neighbors expenditures in shops, prices paid by neighbors, etc. I cluster the standard errors at the city level, because the measure of advantages from regional expansion of chains varies at the city level.[26]

The exclusion restriction is that when chains increase the number of stores in nearby cities, it only affects shops in neighborhoods suitable for chains by increasing the probability of a chain store entering their neighborhood. A possible concern with the IV is that an increase in the number of chain stores in nearby cities is associated with an overall increase in the number of chain stores in this city. If consumers travel outside their neighborhoods, more chains in the city imply that consumers are more likely to purchase in chains, even if chains are not in the neighborhood where they live. City-year fixed effects capture this potential increase in purchases in chains because it affects all the customers in the city.

However, the overall rise in the number of chain stores in the city would still be an issue if it affects customers differentially based on the prevalence of wide streets in the neighborhood they live. It is impossible to test whether this indeed happens. However, it is possible to test whether the instrument correlates with certain variables that potentially affect the likelihood of this phenomenon occurring. In particular, I test whether the instrument correlates with customers' characteristics that likely affect demand and the likelihood of purchasing outside their neighborhood. Figure A.8 shows that there is no relationship between the instrument and household characteristics, such as the number of cars, the probability of having a vehicle, labor income, total income, income per capita, monetary expenses, and household demographics. Since the instrument does not correlate with these variables, it is unlikely that the overall increase in the number of chain stores in the city affects consumers

[26]The robustness section presents results clustering at the neighborhood level, city and year level, city x year level, and city x year and neighborhood level. It also presents results taking into account potential correlation of standard errors across adjacent cities.

differentially by the prevalence of wide streets in their neighborhood.

A related concern is that consumers might purchase in chain stores outside their neighborhood, leading to spillover effects where the chain affected both consumers and shops in the neighborhood of entry and adjacent ones. I re-estimate the main specification using eight alternative neighborhood sizes in 250 meters increments of buffer radius to address this concern. As Figure A.9 shows, the effect on the number of shops stabilizes between 1 and 1.25 km radius, consistent with the neighborhood size of the main specification being large enough to capture the full effect of an additional chain store.

There are still concerns that the IV cannot address. For example, suppose that 7-Eleven has so many stores in a region that it convinces Pepsi to stop selling to shops. Losing Pepsi products would affect shops through a mechanism other than additional chain stores in their neighborhood, violating the exclusion restriction. This example is not a concern because there is no anecdotal evidence of this happening, and the Mexican antitrust authorities ensure this type of practice does not occur.

1.6 Results

The first part of this section presents the estimation of the effects of chains on shops, including the impact on the number of shops, number of entries, number of exits, and performance measures like revenue, profits, and employment at the neighborhood and shop level. The second part presents the effects on neighbors consumption, including expenses on shops, number of visits to shops, probability of visiting shops, expenses by product category, and expenses by product.

Table 1.2: Effect of Chains on Number of Shops

	OLS			2SLS	Reduced Form	First Stage
Dependent Variable:	# of Shops	# of Shops	# of Shops	# of Shops	# of Shops	# of Chain
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Chain Stores	3.33***	-0.56***	-2.01***	-4.60***		
	(0.536)	(0.210)	(0.314)	(0.669)		
Economies of Scale$_{c,t}$ x					-6.49***	1.40***
Chain Suitability$_{m,c}$					(0.583)	(0.157)
Observations	190,664	190,664	190,664	190,664	190,613	190,613
Neighborhood FE		Y	Y	Y	Y	Y
Year x City FE			Y	Y	Y	Y
Clustered SE	City	City	City	City	City	City
Mean Dep. Variable | Chains>0	175	175	175	175	175	6
Mean Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.4	6.4
KP F-statistic				79.79		

Note: The table displays the estimation of Equation 1.3 using 2SLS. Columns 1-3 are OLS estimates (use the number of chain stores as independent variables), and column 4 is the IV estimate. In column 5 and 6 are the reduced form and the first stage estimates. Standard errors are clustered at the city level. The 2SLS models in the paper are estimating using the ivreghdfe command (Correia (2018)).

1.6.1 Effects of Chains on Shops

For each additional chain store in the neighborhood, the number of shops decreases by 4.6. Going from zero to the average number of chain stores in a neighborhood, 6.4, reduces the number of shops by 29 (17%).[27] Column 4 in Table 1.2 contains the second stage results of the main specification (Equation 1.3). Column 1 is an OLS estimation without fixed effects, where the joint determination problem is evident. Markets with higher demand have both more chain stores and more shops. In Column 2, I partially address the issue by introducing neighborhood fixed effects. In Column 3, I further address it by introducing year-city fixed effects. However, the effect in Column 3 still suffers from an upward bias due to neighborhood-level demand shocks that are common for both store formats. I address this issue in Column 4 that presents the 2SLS estimates using the instrument. Columns 5 and 6 report the reduced form and first stage estimates.

Table 1.3: Effect of Chains on Number of Entries and Exits

Dependent Variable:	Number of Entries	Number of Exits	Entry Rate	Exit Rate
	(1)	(2)	(3)	(4)
Number of Chain Stores	-2.03***	-0.80***	-0.003***	0.003***
	(0.523)	(0.228)	(0.001)	(0.001)
Observations	161,103	161,103	158,043	158,006
Neighborhood FE	Y	Y	Y	Y
Year x City FE	Y	Y	Y	Y
Mean Dep. Variable | Chains>0	70	69	0.38	0.41
KP F-statistic	96.53	96.53	95.58	95.62

Note: The table displays the estimation of Equation 1.3 using 2SLS. Standard errors are clustered at the city level.

The reduction in the number of entries is the primary driver of the decrease in the number of shops. For each additional chain store, the number of shop entries decreases by 2, and the number of exits decreases by 0.8. These estimates imply that going from zero to the average

[27] Conditional on having chain stores, neighborhoods have, on average, 6.4 chain stores and 175 shops.

number of chain stores in a neighborhood reduces entries by 19% and exits by 8%. These results are columns 1 and 2 of Table 1.3. The reduction in exits might appear surprising at first. However, as the model highlights, the effect of chains on the number of shop exits is ambiguous. On the one hand, it can make some shops unprofitable, increasing the number of exits. On the other hand, if shops' natural exit rate does not change and chains reduce the total number of shops through decreased entry, the number of shop exits will decreases because there are less shops.[28] The reduction in exits implies that the latter effect dominates the former.

Only observing entries and exits for firms that have been alive long enough to appear in a census is a limitation to this analysis. Two cases arise based on how long the entrant would have survived absent chains. The first case is short entries, where the firm was supposed to enter and exit between censuses. If chains decrease the number of short entries (the same way they reduce the number of regular entries), the effects on entries and exits would be upward biased because short entries are not counted but expected to be more in the "control" group. The second counterfactual is long entries, where the firm was supposed to enter before the census and exit after the census. If chains decrease the lifespan of these entries such that they also exit before the census, the effects on entries and exits would be downward biased because the entry and exit are not counted in the "treated" group.

I also estimate the effect on entry and exit rates and report the results in Columns 3 and 4. Ideally, the entry rate would capture the number of entrants among the potential entrants. I can not construct this rate because the number of potential entrants is unknown. Instead, I use the ratio of new shops to existing shops. Chains reduce the number of shops in the neighborhood, implying that the entry and exit rates will change through the numerator and the denominator. Going from zero to the average number of chain stores in a neighborhood

[28]Only existing shops can exit. Chains reduce the number of entries, reducing the total number of shops and exits as well.

reduces the entry rate by 2 pp (5%) and increases the exit rate by 2 pp (5%). I also estimate the effect on the probability of exit using survival models. Going from zero to the average number of chains in a census tract increases the exit probability of shops between 2.1 and 4.5 pp. Table A.1 reports the estimates of various survival models: Cox, Poisson, and Linear with different combinations of fixed effects and store-level controls.

Figure 1.5: Effects on Shops' Performance

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS where the dependent variable is the inverse hyperbolic sine of the sum or average variable in the row. The average is computed excluding firms that appear for the first and last time in the census. Figure A.10 computes de average without excluding these firms. Standard errors are clustered at the city level.

Consistent with the conceptual framework, Figure 1.5 shows that the negative effects on shops concentrate on the extensive margin. Going from zero to the average number of chain stores in a neighborhood reduces industry (neighborhood) level revenue and profits for shops by 27%.[29] There are similar effects for resale revenue, value-added, profits per worker, and inventories. These include the effect for shops that remain open and the effects through shops that closed or did not open. The shop level effects are less than 1/3 of the industry impact – the average profits and revenue of shops declines by 7%, consistent with the reduction in the number of shops mitigating the negative effects of chains at the shop level.

Interestingly, there is no effect on productivity measured as an output-input ratio, hours worked, total employed, and hours per worker at the shop level. These results might appear conflicting because there is no change in the output-input ratio even though shops have lower revenue and unchanged labor input. However, the measure of inputs is an accounting measure, meaning that it does not capture the owner's opportunity cost (its potential salary). Hence, shops adjusting to lower revenues by decreasing purchases of goods for resale is driving the null effect on productivity measured by the output-input ratio. Shops also reduce their inventories, which is a rational response that we would expect from a sophisticated manager concerned about performance metrics such as inventory turnover.

The average effect on shops can include composition or selection effects. I calculate the averages excluding shops that appear in the census for the first or the last time to reduce the effects from composition and selection. Figure A.10 displays the estimates without this restriction and the results are almost identical.

1.6.2 Effects of Chains on Neighbors' Consumption

Table 1.4 shows that most households continue to purchase in shops, but they frequent them less and spend less. An increase in chain stores from zero to their average number in the neighborhood decreases the probability of neighbors purchasing in shops by 4.5%, equivalently shops retain 95% of the customers that live near the shop. However, the average

[29]Tables A.2 and A.3 present the same results.

Table 1.4: Effect of Chains on Neighbors' Consumption

Dependent Variable: Consumption in Shops	I[Purchase]	Weekly Visits	Weekly Visits	Purchases ($)	Purchases ($)	Food Purchases ($)	Food Purchases ($)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of Chain Stores	-0.004**	-0.034**	-0.037**	-36.35***	-32.06**	-38.80***	-34.70***
	(0.002)	(0.015)	(0.016)	(12.06)	(13.71)	(11.34)	(12.96)
Observations	1,009,356	1,009,356	877,899	1,009,356	877,899	1,009,356	877,899
Year x City FE	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y
Conditional on Purchase in Week			Y		Y		Y
Mean Dep. Variable | Chains>0	0.86	3.9	4.6	2,566	2,967	2,313	2,683
Avg. Chain Stores | Chains>0	9.61	9.61	9.07	9.61	9.07	9.61	9.07
0 to Avg. # Chain Stores	-4.5%	-8.4%	-7.3%	-13.6%	-9.8%	-16.1%	-11.7%
KP F-Statistic	117.53	117.53	111.69	117.53	111.69	117.53	111.69

Note: The table displays the estimation of Equation 1.3 using 2SLS. Standard errors are clustered at the city level. Expenses are in Mexican Pesos (MXN).

number of days neighbors visit shops declines by 8%, of which 90% is from customers that continue to visit shops but do so 7% less.[30] The expenditure in shops declines by 13.6%, and 75% of this effect is from customers that continue to purchase in shops but spend less.[31]

Figure 1.6 displays a stark difference in the effect of chains on neighbors' consumption between non-fresh and fresh products. While losing sales in non-fresh and packed products like sodas, milk, eggs, cigarettes, sweet cookies, and juices, shops retain sales of fresh products like fresh sweet bread, tomatoes, fresh bread, potatoes, onions, and avocados. For products with fresh and non-fresh variations, the reduction in purchases is only for their non-fresh version. Such is the case of packed sweet bread, for which sales decline, while sales for fresh sweet bread and fresh bread do not change. Another example is spicy food. On the one hand, neighbors purchase less packed chilies and salsa from shops, but, on the other, they keep buying serrano, jalapeño, dry chilies, and additional salsa ingredients like tomatoes, green tomatoes, and onion.

Neighbors do not decrease their purchases of fresh goods even though there are 17% fewer shops, implying a potential increase in sales of fresh products per shop. This null effect on revenue of fresh products has relevant implications for the external validity of the findings. In particular, what would happen if the number of chain stores continues growing and it doubles or triples their current number? These findings suggest that shops will continue to lose revenue on non-fresh and standardized products, but they will retain revenue from fresh products, leading to the specialization of shops in fresh products.

There are several potential reasons for shops retaining their sales of fresh products. Shops might have an advantage over chains in offering these products ripe and fresh because of differences in sourcing: shop owners go to the central market or a nearby bakery every day

[30] I multiply the effect conditional on purchasing (-.037) times the share of customers that continue to buy in shops (84%) and divide it by the unconditional effect (-.034).

[31] I multiply the effect conditional on purchasing (-32) times the share of customers that continue to buy in shops (84%) and divide it by the unconditional effect (-36).

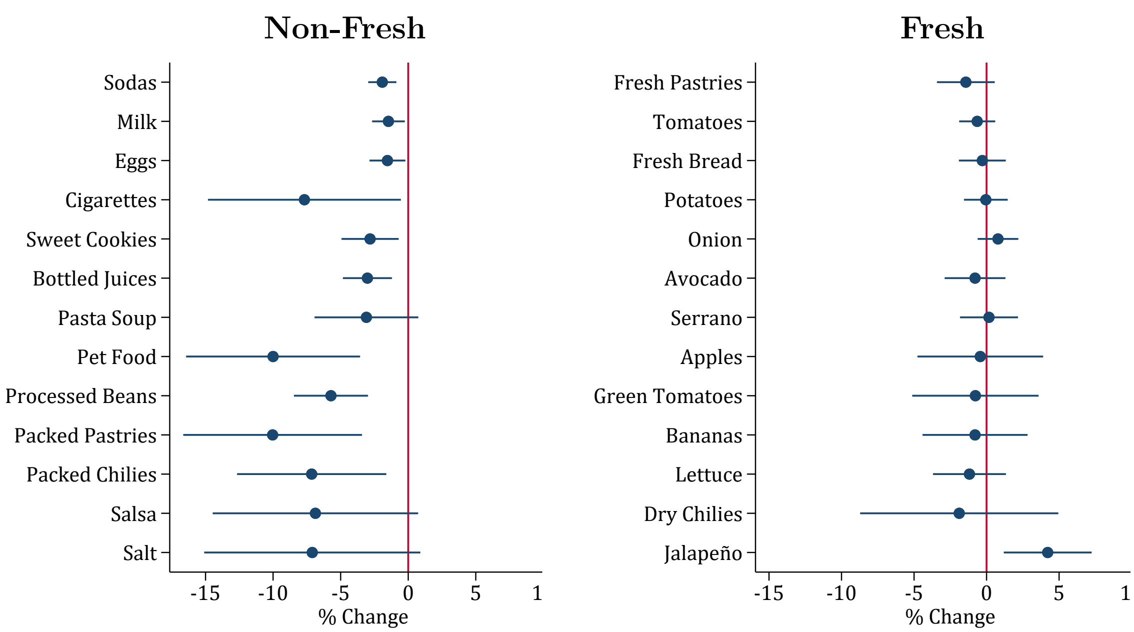


Figure 1.6: Effect on Neighbors Expenditure in Shops

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS replacing the dependent variable with household-level expenditure in pesos for each of the goods. The percentage change is computed by dividing the estimated effect by the household average expense in shops of that product. The effects are for each additional chain store, and on average, there are 9 chain stores in each neighborhood. Goods are sorted from top to bottom by their share of shops' revenue. For non-fresh goods, sodas represent 13% of revenue and salt 0.2%. For fresh goods, sweet bread represent 3% of revenue and jalapeño 0.2%.

and select these products. Freshness and ripeness are even more relevant in a context where consumers are cash and credit constrained and buy products to consume the same day, which is consistent with lower-income households purchasing more often and a larger share of their food in shops (see Figure 1.9). The next section discusses in more detail comparative advantages that allow shops to survive.

1.7 Discussion

This section provides evidence on why shops survive and then discusses potential consumer welfare implications of the increased competition from chains. First, I show that the adverse effects of chains on shops are smaller for owner-operated and small shops. Then, I will present results consistent with anecdotal evidence that suggests that these shops have comparative advantages in building relationships with their customers, facing lower agency costs, and screening their neighbors to provide them with informal credit to buy in the shop. The second subsection discusses potential welfare implications by studying the consumption patterns in chains and shops across the household income distribution.

1.7.1 Shops' Comparative Advantages

I explore whether the effects of chains on shops vary based on shops' size and management type. First, I compare the effect of chains on neighborhood shops and hybrid stores. Hybrid stores are different from shops, because they are less differentiated from chains, which is why they have the same classification code as chains in the Economics Censuses (different from the one of shops). In particular, hybrid shops hire employees, are similar in size to chains stores, have larger catchment areas, and are more likely to provide parking spots.[32] However, hybrid stores are different from chains because their owners only have one store.

[32]Figure A.18 displays an example of a hybrid store.

This is the only analysis in the paper that includes hybrid stores.

Figure 1.7 compares the effects of chains on neighborhood shops and hybrid stores using the interaction of the number of chain stores with a dummy representing hybrid stores as a second endogenous variable and the interaction between the instrument and the same dummy as a second instrument. The data is at the neighborhood, city, time, and shop type level. Hybrid stores are significantly more affected by chains. At the neighborhood level, their drop in profits and value-added are 51% and 53% larger than for neighborhood shops. At the shop level, the differences are starker. The additional reduction for hybrid stores in profits is 142%, in hired employees is 173%, in investment 353%, and between 182% and 273% in inventories.

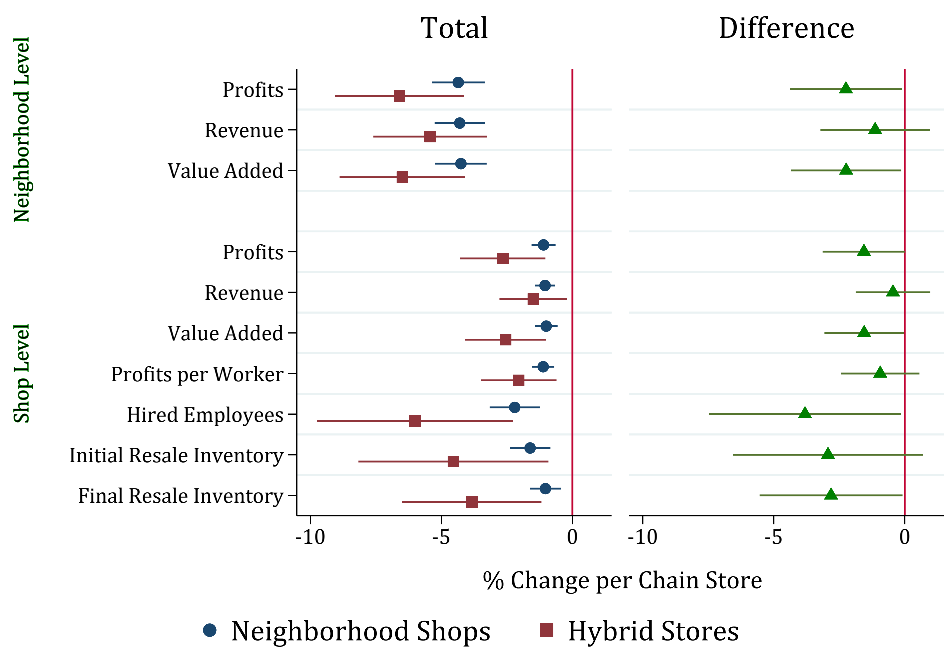


Figure 1.7: Effects of Chains on Shops' Performance by Shop Size

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS but adding i) the interaction of number of chain stores and a dummy variable for whether the average/sum is for a hybrid store to the second stage and ii) the interaction of the instrument and the same dummy to the first stage. The dependent variable in Equation 1.3 is the inverse hyperbolic sine of the row label.

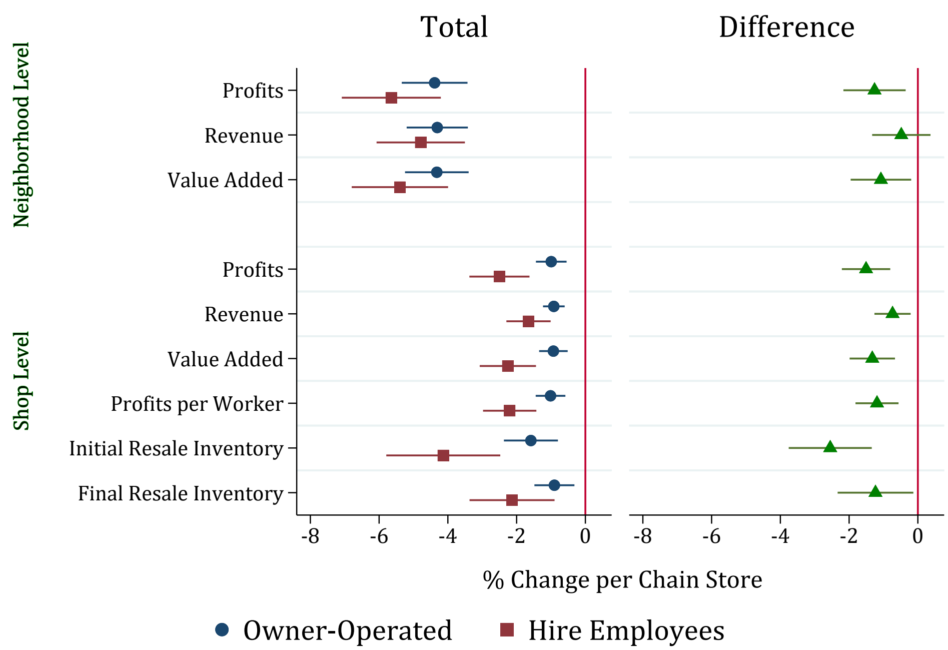


Figure 1.8: Effects of Chains on Shops' Performance by Type of Management

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS but adding i) the interaction of number of chain stores and a dummy variable for whether the average/sum is for an owner-operated shop to the second stage and ii) the interaction of the instrument and the same dummy to the first stage. The dependent variable in Equation 1.3 is the inverse hyperbolic sine of the row label.

Anecdotal evidence discussed in the background section highlights that shops may have comparative advantages over chains stemming from being owner-operated. The customer experience is likely better if customers purchase directly from the owner, who they know and is often a friend and neighbor. These relationships could also allow shops to tailor their product mix to match customers' tastes and offer their neighbors store credit when they can not pay for products. In Figure 1.8, I compare the effect of chains on shops that hire employees (7% of all shops) and on owner-operated ones. Shops that hire employees have a larger decline in neighborhood-level profits by 28%, and at the shop level, in profits by 151%, revenue by 80%, and in profits per worker by 117%.

Owner-operated shops also have lower agency costs because the owners' incentives, different from those of employees, are naturally aligned with what is best for the business because the owner is the residual claimant. Lower agency costs are more advantageous the higher the effort required to perform a task or the harder it is to monitor it. Consistent with shops having advantages from lower agency costs, Figure 1.6 shows that shops retain their sales of fresh products like fresh bread, fruit, and vegetables. These products require higher effort in sourcing because large producers do not deliver them directly to shops, and shop owners need to go to the central market every day to select them. Moreover, these products are not standardized and offering them fresh and ripe requires additional effort. For example, different from a can of Pepsi that has no variation in quality, a banana may be too green or brown. The results in Figure 1.6 also imply that shops specialize in these fresh products because sales of these goods do not decrease while sales of standardized products do.

Relationships between shops and neighbors are likely to be stronger if the owner is also a neighbor. Even though information on where the owner lives is unavailable, I can use whether the shop pays rent or not as a proxy of whether the shop is located in the owner's house or not. Shops located in the owners' houses are different in two ways. First, they are likely to have stronger relationships with the neighborhood because the owner is also a neighbor. Second, they might be physically closer to their neighbors because they are on residential streets. Figure A.14 shows that the negative effects of chains are significantly larger for shops that pay rent. At the neighborhood level, the reduction for shops that pay rent relative to those that do not is 45% larger in profits and 34% larger in revenue. At the shop level, it is 125% larger in profits, 78% larger in revenue, and 125% larger in profits per worker.

Relationships between the shops and the neighbors are hard to measure, but they likely grow with time. Data on how long a customer has lived in a neighborhood is unavailable. However, if I assume that homeowners have lived, on average, longer in the neighborhood, I can use homeownership as a proxy for the strength of the relationship with shops. Consistent with the importance of relationships between shops and customers, I find that controlling for socioeconomic strata, home type, income, and age of the household head, homeowners purchase 13% less in chains and 9% more in shops (see Table A.9).

An alternative to proxy for relationships is using the age of the shop. It is worth pointing out that this proxy is flawed because old shops might have grown old because they are resilient to adversity. Hence, if chains affect older shops less than younger ones, this would be consistent with older shops having an advantage from their neighbors' relationships or being more resilient in general or both. I define old shops as those in the fifth quintile of the age distribution of shops. The average age of young shops is five years, and the average age of old shops is 19 years. Figure A.13 shows that the adverse effects of chains are significantly smaller for old shops. The reduction in profits for young shops, relative to old shops, is five times larger, four times larger for revenue, seven times larger for profits per worker, and five times larger for revenue per worker. Older shops having more and stronger relationships might be the driver of this result. Still, it may also be that older shops have more human capital specific to operating a shop, and they respond better to competition.

Relationships between shop owners and neighbors may also allow shops to better screen their neighbors and offer them informal credit to purchase in the shop. Consistent with shops having a comparative advantage in providing credit to their neighbors, shops supply 16% of all the credit that households use to buy food and beverages (see Figure A.16). Moreover, excluding credit cards, shops provide 78% of credit used to purchase food and beverages. These statistics highlight that in a context where there is limited access to credit, shops' relationships with their neighbors allow shops to solve, even if just partially, the credit market friction that consumers face.

Distance is potentially critical in determining whether the household purchases in a shop or a chain. Households may choose whatever store is closest to them, and because there are more shops than chains, it is more likely that the nearest establishment is a shop. Unfortunately, I do not possess information on how far the shops and chain stores are from households. To circumvent this lack of data, I use the size of the neighborhood as a proxy of how far chain stores are from the consumers and shops. The idea is that in smaller neighborhoods, the additional chain store opening will, on average, be closer to both incumbent shops and households. If distance to the store is important, the negative effect of chains on shops in smaller neighborhoods should be more prominent.

However, smaller neighborhoods might have fewer shops and fewer chains stores, implying that: i) the effect on number of shops might be smaller because there are fewer shops, and ii) the effect on number of shops might be larger because there are less chain stores (if the marginal effect of a chain store is decreasing in the number of chain stores). To take this into account, Figure A.11 reports the effect of an increase of 1% in number of chains stores (relative to the average) on number of shops relative to the average number of shops, where both averages are those corresponding to the neighborhood size decile.[33] The results display a lack of evidence on distance to the establishment being critical in households' decision on

[33]Figure A.12 presents the estimates without taking into account that larger neighborhoods have more shops and chain stores.

whether to purchase in chains or shops.

The evidence in this section suggests that not only do shops survive, but they also have incentives not to grow. On the one hand, when shops grow, hire employees, move outside the owner's house, and target customers outside the neighborhood, they can potentially increase their revenue and profits and exploit advantages from operating at a larger scale. However, this growth will come at the expense of losing comparative advantages that stem from being small and owner-operated, making them less differentiated and more vulnerable to competition from chains.

1.7.2 Potential Consumer Welfare Implications

The entry of chains can have both a positive and negative impact on consumer welfare. Welfare increases for consumers that prefer to purchase in chains, because chains are now more available. However, the expansion of chains reduces the number of shops, potentially reducing access to shops for consumers who prefer purchasing in shops. I focus on these mechanisms, because chains do not have a statistically significant effect on shops' prices (see Figure A.15). Figure 1.9 presents on the left the share of expenses in food and beverages in chains and shops by income decile and on the right the purchase frequency of households in chains and shops also by income decile. Lower-income families spend a larger share of their income in shops and almost nothing in chains; they also visit shops more than twice as often as the highest income decile households. The highest-income families are the ones that purchase in chains the most and most frequently. Hence, these households are the ones who may benefit the most from the expansion of chains.

The key to determine whether lower-income households are worse off with the expansion of chains is to know if they lost access to shops. Figure 1.6 shows that purchases of fresh products in shops do not decline, despite a reduction in number of shops by 17%, suggesting

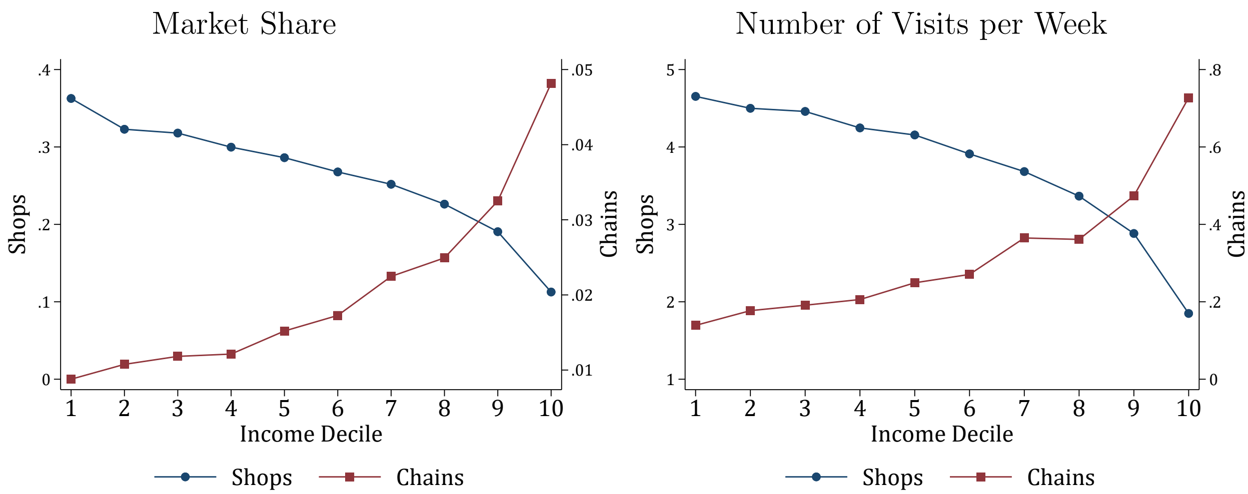


Figure 1.9: Market Shares and Number of Visits by Income Deciles

Source: Income and Expenditure Survey (ENIGH 2018). Food expenses in restaurants are not included.

that households do not lose access to shops. However, even if shops are still available for these households, these families might still be worse off because these shops are now, on average, further away (there are fewer shops), and it is now more costly to get there. These travel costs increments might be small because, based on 2019 data, there is one shop every three blocks.

1.8 Robustness

This section presents robustness checks for alternative IV specifications, neighborhood sizes, controls, and standard errors.

Table 1.5 present the results for alternative IV specifications. All these specifications provide estimates similar and consistent to those of the main specification. The main specification is in column 1. Column 2 presents results using one IV per chain (instead of aggregating across chains), and Column 3 uses a polynomial of the instrument that includes its square and cube. These two specifications provide similar results but have two disadvantages. The first one is that testing the monotonicity assumption is less transparent with multiple instruments, and the second one is that the first stage of the IV is weaker. To construct the economies of scale component of the IV, I aggregate the number of chain stores in second-degree neighboring cities (adjacent and those adjacent to adjacent cities). Columns 4 and 5 present results for using first and third-degree neighboring cities instead. The first stage is stronger the higher the degree of neighboring cities used. Yet, for the disaggregated data, the variation in the number of chain stores in a city explained by the number of chain stores in nearby cities is decreasing in the degree of neighboring cities used (see within R-squared in Table B.1). I use second-degree neighboring cities to balance this trade-off.

To construct the economies of scale component of the IV, I aggregate across chains the square of the number of stores in second-degree neighboring cities and take the square root of the sum. Without the square root, the first stage is weaker due to the instrument having a too long right tail (see Column 6). Alternatively, Column 7 does not square the number of stores at the chain level and subsequently does not take the square root. This measure has a weaker first stage, possibly because it does not capture that chains are more willing to expand when there is less competition.

To construct the suitability for the chains component of the IV, I use the prevalence of wide streets in the neighborhood. Columns 8 to 10 present results using alternative measures of suitability. Columns 8 and 9 create a measure of suitability in two stages. The first stage is a lasso regression of the number of chain stores in each census tract obtained from the 2020 firm directory (DENUE) on explanatory variables. The second stage is to predict the number of chain stores using the lasso selected variables and estimates. This prediction is the measure of suitability used. The variables include sociodemographics at the census tract and municipality level from the 2000 and 2010 population census, street data from open street maps, and municipality fixed effects. The lasso estimations also include the square, cube, and natural logarithm transformation of each variable totaling more than 2,600 variables in

Table 1.5: Robustness - Alternative IV Specifications

	Dependent Variable: # of Neighborhood Shops										
	Main	Per Chain	Squared & Cubed	1st Neighbors	3rd Neighbors	Squared Sum	Sum	Lasso1	Lasso2	Conv 1999	Chain $\text{Stores}_{t\text{-}1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Number of Chain Stores	-4.62***	-4.30***	-4.73***	-4.39***	-4.67***	-4.93***	-4.55***	-3.93***	-3.91***	-4.06***	-4.92***
	(0.673)	(0.514)	(0.643)	(0.726)	(0.599)	(0.858)	(0.682)	(0.512)	(0.538)	(0.562)	(0.674)
Observations	190,664	190,664	190,664	190,664	190,664	190,664	190,664	190,664	190,664	190,664	160,993
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year x City FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Variable | Chains>0	175	175	175	175	175	175	175	175	175	175	175
Mean Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.7
From 0 to Avg. # Conv. Stores	-17.0%	-15.8%	-17.4%	-16.1%	-17.1%	-18.1%	-16.7%	-14.4%	-14.4%	-14.9%	-18.8%
KP F-statistic	80.10	31.93	33.20	79.61	108.59	54.51	71.02	65.09	56.96	81.25	104.41

Note: The table displays the estimation of Equation 1.3 using 2SLS with variations of the instrument. Column 2 presents results using one IV per chain (instead of aggregating across chains), and Column 3 uses a polynomial of the instrument that includes its square and cube. Columns 4 and 5 present results using first and third degree neighboring cities instead of second degree neighboring cities. Column 6 does not take the square root of the square sum of chain stores in nearby cities. Column 7 does not square the number of chain stores in nearby cities before adding them up and does not take the square root of the sum. Columns 8 and 9 create a measure of suitability lasso to select the relevant variables to predict the number of chains stores. In Column 10, I use the number of chain stores and hybrid stores in each neighborhood in 1999 as measure of suitability. Column 11 uses the lagged number of chain stores in nearby cities to construct the instrument.

each analysis. The lasso in column 8 selected 675 variables, and the one in column 9 selected 373 variables. The prevalence of wide roads in the census tract was one of the three variables with the largest magnitude coefficient in both lasso estimations.[34] The difference between Columns 8 and 9 is that column 9 does not include variables from the 2000 population census. In Column 10, I use the number of chain stores and hybrid stores in each neighborhood in 1999 (before more than 90% of the openings of chains) as a suitability measure. The idea behind this specification is that neighborhoods suitable for hybrid stores in 1999 are also suitable for chains in the following two decades.

To capture that chains exploit economies of scale from opening stores in nearby cities, throughout the paper I use the contemporaneous number of convenience chains stores in nearby cities: the spatial correlation in chains' expansion. Regional economies of scale could also be captured using serial correlation instead: chains being more likely to expand in areas where they already have presence. Column 11 uses the lagged number of chain stores in nearby cities to construct the instrument instead of the contemporaneous number. The results are consistent and the first stage is even stronger. However, this is not the preferred specification, because it reduces the sample size by not including observations from 1999.

Table 1.6 presents results adding different sets of controls to the main specification. The economic censuses include the number of establishments for 154 business types other than shops and chains (e.g., restaurants, supermarkets, butchers, hospitals, liquor stores, shoe stores, pet stores, hardware stores, car dealers, banks, schools, and universities). To control for the presence of these businesses, I use a principal components analysis and keep the components with an eigenvalue larger than one as controls in Column 2. Alternatively, Column 3 uses a factor analysis instead and keeps all the factors with an eigenvalue larger than one as controls. The results of columns 2 and 3 are consistent with those of the

[34]This could mean that variation in the variable, or of another variable it is correlated to, has a significant impact on the number of chain stores in the census tract.

Table 1.6: Robustness - Adding Controls

	Dependent Variable: # of Neighborhood Shops									
	Main	Control Other Businesses PCA	Control Other Businesses FA	Nearby Chains Control	Pop. Census Controls	Large Cities Avg. pop 880K	HH Sample	HH Controls	HH Controls PCA	HH Controls FA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of Chain Stores	-4.62***	-4.96***	-4.28***	-4.99***	-3.51***	-6.14***	-6.79***	-6.76***	-6.81***	-6.79***
	(0.673)	(0.995)	(0.959)	(1.000)	(0.721)	(1.607)	(1.243)	(1.239)	(1.250)	(1.251)
Observations	190,664	189,919	189,919	190,664	184,075	36,293	65,061	65,061	65,061	65,061
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year x City FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Variable | Chains>0	175	175	175	175	179	200	210	210	210	210
Mean Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.5	8.1	7.4	7.4	7.4	7.4
From 0 to Avg. # Conv. Stores	-17.0%	-18.2%	-15.7%	-18.3%	-12.8%	-24.9%	-23.7%	-23.6%	-23.8%	-23.7%
KP F-statistic	80.10	88.52	82.32	61.78	69.24	14.67	75.79	77.52	75.63	75.67

Note: The table displays the estimation of Equation 1.3 using 2SLS with alternative controls. Column 2 uses the principal components with eigenvalue larger than one from a PCA to control for the presence of other businesses in the neighborhood (e.g., restaurants, supermarkets, butchers, hospitals, liquor stores, shoe stores, pet stores, hardware stores, car dealers, banks, schools, and universities). Column 3 uses the factors with an eigenvalue larger than one from a factor analysis as controls for the presence of other buisenesses in the neighborhood. Column 4 controls for the number of convenience chain stores around the neighborhood. The measure of chain stores around the neighborhood is the number of chain stores the neighborhood would have if the neighborhood was constructed using a 2km radius circle instead of a 1km minus those in the 1km radius. Column 5 adds controls from the 2000 and 2010 population census, including the average age of household head, household income, hours worked, population, and the number of households. To match the controls to the economic census years, I use linear interpolation and extrapolation. Column 6 presents the results of the main specification but restricting the sample to large cities with an average population of 800,000. Column 7 restricts the sample to neighborhoods for which there is consumption data. Column 8 includes the following household controls: number of inhabitants, men, women, adults, and minors; expenses on clothing, shoes, home, rent, energy, healthcare, public transportation, education; and income, total expenses, and income per capita. Columns 9 and 10 use principal components and factor analyses to control for the same household variables keeping the components or factors with an eigenvalue larger than one.

main specification. Column 4 controls for the number of convenience chain stores around the neighborhood, and the estimate of interest is similar to that of the main specification. The measure of chain stores around the neighborhood is the number of chain stores the neighborhood would have if the neighborhood was constructed using a 2km radius circle instead of a 1km minus those in the 1km radius. Column 5 adds controls from the 2000 and 2010 population census, including the average age of household head, household income, hours worked, population, and the number of households. To match the controls to the economic census years, I use linear interpolation and extrapolation.

Columns 8 to 10 add household controls from the consumption data to the main specification and the results are robust to including these or not. The consumption data, on average, will have households from larger cities living in neighborhoods with more shops and chains. Column 6 presents the results of the main specification but restricting the sample to large cities with an average population of 800,000, and the estimate of interest is similar in magnitude to the one in column 7, which restricts the sample to neighborhoods for which there is consumption data. Column 8 includes the following household controls: number of inhabitants, men, women, adults, and minors; expenses on clothing, shoes, home, rent, energy, healthcare, public transportation, education; and income, total expenses, and income per capita. Columns 9 and 10 use principal components and factor analyses to control for the same household variables keeping the components or factors with an eigenvalue larger than one.

Table A.5 shows the breakdown of the effect of the rise of chains on the number of shops by city size. The magnitude of the effect is larger for bigger cities. For towns with an average population of 14,000, the reduction in the number of shops is 4.3%, it is 15.7% for those with an average population of 260,000, and 24.9% for those with an average population of 880,000 thousand.

Table A.8 presents the results of the estimation of Equation 1.3 but using the natural logarithm of number of shops and number of chains. Column 4 uses the natural logarithm of number of shops as the dependent variable. Column 5 uses the natural logarithm of number of shops as the dependent variable and the natural logarithm of number of chains stores as dependent variable. The results are consistent with those of the main specification. For example, and increase from zero to the average number of chains stores in the neighborhood reduces the number of shops by 20% in the log-linear specification, just three percentage points more than in the main specification.

Throughout the paper, the neighborhood definition was all the census tracts that would fall within 1 km from the center of each census tract. Table A.9 contains the estimates of Equation 1.3, but with alternative distances to construct neighborhoods: 0km (census tract level), 0.25km, 0.5km, 0.75km, 1km, 1.25km, 1.5km, and 2km. The effect of each additional chain store in the neighborhood on the number of shops ranges from a reduction in shops by 2 to 4.8. The negative effect stabilizes at 1 - 1.25 km, consistent with these neighborhood sizes capturing all of the effects of the additional chain store. Table A.6 replicates Table 1.2, but defining the neighborhoods as census tracts. This is the smallest possible neighborhood size. The patterns in the estimates are consistent. The OLS estimates underestimate the negative effect of chains, fixed effects partially reduce the bias, and the IV addresses the bias due to unobservable neighborhood-specific shocks.

I cluster standard errors at the city level throughout the paper. Table A.7 contains standard errors of alternative clustering procedures. Column 2 clusters at the neighborhood and year level. Column 3 clusters at the city and year level. Column 4 clusters at the city-year level. Column 5 clusters at the city-year and neighborhood level. Clustering at the city level results in the largest standard errors, making the main specification the most conservative across clustering specifications.

Another concern in the estimation of standard errors is the potential correlation of unobserved shocks across adjacent cities or adjacent neighborhoods. Because of computational limitations, I can not estimate standard errors that take into account these potential correlations in my full sample. Hence, I run 500 iterations with 5,000 randomly selected neighborhoods where I compute the standard errors at the city level and correcting for the potential correlation of unobserved shocks across adjacent cities and 2^{nd}-degree adjacent cities.[35] The top graph on Figure A.17 shows that clustering at the city level, as done throughout the paper, is the most conservative alternative. In only one out of the five-hundred iterations, the widest confidence interval is not the one clustering at the city level.[36] I repeat an equivalent analysis to take into account the potential correlation in errors across adjacent neighborhoods. The results, displayed at the bottom of Figure A.17, show that clustering at the city level is also more conservative.[37]

1.9 Conclusion

Developing countries have hundreds of millions of microenterprises. As these countries develop, their microenterprises face increased competition from larger and more efficient firms. Standard economic models predict that this process will reallocate resources from low-efficiency firms that downsize and exit to more efficient ones. However, microenterprises in developing countries continue to be overwhelming in number despite facing direct competition from larger firms offering similar and often identical products and services.

This paper contributes to understanding this phenomenon by studying how one of the most prevalent microenterprises, the neighborhood shop, responds to increased competition from convenience chains in Mexico between 1999 and 2019. I assemble rich data and construct

[35] I use the technique proposed by Colella et al. (2019) to account for potential spatial correlation of unobserved shocks and its companion statistical package *acreg*.

[36] The confidence interval is wider clustering at the city level because the correlation of errors is negative across adjacent cities. One example of why this could be the case is that an additional chain store in an adjacent city might have an above-average effect on small cities but a below-average one on larger ones.

[37] The standard errors clustering at the city level are likely larger because they also take into account the potential positive correlation in errors with neighborhoods that are not adjacent but are also in the same city.

a novel instrument to address the endogeneity of chains' expansion. Consistent with a model of differentiated competition between chains and shops, I find that chains reduce the number of shops, primarily through a decrease in shop entry. However, most of them survive, and their customers continue to purchase in shops, but they buy less and less often, particularly less of packaged and standardized goods.

I present evidence consistent with shops having comparative advantages stemming from being small and owner-operated, such as lower agency costs, building relationships with their customers, and offering informal credit. Shops not only survive, but they have incentives not to grow. If shops grow, hire employees, and target customers from outside their neighborhood, they may lose the distinct comparative advantages stemming from being small and owner-operated that allow them to differentiate and survive competition from chains.

In this context, the standard prediction of a more efficient entrant leading reallocation of resources through the exit of less efficient firms does not occur. In particular, the small and less efficient type of firm delivers the most value-added to customers in Mexico by easing cash and credit constraints. These small shops, though less efficient, have comparative advantages in offering ripe products, which cash-constrained customers consume the same day. Moreover, their relationships with their customers allow them to offer informal credit to consumers that lack access to formal one. The combination of demand factors such as credit and cash constraints and the comparative advantages of shops on easing them allow shops to compensate for their disadvantages in scale.

While the estimates are specific to the context of this paper, the insights can be generalized to other industries and countries. The theoretical literature highlights how the relevance of comparative advantages defines industrial organization (Hubbard, 2004). For example, consumers' taste for fresh and ripe products determines the significance of the small businesses' comparative advantage in offering these products and, therefore, their survival. Hence, we can expect that as long as the comparative advantages stemming from being small and owner-operated are larger than those from economies of scale, industries will be fragmented.

Chapter 2

Grandmothers and the Gender Gap in the Mexican Labor Market

The gender gap in employment rate is a core issue in labor markets. This gap widens when women bear children, reflecting the fact that motherhood plays a significant role in its formation (Angrist and Evans, 1998; Waldfogel, 1998; Bertrand et al., 2010; Kleven et al., 2019). Decision makers can more efficiently guide policy to reduce the gender gap when they understand the role of each motherhood-related mechanism affecting employment. These mechanisms include specialization (Becker, 1991), gender roles (O'Neill, 2003; Dhar et al., 2019), personal preferences (Daymont and Andrisani, 1984), and labor market discriminatory demand (Correll et al., 2007).

This paper focuses on the specific mechanism of childcare availability. Parental employment and the amount of nonparental-provided childcare are likely decided simultaneously; hence, estimating the causal relationship between childcare availability and employment is challenging. To overcome this challenge, I use a natural experiment based on the plausibly exogenous timing of death of coinhabiting grandmothers and a stacked triple-difference to disentangle the effect of these deaths due to their impact on childcare availability from their effects through alternative mechanisms. The first difference is a within individual comparison of employment status quarters before and after the death. The second difference compares those who suffered the loss with those who did not. The third difference exploits the discontinuity in childcare need generated by eligibility to attend elementary school by comparing the double difference effect on parents who need more childcare (oldest child not eligible to attend elementary school, less than 6 years old) with that of parents who need it less (oldest child old enough to attend elementary school, 6 years-old or older). The triple-difference captures the effect of the grandmothers' deaths through their impact on childcare by canceling out mechanisms that impact households irrespectively of the oldest child eligibility to attend elementary school.

The natural experiment and each of the differences in the triple-difference strategy addresses a class of issues that threaten identification. The effect of household characteristics, such as values, that might affect the likelihood of women being employed (both mother and grandmother) cancel out with the first difference that compares the same parent quarters before and after the death. Mechanisms that are present for both parents that suffered a death and those who did not (e.g. an economic recession) cancel out with the second difference. The death of the grandmother affecting the labor force participation through alternative forms of home production, such as taking care of the house, cancel out with the third difference that compares mothers whose oldest child is not eligible to attend elementary school to mothers whose oldest child is.[1]

Grandmothers are one of the most important sources of childcare across the globe. For example, in the United States, grandparents look after 24% of the children on a regular basis,[2] in Europe, between 50 and 70% of grandmothers provide childcare in some form within a year.[3] In Mexico, grandmothers are the primary childcare providers. They take care of almost 40 percent of children up to six years old - as much as schools and daycare combined.[4] The availability of grandmother-provided childcare and mothers' employment are positively correlated. In three-generation households, the grandmother is more likely to

[1]The discontinuity in childcare availability generated by eligibility of the oldest child to attend elementary school is unlikely correlated with the home production of the grandmother. The discontinuity of the double difference effect at exactly 6 years of age of the oldest child is appreciated in Figure 2.3.

[2]Source: U.S. Census Bureau (2013)

[3]Figure 1 in Hank and Buber (2009) displays the averages for Spain, Italy, Switzerland, Austria, Greece, Germany, Sweden, France, Netherlands, and Denmark.

[4]See top of Figure A.1 of online appendix. All exhibits with the "A" prefix are in the online appendix.

provide childcare and the mother is more likely to be employed.[5] This paper uses the timing of death of the grandmother to explore whether the relationship between grandmother-provided childcare and mother's employment is causal.

While grandmothers are the primary childcare provider in Mexico, grandfathers rarely provide it.[6] In contrast to the null effect of grandfathers' deaths, grandmothers' deaths, through their impact on childcare, reduce the mothers' employment rate by 12 percentage points (27 percent) on average. This effect is not present for fathers. These findings suggest that it is not only differences across genders in dimensions that remain unchanged with the death of grandmothers (such as preferences, education, experience, or gender roles) that lead to the gender gap in employment.

The evidence suggests that households substitute grandmother-provided childcare with public and private alternatives when public daycares are more available or when private daycares or schools are more affordable. The negative effect of the grandmother's death on mothers' employment is 9 pp smaller if public daycare is one standard deviation more available, 8 pp smaller if private daycare is one standard deviation cheaper, and 9 pp smaller if private schools are one standard deviation cheaper. These heterogeneous effects suggest that even without reducing differences across genders in education, experience, or roles, increasing childcare availability and affordability can significantly reduce the gender gap in employment.

This paper has several advantages over the existing literature that studies the relationship between childcare availability and parental employment: (i) it provides evidence of households substituting the grandmother-provided childcare with private daycare when it is affordable and public daycare when it is available, (ii) quarterly data requires a significantly weaker assumption for causal interpretation (the timing of death of the grandmother being

[5]See bottom of Figure A.1 and Table A.1

[6]See Figure A.1

as good as random), (iii) the panel structure of the data allows to control for both observed and unobserved time-invariant characteristics at the individual level, (iv) the triple-difference disentangles the effect through childcare from the effect through alternative mechanisms (e.g. inheritance, lost income, or household labor), (v) documents that most of the reduction in earned income and hours worked for mothers is driven by a reduction on the extensive margin, and (vi) compares the effect on mothers and fathers. The related literature section discusses the existing literature and the contributions of this paper in more detail.

2.1 Related Literature

Abundant research documents the gender gap in employment and its relationship with motherhood.[7] There has also been progress in identifying the mechanisms through which the gender gap is formed, such as, employer discrimination (Correll et al., 2007) and marital status (Fernandez and Wong, 2014a,b). Within the literature that studies the effect through the childcare mechanism, Jaumotte (2003) uses variation across OECD countries in childcare subsidies, Givord and Marbot (2015) uses a French reform in family allowance, and Lefebvre and Merrigan (2008) uses a new childcare policy implemented in Quebec to estimate the effect of childcare availability on parental employment. While policy changes create variation across time for all households simultaneously, using the timing of death of grandmothers poses an identification advantage because it generates variation across time specific to the household that is improbably correlated with changes in societal values that may drive policy changes.[8]

Within the papers that study the relationship between childcare availability and labor

[7] See, for example, Kühn et al. (2017); Bertrand et al. (2010); Waldfogel (1998); Kleven et al. (2019); Cristia (2008); Agüero and Marks (2008); Jérôme et al. (2017); Angelov et al. (2016); Fernández-Kranz et al. (2013)

[8] If, for example, a policy reform occurs at the same time as debate regarding the policy or gender roles, the event study estimates will include the effect of the policy change as well as the possible effect the debate could have on gender attitudes.

supply, there are several papers that use the availability of grandparents as variation in childcare. Zanella (2017) contains a literature review on the relationship between grandparent availability and parental labor force participation, and concludes that some of the limitations of the existing literature are the lack of studies that are able to address causal identification and whether the results extend to developing countries. Moving forward, I first discuss the papers that use grandparent availability as an instrument for grandparent-provided childcare, then I proceed to those that estimate the relationship between grandparent availability and mother's employment directly.

Posadas and Vidal-Fernandez (2013) (PVF2013) and Arpino et al. (2014) (APT2014) use an instrumental variable (IV) based on whether the grandmother is alive or not, and Aparicio-Fenoll and Vidal-Fernandez (2015) (AFVF2015) and Aparicio Fenoll (2019) use retirement eligibility of grandmothers in Italy and in Europe to instrument for grandparent-provided childcare. The triple-difference estimation used in this paper presents an advantage over these IVs because (i) it leverages on individual fixed effects to control for time-invariant characteristics at the individual and household level,[9] (ii) if the death or the retirement of the grandmother affects mothers' employment rate through a mechanism other than childcare (e.g., income effect from lost grandmother's income or grandmother's household labor), the exclusion restriction would be violated; instead, the triple-difference disentangles the effect through childcare by exploiting the discontinuity in childcare availability generated by eligibility to attend elementary school,[10] and (iii) it relies on a weaker identification assumption than PVF2013 and APT2014: while the IV requires the grandmother being alive or dead to be random, the triple difference only requires the timing (quarter) of death to be random.

[9]Footnote 6 in PVF2013 discusses how the IV with fixed effects would account for both endogeneity and time-varying heterogeneity, but because their estimates are imprecise, these results are only available upon request.

[10]The third difference compares the effect on mothers whose oldest child is eligible to attend elementary school to mothers of those too young to. Effects that are common for both groups, such as inheritance, cancel out.

If household characteristics such as habits, income, or education affect the probability of the grandmother being dead (longevity) and also affect mother's employment probability (Hughes et al., 2007; Chen and Liu, 2011; Di Gessa et al., 2016), the IV estimate would be biased.[11]

PVF2013 also has a fixed effects (FE) specification where the independent variable is whether the grandmother provides childcare. The main advantage of this paper over the direct estimation of the effect of grandmother-provided childcare on mothers' employment with FE in PVF2013 is that the interpretation of the natural experiment and triple-difference is causal, while the interpretation of the FE in PVF2013 is not. As PVF2013 mention, FE by themselves, cannot address reverse causality (whether the grandmother provides childcare because the daughter works or vice versa).

Bratti et al. (2018) (BTS2018) further discusses disadvantages of using the IVs in PVF2013, APT2014, AFVF2015, and Maurer-Fazio et al. (2011)[12] to estimate the causal relationship between grandmother-provided childcare and parental employment. Instead, BTS2018 directly estimates the relationship between female labor force participation and availability of mothers, mothers-in-law, fathers, and fathers-in-law using pension reform-induced changes in retirement eligibility in Italy. While the exclusion restriction is not a concern for BTS2018, the triple-difference advantages (i) and (iii) over PVF2013, APT2014, AFVF2015, and AF2019 are also advantages over BTS2018. If any household characteristic such as education, habits (e.g. nutrition), or income affect both mother's employment and grandmother's longevity, the estimate would be biased. While BTS2018 requires longevity or retirement

[11] For example, in PVF2013 sample, families with deceased maternal grandmothers seem to be more disadvantaged than their counterparts. On the other hand, the first difference of the triple-difference, compares the quarters before the death of the grandmother to the quarters after (within individual variation). Hence, only requiring the timing to be random.

[12] They instrument the presence of grandparents using the mother's and father's age and provincial dummies. The exclusion restriction would be violated if the mother's age affects her employment by a mechanism other than the presence of the grandmother in the household (e.g. experience being correlated with salary and age, and salary affecting labor force participation).

eligibility to be random,[13] this paper only requires the time (quarter) of death to be random. Moreover, the effect captured by BTS2018 does not need to be through the childcare mechanism and might include effects through inheritance or grandmothers' household production. Compton and Pollak (2013) finds a positive correlation between geographical proximity to grandmothers and mothers' labor supply in the U.S.. The timing of a death represents endogeneity of lesser concern than distance to grandparents: while households can choose where to live, they cannot choose the grandmother's time of death.

The effects of childcare availability on mothers' employment are of special interest in developing countries, where the severity of the gender gap is exacerbated due to less progressive attitudes about women in the labor force, gender-based violence, and women having less decision-making power (Jayachandran, 2015).[14] As discussed in more detail by Jayachandran (2021), changing gender norms regarding who is responsible for household work and child care is one way of freeing up women to participate in the labor market, but other alternatives that free up women's time could help as well. For example, childcare availability can enable women to join the labor market despite gender norms that place the burden of childcare on women.

In the context of developing countries, Barros et al. (2013) use a lottery for city daycare in Rio de Janeiro to estimate the effect of childcare availability on mothers' employment. Martínez A. and Perticará (2017) use a randomization on offering after-school care in Chile. Hojman and Lopez Boo (2019) use random assignment of childcare centers across Nicaragua's poorest neighborhoods, and Clark et al. (2019) use randomization of subsidized daycare

[13]The omitted category in the empirical estimation is when the potential provider is dead. Hence, for a causal interpretation of the coefficients, longevity would need to be random. For the difference in coefficients of (i) alive and eligible and (ii) alive and ineligible, randomness in eligibility is required for a causal interpretation.

[14]Moreover, gender inequality, by itself, is considered a barrier to development; in the words of Amartya Sen, "[t]he changing agency of women is one of the major mediators of economic and social change, and its determination as well as consequences closely relate to many of the central features of the development process" (Sen, 1999, p. 202).

in a settlement in Nairobi. This paper further contributes to this literature by studying differences across genders, using variation in the primary source of childcare, using a natural experiment on a national scale, and testing whether the availability and affordability of daycare can mitigate the negative effect of the loss of family-provided childcare.

Khanna and Pandey (2021) estimate the net effect of the death of the coinhabiting mother-in-law on the daughter-in-law labor force participation using a two-wave survey conducted in India in 2005 and 2012. There are several advantages of this paper over Khanna and Pandey (2021): i) the triple difference estimates the effect through childcare instead of the net effect of the death (which includes household labor, inheritance, and other mechanisms), ii) quarterly data allows to test for pretrends, and iii) this paper documents substitution between private and public daycare alternatives and randmother-provided childcare.

Finally, this paper is also related to existing work studying the gender gap and the motherhood penalty in the Mexican labor market. For example, Aguilar Gomez et al. (2019) show that relative to four quarters before the birth of a child, mothers relative to fathers are 12.5 pp less likely to be in the workforce one quarter before the birth, 20 pp less likely to be in the workforce the first quarter after the birth, and 15 pp less likely to be in the workforce four quarters after the birth. Arceo-Gómez and Campos-Vázquez (2014) analyze the gender wage gap in Mexico from 1990 to 2010, finding that while the gap has decreased, it is still 6% in 2010. Calderón (2014) studies the effect of a child care program (Estancias Infantiles para Apoyar a Madres Trabajadoras) on easing burdens on working women, finding that the program increased women's probability of working, reduced the time they devoted to child-rearing and increased their labor incomes.

2.2 Data, the Gender Gap, and the Motherhood Penalty

The main data source is the Mexican National Survey of Occupation and Employment (ENOE). The ENOE is the largest household survey conducted in Mexico, and it is superior to administrative data in this context because it includes both the formal and informal sectors of the economy.[15] Its data collection occurs every quarter in a rotating panel format with five observations per household. The ENOE data used in this paper spans Q1 2005 to Q1 of 2020, a total of 61 surveys (one per quarter). Each survey visits approximately 120,000 households. The survey's demographics section includes information on every member of the household, such as their relationship to the head of household, gender, education, marital status, reason for not living in the household any more (after first survey), access to health care, employment, income, and hours worked.[16]

I map households across surveys using the household id to create a panel with five observations per household. To map individuals across surveys and create an individual-level panel, I use the the line number and validate using date of birth, age, and gender.[17] I focus on three-generation households, because the data provides grandparents' information only if they live in the same household.[18] Within three-generation households, the generation to which each individual belongs to is identified only in terms of their relationship to the house-

[15] Sixty percent of the workers in Mexico work in the informal sector (OIT, 2014). This paper uses the classification of informality used by the Mexican Statistical authority (INEGI): subordinate employees with pay belong to the informal sector if they do not have access to Mexican Social Security. Access to Social Security in Mexico is achieved by being affiliated with the Mexican Social Security Institute or an equivalent. This affiliation guarantees access to benefits, such as health care, disability insurance, housing credit, and a pension plan. INEGI (2014).

[16] The head of household is the individual who is highest in the hierarchy due to being the main economic contributor, the eldest, or the main decision maker (INEGI, 1997).

[17] The line number is generally a within household identifier. Validating using date of birth, age, and gender, 99.9% of observations appear to be correctly identified. I exclude the remaining 0.1% where at least one member of the household is not identified from the analysis.

[18] I consider all first-generation individuals to be grandparents, although they are not necessarily grandparents. They could be, for example, siblings of the grandparents.

hold head, but not in terms of their relationship to other family members.[19] For women, mothers are identified by belonging to the second generation and having children. For men, fathers are identified by belonging to the second generation and being married or coinhabiting with their partner.[20] The death of a coinhabiting grandparent is revealed whenever the respondent answers that the grandparent is not present because he or she passed away.[21]

On average, three-generation households represent 27 million Mexicans and 4.7 million households in Mexico — 23 percent of the total population and 15 percent of the households. Mothers in three-generation households are not identical to those in other households. Mothers in three-generation households are 1.5 pp (2.6%) more likely to live in a large city (population $\geq$ 100,000) and 2.9pp (17%) less likely to be rural (population $\leq$ 2,500) (Table A.1, columns 1 and 2). The lower housing costs in less populated areas may be why three-generation households are less common. In three-generation households, grandmothers are 30pp (80%) more likely to provide childcare, and mothers are 12 pp (34%) more likely to be employed (Figure A.1 and column 3 of Table A.1). Three-generation households' income is 23% higher, but their income per capita is 27% lower (Columns 4 and 5). Mothers in three-generation households are 2.2 years younger (8%), and controlling for age, they have 14% fewer children, have 2% more years of schooling, are 8% more likely to be high school graduates, and 7% more likely to be college graduates (Columns 6-10). Finally, mothers in three-generation households have 12% higher average income and work 35% more hours, but conditioning on being employed, they earn 12% less and work 10% more hours (Columns 11-14).

Because of these differences between mothers in three-generation households and other

[19]For example, suppose the grandfather (first generation) is the head of the household. His children and their spouses are the second generation, and his grandchildren are the third generation. The relationships between the individuals in the second and third generations are not identified; a second generation individual could thus be either a father or an uncle of the third generation.

[20]There is no question on having children for men.

[21]Figure A.2 shows the frequency distribution of the grandparents' ages and Figure A.3 shows the frequency distribution of the age at which the grandparents died.

households, the findings for three-generation households cannot be directly extended to other households. The subsection Bounds for the Average Effect of Grandmothers' Death on Women in the Results Section uses the estimates for three-generation households and two sets of assumptions to bound the average effect of the grandmother's death on women's employment rate in Mexico.

I add additional restrictions to construct the primary estimation sample of three-generation households. To reduce noise in the data, I do not include households where the oldest grandchild is thirty or older, where grandparents are less than forty years old, or more than one grandmother or grandfather died. This restriction reduces the sample by 6%, but ensures that the three-generation households are more standard in their composition. There are three additional more meaningful restrictions i) at most one grandfather and one grandmother (further reduces the sample by 0.7%), ii) at most one mother and one father (further reduces the sample by 21%), and iii) balanced panel with five observations per individual (further reduces the sample by 15%). The robustness section shows that the results are consistent and very similar if lifting any of these restrictions on the sample. The main reason for restrictions i) and ii) is to avoid situations in which the childcare provided by the grandmother who died is replaced by that of another coinhabiting grandmother after the death. Table A.2 compares mothers in the base sample and mothers in samples after lifting these restrictions. Lifting the restrictions has little influence on the composition of mothers.

To construct measures of childcare availability, I use the National Economic Units Statistical Directory (DENUE) of 2015 and the Population Census of 2020.[22] The DENUE lists all the public and private daycare facilities, and the population census provides the number of children up to five years of age living in each municipality. I construct of measure of the availability of public and private daycare at the municipality level by dividing the number

[22]I use the 2015 DENUE because it is the first one that is available, and the 2020 population census because it is the first population census available after the 2015 DENUE.

of daycare facilities of each type by the number of children up to five years of age. I use this measure for 1,479 municipalities that are also covered by the ENOE.

To construct measures of childcare affordability, I use data from the Employment and National Security Survey (ENESS) from 2009, 2013, and 2017. The ENESS is a joint project between the Mexican National Statistics Institute (INEGI) and the Mexican Institute of Social Security (IMSS). The survey occurs every four or five years since 1996 to provide statistical information regarding the coverage and characteristics of social security and health care services in Mexico. As an accompanying module of the ENOE, it covers all the households covered by the ENOE for two out of the three months in the quarter.[23] Hence the ENESS covers roughly two-thirds of the ENOE sample for the quarter.[24] Both ENOE and ENESS are designed to be representative at the state and country level.

The ENESS data includes responses for 209,266 households. These households use public and private daycare providers for 3,991 and 1,177 children under seven years old. The ENESS asks how much the household paid for the service and the number of hours. I use this information to compute the cost per hour of daycare for each child.[25] Then I average at the locality level and by whether the service was public or private.[26] The result is a proxy for the private cost of daycare for 231 localities and the public cost of daycare for 527 localities.[27]

[23]It covers households covered by the ENOE during July and August of 2013 and 2017 and households covered by the ENOE during May and June of 2009.

[24]Specifically, 66.63%, 67.87%, and 62.4% respectively for 2017, 2013, and 2009.

[25]I cap the total hours at 12 hours per day.

[26]I first demean by year to remove year specific variation, such as inflation. I compute two averages: simple average and weighted average using the ENESS probability weights. The results using probability weights are available upon request and very similar in magnitude and statistical significance to those without weights.

[27]The analysis only uses the localities where grandmother's deaths in three-generation households are observed: 209 and 483 localities to estimate heterogeneous effects by the cost of private and public daycare.

2.2.1 Grandparents, Children, and Childcare

The ENESS asks households that are not using a public or private daycare service about their reason for not doing so.[28] Approximately 40 percent responded that they had no need for public or private daycare services, and almost 40 percent responded that either they had no access, they could not afford it, or it was not possible to take or pick up their child (see Figure A.4). Of those who responded that they did not need daycare, more than 90 percent relied on a family member to provide childcare; and specifically, more than 60 percent relied on grandmothers to provide childcare (see Figure A.5).

According to Mexican law, education from preschool to middle school is compulsory, and kids should attend school starting from 3-4 years of age. However, *de facto*, school becomes a relevant "childcare" provider only when kids turn 5-6 and they start elementary school. This may be because even if they go to kindergarten, the parents still consider the grandmother the primary provider. The bottom of Figure A.1 shows parents' response to who takes care of the child when the mother goes to work by the child's age from the ENESS. Until the children are four years old, grandmothers look after 44% of the children while schools look after 6%. However, by the time they are six years old, grandmothers only look after 24% of them while schools look after 52% of them.

2.2.2 The Gender Gap and the Motherhood Penalty

The motherhood penalty in employment, the difference in employment rate between women with children and women without them, forms between the ages of twenty and thirty and remains thereafter.[29] Top of Figure 2.1 displays the motherhood penalty and gender gap in

[28]The question limits the respondent to one answer.

[29]The ENOE classifies the employed into four categories: (i) subordinate workers with pay, (ii) employers, (iii) self-employed, and (iv) workers without pay. This paper considers an individual as employed if he or she is a subordinate worker with pay. This classification includes people who worked at least one hour the previous week and those who did not work but have a job. For the second group, they may not have worked

three-generation households (left) and in Mexico (right). The pattern is similar, but the gaps are narrower in three-generation households because of a higher employment rate of women with children between ages of twenty and forty. This is consistent with the findings discussed in the Related Literature section for other countries: the availability of the grandmother is positively correlated with mother's employment. The next section explains how the triple-difference estimation addresses whether this correlation is causal.

2.3 Empirical Strategy

The timing of death of grandmothers provides variation to childcare availability that identifies its effect on mothers' labor supply. The first empirical specification is a triple-difference. The first difference compares mothers' employment status before and after the death of the grandmother. The second difference compares mothers that suffered a loss to those who did not. Since the death of the grandmother may affect the labor supply through several mechanisms, the third difference disentangles the effect of the death due to its impact on childcare from its effect through alternative mechanisms by comparing the the double-difference effect for mothers whose oldest child is eligible to attend elementary school to those whose oldest child is not. Childcare is scarcer and needed more when children cannot attend elementary school; the triple-difference captures the effect that the death of the grandmother has on mothers of young children but not on mothers of older children, the childcare mechanism.

I use individual fixed effects to control for both observable and unobservable mother-grandmother-household time invariant characteristics that could correlate with both the timing of death of the grandmother and the mother's labor supply. Locality-year-quarter fixed effects control for locality-specific shocks to the labor market, for example a city-specific

while employed because of, for example, a strike, suspension, training, vacations, or personal days. The robustness section repeats the main analyses considering working as any of the first three categories of the employed according to the ENOE; the findings are consistent.

boost in government spending. Young child-year-quarter fixed effects control for shocks that are specific to children's age, for example, a nationwide education reform or a new public daycare policy. Grandmother died-year-quarter fixed effects control for pre-existing differences between households where the grandmother will die during the survey period and those where she will not. Ten alternative specifications gradually reducing what the fixed effects control for are also reported in Table 2.1. Equation 2.1 is the main specification and β_2, the triple-difference estimate, is the parameter of interest:

$$\begin{aligned} Employed_{i,l,t} \ = \ & \beta_1 Post_{i,l,t} \times Death^{GM}_{i,l} \ + \ \beta_2 Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l} \\ & + \phi_i \ + \ \zeta_{l,t} \ + \ \gamma_{t,YoungChild} \ + \ \eta_{t,DeathGM} \ + \ \varepsilon_{i,l,t} \end{aligned} \tag{2.1}$$

Where $Employed_{i,l,t}$ takes the value of 1 if mother i living in locality l is employed at time (year-quarter) t and 0 otherwise, $Death^{GM}_{i,l}$ is a dummy variable that takes the value of 1 if the mother suffered the death of the grandmother at any point through the span of the surveys and 0 otherwise, $Post_{i,l,t}$ takes the value of 1 for every period after the death of the grandmother and 0 otherwise, $YoungChild_{i,l}$ indicates that the oldest child in the household is young, ϕ_i is the individual fixed effect, $\zeta_{l,t}$ is the year-quarter-locality fixed effect, $\gamma_{t,YoungChild}$ is the year-quarter-young child fixed effect, $\eta_{t,DeathGM}$ is the year-quarter-grandmother died fixed effect. All the lower level interactions are captured by the fixed effects.

The main specification uses an age cutoff of the oldest child of at most 5 years old to be considered a young child.[30] The three main reasons to use the 5-years-old cutoff are that: (i) it exploits a discontinuity of childcare availability by separating children that can, and by law should, attend primary schools from younger children, (ii) it is consistent with

[30] An alternative cutoff could be at most 1 year, because many childcare facilities do not accept children younger than 2 (Profeco, 2004). This cutoff reduces the number of observation too much.

governmental classification of children by age, and (iii) it presents an advantage over using a cutoff at a younger age by increasing the size of the "treatment" group.[31]

One of the most common concerns in the literature is the grandmother affecting labor force participation of the mother through alternative mechanisms as inheritance, sickness, or household labor. For example, the income effect of inheritance may increase leisure consumption. I use the discontinuity in childcare availability generated by eligibility to attend elementary school to disentangle the childcare mechanism from this other mechanisms that are unlikely discontinuous at exactly the age of 6. As long as mothers whose oldest child is 4 years-old and those whose oldest child is 6 years-old are as likely to receive an inheritance, or the grandmother is as likely to supply household labor or to be sick, then these effects will cancel out with the third difference and the remaining effect is the effect through the childcare mechanism.

To test the discontinuity generated by the availability of elementary school, I estimate the triple-difference effect by age bracket of the oldest child. In this specification, the dummy variable for having a young child in the household, $YoungChild_i$, is replaced by 3 dummy variables indicating the age bracket of the oldest child in the household: (i) at most 3 years old, $YoungChild_{i,l,1}$, (ii) between 4-5, $YoungChild_{i,l,2}$, (iii) between 6-10, $YoungChild_{i,l,3}$. The omitted category is when the oldest grandchild is older than 10 and it is captured by β_1. The estimated equation is:

$$
\begin{aligned}
Employed_{i,l,t} \;=\; & \beta_1 Post_{i,l,t} \times Death^{GM}_{i,l} \;+\; \sum_{k=1}^{k=3} \beta_{2,k} Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l,k} \\
& + \phi_i \;+\; \zeta_{l,t} \;+\; \gamma_{t,YoungChild} \;+\; \eta_{t,DeathGM} \;+\; \varepsilon_{i,l,t}
\end{aligned}
\tag{2.2}
$$

I use an event study design to test for common trends in employment prior to the grandmother's death in households with young children and in households with older children.

[31] In governmental classification, 0-2 years is initial education and 3-5 is preschool (Profeco, 2004).

This design also measures the persistence of the effect by including an estimate for each period after the death. The event study equation is built from Equation 2.1, but adds a time index s, which is the time relative to the death of the grandmother. Since each individual is observed for five periods, $s \in \{-4, -3, -2, -1, 1, 2, 3, 4\}$. Period s=-1 is the last period before the death and period s=1 is the first period after the death. The period $s = -1$ is the omitted category in the estimation. The estimated equation is the following:

$$\begin{aligned} Employed_{i,l,t,s} \; = \; & \sum_{s=-4}^{s=4} \Big(\beta_{1,s} Post_{i,l,t,s} \times Death^{GM}_{i,l} \; + \; \beta_{2,s} Post_{i,l,t,s} \times Death^{GM}_{i,l} \times YoungChild_{i,l}\Big) \\ & + \phi_i \; + \; \zeta_{l,t} \; + \; \gamma_{t,YoungChild} \; + \; \eta_{t,DeathGM} \; + \; \varepsilon_{i,l,t} \end{aligned} \tag{2.3}$$

If childcare availability and gender roles jointly contribute to the formation and persistence of the gender gap, the triple-difference negative effect on employment probability would be larger for mothers than for fathers. Equations 2.1 and 2.3 are modified to include a quadruple difference resulting on equations 2.4 and 2.5, where $Mother_{i,l}$, takes a value of 1 if the second generation individual is a mother. All fixed effects, except individual, are interacted with gender of the parent.

$$\begin{aligned} Employed_{i,l,t} \; = \; & \beta_1 Post_{i,l,t} \times Death^{GM}_{i,l} \; + \; \beta_2 Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l} \\ & + \beta_3 Post_{i,l,t} \times Death^{GM}_{i,l} \times Mother_{i,l} \; + \; \beta_4 Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l} \times Mother_{i,l} \\ & + \phi_i \; + \; \zeta_{l,t,Gender} + \; \gamma_{t,YoungChild,Gender} \; + \; \eta_{t,DeathGM,Gender} \; + \; \varepsilon_{i,l,t} \end{aligned} \tag{2.4}$$

$$Employed_{i,l,t,s} = \sum_{s=-4}^{s=4} \big(\beta_{1,s} Post_{i,l,t,s} \times Death_{i,l}^{GM} + \beta_{2,s} Post_{i,l,t,s} \times Death_{i,l}^{GM} \times YoungChild_{i,l} + \beta_{3,s} Post_{i,l,t,s} \times Death_{i,l}^{GM} \times Mother_{i,l} + \beta_{4,s} Post_{i,l,t,s} \times Death_{i,l}^{GM} \times YoungChild_{i,l} \times Mother_{i,l}\big) + \phi_i + \zeta_{l,t,Gender} + \gamma_{t,YoungChild,Gender} + \eta_{t,DeathGM,Gender} + \varepsilon_{i,l,t,s} \tag{2.5}$$

Grandfathers are significantly less likely to provide childcare than grandmothers. While grandmothers provide almost 40 percent of total childcare, grandfathers are not even an explicit option in the ENESS and fall on the category of other family members. Other family members provide in total close to 20 percent of childcare (see Figure A.1). The death of a grandfather is used as a placebo in the robustness section, where the specifications described in this section for the death of the grandmother are estimated for the death of a grandfather. If the triple-difference is indeed capturing the childcare availability mechanism, the triple-difference effect should not be present (or be smaller) when a grandfather dies, because a grandfather does not provide childcare as much nor as often as the grandmother.

The main specification of the empirical strategy uses the age of the oldest child in the triple difference to separate households that need more childcare from those that need it less. Previous literature has estimated the effect of childcare availability on mothers' employment rate and heterogeneous effects based on cutoffs built using the age of the youngest child (Posadas and Vidal-Fernandez, 2013; Arpino et al., 2014; Bratti et al., 2018; Compton and Pollak, 2013). Other papers look at the time of birth of the first child (Kleven et al., 2019). I use the age of the oldest instead of the youngest child because the Mexican context is different than that of developed countries that have been the focus of previous literature.

In Mexico, it is widespread for children to provide care for other children in the household. Based on time allocation from the ENOE, when there is a child up to five years old in the household, 26% of 17-years-olds spend some time exclusively providing care without pay.[32]

[32]Source: Author calculations using ENOE Q1 2005 to Q1 2020 based on the following question: During

This is true even for younger children; 16% of 12-year-olds provide care when there is a child up to 5 years old in the household. The time allocation question is only available for 12-year-olds and older, but even younger children are likely to provide care. Based on the ENESS question, who takes care of the child when the mother goes to work, more than 1% of six-year-olds are left alone when their mothers go to work.[33] Hence, the split between households that need more childcare and those that need it less is more reasonable when using the age of the oldest grandchild because the households classified as those that need more childcare do not have older siblings that can provide care. However, specifications using the age of the youngest child are also valid, and I present these estimates in the Robustness Section.

2.4 Results

This section has three subsections. The first subsection presents the estimates for the sample of mothers belonging to the second generation in three-generation households. This section documents that (i) the death of the grandmother, through its impact on childcare (second difference), reduces the probability of being employed of mothers by 12 percentage points (27%), (ii) the effect is economically and statistically significant as long as the oldest child is not old enough to attend elementary school, (iii) the effect is persistent for at least 4 quarters after the death, (iv) mothers' income decreases 53% and hours worked decrease 30% — driven mostly by a reduction in the extensive margin, and (v) the death of the grandmother, through its impact on childcare (second difference), reduces the probability of being employed for mothers by 15 percentage points more than for fathers (quadruple difference).

last week, how much time did you spend exclusively taking care without pay of children, elderly, sick, or disabled?

[33] This question is only available for children up to six years old.

The second subsection uses the estimate of the average effect of the grandmother's death in three-generation households and two sets of assumptions to create bounds on the average effect of the grandmother's death on women in Mexico. This exercise implies that the average effect of grandmothers' death on womens' employment in Mexico ranges between 1.8pp and 5.3pp (4.6% - 13.5%).

The third subsection presents heterogeneous effects of the grandmother's death on mothers' employment. The negative effect of the grandmother's death on mothers' employment is 9 pp smaller if public daycare is one standard deviation more available, 8 pp smaller if private daycare is one standard deviation cheaper, and 9 pp smaller if private schools are one standard deviation cheaper. This section also shows that the negative effect of the grandmother's death is larger when the maternal grandmother dies, increasing in the hours the grandmother used to provide care, and significantly smaller if there are male grandchildren.

2.4.1 The Effect on Employment, Hours Worked, and Earned Income

The estimates of equation 2.1 are displayed in Panel A of Table 2.1. The results of the main specification, with individual, locality-year-quarter, young child-year-quarter, and grandmother died-year-quarter fixed effects (FE), are in column one: the death of the grandmother, through its impact on childcare, reduces mothers' employment rate by 12.4 pp (p-value = .00005). Columns 2-11 display alternative specifications gradually reducing what the fixed effects control for. The triple-difference estimates of the reduction in employment rate with the different combinations of FE range between 7.5 and 12.4 pp.

For the four quarters before the death of the grandmother, both mothers of children at most five years old and mothers of children older than five have a similar flat trend in their employment rate, which is not statistically different from its level in the quarter just before the grandmother's death (see bottom of Figure 2.2). After the death of the grandmother, while there is no effect on mothers of older children, the employment rate of mothers of children five years old or younger declines between 11 and 17 percentage points for the next four quarters after the death. The difference between these two groups of mothers — the triple-difference effect — is statistically significant for the four periods after the death of the grandmother (see top of Figure 2.2). In the triple-difference figure (top of Figure 2.2), the omitted category is $t = -1$, hence this coefficient is not estimated. Similarly, plotting the two double differences (bottom of Figure 2.2), $t = -1$ is the omitted category and there is one additional coefficient for the older children households that is captured by the grandmother died-year-quarter fixed effect.[34]

Relative to mothers of children older than 10 years, the death of the grandmother, through its impact on childcare, reduces the probability of being employed for mothers whose oldest child is at most 3 years old or between 4 and 5 years old by 15 and 12 percentage points, respectively (see Figure 2.3). The negative effect of the death of the grandmother fades away if the oldest child is old enough to attend elementary school or older. This exercise documents a clear discontinuity in the effect of the grandmothers' death at the time when the oldest child is eligible, and by law required, to attend elementary school.

The effect on the probability of being employed is the net effect from transitions across full-time employment, part-time employment, and unemployment. In the sample, 43% of mothers are employed, 34% work full-time (more than 30 hours per week), and 9% work part-time (less than 30 hours per week). Columns 1 and 2 of Table 2.2 display the effect of grandmothers' death on the full-time and part-time employment rates of mothers. Through its impact on childcare availability, the grandmother's death reduces the probability of being employed full-time by 8.5 pp (25%) and the probability of being part-time employed by 3.9

[34]If the triple-difference was not staggered (if all grandmothers had died in the same quarter), all the coefficients for the older children households would be captured by the grandmother died-year-quarter fixed effect (instead of one of them, when it is staggered).

pp (40%).

Even though more mothers employed full-time leave the labor force than part-time employed ones, they are not more likely to do so. The reduction in the probability of being full-time employed is 8.5pp and 3.9pp for part-time employed. However, in relative terms, the latter is larger because 34.4% of women are full-time employees and 9.6% of women are part-time employees. Moreover, mothers employed full-time left their jobs instead of switching to part-time employment. Column 5 shows the effect on the probability of part-time employment for the subsample of mothers employed full-time in the first survey wave. These mothers do not transition from full-time employment to part-time employment. Their probability of being part-time employees decreases by 5.8pp (not statistically significant at conventional levels).

The shock to childcare availability also affects hours worked and earned income. The grandmother's death, through its impact on childcare, reduces weekly hours worked for mothers by 30% and earned income by 53% (see columns 3 and 5 of Table 2.2). These effects include both the extensive and intensive margin. The extensive margin is from mothers that went from employed to unemployed, and the intensive margin is from mothers that continue to be employed but for fewer hours or with a lower wage. Columns 4 and 6 display the results for the intensive margin – restricting to the sample of mothers with strictly positive income and hours worked. The effect through the intensive margin is a reduction in hours worked by 12% and in earned income by 26%, but both of these effects are not statistically significant. The results are consistent with a lack of flexibility in the labor market and mothers being pushed out of the labor market when losing the grandmother-provided childcare.

The motherhood penalty in Mexico, the difference in employment rate between women with children and without children, is 17, 22, and 14 percentage points at ages of twenties, thirties and forties, respectively (see top of Figure 2.1). This section's estimate of the effect of the grandmother's death, through its impact on childcare, is a 12 percentage points reduction in employment rate. Keeping preferences, socioeconomic constraints, gender roles, and discriminatory demand fixed, a reduction to childcare availability results in a reduction of mothers' employment by a magnitude larger than half the entire motherhood penalty.

If a lack of childcare availability and a parent-gender component are jointly contributing to the formation of the gender gap in employment, the death of the grandmother, through its impact on childcare, would have a larger negative effect on mothers' employment than on fathers'. Panel B of Table 2.1 compares the triple-difference effect for fathers to that of mothers using a quadruple difference. The effect of the grandmother's death, through the childcare mechanism, is 14.7 pp larger reduction in employment rate for mothers than for fathers. Columns 2-11 contain estimates Columns 2-11 display alternative specifications gradually reducing what the fixed effects control for; the estimates of the coefficient of interest are consistent across specifications and the quadruple difference estimate ranges between 7.4 and 14.8 pp. For the four quarters before the death of the grandmother, the employment rate of each of the four subgroups (men and women in households with young and with older children) has a flat trend and is not statistically different from its level in the last period before the death, see Figure 2.4). After the death of the grandmother, only mothers in households where the oldest child is less than five years have an economically and statistically significant drop in employment rate.

The findings are consistent with mothers having a greater share of the responsibility for childcare provision. The Mexican National Bureau of Statistics implicitly acknowledged these asymmetries. For example, in the ENESS, question 22 reads, "[w]hen the mother of [name of child] goes to work, the child stays with?" There is no equivalent question for when the father goes to work. Moreover, for the possible answers to this question, the grandmother is an explicit option, but it was not until the 2013 survey that the father was included as an explicit possible answer. Grandfathers have never been included as an explicit option (INEGI, 2009; INEGI, 2013).

The gender gap in employment in Mexico, the difference in employment rate between women and men, is at its maximum size during ages twenties, thirties, and forties, ranging between 24 and 30 percentage points (see Figure 2.1). This section's estimate of the differential effect on employment across genders of the grandmother's death, through its impact on childcare, is 15 percentage points.

2.4.2 Bounds for the Average Effect of Grandmothers' Death on Women

The estimated 12pp decrease in mothers' employment rate after the grandmother's death is based on three-generation households. Using two sets of assumptions, I create bounds for the average effect for women in Mexico. The average effect is a weighted average of the effect on one, two, three, and more than three generations households. The following equation denotes this weighted average, where i denotes the number of generations in a household and Share_i is the share of households with i generations:[35]

$$\text{Average Effect} = \sum_{i=1}^{4} \text{Effect}_i \times \text{Share}_i \tag{2.6}$$

The effect is only known for three-generation households (i=3). I use two sets of assumptions to determine the effect on other households. For the lower bound, I assume that grandmothers' deaths only affect three-generation households. Hence the lower bound is 0 x 23% + 0 x 61% + 12 x 15% + 0 x 1% = 1.8pp. For the upper bound, I use two assumptions: i) conditionally on the grandmother providing childcare, the negative effect of grandmothers' deaths through childcare is the same for mothers in two- and three-generation households, and ii) there is no effect through childcare when grandmothers do not provide childcare. The effect on mothers of three-generation households where the grandmother provides childcare

[35] Households with four or more generations are represented by i=4.

is given by the average effect on three-generation households divided by the share of households where the grandmother provides childcare (12pp/57% = 21pp). The average effect on mothers of two-generation households is the effect on mothers of three-generation households where the grandmother provides childcare times the probability of the grandmother providing childcare in two-generation households (21pp x 27% = 5.7pp). Hence the upper bound of the average effect is 5.27pp (0 x 23% + 5.7 x 61% + 12 x 15% + 0 x 1%). This exercise implies that the average effect of grandmothers' death on womens' employment in Mexico ranges between 1.8pp and 5.3pp (4.6% - 13.5%).[36]

2.4.3 Heterogeneous Effect of the Grandmother's Death

To measure heterogeneous effects, two additional coefficients are estimated. These coefficients are those on the interaction between the variable for which heterogeneous effects are estimated, $Z_{i,l}$, and the variables of Equation 2.1. The estimating equation is the following:

$$\begin{aligned} Employed_{i,l,t} \; &= \; \beta_1 Post_{i,l,t} \times Death^{GM}_{i,l} \; + \; \beta_2 Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l} \\ &+ \beta_3 Post_{i,l,t} \times Death^{GM}_{i,l} \times Z_{i,l} \; + \; \beta_4 Post_{i,l,t} \times Death^{GM}_{i,l} \times YoungChild_{i,l} \times Z_{i,l} \\ &+ \phi_i \; + \; \zeta_{l,t} \; + \; \gamma_{t,YoungChild} \; + \; \eta_{t,DeathGM} \; + \; \varepsilon_{i,l,t} \end{aligned} \tag{2.7}$$

Heterogeneity by Availability of Daycare

To create a measure of public and private daycare availability, I divide the number of public and private daycares in the municipality (from DENUE) by the number of children up to five years old (from the Population Census). I use this measure for 1,479 municipalities for which the ENOE also has data. If the availability of daycare is correlated with other variables,

[36]The estimation of the range includes mothers and women in general, not including women in one-generation households (non-mothers) would imply a range of 2.3pp - 6.8pp (the share of two- and three-generations households would increase by a factor of 1.29).

such as income, there is a risk that I capture heterogeneity by income instead of capturing heterogeneity by the availability of daycare. I address this concern by using an additional measure of daycare availability not driven by the average income, size, or share of working mothers. To construct this measure, I regress the average cost of daycare (either public or private) on the share of employed mothers, dummies for quintiles of average income, and dummies for quintiles of population, using the following estimating equation:

$$Availability_l = \beta_0 + \beta_1 ShareEmpMothers_l + \sum_{j=1}^{4} \psi_j Income_l^j + \sum_{j=1}^{4} \Gamma_j Population_l^j + \epsilon_l \quad (2.8)$$

The residual of the previous estimation is the measure daycare availability that is not explained by mothers' employment rate, income, and population.

The negative effect of the grandmother's death on mothers' employment is 9 pp smaller if public daycare is one standard deviation more available (Table 2.3, columns 2 and 3). Using the observed measure of daycare availability instead of the residual one leads to very similar results (columns 5 and 6). This finding is consistent with substitutability between the grandmother-provided childcare and public daycare if public daycare is available enough. At least two mechanisms could drive this substitution: i) when the grandmother dies, mothers in locations where public daycare is more available substitute grandmother-provided childcare with public daycares to continue to be employed, or ii) mothers in locations where public daycare is more available use public daycare more and grandmother-provided childcare less, hence the smaller effect. There is no heterogeneity by the availability of private daycare (columns 1, 3, 4, and 6). This finding suggests that increasing the availability of public daycare, can, as a stand-alone policy, significantly increase female employment and contribute to closing the gender gap.

Heterogeneity by Affordability of Daycare

To create a measure of daycare affordability, I average the hourly cost and total cost of daycare in the locality using data from the ENESS.[37] The ENESS includes the childcare alternative that households use, how much they pay, and for how many hours. The implicit assumptions of using these average total price and price per hour are that: i) the price paid by households that use private and public daycare is representative of the price that households that do not use these alternatives would pay, and ii) that the average computed from the ENESS respondents is informative of the cost level of daycare alternatives in the locality. To avoid issues related to the measure of daycare affordability capturing income, population, or share of mothers working, I also use a residualized measure estimated using Equation 2.8. This measure is equivalent to the one used for childcare availability in the previous section.

The negative effect of the grandmother's death on mothers' employment is 8 pp smaller if private daycare is one standard deviation cheaper (Table 2.4, columns 1 and 3). This heterogeneity is robust to using the residual and the actual affordability measures (columns 7 and 9). The result is also robust to using the hourly and total costs (columns 4, 6, 10, and 12). Public daycare has no equivalent heterogeneity (columns 2, 3, 5, 6, 8, 9, 11, and 12). There are two considerations with the public daycare cost measure: lack of price variation and capacity constraints (no vacancies) (Huerta, 2011). Public daycare is mostly free: more than one-fourth of the localities have an average cost of 0, and 96 percent have an average hourly cost below 0.33 USD.[38] Moreover, even if there was price variation for public daycare, it might not necessarily be a measure of how accessible it is because there are no vacancies.

[37] A locality in Mexico is any place in the country with one or more dwellings, inhabited or not, this place must be recognized by a name given by law or custom (INEGI, 2018). According to the 2010 Population Census, there are 3,647 urban localities (more than 2,500 inhabitants), with an average number of inhabitants of 23,656. Localities are the smallest geographical unit for which daycare costs are available.

[38] The exchange rate used to calculate is: 1USD = 15 MXN

On the other hand, private daycare price varies more because it is unregulated.[39]

The average cost of daycare may also capture the overall level of childcare costs in the locality, not only daycares. If this was the case, the heterogeneity that I find in the cost of daycare should also exist in other childcare alternatives. A common response to who takes care of children up to six years old when the mother goes to work in the ENESS is that the children go to school (21%, see Figure A.1). These schools may be public or private, but the ENESS keeps them in the same category when asking about the price paid. To create measures of affordability of schools and separate the cost of private ones, I compute two averages. The first one is the average cost paid for schools in the locality. This average includes both public and private schools. Since public schools are primarily free and private schools cost, I also use the average conditional on reporting a strictly positive price. This average will not capture free public schools but the cost of private schools.

I find that the negative effect of the grandmother's death on mothers' employment is 9 pp smaller if private schools are one standard deviation cheaper. This result stands irrespective of using the total or hourly cost (Table A.3, columns 2 and 4) and of using the residual or observed cost measure (columns 6 and 8). This heterogeneity is smaller and not statistically significant when using the school cost that includes both public and private schools (columns 1, 3, 5, and 7).[40] These results are consistent with those using daycare prices and share the same conclusion: the negative effect of the grandmother's death is smaller in locations with lower private childcare costs.

These estimates are consistent with those in the literature of other developing countries. Barros et al. (2013) finds that winning a child care slot in Rio de Janeiro increases

[39]There were less than 8 percent of localities with an average private daycare cost of 0, and 70 percent have an average hourly cost above 0.33 USD.

[40]I also estimated heterogeneity by the availability of daycares from the Estancias Infantiles para Apoyar a Madres Trabajadoras Program. I found that one standard deviation increase in the availability of these daycares (measured as estancias infantiles / number of children up to five years old) is associated with an 1-2pp smaller negative effect of grandmothers' death on mothers' employment. However, this result is measured imprecisely and is not statistically significant. This analysis is available upon request.

the mothers' employment probability by 10pp (27%). Hojman and Lopez Boo (2019) find that mothers' probability of working outside the household increases by 14 pp when receiving access to subsidized day care, and Halim et al. (2017) find that the expansion of public preschools in Indonesia increased the employment rate for women with preschool-age children.

Heterogeneity by the Grandmother's Side

Figure 2.5 shows that the negative effect of the grandmother's death is significantly larger if the maternal grandmother dies rather than the paternal one.[41] The total effect of the death of the paternal grandmother is a reduction in mothers' employment by 6pp. However, the impact through childcare is only 3 pp and not statistically significant. On the other hand, the maternal grandmother's death reduces mothers' employment by 21pp. The effect through childcare is a reduction by 17pp (difference between the effect when the oldest is at most five years old vs. when the oldest is older), 14pp larger than the effect of the paternal grandmother's death. This result is consistent with previous results in the development literature. For example, Duflo (2003) finds that there is an effect of grandmother's pension eligibility on weight for height of South African granddaughters only if the mother's mother is who becomes eligible.

Other Heterogeneity

This section presents a heterogeneity analysis of the effect of the grandmother's death on mothers' employment probability by the number of hours the grandmother provided care, the number of children the mom has, the number of grandkids in the household, the number

[41] Within the household, I identify the maternal grandmother relative to the household head as follows: i) the mother-in-law of a dad (household head), ii) the spouse of the grandfather (household head) who has a daughter that is a mother, iii) grandmother (household head) who had a daughter that is a mother, and iv) the mother of a mother (household head).

of grandkids under six years old, the number of male grandchildren, household income, mothers' income, mothers' hours worked, employment type (formal/informal), and mothers' education. I use the responses from the first wave so that the heterogeneity analysis can be interpreted as heterogeneity by ex-ante characteristics.

If childcare and not other forms of home production indeed drive the negative effect that I estimate, then the effect should increase as the grandmother's hours exclusively providing care increase. The ENOE contains a question regarding time allocated to care for others: "During last week, how much time did you spend exclusively taking care without pay of children, elderly, sick, or handicapped?" I use this question to present heterogeneity results in Column 1 of Table 2.5 by the time the grandmother provided care. One standard deviation increase in the number of hours the grandmother provided care (10.8 hrs) in the first survey wave is associated with a further reduction of mothers' employment rate by 9 pp, almost doubling the negative effect of the grandmother's death on mothers. This heterogeneity is not present in households where the oldest grandchild is older than five years.

I also present heterogeneity analysis for the number of children the mother has, the number of grandchildren in the household, and the number of grandchildren under six years of age. There is no statistically significant heterogeneity in all of these measures (Table 2.5, columns 2-4). However, there is significant heterogeneity in the gender composition of the grandchildren. If there are no male grandchildren, the grandmother's death, through its impact on childcare, reduces the mother's employment rate by 18.3pp, but for each male grandchild, this negative effect declines on average by 9pp (Column 5). Column 6 presents an alternative to measure the same heterogeneity: when there are no male grandchildren, the negative effect of the grandmother's death is 20.9pp, but if there is at least one male, this negative effect declines to 5.7pp. While fully characterizing the heterogeneity by gender of grandchildren is beyond the scope of this paper, this result is consistent with a society in which protecting and looking after girls is more important than protecting boys.

I find no heterogeneity by household income, mothers' income, and mothers' hours worked (Columns 7-9). However, there is economically significant heterogeneity by whether the mother was employed in the formal sector or not conditionally on being employed in the first survey wave. In the sample of mothers employed in the first survey wave, the grandmother's death, through its impact on childcare, reduces mothers' employment probability by 34.5pp for those in the informal sector (Column 10). However, this negative effect is 21.8pp smaller for those in the formal sector (yet this difference is not statistically significant, p-value = .15).

In terms of education, Column 11 and 12 show that in absolute terms, the negative effect is much more extensive for more educated mothers. However, relative terms may be more informative because more educated mothers are more likely to be employed. In particular, the grandmother's death reduces the employment rate for mothers without high school by 5.8pp (17%) and for those without college by 10.6pp (27%). The negative effect for those with high school is a reduction in employment rate by 19.1pp (36%) and for those with college by 19.3pp (31.4%). The grandmother's death reduces mothers' employment probability irrespective of their education level.

2.5 Robustness

This section is divided in two subsections: (i) alternative specifications, and (ii) the grandfather's death. The first subsection includes variations to the main specification: using an unbalanced panel, not restricting the maximum number of grandparents or parents in the household, broadening the definition of employment, using only the deaths of young grandmothers, using the age of the youngest child instead of the oldest, and estimating a double-difference only with the sample of parents who lived in a household where the grandmother died. The results are robust to all these alternative specifications. Since a grandfather is significantly less likely to provide childcare, the effect of a grandfather's death, through its impact on childcare, should be smaller (if any); this is documented empirically in the second subsection.

2.5.1 Alternative Specifications

Table 2.6 contains the main specification and ten alternative specifications. The results are robust to all these alternative specifications. The triple difference effect for the death of the grandmother on mothers, through childcare, ranges between a reduction of 8.7 to 16.3 percentage points in the employment rate, and the quadruple difference effect (the additional effect on mothers relative to fathers) ranges between an additional reduction of 6.0 to 21.2 percentage points.[42] Column 2 presents the estimates for the unbalanced panel, which includes households that responded to the ENOE less than five times. Instead of only including households with at most one grandmother and one grandfather, Column 3 allows for any number of first-generation individuals. Instead of only including households with at most one mother and one father, Column 4 allows for any number of fathers and mothers in the household. Column 5 broadens the definition of employed to also include employers, working on your own, and unpaid jobs. Column 6 broadens the definition of employed to also include employers and working on your own.

Throughout the paper, all the observed deaths of grandmothers are used to identify the effect of childcare availability on parents employment rate. Alternatively, I could use only the deaths of young grandmothers, whose death might be more unexpected. Columns 7 and 8 replicate the main estimation but using only the deaths of grandmothers at most 60 and 70 years old.

Column 9 repeats the estimation but using ENOE's probability weights that account, among many other things, for non-response. The estimates of interest are very similar in

[42] The ranges are for all specifications where the age of the oldest child is used as cutoff.

magnitude (within one standard error) and significance. I do not use this specification as the main one because INEGI designed these weights to make the survey representative at that quarter's state and country level. These weights are not necessarily representative of subsamples (three-generation households and three-generation households where the grandmother died). Moreover, the weights were designed to provide quarter-by-quarter snapshots of the labor market and not average effects on subsamples across years.

To disentangle the effect that the grandmother's death has through its impact on childcare from alternative mechanisms, the empirical strategy splits parents by the age of the oldest child. Alternatively, it is possible to use the age of the youngest child. One disadvantage of using the age of the youngest child is that the analysis would not restrict the presence of older children, who could provide childcare. Column 10 replicates the analysis but using the age of the youngest child instead of the oldest. The results are robust to using the oldest or youngest child's age, but as expected, since the specification of the youngest child allows for an additional childcare alternative (siblings), the effects are smaller. To show that older siblings providing care are substitutes for grandmother-provided care, Table A.4 presents the effect of the death of the grandmother on the amount of time that older grandchildren spend providing care. I estimate this effect for children ages 12 to 15, 12 to 18, and 12 to 21 in households where the youngest child is at most five years old, and the grandmother is less than 70.[43] The reason to estimate for grandmothers less than 70 is to avoid, to some extent, grandchildren providing care for the elderly.[44] The grandmother's death increases the amount of time older grandchildren spend providing care between 94 and 112%, and the probability of them providing care by 6 to 7pp (58 to 65%).

One of the three differences used in the triple-difference estimation, is comparing parents in households where the grandmother died vs households where she did not. Alternatively, I

[43]INEGI only asks the question for the population 12 years old and older.

[44]The question adds up the time spent providing care for children, the elderly, the sick, and the disabled.

could estimate a double-difference in the sample where the grandmother died (before vs after the death and young vs old children). A disadvantage of this alternative is the loss of precision from not estimating as precisely the time effects. Column 11 contains the double-difference estimations; the results are consistent with the estimates from the main specification.

2.5.2 The Grandfather's Death

Since a grandfather is less likely to provide childcare, the effect of the death of a grandfather, through its impact on childcare, should be smaller, if any. Top of Figure 2.6 displays the triple-difference estimates of Equation 2.3, but using a grandfather's death instead of the grandmother's. The death of a grandfather has no effect, through the childcare mechanism, on the employment rate of mothers.

2.6 Conclusion

Reducing the gender gap and the motherhood penalty in employment is a critical challenge in labor markets across the globe. Even though the gaps and their relationship with motherhood are well documented, we know less about the relative importance of each mechanism and its causal effect on employment. Innovative identification strategies, including natural experiments, allow researchers to disentangle the role of individual mechanisms in the formation of the gender gap.

This paper uses panel data, a natural experiment, and both a triple and a quadruple difference to estimate the effect of childcare availability on parents' employment rate. The evidence is consistent with the main driver of the gender gap and the motherhood penalty in labor force participation in Mexico being the combination of the lack of childcare availability and gender-asymmetric responsibility for childcare provision. A coinhabiting grandmother's death, through its impact on childcare availability, reduces the employment rate by 15 percentage points more for mothers than for fathers. This magnitude accounts for more than a half of the gender gap in employment in Mexico. Moreover, the death of the grandmother, through its impact on childcare, reduces the employment rate of mothers by 12 percentage points (27 percent); the effect accounts for more than half the entire motherhood penalty in Mexico. Even without changing preferences, socioeconomic constraints, and gender roles, increasing childcare availability can drastically reduce both the motherhood penalty and the gender gap.

In the short term, increasing the availability of childcare can have a significant effect on increasing mothers' labor force participation, which in turn can contribute to reshaping gender roles in the long term. Working women today can increase the opportunities for working women tomorrow by changing societal gender attitudes and perceptions, and by increasing the aspirations and educational attainment for girls.[45]

[45] Beaman et al. (2009) finds that exposure to a female chief councilor improves perceptions of female effectiveness as leaders and weakens gender-roles stereotypes in the public and domestic spheres. Beaman et al. (2012) finds that female leadership in village councils raises the aspirations and educational attainment for girls in India.

2.7 FIGURES

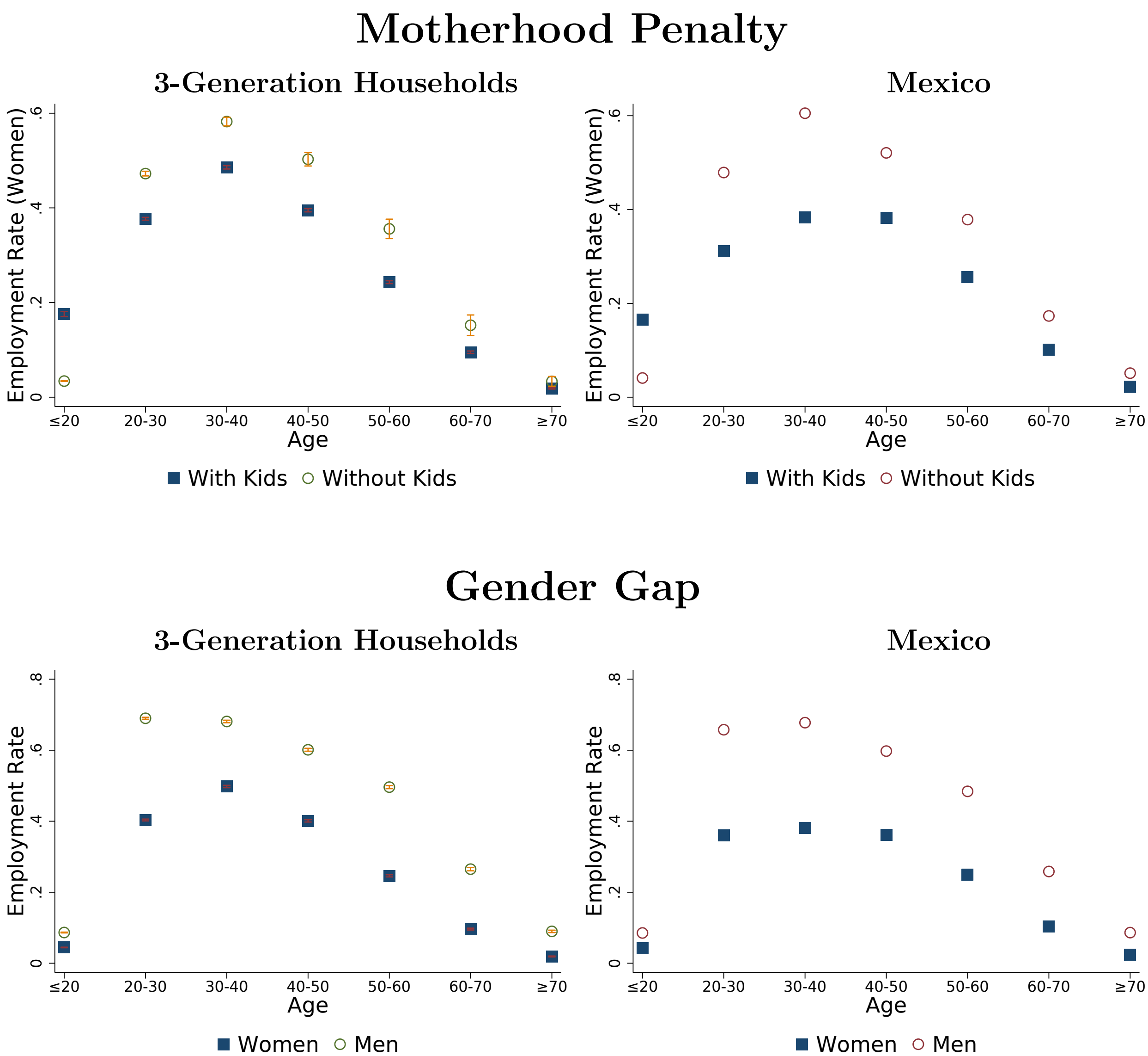


Figure 2.1: The Motherhood Penalty and the Gender Gap

Source: ENOE (Q1 2005 - Q1 2020)

Note: The graph displays the employment rate by age. The figures on the left include only three-generation households. The figures on the right include the full sample and use probability weights to obtain country-level representation. The figures on the top compare women with children to women without them. The figures on the bottom compare men to women.

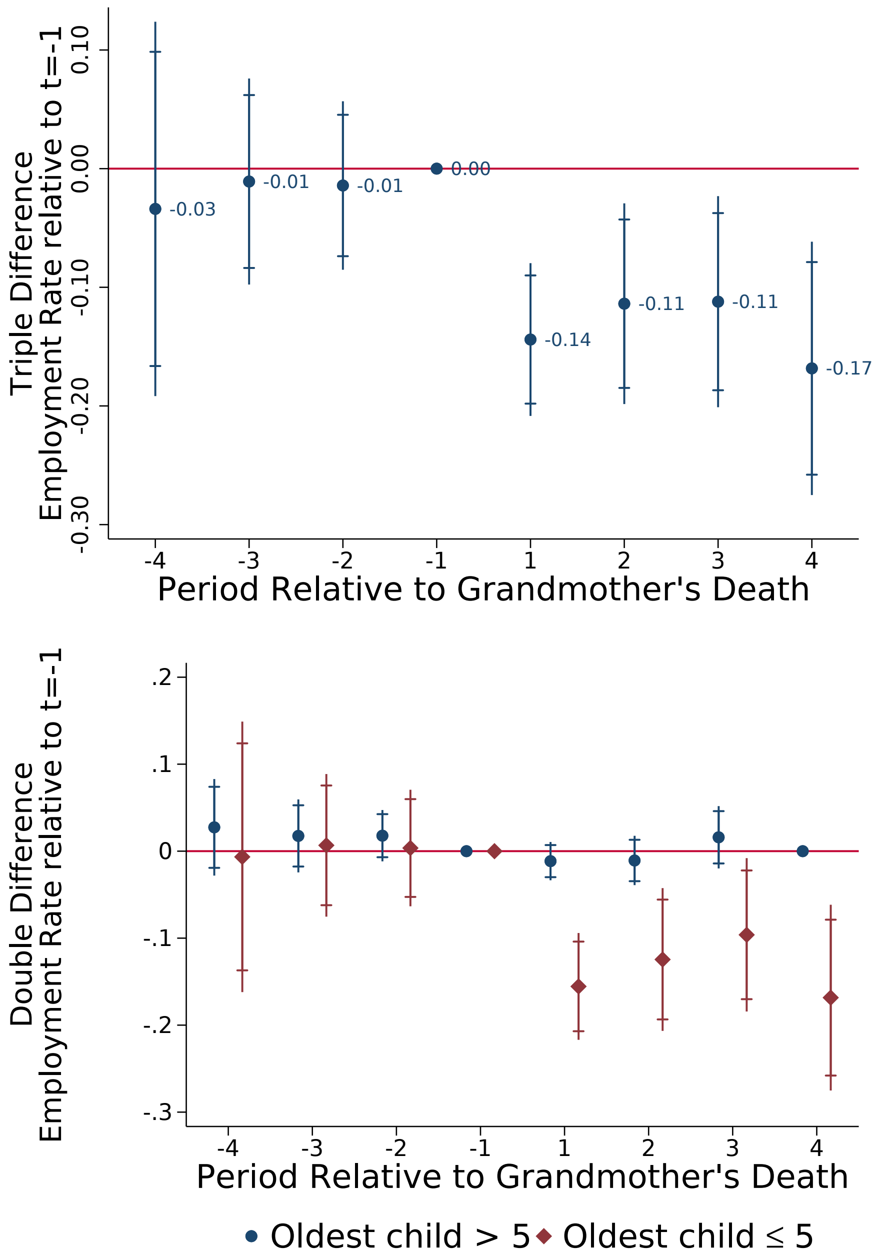


Figure 2.2: Event Study - Grandmother's Death and Mothers' Employment

Note: The graph displays the point estimate and the 90% and 95% confidence interval of the effect of the death of a grandmother on employment for mothers by quarter relative to the quarter just before the death, estimated using Equation 2.3. A household with young children is a household where the oldest child is at most 5 years old. The chart on the top is the double difference estimate, and the chart on the bottom is the first difference estimate. The sample includes mothers between 20 and 50 years old and living in three-generation household with five observations in the panel, one grandmother or one grandfather or both, the grandmother is at least forty years old, the oldest grandchild is at most thirty years old, the first generation is weakly older than the second generation, and the second generation is weakly older than the third generation. Standard errors are clustered at the household level.

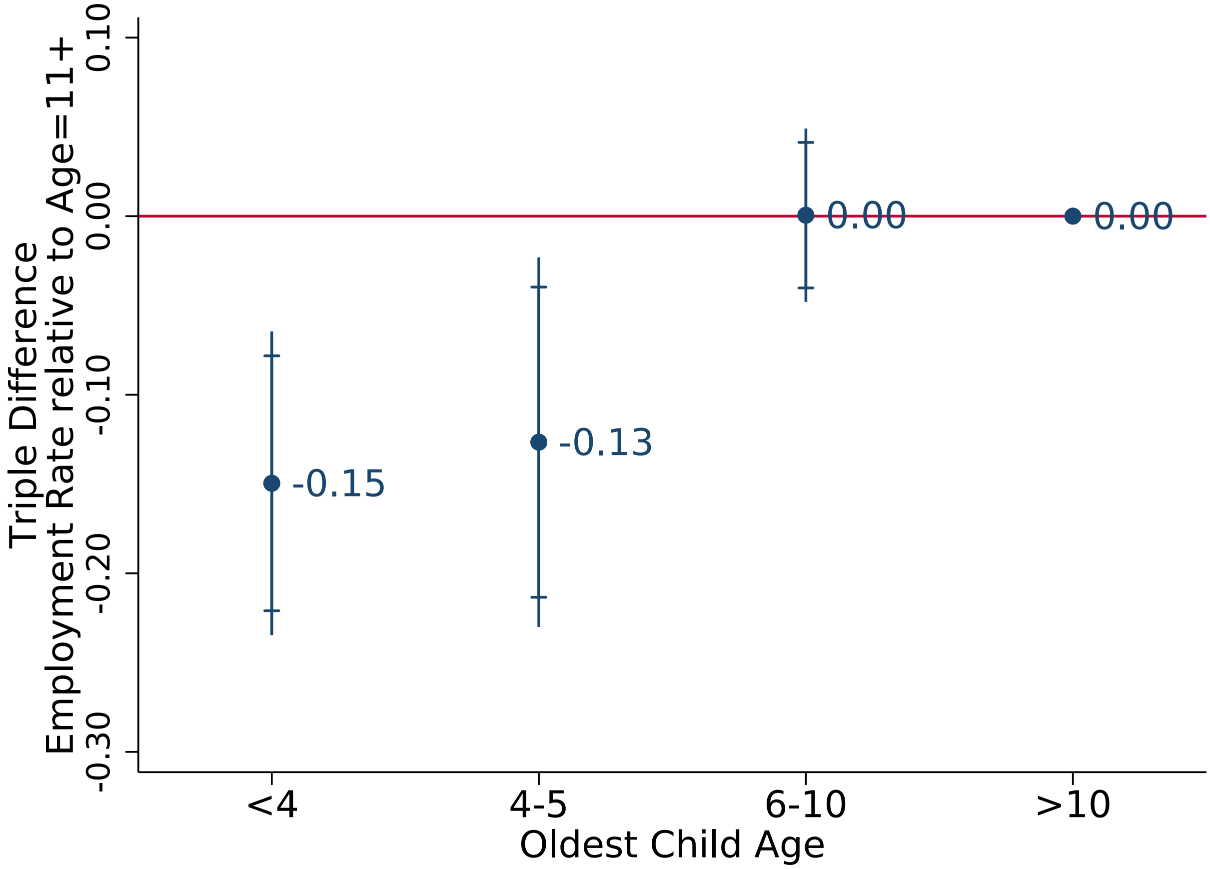


Figure 2.3: Grandmother's Death and Mothers' Employment by Age of the Oldest Child

Note: The graph displays the point estimate and the 90% and 95% confidence intervals of the additional effect that the death of a grandmother has on mothers' employment rate by age of the oldest child in the household relative to when the oldest child in the household is older than 10. The plotted coefficients are $\beta_{2,1}$, $\beta_{2,2}$, $\beta_{2,3}$ of Equation 2.2. The same sample as in Figure 2.2 is used. Standard errors are clustered at the household level.

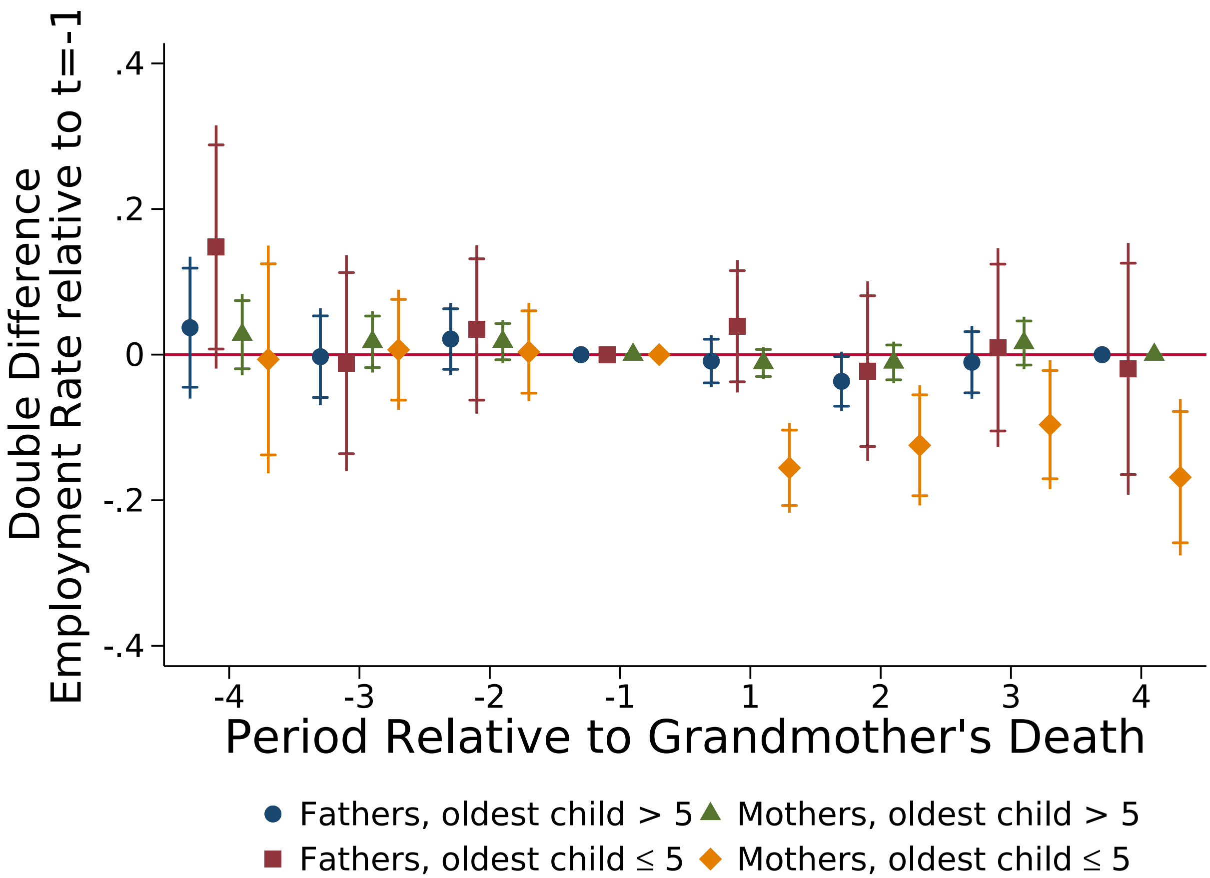


Figure 2.4: Event Study of Grandmother's Death (Mothers) for Mothers and Fathers

Note: The graph displays the point estimate and the 90 and 95% confidence interval of the effect that the death of a grandmother has on the employment rate for mothers and fathers estimated using Equation 2.5. A household with a young children is a household where the oldest child is at most 5 years old. The confidence intervals are computed using standard errors clustered at the household-level. The sample includes mothers and fathers between 20 and 50 years old and living in three-generation household with five observations in the panel, one grandmother or one grandfather or both, the grandmother is at least forty years old, the oldest grandchild is at most thirty years old, the first generation is weakly older than the second generation, and the second generation is weakly older than the third generation. Mothers are identified by belonging to the second generation and having children, and fathers are identified by belonging to the second generation and being married or coinhabiting with their spouse.

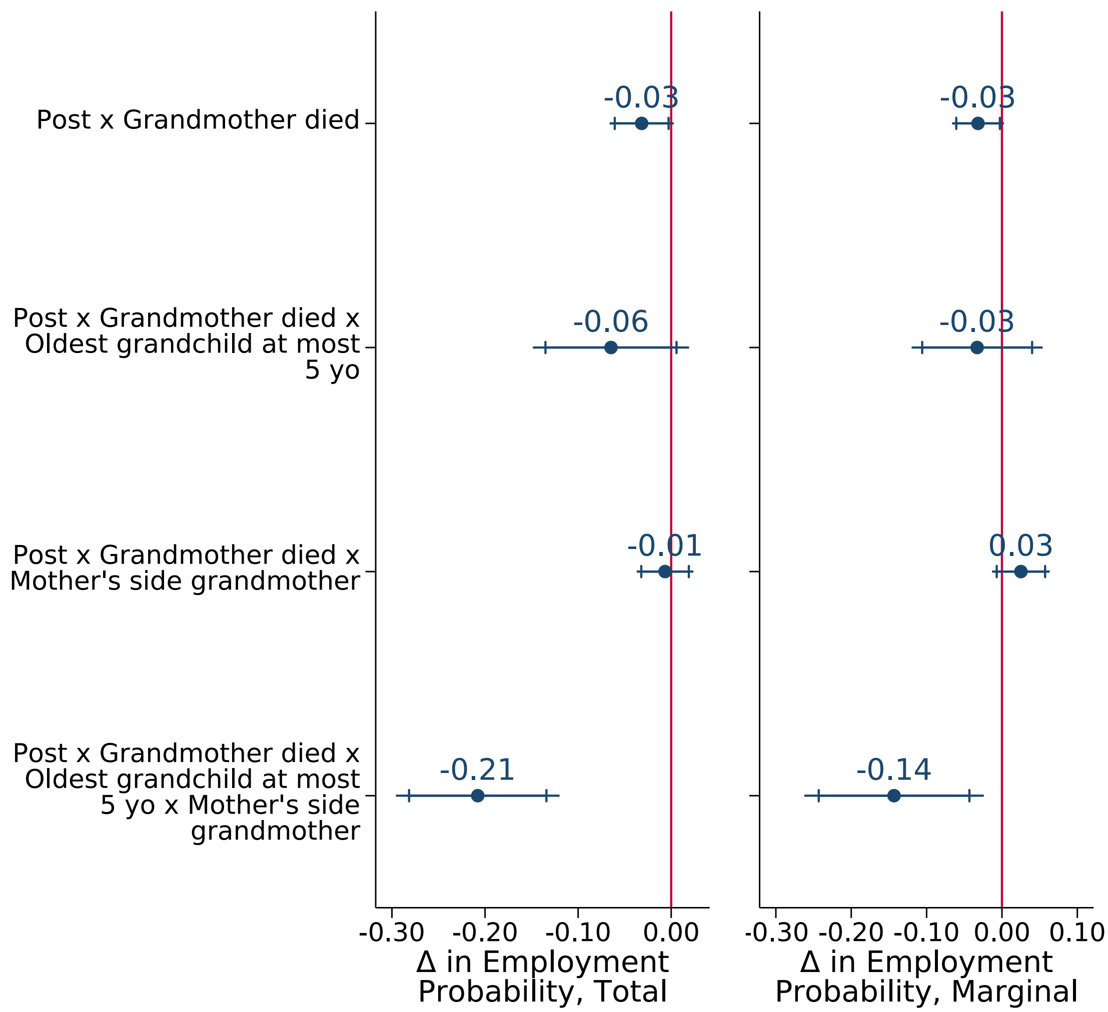


Figure 2.5: Heterogeneity by Grandmother's Side (Maternal vs. Paternal)

Note: The figure displays the total and marginal effect of the death of grandmothers on mothers' employment probability. The coefficients are estimated using equation 2.1, but adding an interaction of the first two terms with a dummy that indicates whether the grandmother who died was the mother's side.

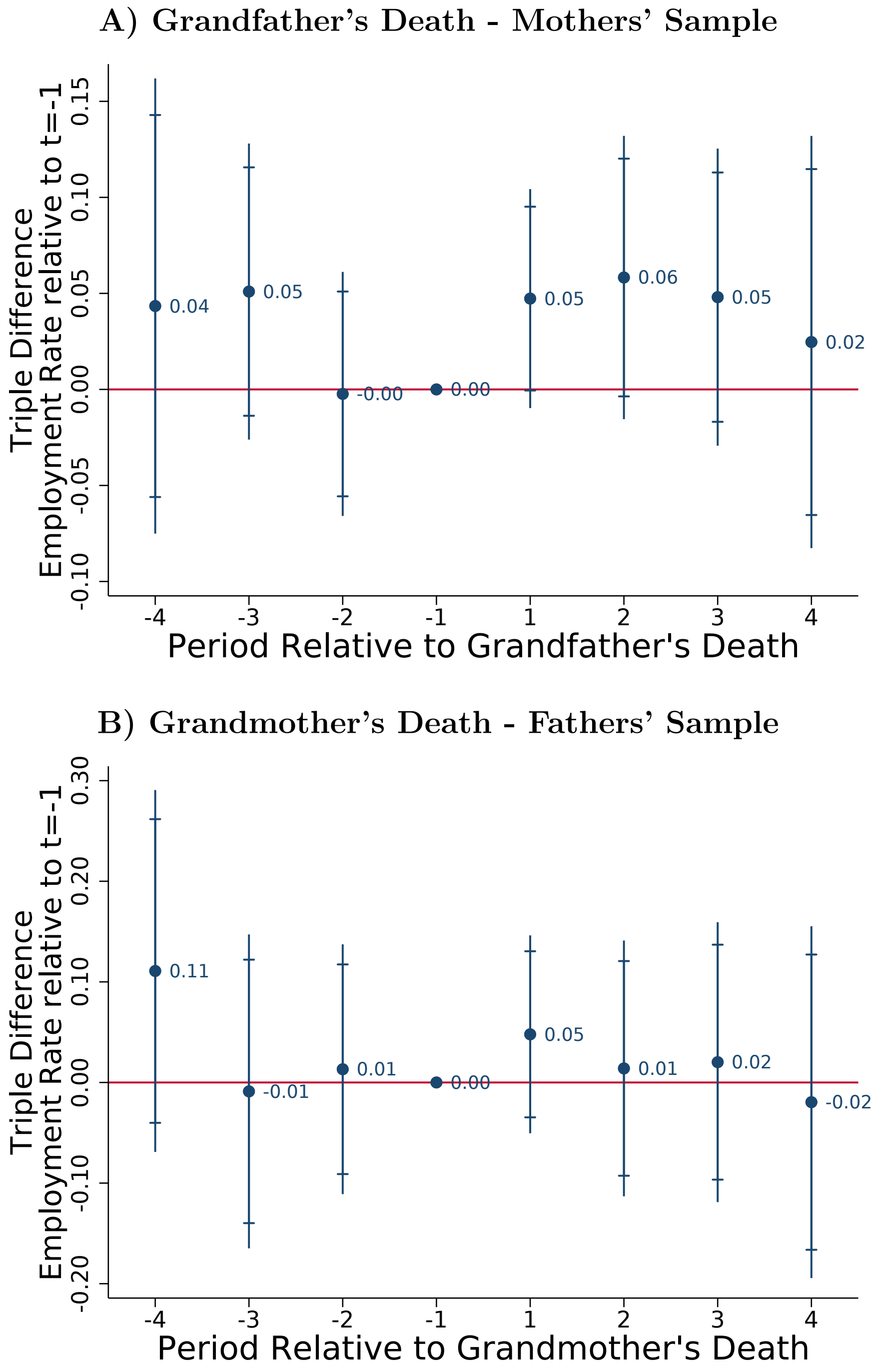


Figure 2.6: Event Study of Grandfather's Death (Mothers) and Grandmother's Death (Fathers)

Note: The top (bottom) graph displays the point estimate and the 90 and 95% confidence interval of the effect that the death of a grandfather (grandmother) has on the employment rate of mothers (fathers) by period relative to the period just before the death. A household with a young children is a household where the oldest child is at most 5 years old. Standard errors are clustered at the household level. The same sample as in Figure 2.4 is used.

2.8 TABLES

Table 2.1: Grandmother's Death and Employment Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A) Mothers											
Post x Grandmother died	-0.0154	-0.0223*	-0.0159	0.00740	-0.0172	0.00642	-0.0175	0.00355	-0.00990	-0.0290**	-0.0307**
	(0.0133)	(0.0123)	(0.0133)	(0.00960)	(0.0120)	(0.00959)	(0.0120)	(0.00852)	(0.00848)	(0.0140)	(0.0151)
Post x Grandmother died x Oldest grandchild at most 5 years old	-0.124***	-0.107***	-0.121***	-0.122***	-0.0916***	-0.118***	-0.0888***	-0.0905***	-0.0870***	-0.0798**	-0.0754*
	(0.0307)	(0.0278)	(0.0307)	(0.0313)	(0.0279)	(0.0312)	(0.0278)	(0.0281)	(0.0280)	(0.0401)	(0.0427)
N	484,464	484,464	484,464	484,464	484,464	484,464	484,464	484,464	484,464	484,244	484,464
Panel B) Mothers and Fathers											
Post x Grandmother died	-0.00971	-0.00930	-0.00428	0.0201	-0.0104	0.0318**	-0.00519	0.0161	-0.0221	-0.00131	0.00266
	(0.0219)	(0.0211)	(0.0219)	(0.0153)	(0.0204)	(0.0153)	(0.0204)	(0.0137)		(0.0186)	(0.0189)
Post x Grandmother died x Oldest grandchild at most 5 years old	0.0232	0.0267	-0.0116	0.0265	0.0381	-0.00869	0.00567	0.0412	0.0102	-0.00233	-0.00179
	(0.0418)	(0.0371)	(0.0416)	(0.0423)	(0.0344)	(0.0422)	(0.0344)	(0.0354)	(0.0353)	(0.0475)	(0.0474)
Post x Grandmother died x Mother	-0.00569	-0.0130	-0.0116	-0.0127	-0.00682	-0.0254	-0.0123	-0.0125	0.0122	-0.0277	-0.0333
	(0.0252)	(0.0243)	(0.0252)	(0.0183)	(0.0234)	(0.0183)	(0.0234)	(0.0163)	(0.0162)	(0.0234)	(0.0244)
Post x Grandmother died x Oldest grandchild at most 5 x Mother	-0.147***	-0.134***	-0.109**	-0.148***	-0.130***	-0.110**	-0.0945**	-0.132***	-0.0973**	-0.0774	-0.0736
	(0.0489)	(0.0431)	(0.0486)	(0.0492)	(0.0405)	(0.0490)	(0.0404)	(0.0413)	(0.0411)	(0.0595)	(0.0625)
N	743,733	743,733	743,733	743,733	743,733	743,733	743,733	743,733	743,733	743,215	743,733
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	-
Year - Quarter - Locality - Gender FE	Y	-	Y	Y	-	Y	-	-	-	Y	Y
Year - Quarter - Young Child - Gender FE	Y	Y	-	Y	Y	-	-	Y	-	-	-
Year - Quarter - Grandmother Died - Gender FE	Y	Y	Y	-	Y	-	Y	-	-	-	-
Year - Locality - Gender FE	-	Y	-	-	-	-	-	-	-	-	-
Age - Gender FE	-	-	-	-	-	-	-	-	-	Y	-
Household composition - Gender FE	-	-	-	-	-	-	-	-	-	Y	-
Household income - Gender FE	-	-	-	-	-	-	-	-	-	Y	-
Education - Gender FE	-	-	-	-	-	-	-	-	-	Y	-

Note: All models estimate the coefficients of lower level interactions if they are not captured by the fixed effects. The sample includes "mothers" and "fathers" of the second generation between 20 and 50 years of age living in three-generation households: females with children are classified as mothers and males that are married or coinhabiting are classified as fathers. In panel B) the fixed effects are interacted with the gender of the parent (Panel A only has mothers). The Age x Gender fixed effect (FE) uses 5-year age brackets. The HH Composition FE is the interaction of the number of members in the second generation, in the third generation, and in the household. GF age and GM age are the Grandfather and Grandmother age FE. Income FE is the decile of per capita family income. The Education x Gender FE is the maximum level of education interacted by gender. Households included in the sample have 5 observations, one grandmother or one grandfather or both, the grandmother is at least forty years old, the oldest grandchild is at most thirty years old, the first generation is weakly older than the second generation, and the second generation is weakly older than the third generation. Standard errors are clustered at the household level. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Correia (2016) is used to estimate high-dimensional FE.

Table 2.2: Grandmother's Death, Employment, Hours Worked, and Earned Income

	Employment					Hours Worked		Earned Income	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable:	Full-time	Part-time	Part-time	Part-time	Part-time	Intensive + Extensive	Intensive	Intensive + Extensive	Intensive
Post grandmother death	-0.0253*	0.00989	0.0316	-0.0266	0.0440	0.00306	0.0357	0.0145	0.132
	(0.0151)	(0.0131)	(0.0287)	(0.0851)	(0.0297)	(0.0718)	(0.0578)	(0.151)	(0.107)
Post grandmother death x	-0.0850***	-0.0391*	-0.126*	-0.339**	-0.0575	-0.355**	-0.127	-0.763**	-0.296
oldest grandchild at most 5	(0.0289)	(0.0233)	(0.0647)	(0.152)	(0.0499)	(0.164)	(0.106)	(0.339)	(0.180)
Sample	Full	Full	Employed First Wave	Part-time First Wave	Full-time First Wave	Income and Hours Avail.	Employed	Income and Hours Avail.	Employed
N	484,454	484,454	196,376	35,894	148,905	393,456	123,942	393,456	123,942

Note: The table displays the marginal effect of the grandmother's death on the probability of being full-time employed, half-time employed, and the inverse hyperbolic sine of earned income and hours worked. For columns 6 and 8 only observations with either both strictly positive hours worked and earned income or both hours worked and earned income equal to zero are included. Columns 7 and 9 include observations with strictly positive hours worked and earned income. Hours worked and income are winsorized at the 5% level from each tail. Standard errors clustered at the household level. Part-time employment is 30 hours or less per week, and Full-time employment is more than 30 hours a week.

Table 2.3: Heterogeneity by Daycare Availability

	Residual			Observed		
	(1)	(2)	(3)	(4)	(5)	(6)
Post grandmother death	-0.0155	-0.0155	-0.0156	-0.0160	-0.0156	-0.0161
	(0.0133)	(0.0133)	(0.0133)	(0.0132)	(0.0133)	(0.0132)
Post x Grandmother Died x Oldest Grandchild at most 5 years old	-0.125***	-0.123***	-0.122***	-0.123***	-0.121***	-0.120***
	(0.0312)	(0.0304)	(0.0307)	(0.0306)	(0.0301)	(0.0299)
Post x Grandmother Died x Private Daycares per Child	0.0153		0.0151	0.00870		0.00820
	(0.00968)		(0.00969)	(0.0095)		(0.00955)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Private Daycares per Child	-0.0120		-0.0237	-0.0163		-0.0352
	(0.0305)		(0.0309)	(0.029)		(0.0302)
Post x Grandmother Died x Public Daycares per Child		0.00353	0.00198		0.00435	0.00325
		(0.00737)	(0.00698)		(0.00771)	(0.0074)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Public Daycares per Child		0.0860**	0.0910**		0.0781**	0.0948**
		(0.039)	(0.0401)		(0.0376)	(0.0408)
N	483,425	483,425	483,425	484,454	484,454	484,454
Individual FE	Y	Y	Y	Y	Y	Y
Year - Quarter - Locality FE	Y	Y	Y	Y	Y	Y
Year - Quarter - Young Child FE	Y	Y	Y	Y	Y	Y
Year - Quarter - Grandmother Died FE	Y	Y	Y	Y	Y	Y

Note: The table displays heterogeneity of the marginal effect of the grandmother's death on mother's employment by daycare availability, estimated using Equation 2.7. Daycare per child is calculated by dividing the number of daycare facilities by the number of children up to five years old in the municipality. Daycares per child are standardized. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Standard errors are clustered at the household level.

Table 2.4: Heterogeneity by Daycare Affordability

	Residual						Observed					
	Hourly Cost			Total Cost			Hourly Cost			Total Cost		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post grandmother death	-0.0161	-0.0141	-0.0133	-0.0158	-0.0141	-0.0130	-0.0164	-0.0145	-0.0137	-0.0159	-0.0144	-0.0132
	(0.0156)	(0.0149)	(0.0156)	(0.0156)	(0.0149)	(0.0156)	(0.0156)	(0.0149)	(0.0156)	(0.0156)	(0.0149)	(0.0157)
Post x Grandmother Died x Oldest Grandchild at most 5 years old	-0.110***	-0.122***	-0.110***	-0.112***	-0.122***	-0.112***	-0.111***	-0.123***	-0.112***	-0.113***	-0.123***	-0.114***
	(0.0371)	(0.0371)	(0.0375)	(0.0385)	(0.0370)	(0.0388)	(0.0380)	(0.0374)	(0.0386)	(0.0390)	(0.0374)	(0.0398)
Post x Grandmother Died x Cost of Public Daycare		0.00592	0.0105		0.00316	0.00562		0.00198	0.00362		0.00372	0.00501
		(0.0111)	(0.0138)		(0.0106)	(0.0132)		(0.0113)	(0.0139)		(0.0108)	(0.0134)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Cost of Public Daycare		-0.00684	-0.00512		-0.00631	-0.00190		0.000515	0.00781		0.00480	0.0132
		(0.0332)	(0.0371)		(0.0299)	(0.0326)		(0.0359)	(0.0396)		(0.0328)	(0.0357)
Post x Grandmother Died x Cost of Private Daycare	-0.00813		-0.00627	0.000904		0.00308	-0.000468		0.00266	0.00498		0.00761
	(0.0114)		(0.0115)	(0.0118)		(0.0121)	(0.0115)		(0.0116)	(0.0123)		(0.0127)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Cost of Private Daycare	-0.0767**		-0.0790**	-0.0648**		-0.0676**	-0.0601*		-0.0649*	-0.0508		-0.0563*
	(0.0370)		(0.0370)	(0.0330)		(0.0327)	(0.0352)		(0.0354)	(0.0344)		(0.0340)
N	316,832	354,865	312,892	316,832	354,865	312,892	317,239	355,928	313,279	317,239	355,928	313,279
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Locality FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Young Child FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Grandmother Died FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table displays heterogeneity of the marginal effect of the grandmother's death on mother's employment by daycare affordability, estimated using Equation 2.7.. Public and private daycare costs are standardized. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Standard errors are clustered at the household level.

Table 2.5: Other Heterogeneity

Heterogenous Effect on Employment Probability by:	(1) Hours Grandmother Provided Care†	(2) # of Kids	(3) # of Grandkids	(4) # of Grandkids < 6	(5) # of Male Grandkids	(6) 1+ Male Grandkids	(7) Household Income†	(8) Income†	(9) Hours†	(10) Formal Employment	(11) High School +	(12) College +
Post grandmother death	-0.0197	-0.0286	-0.0148	-0.0196	-0.0111	0.00472	-0.0164	-0.0199	-0.0187	-0.0879**	-0.0360**	-0.0161
	(0.0158)	(0.0228)	(0.022)	(0.0143)	(0.0172)	(0.02)	(0.0133)	(0.0241)	(0.0239)	(0.0361)	(0.0157)	(0.0139)
Post x Grandmother Died x Oldest Grandchild at most 5 years old	-0.122***	-0.134**	-0.172**	-0.167**	-0.183***	-0.209***	-0.129***	-0.200**	-0.201**	-0.345***	-0.0581	-0.106***
	(0.0305)	(0.062)	(0.0678)	(0.0659)	(0.0436)	(0.0493)	(0.0316)	(0.0822)	(0.0828)	(0.125)	(0.0409)	(0.0332)
Post x Grandmother Died x Z	-0.0126	0.00529	-0.000279	0.0116	-0.00391	-0.0270	-0.00912	0.0204	-0.00697	0.0814**	0.0490***	0.00488
	(0.0205)	(0.00768)	(0.00796)	(0.0155)	(0.00936)	(0.021)	(0.00882)	(0.0128)	(0.0187)	(0.0368)	(0.0189)	(0.0272)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Z	-0.0904**	0.0117	0.0363	0.0243	0.0897**	0.152**	-0.0319	0.0277	-0.0189	0.218	-0.133**	-0.0866
	(0.0446)	(0.0362)	(0.0447)	(0.0466)	(0.0369)	(0.0621)	(0.04)	(0.0512)	(0.0867)	(0.15)	(0.0605)	(0.0823)
N	484,454	484,454	484,454	484,454	484,454	484,454	484,454	175,210	175,210	196,376	484,454	484,454
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Y - Q - Locality FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Y - Q - Young Child FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Y - Q - Grandmother Died FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table displays heterogeneity of the marginal effect of the grandmother's death on mother's employment by the variable on the column header, estimated using Equation 2.7. Variables with a † at the end are standardized. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Standard errors are clustered at the household level. Y-Q stands for Year-Quarter.

Table 2.6: Effect on Employment Probability: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Base	Unbalanced	Any # of grandparents	Any # of parents	Any type of work	Any paid work	Grandmother ≤ 60	Grandmother ≤ 70	ENOE weights	Youngest ≤ 5	DiD
Panel A) Mothers											
Post x Grandmother died	-0.0154	-0.0216*	-0.0150	-0.0270**	0.00309	0.00317	-0.0263	-0.0303*	-0.0366*	-0.0131	0.0131
	(0.0133)	(0.0126)	(0.0130)	(0.0123)	(0.0163)	(0.0158)	(0.0290)	(0.0172)	(0.0201)	(0.0146)	(0.0283)
Post x Grandmother died x Oldest	-0.124***	-0.0986***	-0.129***	-0.100***	-0.0909**	-0.101***	-0.163***	-0.137***	-0.147***	-0.0440**	-0.161**
grandchild at most 5 years old	(0.0307)	(0.0273)	(0.0306)	(0.0285)	(0.0371)	(0.0356)	(0.0548)	(0.0369)	(0.0447)	(0.0191)	(0.0672)
N	484,454	561,119	487,651	620,172	484,454	484,454	484,454	484,454	484,454	484,454	2,561
Panel B) Mothers and Fathers											
Post x Grandmother died	-0.00971	0.00449	-0.00520	-0.00313	0.000665	-0.00409	-0.0308	-0.0237	0.0531*	-0.0353	-0.000486
	(0.0219)	(0.0208)	(0.0218)	(0.0215)	(0.0170)	(0.0180)	(0.0396)	(0.0260)	(0.0298)	(0.0242)	(0.0433)
Post x Grandmother died x Oldest	0.0232	0.0200	0.0244	0.0337	0.0198	0.00551	0.00390	0.0754	0.0374	0.0589**	0.0298
grandchild at most 5 years old	(0.0418)	(0.0395)	(0.0422)	(0.0383)	(0.0341)	(0.0356)	(0.0739)	(0.0478)	(0.0563)	(0.0285)	(0.0957)
Post x Grandmother died x Mother	-0.00569	-0.0261	-0.00985	-0.0239	0.00242	0.00726	0.00449	-0.00656	-0.0898**	0.0222	0.0136
	(0.0252)	(0.0240)	(0.0251)	(0.0242)	(0.0232)	(0.0238)	(0.0507)	(0.0315)	(0.0354)	(0.0284)	(0.0507)
Post x Grandmother died x Oldest	-0.147***	-0.119***	-0.154***	-0.134***	-0.111**	-0.107**	-0.167*	-0.212***	-0.185***	-0.103***	-0.190*
grandchild at most 5 x Mother	(0.0489)	(0.0454)	(0.0499)	(0.0451)	(0.0487)	(0.0495)	(0.0862)	(0.0568)	(0.0697)	(0.0341)	(0.103)
N	743,723	861,568	749,231	933,821	743,723	743,723	743,723	743,723	743,723	743,723	3,591
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter-Young Child-Gender FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
Year-Quarter-GM Died-Gender FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
Year-Quarter-Locality-Gender FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table displays the marginal effect ofthe grandmother's death on the employment rate of mothers (Panel A) and the employment rate of mothers and fathers (Panel B). Column 1 is the main specification. Column 2, Unbalanced, the restriction of observing the household for five surveys is dropped. Column 3, Any number of grandparents, allows for any number of members of the first generation of the household. Column 4, any number of parents, allows for any number of members of the second generation of the household. In Column 5, Any Work, the dependent variable takes the value of one if the individual is a subordinate and paid employee, an employer, works on his/her own, or works without pay. In Column 6, Any paid Work, the dependent variable takes the value of one if the individual is a subordinate and paid employee, an employer, or works on his/her own. In Column 7 and 8, the Grandmother died dummy takes the value of 1 only if the grandmother that died was under 60 or 70 years old. In Column 9, ENOE weights, the estimation uses the probability weights available in the ENOE. In Column 10, Youngest ≤ 5, the dummy Young Children takes the value of 1 in the youngest child in the household is at most 5 years old. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. The numbers in parenthesis are the standard errors clustered at the household-level.

Bibliography

Aguilar Gomez, S., E. Arceo-Gomez, and E. Toledo (2019, 01). Inside the black box of child penalties. *SSRN Electronic Journal*.

Agüero, J. M. and M. S. Marks (2008). Motherhood and female labor force participation: Evidence from infertility shocks. *The American Economic Review, Papers and Proceedings 98*, 500–504.

Alfaro-Urena, A., I. Manelici, and J. Vasquez (2019). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *Working Paper*.

Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics 34*, 545–579.

Angrist, J. D. and W. N. Evans (1998). Children and their parents' labor supply: Evidence from exogenous variation in family. *The American Economic Review 88*, 450–477.

Aparicio Fenoll, A. (2019). The uneven impact of women's retirement on their daughters' employment. *Review of Economics of the Household*.

Aparicio-Fenoll, A. and M. Vidal-Fernandez (2015, 03). Working women and fertility: the role of grandmothers' labor force participation. *CESifo Economic Studies 61*(1), 123–147.

Arceo-Gómez, E. O. and R. M. Campos-Vázquez (2014, jul.). Evolución de la brecha salarial de género en méxico. *El Trimestre Económico 81*(323), 619–653.

Arpino, B., C. D. Pronzato, and L. P. Tavares (2014). The effect of grandparental support on mothers' labour market participation: An instrumental variable approach. *European Journal of Population 30*, 369–390.

Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2017, 03). Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan*. *The Quarterly Journal of Economics 132*(3), 1101–1164.

Atkin, D., D. Donaldson, I. Rasul, M. Teachout, E. Verhoogen, and C. Woodruff (2019). Firms, trade, and productivity. *IGC International Growth Center, Evidence Paper*.

Atkin, D., B. Faber, and M. Gonzalez-Navarro (2018). Retail globalization and household welfare: Evidence from mexico. *Journal of Political Economy 126*(1), 1–73.

Atkin, D., A. K. Khandelwal, and A. Osman (2017, 02). Exporting and Firm Performance: Evidence from a Randomized Experiment*. *The Quarterly Journal of Economics 132*(2), 551–615.

Bao, C. G. and M. X. Chen (2018, October). Foreign rivals are coming to town: Responding to the threat of foreign multinational entry. *American Economic Journal: Applied Economics 10*(4), 120–57.

Barros, R. P. d., P. Olinto, T. Lunde, and M. Caralho (2013). The impact of access to free childcare on women's labor market outcomes : evidence from a randomized trial in low-income neighborhoods of rio de janeiro. *Gender Impact : the World Bank's Gender Impact Evaluation Database.*.

Basker, E. (2005, 02). Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion. *The Review of Economics and Statistics 87*(1), 174–183.

Basker, E. and M. Noel (2009). The evolving food chain: Competitive effects of wal-mart's entry into the supermarket industry. *Journal of Economics & Management Strategy 18*(4), 977–1009.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics 124*, 1497–1540.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science 335*, 582–586.

Becker, G. (1991). *A Treatise on the Family*. Harvard University Press.

Bergquist, L. F. and M. Dinerstein (2020, December). Competition and entry in agricultural markets: Experimental evidence from kenya. *American Economic Review 110*(12), 3705–47.

Bertrand, M., C. Goldin, and L. F. Katz (2010, July). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics 2*(3), 228–55.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2012, 11). Does Management Matter? Evidence from India *. *The Quarterly Journal of Economics 128*(1), 1–51.

Bratti, M., T. Frattini, and F. Scervini (2018, October). Grandparental availability for child care and maternal labor force participation: pension reform evidence from Italy. *Journal of Population Economics 31*(4), 1239–1277.

Bresnahan, T. F. and P. C. Reiss (1990). Entry in monopoly markets. *The Review of Economic Studies 57*(4), 531–553.

Bresnahan, T. F. and P. C. Reiss (1991). Entry and competition in concentrated markets. *Journal of Political Economy 99*(5), 977–1009.

Bruhn, M., D. Karlan, and A. Schoar (2018, apr). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in mexico. *Journal of Political Economy 126*(2), 635–687.

Busso, M., O. Fentanes, and S. Levy (2018). The longitudinal linkage of mexico's economic census 1999-2014. *IDB Technical Note ; 1477*.

Busso, M. and S. Galiani (2019, January). The causal effect of competition on prices and quality: Evidence from a field experiment. *American Economic Journal: Applied Economics 11*(1), 33–56.

Calderón, G. (2014, Apr). The Effects of Child Care Provision in Mexico. Working Papers 2014-07, Banco de México.

Chen, F. and G. Liu (2011, 12). The Health Implications of Grandparents Caring for Grandchildren in China. *The Journals of Gerontology: Series B 67B*(1), 99–112.

Clark, S., C. W. Kabiru, S. Laszlo, and S. Muthuri (2019, August). The Impact of Childcare on Poor Urban Women's Economic Empowerment in Africa. *Demography 56*(4), 1247–1272.

Coca-Cola Mexico (2020, August). Hablemos de tienditas. `http://www.coca-colamexico.com.mx/noticias/comunidad/hablemos-de-tienditas`. Accessed: 2021-10-08.

COFECE (2020). Comisión federal de competencia económica, estudio de competencia en el canal moderno del comercio al menudeo de alimentos y bebidas.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019, Aug). Inference with Arbitrary Clustering. (12584).

Compton, J. and R. A. Pollak (2013). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics 79*, 72–90.

Correia, S. (2016). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Technical report. Working Paper.

Correia, S. (2018, September). IVREGHDFE: Stata module for extended instrumental variable regressions with multiple levels of fixed effects. Statistical Software Components, Boston College Department of Economics.

Correll, S. J., S. Benard, and I. Paik (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology 112*, 1297–1339.

Cristia, J. P. (2008). The effect of a first child on female labor supply evidence from women seeking fertility services. *Journal of Human Resources 43*, 487–510.

Daymont, T. N. and P. J. Andrisani (1984). Job preferences, college major, and the gender gap in earnings. *The Journal of Human Resources 19*(3), 408–428.

De Mel, S., D. McKenzie, and C. Woodruff (2008, 11). Returns to Capital in Microenterprises: Evidence from a Field Experiment*. *The Quarterly Journal of Economics 123*(4), 1329–1372.

Dhar, D., T. Jain, and S. Jayachandran (2019). Intergenerational transmission of gender attitudes: Evidence from india. *The Journal of Development Studies 55*(12), 2572–2592.

Di Gessa, G., K. Glaser, and A. Tinker (2016). "the impact of caring for grandchildren on the health of grandparents in europe: A lifecourse approach". *"Social Science and Medicine" "152"*, "166 – 175".

Duflo, E. (2003, 06). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review 17*(1), 1–25.

Fafchamps, M., D. McKenzie, S. Quinn, and C. Woodruff (2014). Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in ghana. *Journal of Development Economics 106*, 211–226.

Fernandez, R. and J. Wong (2014a). Divorce risk, wages and working wives: A quantitative life-cycle analysis of female labour force participation. *The Economic Journal 124*, 319–358.

Fernandez, R. and J. Wong (2014b). Unilateral divorce, the decreasing gender gap, and married women's labor force participation. *American Economic Review 104*, 342–347.

Fernández-Kranz, D., A. Lacuesta, and N. Rodríguez-Planas (2013). The motherhood earnings dip evidence from administrative records. *The Journal of Human Resources 48*, 169–197.

Givord, P. and C. Marbot (2015). Does the cost of child care affect female labor market participation? an evaluation of a french reform of childcare subsidies. *Labour Economics*.

Gonzalez Sanchez, R. and M. Gaytán (2015, 12). Análisis de la elección del consumidor entre tiendas de conveniencia y tiendas de abarrotes en colima. uso del modelo de regresión multinomial logit. *Paradigma económico 7*, 27 – 46.

Halim, D., H. Johnson, and E. Perova (2017, March). Could Childcare Services Improve Women's Labor Market Outcomes in Indonesia? World Bank Publications - Reports 31484, The World Bank Group.

Haltiwanger, J., R. Jarmin, and C. Krizan (2010). Mom-and-pop meet big-box: Complements or substitutes? *Journal of Urban Economics 67*(1), 116–134. Special Issue: Cities and Entrepreneurship.

Hank, K. and I. Buber (2009). Grandparents caring for their grandchildren: Findings from the 2004 survey of health, ageing, and retirement in europe. *Journal of Family Issues 30*(1), 53–73.

Hausman, J. and E. Leibtag (2007). Consumer benefits from increased competition in shopping outlets: Measuring the effect of wal-mart. *Journal of Applied Econometrics 22*(7), 1157–1177.

Hojman, A. and F. Lopez Boo (2019). Cost-effective public daycare in a low-income economy benefits children and mothers. *IDB WORKING PAPER SERIES Nº IDB-WP-1036*.

Holmes, T. J. (2011). The diffusion of wal-mart and economies of density. *Econometrica 79*(1), 253–302.

Hsieh, C.-T. and B. A. Olken (2014, September). The missing "missing middle". *Journal of Economic Perspectives 28*(3), 89–108.

Hubbard, T. N. (2004). Affiliation, integration, and information: ownership incentives and industry structure. *The Journal of Industrial Economics 52*(2), 201–227.

Huerta, A. (2011). Problemática de las madres con hijos pequeños para acceder o permancer en el mercado laboral.

Hughes, M. E., L. J. Waite, T. A. LaPierre, and Y. Luo (2007, 03). All in the Family: The Impact of Caring for Grandchildren on Grandparents' Health. *The Journals of Gerontology: Series B 62*(2), S108–S119.

INEGI (1997). Población.

INEGI (2009). Encuesta nacional de empleo y seguridad social. `http://www.beta.inegi.org.mx/proyectos/enchogares/modulos/eness/2009/default.html`.

INEGI (2013). Encuesta nacional de empleo y seguridad social. `http://www.beta.inegi.org.mx/proyectos/enchogares/modulos/eness/2013/default.html`.

INEGI (2014). La informalidad laboral encuesta nacional de ocupación y empleo marco conceptual y metodológico. `http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/contenidos/espanol/bvinegi/productos/metodologias/ENOE/ENOE2014/informal_laboral/702825060459.pdf`.

INEGI (2018). *Manual de Cartografía Geoestadística*.

Jaumotte, F. (2003, December). Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries. OECD Economics Department Working Papers 376, OECD Publishing.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics 7*(1), 63–88.

Jayachandran, S. (2021, jul). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review 69*(3), 576–595.

Jensen, R. and N. H. Miller (2018, December). Market integration, demand, and the growth of firms: Evidence from a natural experiment in india. *American Economic Review 108*(12), 3583–3625.

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

Jérôme, A., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*, 293–337.

Karlan, D., R. Knight, and C. Udry (2015). Consulting and capital experiments with microenterprise tailors in ghana. *Journal of Economic Behavior and Organization 118*, 281–302. Economic Experiments in Developing Countries.

Khanna, M. and D. Pandey (2021). Reinforcing gender norms or easing houseworkburdens? the role of mothers-in-law in determiningwomen's labor force participation. *Working Paper*.

Kleven, H., C. Landais, and J. E. Søgaard (2019, October). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kühn, S., R. Horne, and S. Yoon (2017). *World employment and social outlook : trends for women 2017*. Geneva: ILO.

Lefebvre, P. and P. Merrigan (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from canada. *Journal of Labor Economics 26*(3), 519–548.

Macchiavello, R. and A. Morjaria (2020, 12). Competition and Relational Contracts in the Rwanda Coffee Chain*. *The Quarterly Journal of Economics 136*(2), 1089–1143.

Martínez A., C. and M. Perticará (2017). Childcare effects on maternal employment: Evidence from Chile. *Journal of Development Economics 126*(C), 127–137.

Matsa, D. A. (2011, 08). Competition and Product Quality in the Supermarket Industry*. *The Quarterly Journal of Economics 126*(3), 1539–1591.

Maurer-Fazio, M., R. Connelly, L. Chen, and L. Tang (2011). Childcare, eldercare, and labor force participation of married woman in urban china, 1982-2000. *The Journal of Human Resources 46*, 261–294.

McKenzie, D. and C. Woodruff (2015, August). Business Practices in Small Firms in Developing Countries. (21505).

McKenzie, D. and C. Woodruff (2017). Business practices in small firms in developing countries. *Management Science 63*(9), 2967–2981.

OIT (2014). Organización internacional del trabajo, el empleo informal en méxico: situación actual, políticas y desafíos.

O'Neill, J. (2003). The gender gap in wages, circa 2000. *The American Economic Review 93*(2), 309–314.

Posadas, J. and M. Vidal-Fernandez (2013). Grandparents' childcare and female labor force participation. *IZA Journal of Labor Policy 2:14*.

Profeco (2004). Procuraduría federal del consumidor, para el cuidado de sus pequeños... guarderÍas.

Royston, P. and P. C. Lambert (2011). *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*. StataCorp LP.

Sen, A. (1999). *Development as Freedom*. Oxford University Press.

SME Finance Forum (2019). Msme economic indicators.

U.S. Census Bureau (2013). Who's minding the kids? child care arrangements: Spring 2011.

Waldfogel, J. (1998). Understanding the "family gap" in pay for women with children. *The Journal of Economic Perspectives 12*, 137–156.

Zanella, G. (2017). How does grandparent childcare affect labor supply? *IZA World of Labor 337*.

Appendix A

Appendix Tables and Figures

Surviving Competition:

Neighborhood Shops vs Convenience Chains

A.1 Appendix: Tables and Figures

A.1.1 Tables

Table A.1: Effect of Chains on Shop Survival

	Dependent Variable:		Dependent Variable: Store Level (Exit=1)							
	Cox		Poisson				OLS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of Chain Stores	0.047***	0.039***	0.047***	0.039***	0.030***	0.030***	0.021***	0.017***	0.013***	0.014***
	(0.0012)	(0.0012)	(0.0008)	(0.0008)	(0.0009)	(0.0019)	(0.0004)	(0.0004)	(0.0005)	(0.0009)
Observations	1,892,525	1,643,883	1,892,525	1,643,883	1,643,612	1,641,019	1,892,525	1,643,883	1,643,873	1,643,184
Store Controls		Y		Y	Y	Y		Y	Y	Y
Year x City FE					Y	Y			Y	Y
Neighborhood FE						Y				Y
Mean Dep. Variable | Chains>0	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
Mean Chain Stores | Chains>0	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5

Note: The table displays the estimation of survival models. Column 1 and 2 are Cox survival models. Column 3-5 are survival models estimated using a Poisson and age of establishment fixed effect measured by the number of censuses the establishment has been open. Hazard ratios of Cox models and Poisson models after splitting on all observed failure times are identical (Royston and Lambert, 2011, Section 4.5). Hence, the coefficients of columns 1-2 and 3-4 are identical, but the standard errors reflect the differences in underlining assumptions of each method. Columns 6-8 are OLS estimates with age of establishment fixed effects.

Table A.2: Industry (Neighborhood) Level Effects

	Dependent Variable: Inverse Hyperbolic Sine of Total												
	Profits	Resale Revenue	Revenue	Purchases for Resale	Value Added	Profits per Worker	Initial Resale Inventory	Final Resale Inventory	Hours per Employed	Hours Worked	Total Employed	Output/ Input Ratio	Fixed Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Number of Chain Stores	-0.044	-0.044	-0.043	-0.044	-0.043	-0.044	-0.047	-0.043	-0.033	-0.033	-0.032	-0.033	-0.033
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
Observations	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152	189,152
Year x City FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Clustered SE	City	City	City	City	City	City	City	City	City	City	City	City	City
Average Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4
Effect from 0 to Avg. # Chain Stores	-27.4%	-27.4%	-27.1%	-27.9%	-26.9%	-27.5%	-29.7%	-27.0%	-21.1%	-20.7%	-20.2%	-20.7%	-20.9%
KP F-statistic	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08	73.08

Note: The table displays the estimation of Equation 1.3 using 2SLS. The dependent variable is the inverse hyperbolic sine of the total of the variable in each column header. The total is the sum across all shops in the neighborhood.

Table A.3: Shop Level Effects

Sample: Non-Entry and Non-Exit Shops	Dependent Variable: Inverse Hyperbolic Sine of Average												
	Profits	Resale Revenue	Revenue	Purchases for Resale	Value Added	Profits per Worker	Initial Resale Inventory	Final Resale Inventory	Hours per Employed	Hours Worked	Total Employed	Output/ Input Ratio	Fixed Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Number of Chain Stores	-0.012	-0.009	-0.009	-0.010	-0.011	-0.012	-0.013	-0.012	-0.001	-0.001	0.000	0.000	-0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.000)	(0.001)	(0.001)	(0.000)	(0.004)
Observations	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482	185,482
Year x City FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Clustered SE	City	City	City	City	City	City	City	City	City	City	City	City	City
Average Chain Stores | Chains>0	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5
Effect from 0 to Avg. # Chain Stores	-7.55%	-6.08%	-5.78%	-6.73%	-7.01%	-7.54%	-8.68%	-7.57%	-0.61%	-0.39%	0.24%	-0.27%	-2.71%
KP F-statistic	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77	71.77

Note: The table displays the estimation of Equation 1.3 using 2SLS. The dependent variable is the inverse hyperbolic sine of the average of the variable in each column header. The average is across all the shops in the neighborhood that do not appear in the census tract for the first or last time.

Table A.4: Effects on Performance of Hybrid Stores

	Intensive + Extensive Margin Effect Dependent Variable: IHS of Sum of				Intensive Margin Effect Dependent Variable: IHS of Average				
	Resale Revenue	Revenue	Profits	Profits per Worker	Resale Revenue	Revenue	Profits	Profits per Worker	Output/ Input
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of Chain Stores	-0.0538	-0.0540	-0.0592	-0.0518	-0.0142	-0.0145	-0.0196	-0.0136	-0.0010
	(0.0108)	(0.0107)	(0.0109)	(0.0097)	(0.006)	(0.006)	(0.007)	(0.005)	(0.0010)
Observations	151,470	151,470	151,470	151,470	151,470	151,470	151,470	151,451	151,443
Year x City FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Neighborhood FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Clustered SE	City	City	City	City	City	City	City	City	City
Effect from 0 to Avg. # Chain Stores	-35.08%	-35.23%	-38.56%	-33.83%	-9.47%	-9.62%	-13.01%	-9.02%	-0.68%
KP F-statistic	68.66	68.66	68.66	68.66	68.66	68.66	68.66	69.65	68.70

Note: The table displays the estimation of Equation 1.3 using 2SLS replacing the dependent variable with the inverse hyperbolic sine of the sum or average resale revenue, revenue, profits, profits per worker, and output to input ratio of hybrid stores.

Table A.5: Effect by City Size

	Dependent Variable: # of Neighborhood Shops			
	IV All Urban	Towns Avg. pop 14K	Cities Avg. pop 262K	Large Cities Avg. pop 880K
	(1)	(2)	(3)	(4)
Number of Chain Stores	-4.60***	-1.66**	-4.08***	-6.14***
	(0.669)	(0.718)	(0.616)	(1.607)
Observations	190,664	90,357	65,759	36,293
# of Cities	1,961	1,813	120	29
Neighborhood & Year x City FE	Y	Y	Y	Y
Mean Dep. Variable | Chains>0	175	124	191	200
Mean Chain Stores | Chains>0	6.4	3.2	7.3	8.1
Effect from 0 to Avg. # Conv. Stores	-16.9%	-4.3%	-15.7%	-24.9%
KP F-statistic	79.79	111.80	115.72	14.67

Note: The table displays the estimation of Equation 1.3 using 2SLS splitting the sample by town size.

Table A.6: Robustness - Census Tract Level Effect

	OLS			2SLS	Reduced Form	First Stage
Dependent Variable:	# of Shops	# of Shops	# of Shops	# of Shops	# of Shops	# of Chain
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Chain Stores	0.03	-0.29***	-0.58*** .0	-2.07***		
	(0.195)	(0.074)	(0.059)	(0.326)		
Economies of Scale$_{c,t}$ x					-0.23*** .0	0.11***
Chain Suitability$_{m,c}$					(0.033)	(0.012)
Observations	190,614	190,614	190,614	190,614	190,563	190,563
Neighborhood FE		Y	Y	Y	Y	Y
Year x City FE			Y	Y	Y	Y
Clustered SE	City	City	City	City	City	City
Mean Dep. Variable | Chains>0	13	13	13	13	13	2
KP F-statistic				93.61		

Note: The table replicates Table 1.2, but defining the neighborhoods as census tracts. This is the smallest possible neighborhood size.

Table A.7: Robustness - Alternative Standard Errors

	Dependent Variable: # of Neighborhood Shops				
	(1)	(2)	(3)	(4)	(5)
Number of Chain Stores	-4.62***	-4.62***	-4.62***	-4.62***	-4.62***
	(0.673)	(0.236)	(0.635)	(0.457)	(0.417)
Observations	190,664	190,664	190,664	190,664	190,664
Neighborhood FE	Y	Y	Y	Y	Y
Year x City FE	Y	Y	Y	Y	Y
Clustered SE	City	Neighborhood Year	City Year	City x Year	City x Year Neighborhood
Mean Dep. Variable | Chains>0	175	175	175	175	175
Mean Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.4
From 0 to Avg. # Conv. Stores	-17.0%	-17.0%	-17.0%	-17.0%	-17.0%
Underid. KP LM stat	30.73	3.21	3.00	94.59	89.93
Weak ID F statistic	35,658	35,658	35,658	27,463	35,658
KP F-statistic	80.10	123.65	56.24	174.24	211.74

Note: The table displays the estimation of Equation 1.3 using 2SLS clustering the standard errors at different levels.

Table A.8: Robustness - Logarithmic Specifications

	Dependent Variable: # of Neighborhood Shops					
	OLS	OLS	OLS	2SLS	2SLS (Log-Linear)	2SLS (Log-Log)
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Chain Stores	3.33***	-0.56***	-2.01***	-4.60***	-0.03***	-0.76***
	(0.536)	(0.210)	(0.314)	(0.669)	(0.003)	(0.118)
Observations	190,664	190,664	190,664	190,664	190,664	190,664
Neighborhood FE		Y	Y	Y	Y	Y
Year x City FE			Y	Y	Y	Y
Mean Dep. Variable | Chains>0	175	175	175	175	175	175
Mean Chain Stores | Chains>0	6.4	6.4	6.4	6.4	6.4	6.4
From 0 to Avg. # Conv. Stores	12.2%	-2.1%	-7.4%	-16.9%	-20.0%	
KP F-statistic				79.79	80.29	35.80

Note: The table displays the estimation of Equation 1.3. Column 4 uses the natural logarithm of number of shops as the dependent variable. Column 5 uses the natural logarithm of number of shops as the dependent variable and the natural logarithm of number of chains stores as dependent variable.

Table A.9: Purchases in Store Types by Home Ownership

	Dependent Variable: $ Purchases in Store													
	Convenience Chains							Corner Shops						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
I[Owns Living Space=1]	-75.4	-54.3	-54.2	-44.2	-66.2	-60.9	-26.9	137.4	90.3	90.2	11.9	167.7	81.1	196.1
	(17.9)	(14.2)	(14.2)	(13.8)	(14.2)	(13.9)	(14.3)	(40.8)	(34.7)	(34.5)	(32.8)	(32.2)	(31.5)	(29.9)
Mean Dependent Variable	193	193	193	193	193	193	193	2,128	2,128	2,128	2,128	2,128	2,128	2,128
Observations	36,016	35,966	35,966	35,966	35,966	35,966	35,966	36,016	35,966	35,966	35,966	35,966	35,966	35,966
Cluster SE	City	City	City	City	City	City	City	City	City	City	City	City	City	City
Census Tract FE		Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y
HH Socioeconomic Strata FE			Y			Y	Y			Y			Y	Y
HH Dwelling Type FE				Y		Y	Y				Y		Y	Y
HH Income pc decile FE					Y	Y	Y					Y	Y	Y
HH Head Age Contol							Y							Y

Note: The table displays the estimates of regressing the expenditure at the household level in shops / chains on a dummy variable of whether the household owns its home using data from ENIGH 2018.

A.1.2 Figures

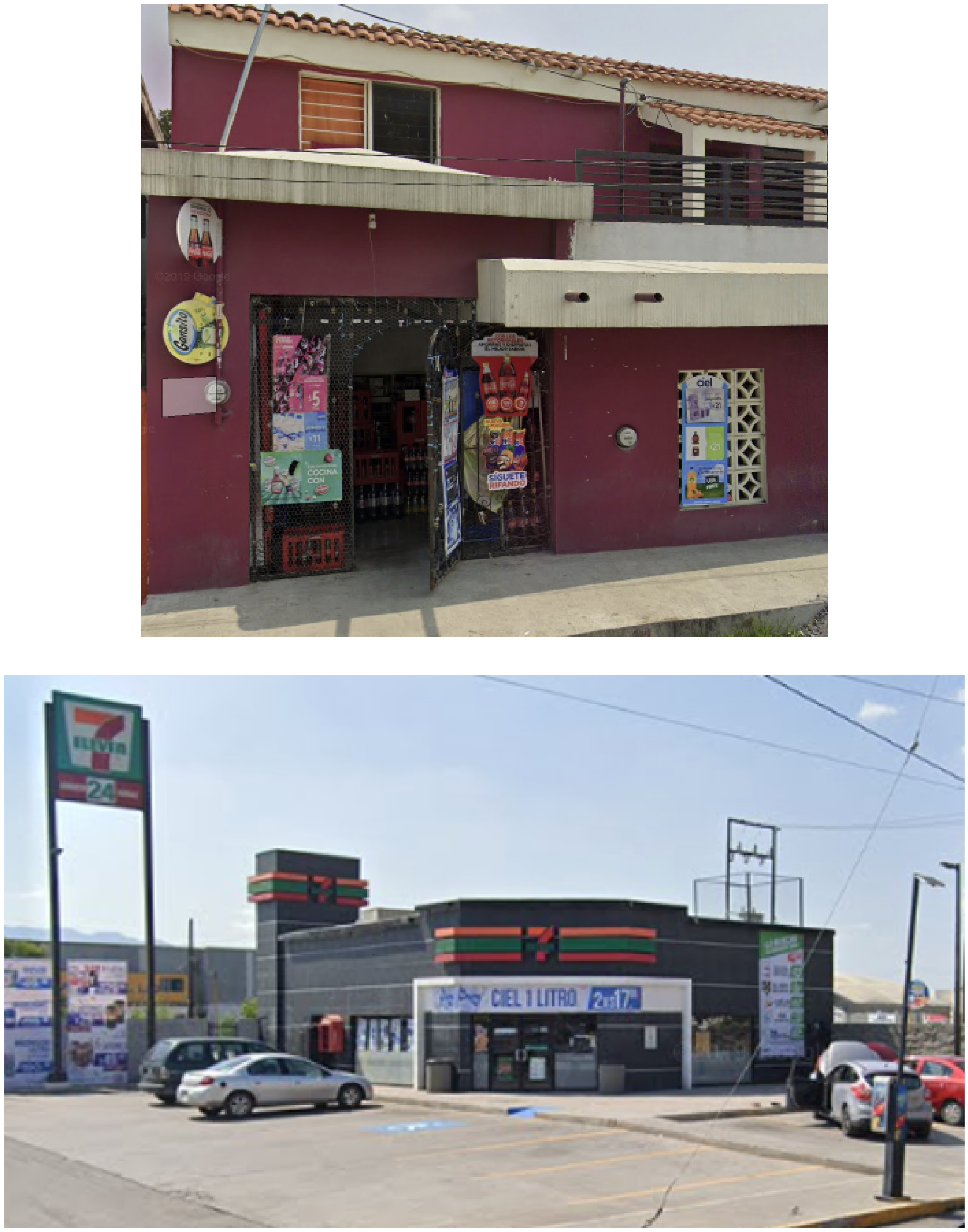


Figure A.1: Shops and Chains

Source: Google Maps
Note: The figure contains an example of a shop (left) and a chain store (right) in Saltillo, Mexico. These two stores are within 15 meters of each other.

1999 **2019**

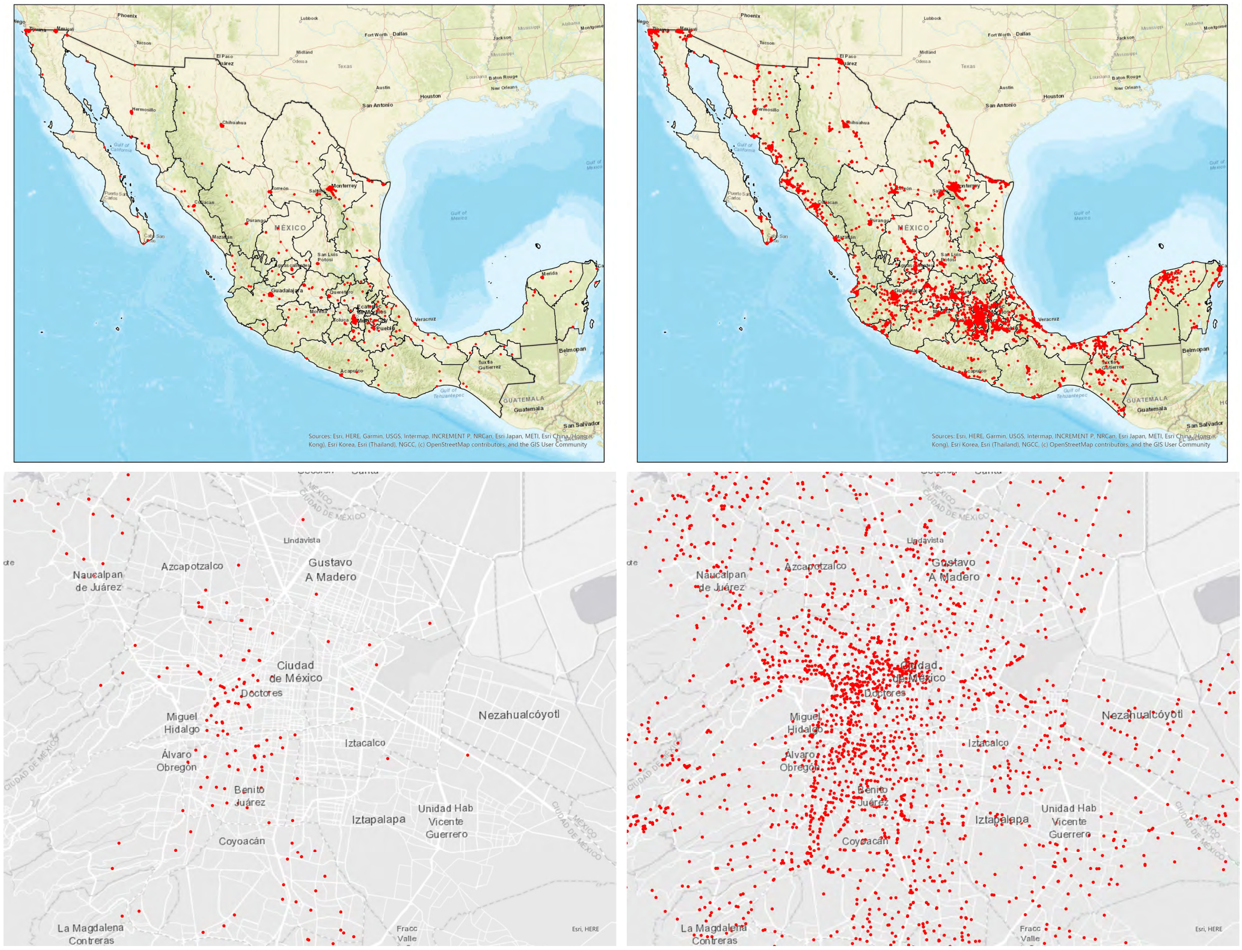
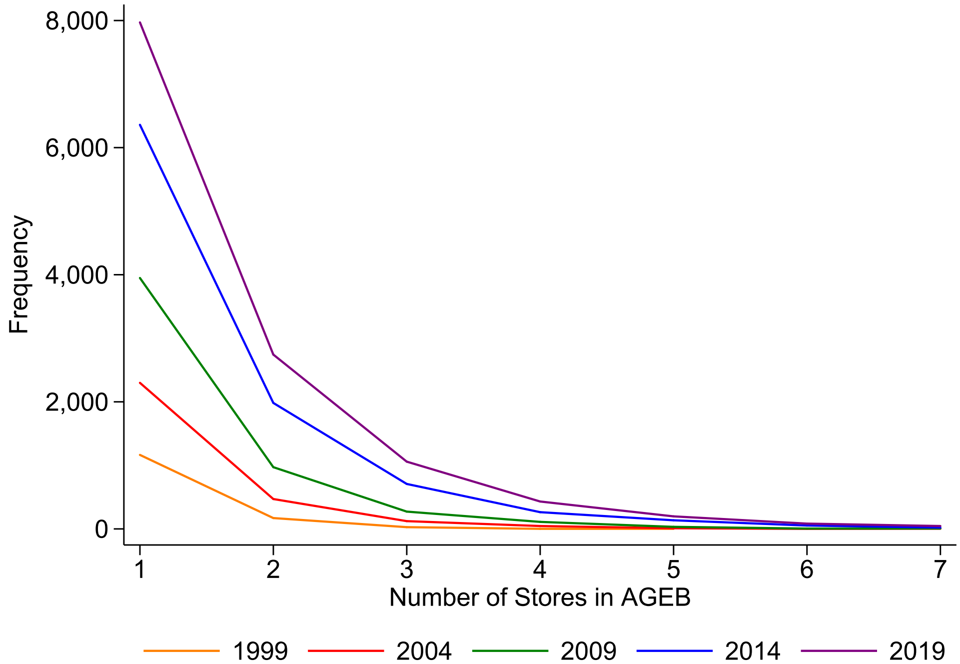


Figure A.2: Chain Stores

Note: The maps display the location of chain stores. A chain store is a store that belongs to a chain with more than 100 stores. Locations for 1999 are an approximate using the 1999 Economic Census Data. Locations for 2019 are obtained from DENUE 2020.

A) AGEBs Distribution by Number of Chain Stores



B) AGEBs Distribution by Number of Shops

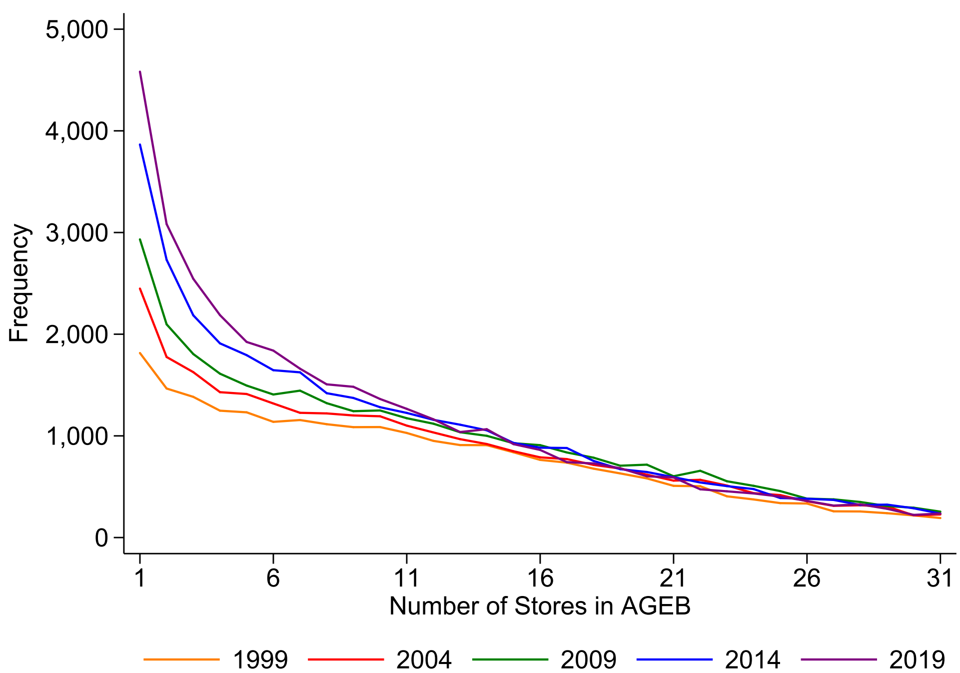


Figure A.3: Frequency Distribution by Number of Shops and Chain Stores

Note: The distributions of AGEBs by number of stores are computed using data from the 1999, 2004, 2009, and 2014 Economic Censuses. The AGEBs distribution by number of chain stores is conditional on the AGEB having at least one chain store.

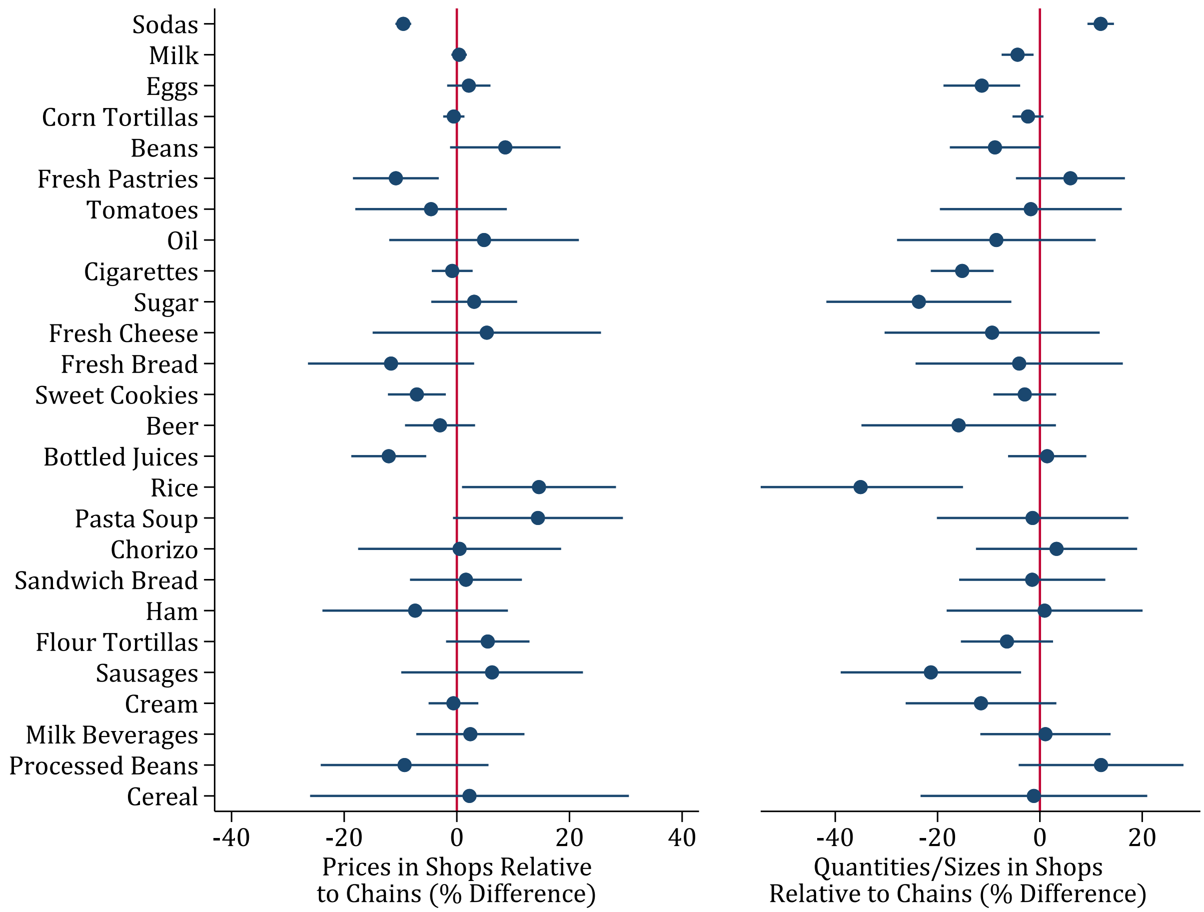


Figure A.4: Prices and Quantities/Sizes Differences Between Chains and Shops

Note: The figure displays the differences in prices and quantities/sizes between purchases in Chains and Shops. The standard errors are clustered at the city level and the estimation includes household fixed-effects. Prices are per unit, for example, sodas and other beverages is price per litter and beans, tomatoes, and rice is price per kilogram.

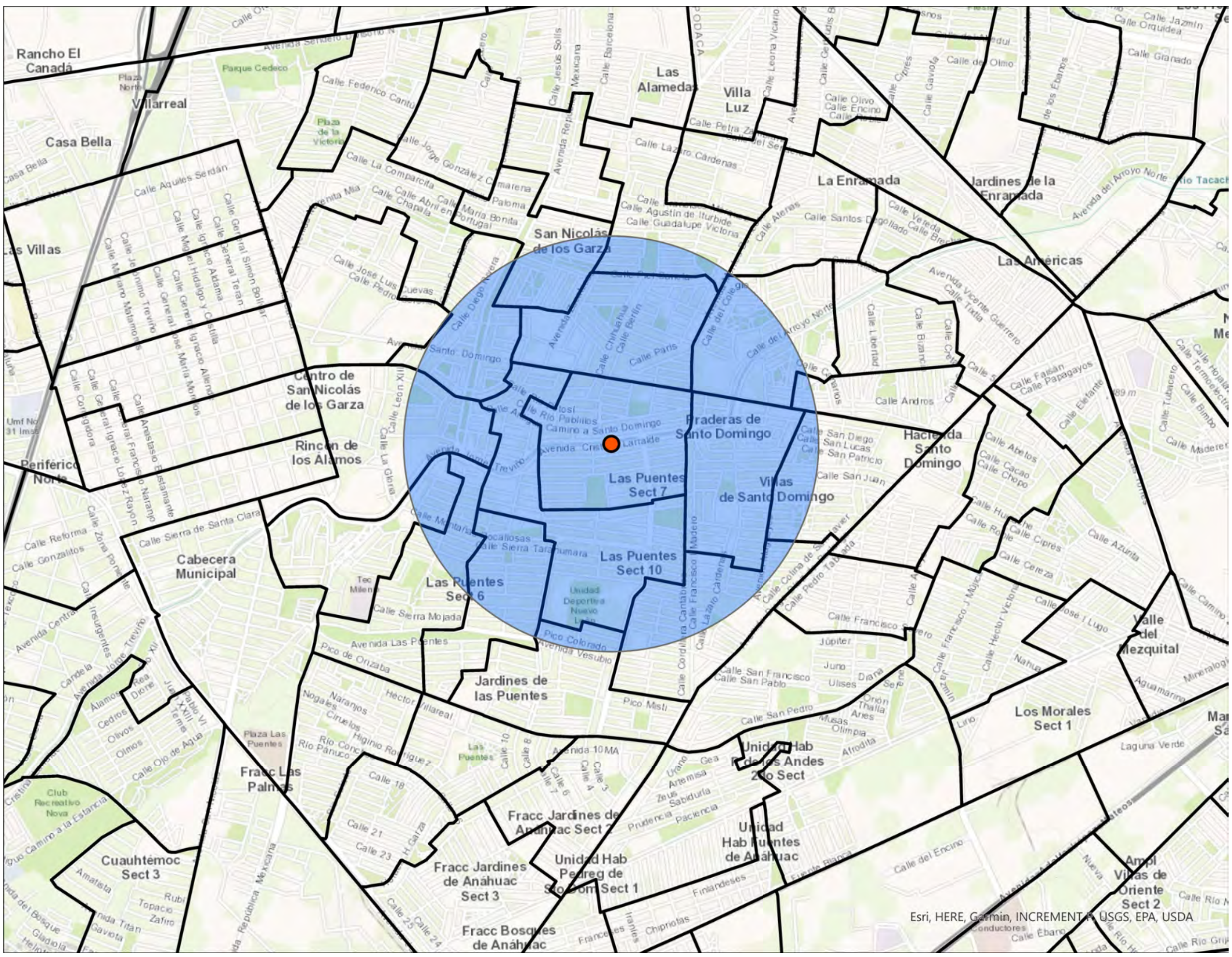


Figure A.5: Market Definition

Note: The map displays a 1km-radius circle with centered at the center of the AGEB. All the AGEBs that intersect with the circle define a neighborhood. The AGEBs shape and location is obtained from INEGI Marco Geostadístico.

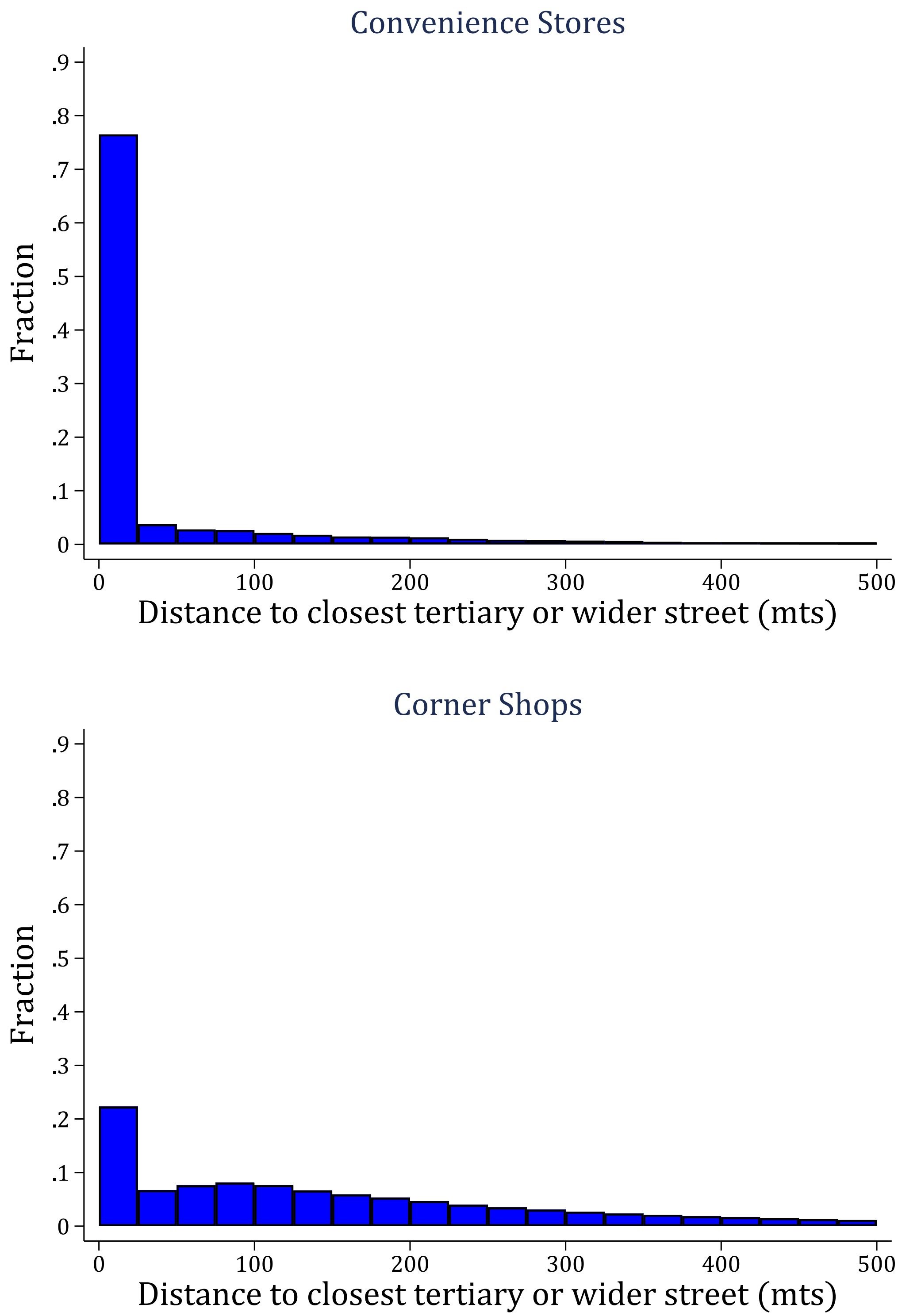


Figure A.6: Distance to Closest Wide Street

Note: The graphs display the distribution of distance from the store to the closest wide street. A wide street is defined as a street that is classified as trunk, primary, secondary, or tertiary by Open Street Maps. Streets location and type is obtained from Open Street Maps and stores locations are obtained from DENUE 2020.

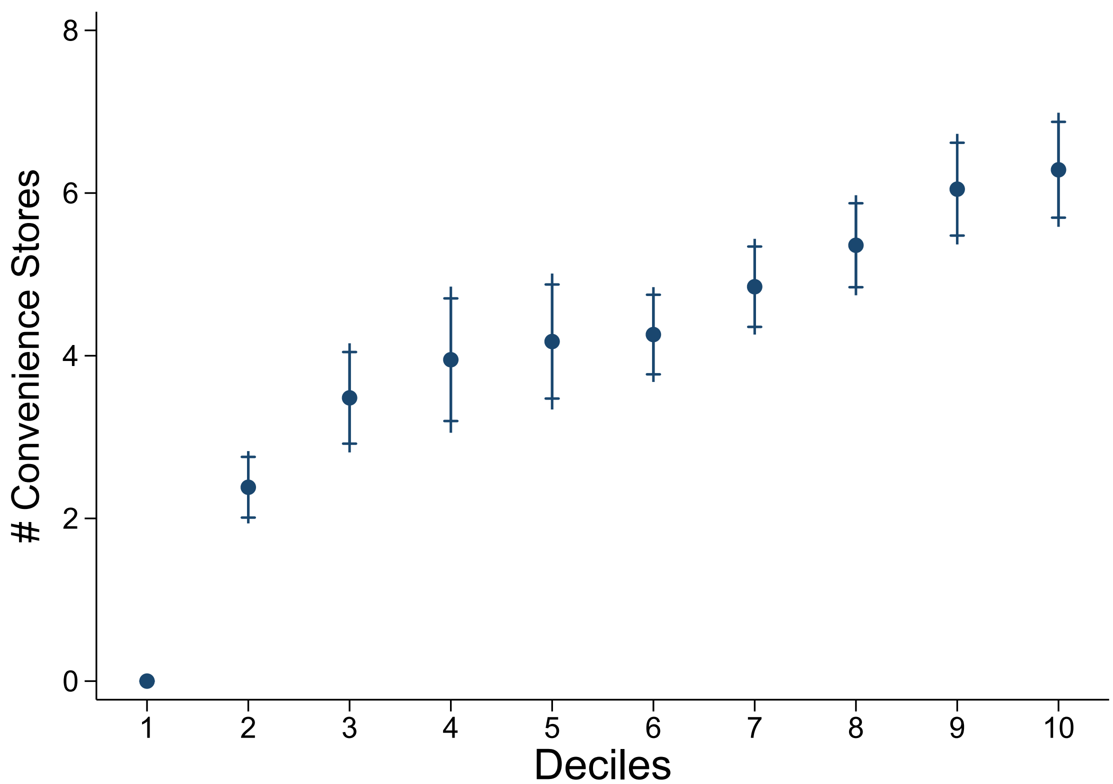


Figure A.7: Relationship between the number of chain stores and the instrument

Note: The figure displays the relationship between the instrument and the number of chain stores in the neighborhood. The figure displays estimates and 90 and 95% confidence intervals from a regression where the dependent variable is the number of chain stores in a neighborhood and the independent variables are dichotomous variables that take the value of 1 for each of the deciles 2 through 10 of the instrument. The estimation includes year-city and neighborhood fixed effects.

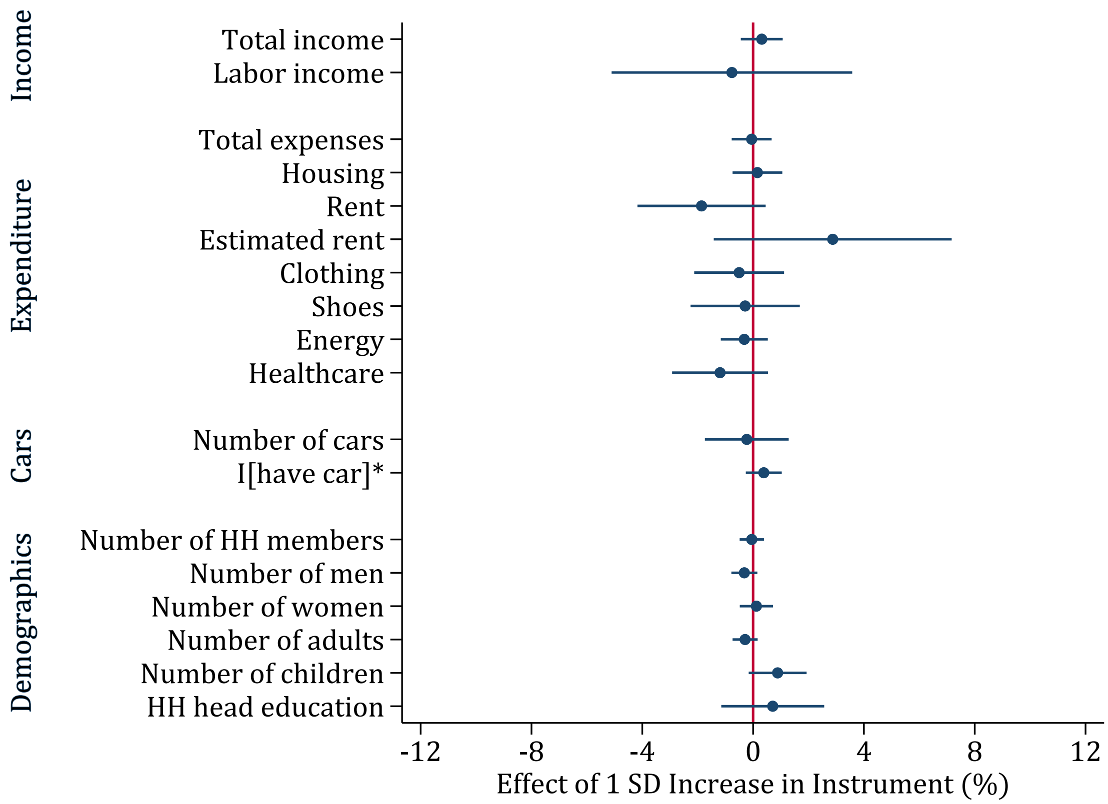


Figure A.8: Placebo - Relationship Between the Instrument and Household Characteristics

Note: The figure displays the estimates of regressing household characteristics on the instrument. Household characteristics vary are at the household level and the instrument varies at the neighborhood-year level.

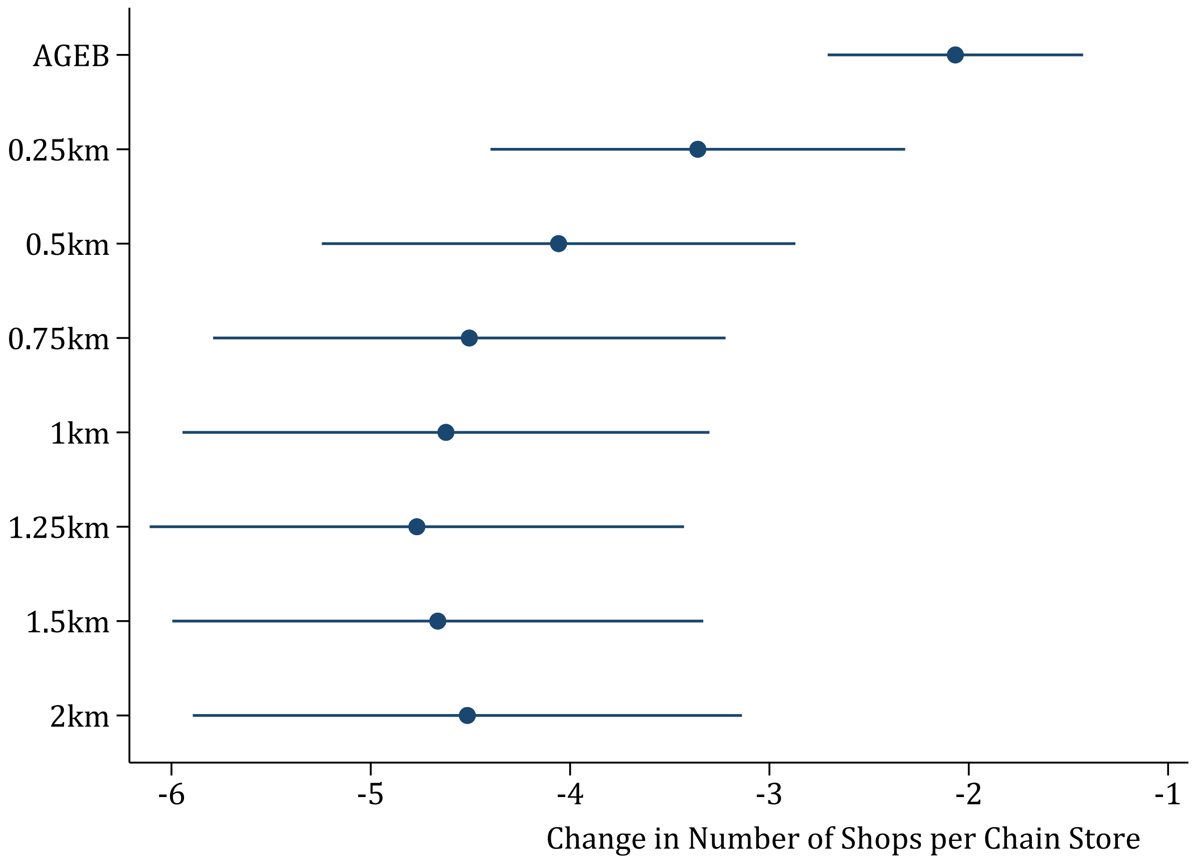


Figure A.9: Robustness - Alternative Neighborhood Definition

Note: The figure displays the estimation of Equation 1.3 using 2SLS with alternative neighborhood definitions. In row 1, the neighborhood is defined at the census tract level. In row 2, all the census tracts that are within 0.25km of a census tract center constitute a neighborhood. The neighborhood size increases all the way to row 8, where all the census tracts within 2km of the center of a census tract constitute a neighborhood.

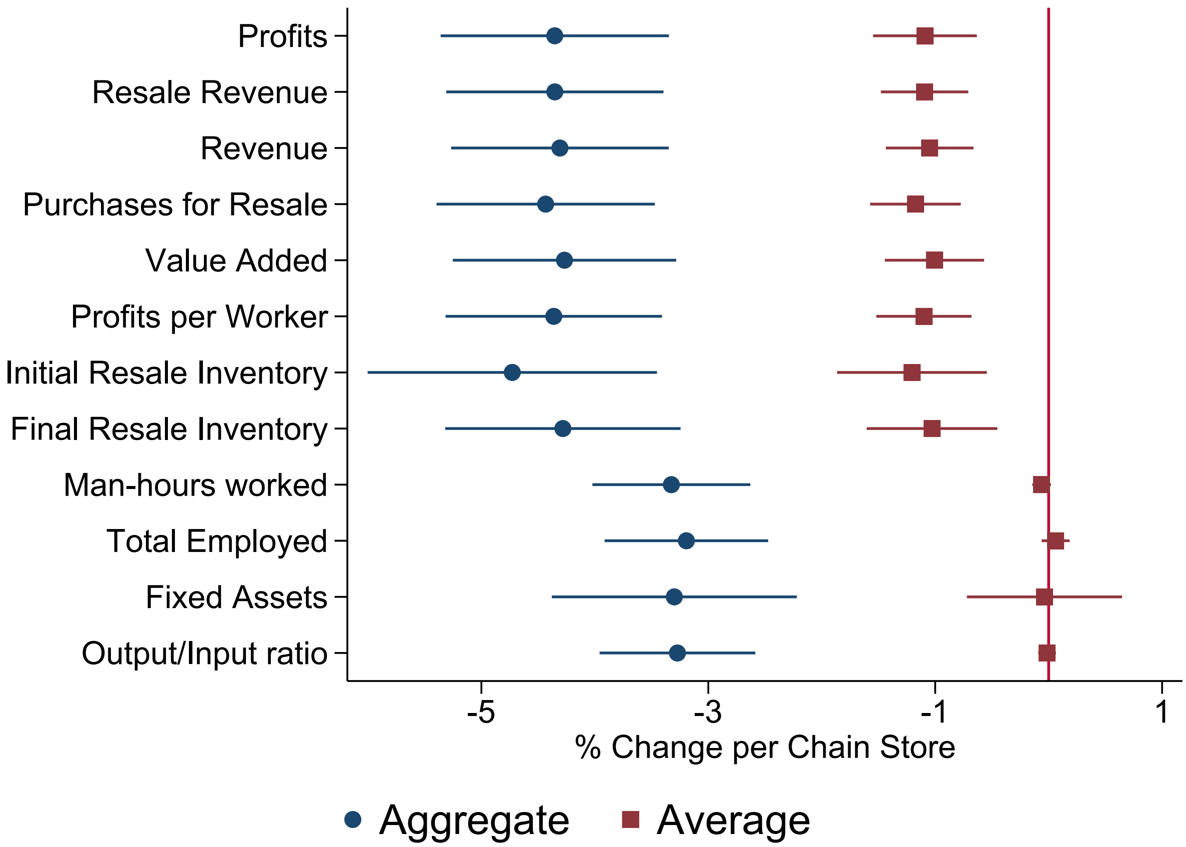


Figure A.10: Effects on Shops' Performance

Note: The figure displays the estimation of Equation 1.3 using 2SLS where the dependent variable is the inverse hyperbolic sine of the sum or average resale revenue, revenue, profits, value added, profits per worker, Initial resale inventory, final resale inventory, man-hours worked, total employed, fixed assets, and output/input ratio.

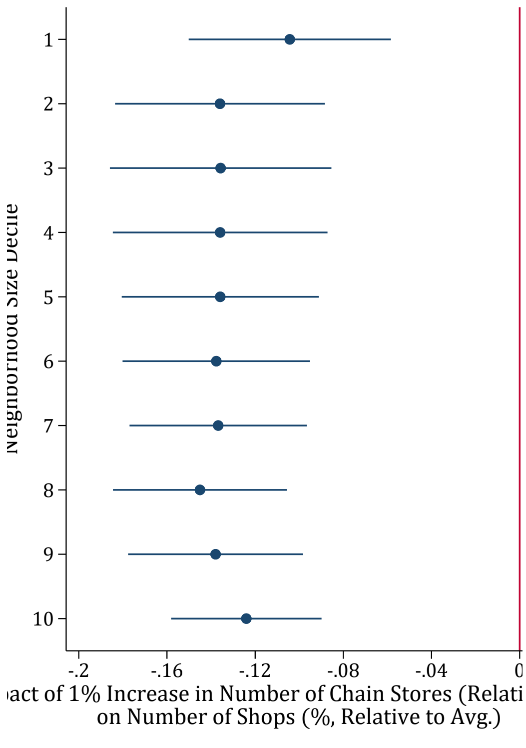


Figure A.11: Effects of Chains on Shops by Neighborhood Size

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS but interacting both the instrument and the number on convenience chain stores with dummies for each decile of neighborhood sizes. The first decile contains the smallest neighborhoods. Deciles are constructed within city.

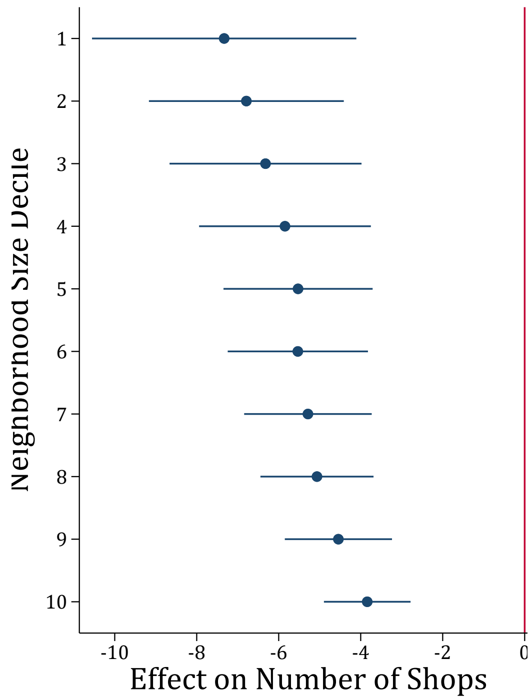


Figure A.12: Effects of Chains on Shops by Neighborhood Size

Note: The figure displays the estimation and 95% confidence intervals of Equation 1.3 using 2SLS but interacting both the instrument and the number on convenience chain stores with dummies for each decile of neighborhood sizes. The first decile contains the smallest neighborhoods. Deciles are constructed within city.

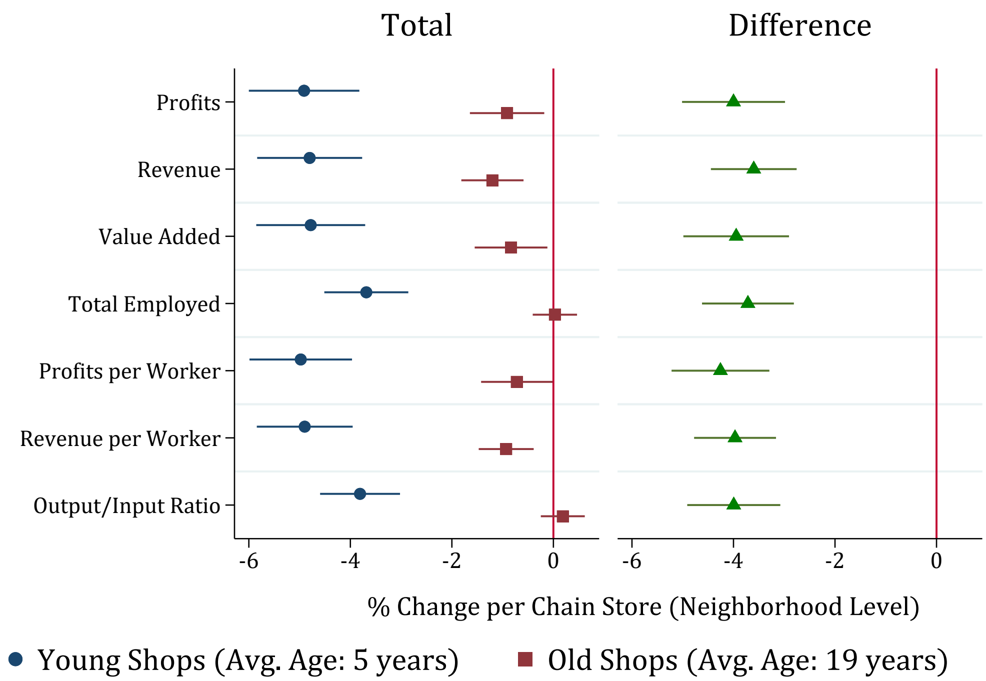


Figure A.13: Effects of Chains on Shops' Performance by Shop Age

Note: The figure displays the estimation of Equation 1.3 using 2SLS but adding i) the interaction of number of chain stores and a dummy variable for whether the average/sum is for shops in the fifth age quintile and ii) the interaction of the instrument and the same dummy to the first stage. The dependent variable in Equation 1.3 is the inverse hyperbolic sine of the row label.

Figure A.14: Effects of Chains on Shops' Performance by Whether they Pay Rent

Note: The figure displays the estimation of Equation 1.3 using 2SLS but adding i) the interaction of number of chain stores and a dummy variable for whether the average/sum is for shops that pay rent and ii) the interaction of the instrument and the same dummy to the first stage. The dependent variable in Equation 1.3 is the inverse hyperbolic sine of the row label.

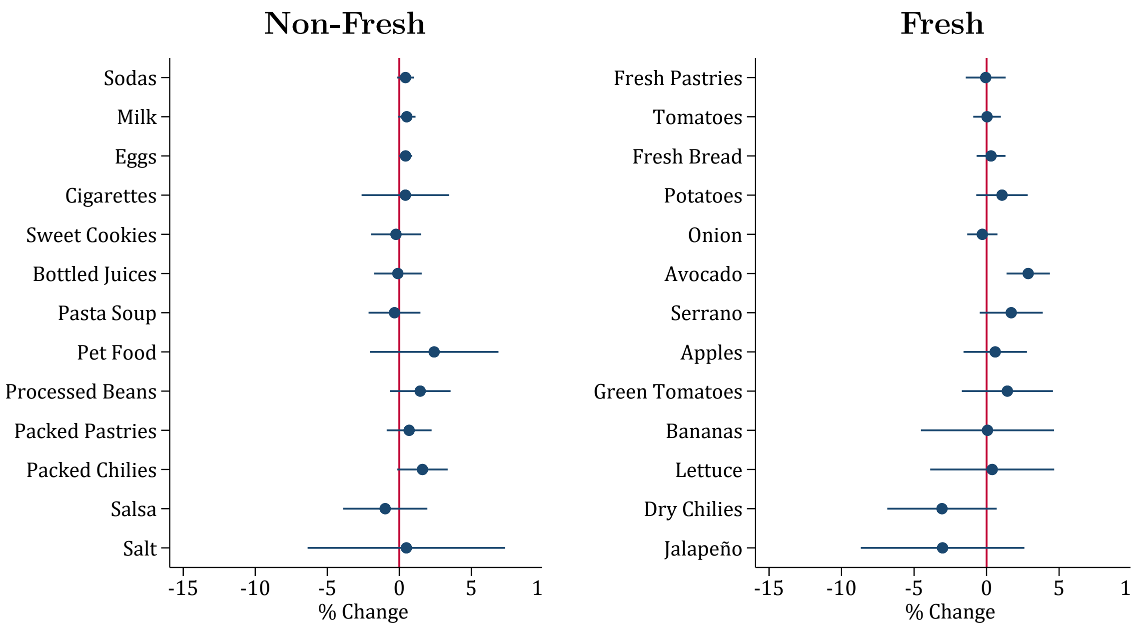


Figure A.15: Effect on Shops' Prices

Note: The figure displays the estimation of Equation 1.3 using 2SLS replacing the dependent variable with household-level price paid in pesos for each unit of the goods. The percentage change is computed by dividing the estimated effect by the household average product price in shops. The effects are for each additional chain store, and on average, there are 9 chain stores in each neighborhood. Goods are sorted from top to bottom by their share of shops' revenue. For non-fresh goods, sodas represent 13% of revenue and salt 0.2%. For fresh goods, sweet bread represent 3% of revenue and jalapeño 0.2%.

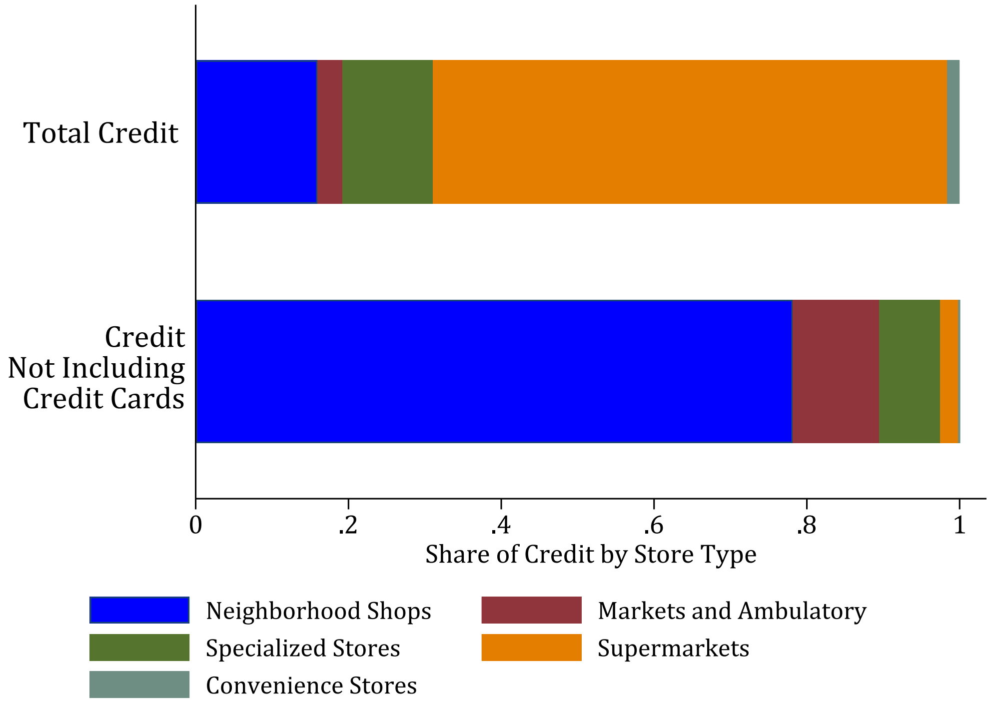
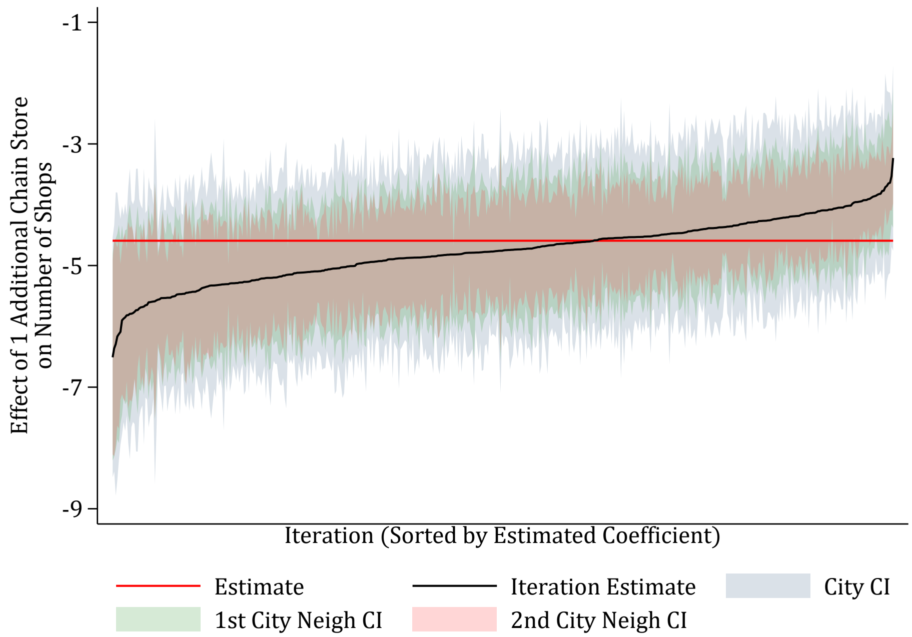
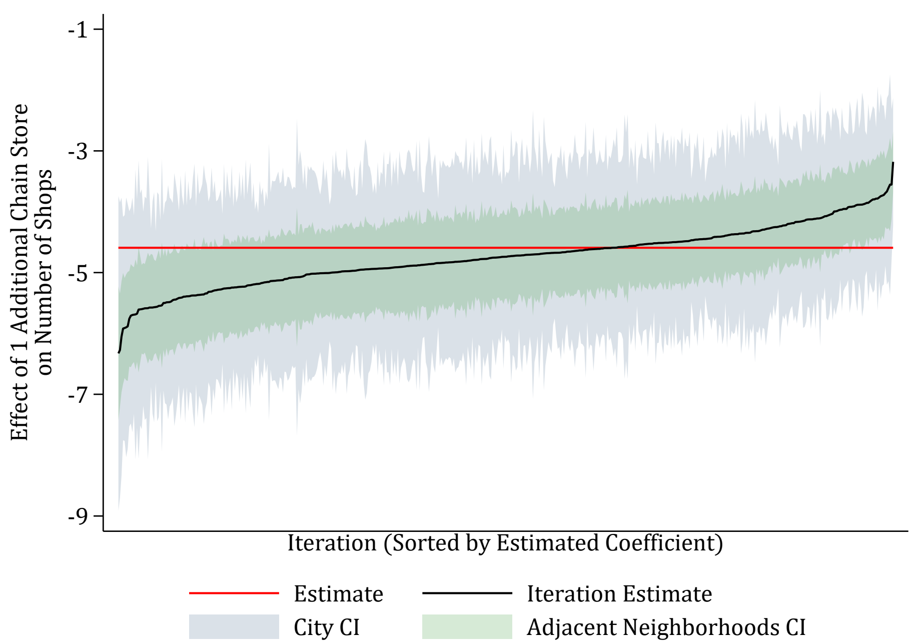


Figure A.16: Share of Credit Offered by Establishment Type

Note: The figure displays the share of credit offered by each establishment type for purchases of food and beverages computed using ENIGH 2018.



A) Spatial Correlation in Errors Across Adjacent Cities



B) Spatial Correlation in Errors Across Adjacent Neighborhoods

Figure A.17: Addressing Potential Spatial Correlation in Standard Errors

Note: The figure displays the estimation of Equation 1.3 using 2SLS. Each iteration contains a random sample of 5,000 markets, and in total there are 500 iterations in each figure. In Figure A), standard errors are clustered at the city level, corrected for the potential correlation of unobserved shocks across adjacent cities, and corrected for the potential correlation of unobserved shocks across 2^{nd}-degree adjacent cities. In Figure B), standard errors are clustered at the city level and corrected for the potential correlation of unobserved shocks across adjacent neighborhoods. I use the technique proposed by Colella et al. (2019) to account for potential spatial correlation of unobserved shocks and its companion statistical package *acreg*.

Figure A.18: Hybrid Stores

Source: Google Maps
Note: The figure contains an example of a hybrid store. Hybrid stores share the same establishment type code as Chains, but different from Chains, the owners only have one store.

A.2 Appendix: Zeroth Stage

Chains exploit economies of scale arising from stores in nearby cities sharing distribution, monitoring, marketing, and overhead costs. As a result, chains will open stores in cities close to each other. Figure 1.3 shows this within-firm spatial correlation in store openings between 2016 and 2020.

To quantify the importance of economies of scale I estimates the relationship between the number of chain stores that chain f has in cities adjacent to city c at time t and the number of chain stores that f has in city c. The coefficient of interest, β, tests for spatial correlation in the number of chain stores after controlling for firm-time, city-time, and firm-city fixed effects. I interpret this spatial correlation as economies of scale.[1]

$$\#Stores_{f,c,t} = \eta_{f,t} + \mu_{c,t} + \zeta_{f,c} + \beta \#StoresNearbyTowns_{f,c,t} + \epsilon_{f,c,t} \tag{A.1}$$

The results presented in Table B.1, estimated using equation A.1, show that across all specifications, there is strong evidence of economies of scale: number of same-chain stores in towns nearby to town c are positively correlated with number of same-chain stores in town c. Columns 1-4 use 2^{nd} degree neighbors (adjacent towns and towns adjacent to these), columns 5-6 use 1^{st} degree neighbors (adjacent towns), and columns 7-8 use 3^{rd} degree neighbors (adjacent towns and towns adjacent to these and towns adjacent to these). Column 4 is the preferred specification because it uses 2^{nd} degree adjacent cities (same as the IV), and includes all the fixed effects combinations. Economies of scale matter: 19 additional same-chain stores in nearby cities translate to one more store in the city – accounting for 9% of the variation in the number of stores each chain has in a city.[2]

[1] If there was no cost sharing leading to economies of scale, there would be no reason for chains to open more stores in nearby cities rather than in other cities.

[2] The 9% is obtained by computing the within R-squared. It is the R-squared after demeaning each variable with respect to the fixed effects.

Table B.1: Same-Chain Economies of Scale

	Dependent Variable: # of Stores in City							
Nearby Cities:	2nd Degree Adjacent Cities				Adjacent Cities		3rd Degree Adjacent Cities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Stores Nearby	0.045***	0.047***	0.047***	0.046***	0.122***	0.115***	0.023***	0.022***
Cities (same chain)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Sample Size	416,704				416,398		416,398	
Clustered SE	City	City	City	City	City	City	City	City
Year, City, & Firm FE	Y							
Firm x City FE		Y	Y	Y	Y	Y	Y	Y
Year x Mun FE			Y	Y	Y	Y	Y	Y
Year x Firm FE				Y		Y		Y
R-squared	0.159	0.717	0.730	0.730	0.738	0.739	0.718	0.718
Within R-squared	0.105	0.150	0.152	0.111	0.176	0.140	0.115	0.073

Note: The table displays the estimation of Equation A.1. For columns 1-4, Nearby Towns are the adjacent towns, for columns 5-6 Nearby Towns are the adjacent towns and those adjacent to these, and for columns 7-8 Nearby Towns are the adjacent towns, those adjacent to these, and those adjacent to the adjacent towns.

Table B.2: Cross-Chain Economies of Scale

	Dependent Variable: # of Stores in City						
Nearby Cities:	2nd Degree			Adjacent Cities		3rd Degree	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of Stores Nearby Cities	0.000***	-0.001***	-0.001***	-0.001***	-0.003***	0.000***	-0.001***
(different chain)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample Size	13,751,232			13,751,232		13,751,232	
Clustered SE	City	City	City	City	City	City	City
Year, City, & Firmj FE	Y	Y		Y		Y	
Firm k FE		Y		Y		Y	
Firm j x City & Firm k x City FE			Y		Y		Y
Firm j x Year & Firm k x Year FE			Y		Y		Y
Year x Mun FE			Y		Y		Y
R-squared	0.060	0.060	0.696	0.060	0.696	0.060	0.696
Within R-squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The table displays the estimation of Equation A.2. For columns 1-3, Nearby Towns are the adjacent towns, for columns 4-5 Nearby Towns are the adjacent towns and those adjacent to these, and for columns 6-7 Nearby Towns are the adjacent towns, those adjacent to these, and those adjacent to the adjacent towns.

The previous analysis tests the existence and importance of economies of scale in determining the time and city of opening of same-chain stores. The following analysis tests whether these economies of scale are indeed firm specific. If all chains enter the same cities at the same time, this would be likely driven by something other than economies of scale. The following equation tests for cross-firm economies of scale, which should not exist if economies of scale are indeed firm-specific and driven by cost sharing within firm. The coefficient of interest, β, estimates the relationship between the number of stores chain g has in cities nearby to city c at time t and the number of stores that chain f (a competitor) has in city c at time t after controlling for firm(f)-time, firm(g)-time, city-time, firm(f)-city, and firm(g)-city fixed effects.

$$\#Stores_{f,c,t} = \eta_{f,t} + \mu_{c,t} + \zeta_{f,c} + \gamma_{g,t} + \delta_{g,c} + \beta \#StoresNearbyTowns_{g,c,t} + \epsilon_{f,c,t} \tag{A.2}$$

Economies of scale are firm-specific: the positive correlation in Table B.1 dissipates when using the number of different-chain stores (competitors) in nearby cities, and the number of competitors in nearby cities account for less than 0.0001% of the variation in the number of stores each chain has in a city. The results are in Table B.2. Across all specifications, there is no evidence of cross-firm economies of scale. Moreover, there is a small pro-competitive effect: a negative relationship between the number of stores a competitor g has in towns adjacent to town c and the number of stores chain f has in town c.

A.3 Two-Dimensional Amenities Differentiation

The goal of the second organizing framework is to incorporate heterogeneity in shops and consumers to illustrate that stores that are located closer to chains in the amenities space (less differentiated) will suffer more from the entry of chains. The framework also shows that as a result of the entry of chains, shops specialize in their comparative advantages. This is a result of chains stealing the customers for which the comparative advantages of shops are relatively less important.

The framework is a very simplified two-dimensional Hotelling. There are two significant simplifications: i) firms offer the same price, and ii) firms location is determined by its type. The vertical dimension, *convenience*, includes a bundle of parking, acceptance of diverse payment methods, location close to a wide street, and uniformity. Moving up in the space translates to more of the amenities in the bundle. The horizontal dimension, *shop experience*, includes a bundle of store credit, product freshness, relationships with the owner, and gathering place. Moving to the right in the space translates to more of the amenities in the bundle.

Customers are located in the space based on their ideal combination of amenities. Panel A) of Figure A.19 displays three examples of customers. The customer in the lower left of the space is a customer whose optimal combination of amenities is low in both dimensions: she dislikes walking all the way to a wide street (low vertical dimension) and dislikes personal interactions (low horizontal dimension). Panel B) displays firms location. There are three types of incumbents: hybrid stores, young shops, and old shops. Hybrid stores differ from shops because they are bigger, hire employees, target customers from outside the neighborhood, and sometimes offer parking. These combination of amenities locate them northwest of shops and southeast of chains. Building relationships, knowing which customers are credit worthy, and sourcing fresh products are all activities that take time to perfect. Older shops

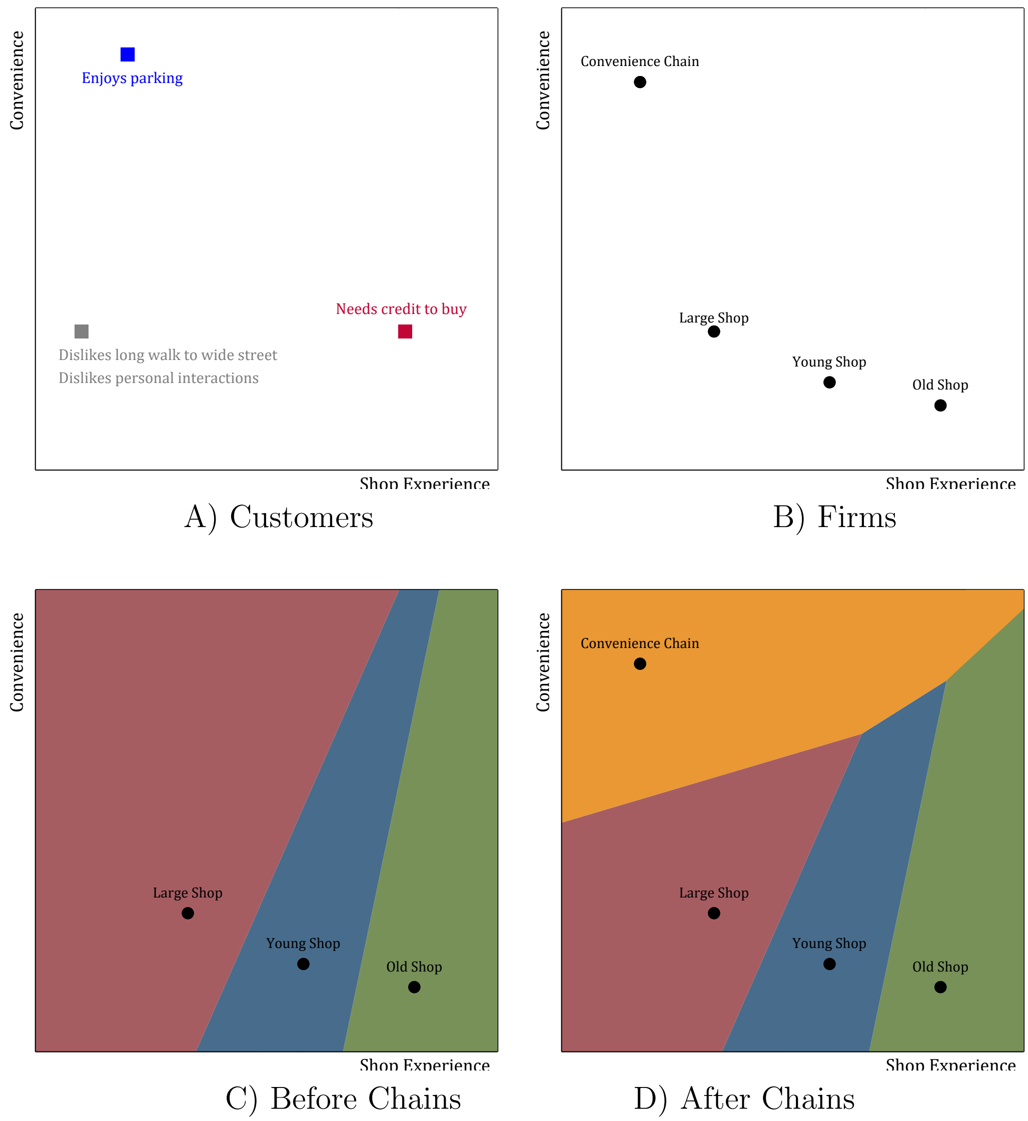


Figure A.19: Two-Dimensional Amenities Differentiation

Note: Figure A) displays customers located in the two-dimensional amenities space based on ideal combination of amenities. Figure B) displays the location of firms based on the value proposition that they offer. Figure C) displays the split of customers before the entry of chains. Figure D) displays the split of customers after the entry of chains.

are better able to differentiate in these dimensions than younger ones, because they had more time to perfect this activities. This locates older shops to the right of younger ones.

Panel C) displays the market split between hybrid stores, young shops, and old shops before the entry of chains. Panel D) displays the market split after the entry of chains. The first takeaway is that hybrid stores lose more customers than regular shops because they are located closer in the amenities space to chains. The second takeaway is that young shops lose more customers than old ones because they are less differentiated. This argument is similar to the one of McKenzie and Woodruff (2015), where better managed small firms will face less competition because they will market differently or differentiate their products.[3] The last takeaway is in terms of customer composition: the customers that leave shops are those for who the comparative advantages of shops are relatively less important. Keeping fixed the preference for amenities in the vertical dimension, shops lose the customers that want Coca-Colas, but retain those who want tomatoes.

[3]McKenzie and Woodruff (2015) introduce this idea to explain how endogeneity can result in competition being negatively correlated to management practices for small firms (in strong contrast to the results for large firms).

Appendix B

Appendix Tables and Figures

Grandmothers and the Gender Gap in the Mexican Labor Market

B.1 Online Appendix

Table A.1: Mothers in 2 vs 3-Generation Households

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Dependent Variable	Large City	Rural	Employed	HH Income	HH Income (pc)	Age	# of Children	Years of Schooling	Highschool +	College+	Earned Income	Earned Income	Hours Worked	Hours Worked
3-Generation Household	0.0147**	-0.0289***	0.117***	1729.9***	-483.2***	-2.201***	-0.261***	0.266***	0.0416***	0.0159***	238.7***	-646.7***	5.513***	3.609***
	(0.00646)	(0.00350)	(0.00334)	(101.7)	(20.87)	(0.0425)	(0.00531)	(0.0490)	(0.00600)	(0.00450)	(23.99)	(48.08)	(0.161)	(0.204)
Mean Dependent Variable	0.57	0.17	0.35	7,550	1,787	28.7	1.8	10.7	0.51	0.22	1,956	5,441	14.9	40.0
Sample Size	267,593	267,593	260,416	260,416	260,416	260,416	260,416	260,162	260,416	260,416	231,267	63,211	231,267	63,211
Sample	All	All	All	All	All	All	All	All	All	All	All	Employed	All	Employed
Quarter FE	Y	Y												
Quarter x Locality FE			Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age FE							Y	Y	Y	Y	Y	Y	Y	Y

Note: The table displays the differences between mothers that live in two-generation households and those that live in three-generation households. Mothers of ages between 20 and 50 living in households were the oldest member of the youngest generation is less than 10-years-old are included. The employed sample includes only mothers that are employed and have strictly positive income and hours worked. Standard errors clustered at the locality level.

Table A.2: 3-Generation Households Alternative Samples

	(1)	(2)	(3)	(4)
Sample	Base	Unbalanced	Any # of grandparents	Any # of parents
Observations	484,454	561,119	488,286	620,172
Employed	0.44	0.44	0.44	0.44
Any work	0.56	0.56	0.56	0.56
Any paid work	0.53	0.53	0.53	0.53
Age	32.0	31.9	32.0	31.6
Number of Kids	1.91	1.91	1.91	1.88
Hours Employed	21.75	21.79	21.75	21.76
Income x Hour	11.62	11.74	11.62	11.20
Income	1,822	1,843	1,823	1,770
Formal Employment	0.31	0.31	0.31	0.30
Household Size	5.76	5.73	5.76	6.48

Note: The table displays descriptive statistics of alternative samples of 3-generation households. Column 1 is the main specification. Column 2, Unbalanced, lifts the restriction of observing the individual for five surveys. Column 3, Any number of grandparents, allows for any number of members of the first generation of the household. Column 4, any number of parents, allows for any number of members of the second generation of the household.

Table A.3: Heterogeneity by School Affordability

	Residual				Observed			
	Hourly Cost		Total Cost		Hourly Cost		Total Cost	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post grandmother death	-0.0171	-0.0211	-0.0173	-0.0210	-0.0185	-0.0209	-0.0186	-0.0207
	(0.0140)	(0.0148)	(0.0140)	(0.0148)	(0.0140)	(0.0148)	(0.0140)	(0.0148)
Post x Grandmother Died x Oldest Grandchild at most 5 years old	-0.127***	-0.121***	-0.124***	-0.113***	-0.125***	-0.118***	-0.123***	-0.111***
	(0.0340)	(0.0373)	(0.0333)	(0.0367)	(0.0338)	(0.0371)	(0.0333)	(0.0370)
Post x Grandmother Died x Cost of School	-0.0202**		-0.0227**		-0.0136		-0.0153	
	(0.00951)		(0.0102)		(0.00971)		(0.0102)	
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Cost of School	-0.0355		-0.0273		-0.0258		-0.0283	
	(0.0528)		(0.0431)		(0.0515)		(0.0425)	
Post x Grandmother Died x Cost of Private School		0.00105		0.00333		0.00444		0.00764
		(0.0115)		(0.0121)		(0.0116)		(0.0122)
Post x Grandmother Died x Oldest Grandchild at most 5 years old x Cost of Private School		-0.0892***		-0.0914***		-0.0838**		-0.0863***
		(0.0337)		(0.0304)		(0.0363)		(0.0319)
N	421,099	353,526	421,144	353,526	423,961	354,556	424,006	354,556
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Locality FE	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Young Child FE	Y	Y	Y	Y	Y	Y	Y	Y
Year - Quarter - Grandmother Died FE	Y	Y	Y	Y	Y	Y	Y	Y
# of localities to estimate residuals	1758	615	1761	615	-	-	-	-

Note: The table displays heterogeneity of the marginal effect of the grandmother's death on mother's employment by school affordability, estimated using Equation 2.7. School costs are standardized. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Standard errors are clustered at the household level.

Table A.4: Grandmother's Death and Older Grandchildren's Time Spent Providing Care

Dependent Variable:	Time Spent Providing Care			𝟙[Spent Time Providing Care]		
	(1)	(2)	(3)	(4)	(5)	(6)
	12-15 Years Old	12-18 Years Old	12-21 Years Old	12-15 Years Old	12-18 Years Old	12-21 Years Old
Post x Grandmother Died	0.756**	0.718**	0.663**	0.0713**	0.0680**	0.0614**
	(0.348)	(0.328)	(0.326)	(0.0328)	(0.0308)	(0.031)
N	46,639	62,366	67,896	46,639	62,366	67,896
Individual FE	Y	Y	Y	Y	Y	Y
Locality FE	Y	Y	Y	Y	Y	Y
Year - Quarter FE	Y	Y	Y	Y	Y	Y

Note: The table displays the effect of the grandmother's death on the inverse hyperbolic sine of the time older grandchildren spend providing care. The sample is grandchildren 12-15, 12-18, or 12-21 years old in households where the youngest grandchild is up to five years old. Time allocation providing is the response to the following question from ENOE Q1 2005 to Q1 2020: During last week, how much time did you spend exclusively taking care without pay of children, elderly, sick, or handicapped? Columns 1-3 further restrict the sample to households where the grandmother was up to 70 years old. The number of stars indicates the significance level at which the coefficient is statistically significant: .01, .05, and .1 for three, two, and one stars, respectively. Standard errors are clustered at the household level.

B.1.1 Online Appendix Figures

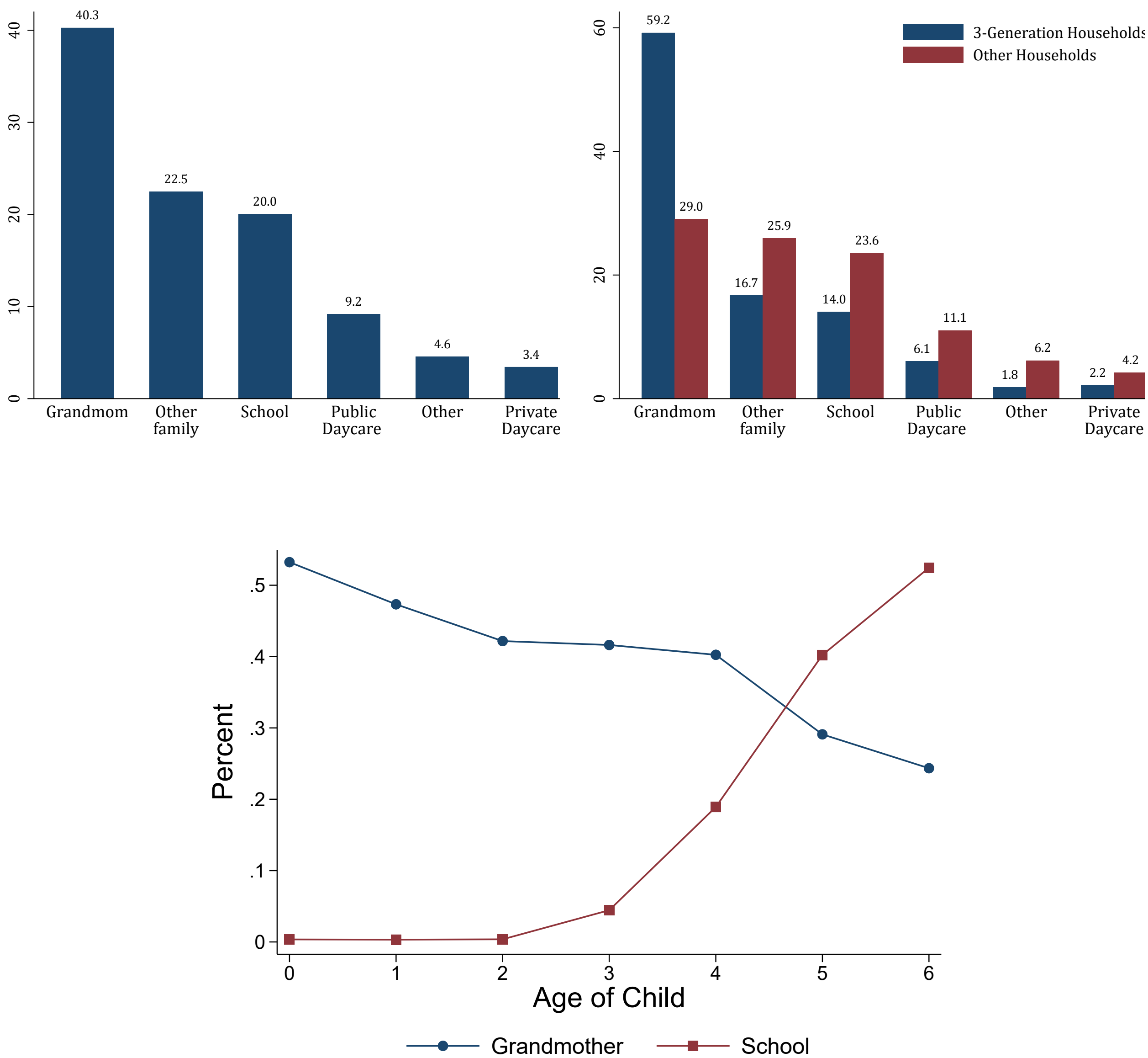


Figure A.1: When the mother goes to work, who takes care of the child?

Source: ENESS 2009, 2013
Note: The surveys include responses for children between age 0 and 6 years. Other includes non-family members and leaving the child alone. Children that go with their mothers to work or whose mothers do not work are not included. ENESS 2017 is not included because the grandmother option was replaced by grandparents.

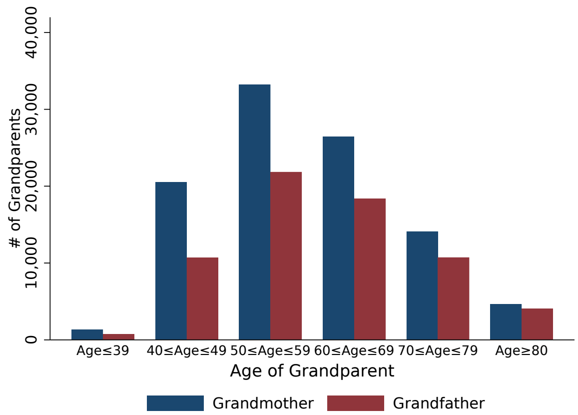


Figure A.2: Age of Grandparents
Source: ENOE (Q1 2005 - Q1 2020)
Note: The sample includes three-generation households.

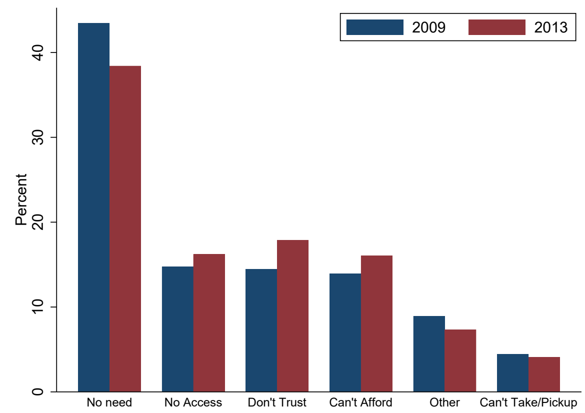


Figure A.4: Why are you not using daycare?
Source: ENESS (2009, 2013)

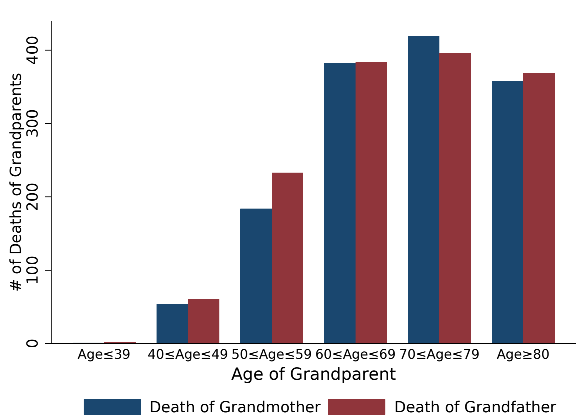


Figure A.3: Deaths of Grandparents
Source: ENOE (Q1 2005 - Q1 2020)
The sample includes three-generation households.

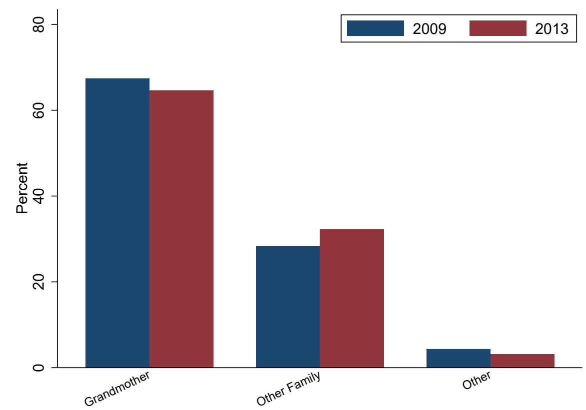
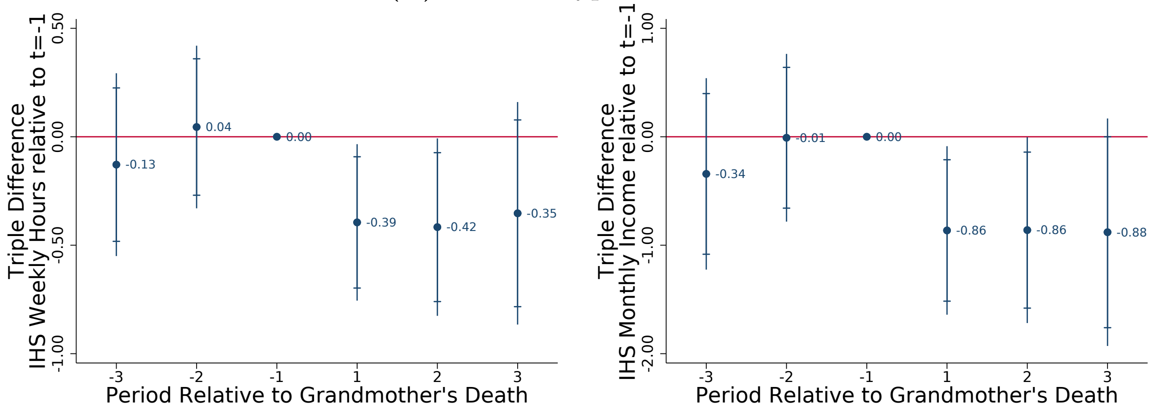
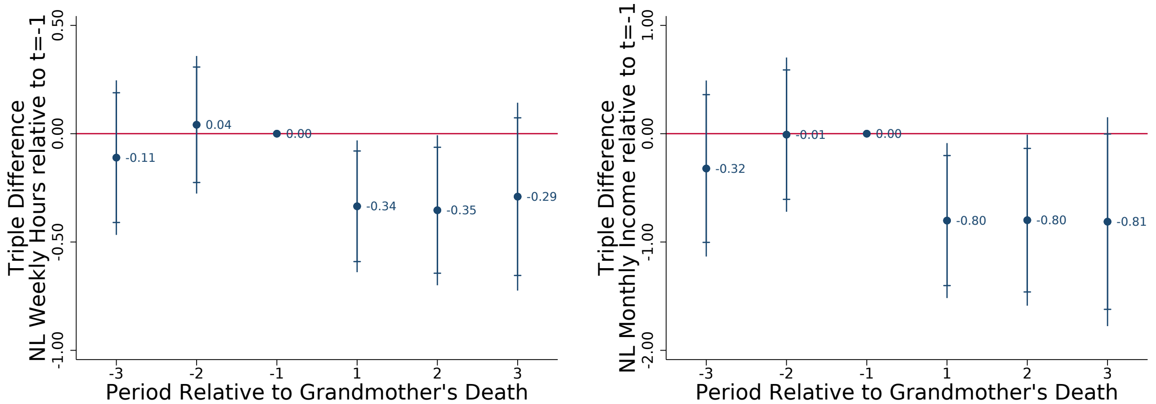


Figure A.5: If there is no need for daycare, who takes care of the child?
Source: ENESS 2009, ENESS 2013

(A) Inverse Hyperbolic Sine



(B) Natural Logarithm Transformation



(C) No Transformation

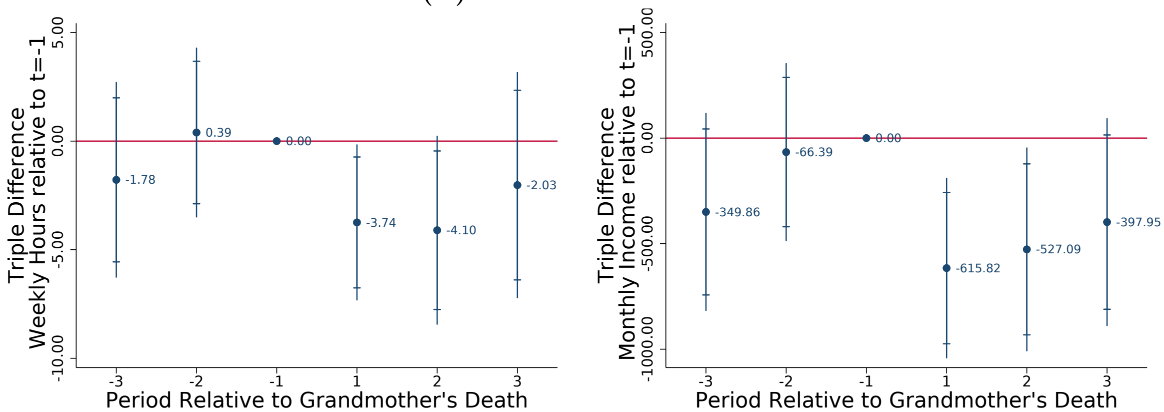


Figure A.6: Earned Income and Hours Worked - Mothers Sample

Note: The graphs display the point estimate and the 90% and 95% confidence interval of the effect that the death of the grandmother has on earned income and hours worked of mothers. The estimation is based of Equation 2.3, but replaces the dependent variable with earned income or hours worked. Income is winsorized at a 5% level from each tail, excluding 0's. Hours worked is winsorized at a 5% level from the right tail. Only observation where both earned income and hours worked are positive or both are zero are included. Panel A) presents results for the inverse hyperbolic sine transformation, Panel B) the natural logarithm, and Panel C) no transformation. The confidence intervals are computed using standard errors clustered at the household-level.