EL COLEGIO DE MÉXICO
CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN ECONOMÍA

PRICE DISPERSION AND COMPETITION: EVIDENCE FROM THE MEXICAN SUPERMARKET INDUSTRY

CRISTIAN EDUARDO GUDIÑO GARCIA

PROMOCIÓN 2021 – 2023

ASESOR:

DR. RAYMUNDO MIGUEL CAMPOS VÁZQUEZ

AGOSTO 2023

Agradecimientos

En primer lugar, quiero agradecer a quienes ya no están, pero son el punto de partida más cercano de mi fuente de vida: mis abuelos. Quiero agradecerles por educar a mis padres con todo su amor y ser una de mis fuentes de inspiración. Además, quiero hacer un agradecimiento especial a mi abuelo Santos por enseñarme el amor a la vida, el béisbol y quien sé que siempre me acompaña en mis logros y caídas.

Quiero agaradecer a mi fuente de amor, apoyo y vida: mis padres. Les agradezco eternamente por dejarme perseguir mis sueños, apoyarme en cada una de mis decisiones y cuestionarme hacia dónde quiero ir. A mi madre, Mónica, quien siempre me ha dado todo su amor y nunca me ha dejado solo. A mi padre, Leonel, quien me enseña a no rendirme ni traicionarme, ser leal y siempre hacer lo correcto.

A mis hermanas, les agradezco su incondicionalidad porqué incluso cuando la vida nos separe a la distancia, sé que estarán más cerca de lo que creo. A hermana Fernanda quien me enseña que la vida deber ser sencilla y siempre divertirme. A mi hermana Diana y mi sobrino quienes me enseñan a superar la adversidad y ser mejor.

A mis amigos, Josselyn, Archundia y Lino, con quienes compartí la gran mayoría del tiempo. Quienes me enseñaron que vale más la honestidad y honradez, quienes se esforzaron al máximo durante la Maestría, y quienes nunca me abandonaron. A mis amigos Fernando, Charleston y Daniel, quienes siempre han estado para mí, sin importar la distancia que nos separe, *El Templo del Saber* siempre nos mantendrá unidos. A todos mis amigos de la generación 2021 - 2023 con quienes compartí el salón de clases, les agradezco por las enseñanzas, las risas, los enojos, la desesperación, y la felicidad infinita: *La Familia*, *Los Chiefs*.

Al Dr. Raymundo Campos por su pericia, paciencia, comprensión y confianza indondicional, gracias por guiarme y dirigirme, y enseñarme lo más importante:*siempre atarme al mástil y no escuchar las sirenas.* Su exigencia me ha dotado de disciplina en la vida la cual ha sido indispensable. Sin su apoyo este trabajo no hubiera sido posible.

A la Dra. Diana Terrazas por su apoyo incondicional, por las enseñanzas y luz que en cada plática he podido tener. Su motivación ha sido esencial en cada paso que doy. *Que la fuerza la acompañe siempre.* Además, quiero agradecer al Dr. Carlos Reyes quien ha confiado en mí desde el primer momento y ha sumado motivación a esforzarme y ser mejor todos los días.

Al Dr. Julen Berasaluce por confiar en mí y dar la oportunidad de aprender, y enseñar con pasión. Al Dr. Edwin vanGameren por cada plática. A Chelita, Karine, Elena y Olivia quienes siempre me apoyaron y auxiliaron.

Finalmente, quiero agradecer a El Colegio de México y El Centro de Estudios Económicos por haberme dado la oportunidad, por segunda ocasión, de ser egresado en Economía por esta institución que llamo *El Templo del Saber*.

"Last, but not least, I want to thank me,
I want to thank me for believing in me,
I want to thank me for doing all this hard work,
I want to thank me for having no days off,
I want to thank me for never quitting,
I want to thank me for always being a giver
and trying to give more than I receive,
I want to thank me for trying to do more right than wrong,
I want to thank me for just being me at all times"

Snoop Dogg

Abstract

In this work, I combine two different approaches to analyze the price dispersion in the Mexican Supermarket Industry, and to estimate the uniform pricing that supermarket chains charge across their stores as well as the degree of price dispersion between urban areas – poor and rich. These methodologies are based on models proposed by DellaVigna and Gentzkow (2019) and Eizenberg et al. (2021), which are able to discriminate quite well between urban areas with monopolistic competition and areas with more than three competitors, and are applied to the supermarket industry in Mexico.

Furthermore, to provide a theoretical framework, this work contains a brief literature review of the theory of monopolistic competition models, search models, and the Hotelling model; and the applied literature on other markets. The theoretical models suggest that consumers in poor neighborhoods have different transport costs from those in rich areas, as well as different search costs – poor areas have limited access to price information. This cost asymmetry allows supermarket chains to discriminate between regions, charging higher prices in those areas with higher costs in price information, and transport.

The empirical model is estimated for the Mexican Supermarket Industry from January 2017 to December 2022, considering 23 products of regular consumption, and using as primary data sources the Who is Who on the Prices (QQP) and the National Deprivation Index (NDI), the empirical evidence suggests the following: although retail prices differ across time, there is no variation across stores within a supermarket chain; price difference within chains is 0.003 log points, but difference between is 0.001 log points, which is explain by local competition dynamics; there is a negative relationship between population and retail prices, and a positive relationship between deprivation index and prices.

Cristian E. Gudiño García

Contents

Price Dispersion and Competition: Evidence from the Mexican Supermarket Industry

Cristian E. Gudiño Garcia

El Colegio de México

Centro de Estudios Económicos

1 Introduction

It is considered that the adjustments of local retail prices to economic conditions are central to economic policy questions because differences in local retail prices between rich and poor areas may exacerbate income inequality (DellaVigna and Gentzkow, 2019; Alcott et al., 2019). Indeed, the distribution of retail prices could be affected by spatial competition frictions between stores resulting in a bias of the standard inequality measures (Eizenberg et al., 2021; Moretti, 2013). Furthermore, in the modern retail market, firms are innovating to improve their price strategies making it easier to differentiate between different groups of consumers charging different prices for identical products (Chevalier and Kashyap, 2019; Nakamura, 2008).

In this work, I explore the determinants of retail prices in the Mexican supermarket industry in terms of demographics as well as the extent of uniform pricing within supermarket chains. On the one hand, using Fixed Effects (FE) models, I investigate the effects of population and deprivation index of the urban areas where retail stores are established on the retail prices of 23 products of regular consumption by households in Mexico. On the other hand, using a measure of pricing similarity proposed by DellaVigna and Gentzkow (2019), I estimate the extent of uniform pricing within supermarket chains for all the stores, and for those stores in the bottom, and top National Deprivation Index (NDI).

The main analysis is based on store-level data from the Who is Who on the Prices (QQP) retail panel of more than $2,000$ products. In addition, since the objective of this work is to analyze price patterns on basic geostatistical area level, data from the National Population Committee (CONAPO) of Deprivation Index were used to add demographics and socioeconomic variables in the QQP dataset. I use the median price of products by month over 7 years, from 2016 to 2022, in six supermarket chains: Walmart, Soriana, Chedraui, Comercial Mexicana, LEY, and HEB.

The first set of results in Section 5.2 underlines the extent of uniform pricing within and between supermarket chains. The price variation between chains is smaller than the variation in within stores of the same chain[1]. On average, regular prices differ in 0.001 log points between chains, but within chains the difference is in the range of 0.003 log points. Nonetheless, prices are similar within supermarket chains even for stores that face different economic conditions and are geographically separated (DellaVigna and Gentzkow, 2019).

The second set of results in Section 5.3 shows price differentials across urban areas explained by socioeconomic and demographics. Regular prices are affected by population density and deprivation level of the areas where stores are established. Consumers in more populated areas with low deprivation level are charged lower prices. In contrast, those residents of less populated areas with high deprivation level are charged higher retail prices. An increase of 1 percentage point in population implies a decrease of 0.003 percentage points in retail prices.

Both empirical findings are consistent with prior works, such as DellaVigna and Gentzkow (2019), Eizenberg et al. (2021), and Chevalier and Kashyap (2019). The relevance of this work resides in the economic implications of uniform pricing and price differences across areas. On the one hand, they may exacerbate income inequality. On the other hand, uniform pricing could dampen the response of retail prices to local economic shocks (Cawley et al., 2018; Leung, 2018 cited by DellaVigna and Gentzkow (2019)).

Although this work is not the first document to analyze uniform pricing or price differentials across urban areas, it is one of the first steps toward a better understanding of price patterns in the Mexican supermarket industry. Hence, the works proceeds as follows: Section 2 presents a brief literature review of prior related works, Section 3 presents a description of the supermarket industry in Mexico, and Section 4 describes a theoretical framework. Section 5 presents the data, the empirical strategies, and the empirical findings; while the Section 6 concludes.

[1] *Cf.* DellaVigna and Gentzkow (2019). They found an opposite effect, where pricing similarity is smaller in within chains than in between.

2 Related Literature

Differences in retail prices across urban areas have attracted attention in the economic literature, mainly because of their economic implications on inequality measures (Moretti, 2013). According to Eizenberg et al. (2021), starting literature with Caplovitz (1967) (cited by Eizenberg et al., 2021) has attempted to understand whether *"The Poor Pay More"*. However, there are mixed empirical findings. On the one hand, some works such as MacDonald and Paul (1991), Chung and Myers (1999), and Eizenberg et al. (2021) have found that retail prices charged in poor neighborhoods are higher than those charged in rich areas. On the other hand, Richburg Hayes (2000) and Aguiar and Hurts (2007) report that prices in paid by high-income households are higher.

MacDonald and Paul (1991) compared the price of a sample of goods across 322 supermarkets in the United States, showing that prices in suburban areas were 4% lower than in areas were poorer population lived. In contrast, Richburg Hayes (2000) reports that prices charged in rich zip codes are significantly higher compared with those prices paid in poor zip codes.

In the theoretical contributions, some works have considered spatial frictions (Hotelling, 1929; Salop, 1979). In this way, several empirical papers have taken an approach to study these spatial frictions in different industries (Adams and Williams, 2019; D. R. Davis et al., 2019; P. Davis, 2006). Nonetheless, Chintagunta et al. (2003) and Dubois and Jódar-Rosell (2010) study spatial frictions and competition in supermarket industry, using a discrete demand model, and a supply-side model they identify heterogeneous marginal costs, which results in differences in the retail prices due to the spatial frictions.

Uniform pricing in supermarket industry echoes puzzles in other markets such as soft drinks, movie tickets and rental cars (McMillan, 2007; Orbach and Einav, 2007; Cho and Rust, 2010). However, in these markets uniform pricing is different because prices are fixed across separate markets instead of across multiple products in the same market (DellaVigna and Gentzkow, 2019).

Some works consider that firms take uniform pricing as a price strategy to have a clean brand image (Eizenberg et al., 2021) or to maximize their benefits reaching and optimal price (Adams and Williams, 2019). Cavallo (2017) argues that retail stores with physical location are not the only stores where we can find uniform pricing, but the retailer website set uniform pricing. Other prior work argue that firm characteristics have significant effects on retail price variation between and within chains (Nakamura, 2008).

Finally, most of the prior works argue that uniform pricing is constant in some markets. Chevalier and Kashyap (2019) and Coibion et al. (2015) consider that retail prices paid respond to macroeconomic shocks even when they are constant. Furthermore, this article is related with works in behavioral industrial organization such as DellaVigna and Malmendier (2004), DellaVigna and Gentzkow (2019), and Eizenberg et al. (2021).

3 Supermarket Industry

A supermarket is defined by the North American Industry Classification System (NAICS) as a retail food and beverage establishment formed by economic agents who sell these goods without further processing to final consumers[2]. Based on the technology used, these agents are classified into three distribution channels: modern, electronic, and traditional.

According to the Federal Economic Competition Commission (COFECE, 2020), the modern channel is formed by chains of sell-service shops which use technology to manage stocks and record sales, in addition to having a wide variety of products and high sales volumes; their service times are long, as they hire employees for different shifts. Supermarket chains in the modern channel offer consistency, quality, and diversity of products imposing higher standards on their suppliers than those in the traditional channel, such as Official Mexican Standards (NOM)[3]. These standards imply that producers who wish to trade their products on this channel fix higher prices due to the additional costs to guarantee a homogeneous quality level in the products.

In Mexico, supermarket chains organized their stores in formats considering factors such as the socioeconomic level of customers, diversity of products, retail area, and number of employees. On the one hand, discount stores (TDBE[4]) have less than $1,500m^2$ of surface, between 11 and 30 employees, and offer less than 15,000 products. On the other hand, supermarkets (BSHM[5]) have between $2,500m^2$ and $10,000m^2$ of area, 51 to 250 employees, and sell more than 20,000 goods (COFECE, 2020).

The regulatory authority, COFECE (2020), argues that the BSHM formats have as a common characteristic to attract those customers of medium and high income, then when they are close, exert competitive pressures on each other. In contrast, the TDBE are auster establishments with smaller retail areas, and less variety of products, having as potential customers those who have lower income, then they do not exert competitive pressures.

At the national level, although there are 79 chains, each with more than 20 shops, only three chains operate in the entire country: Walmart, Soriana, and Chedraui – the latter has no presence in some states. Since Walmart has 40% of BSHM stores in Mexico, followed by Soriana with 20%, and Chedraui with 8%, the supermarket industry appears to be concentrated not only on a national level, but local markets could be since Casa Ley and HEB have 3% and 2% of BSHM stores, respectively.

Under the assumption that competition in the retail market is local, consumers minimize time and transport costs by going to the nearest supermarket, and they go to the second nearest when they get better prices, more diversity of products, or other services that compensate for the cost of transport and

[2]This definition is based on the NAICS 2018. NAICS was developed under the auspices of Statistics Canada (SC), the Office of Management and Budget (OMB), and the National Institute of Statistics and Geography (INEGI).

[3]Castillo et al. (2018) cited by COFECE (2020). Supermarket chains imposed some practices to guarantee quality of products, such as cooling chamber for transport of products.

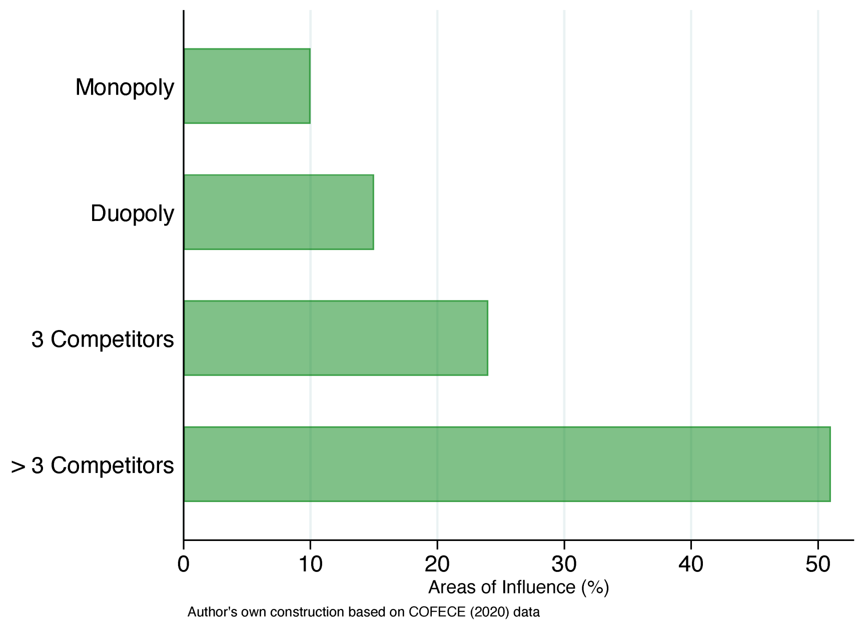
[4]Discount Stores and Express Stores.

[5]Stores, Supermarket, Hypermarket, and Megamarket.

time. As is expected, the proximity of stores determines the price competitive pressure between them. According to COFECE (2020), the radius of the areas of influence, the area around a shop, represents the kilometers that a person can travel by vehicle in a time of 15 minutes, this time is shorter the more population density there is.

Walmart is the chain that faces the least competition in the market, since in half of the influence areas of its established BSHM shops there are no other competitors present as the municipalities where it is located are lowly populated and developed. In fact, in 21% of the Walmart influence areas it faces monopolistic competition (COFECE, 2020).

Figure 1: AREAS OF INFLUENCE OF BSHM BY NUMBER OF COMPETITORS



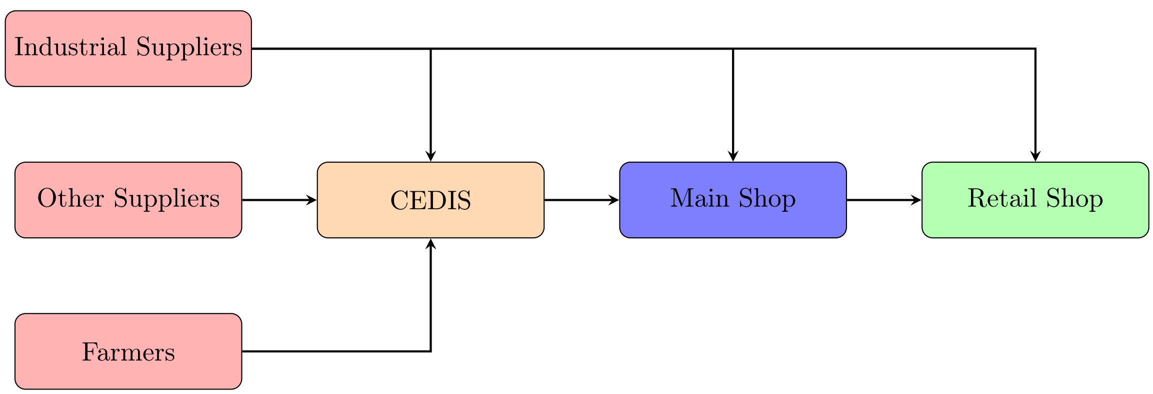
This figure shows the percentage of areas of influence of BSHM shop formats by number of competitors in Mexico based on data from COFECE (2020).

Based on the above chart, Figure 1, we can say that 10% of the areas of influence of BSHM format stores are monopolistic, where there is only one supermarket chain established; while in the 15%, there are two supermarket chains. Likewise, in 24% of the areas of influence, there are three competitors, whereas in the remaining areas, there are more than three competitors.

Although competition is local, decisions that influence dynamic competition are made at a regional level. At this level, supermarket chains make decisions to establish their distribution networks, which enables the creation of networks of shops to expand into local markets within that region (COFECE, 2020). Then, as chains expand, they gain cost advantages, which can be entry barriers or expansion barriers for their competitors by reducing the profitability of potential entrants.

In the modern channel, products arrive at retail shops in three different ways: suppliers, distribution centres (CEDIS), and from another nearby shop of the same chain, the "main shop". According to COFECE (2020), CEDIS and retail shops constitute an important entry barrier to new chains and the expansion of those chains of smaller size. The established chains can immediately cover lower costs associated with increases in demand since they have to increase only investment in supply and stock rotation with the infrastructure they have instead of building a new CEDI.

Figure 2: DISTRIBUTION CHAIN IN THE MODERN CHANNEL



Source: Author's own construction based on data from COFECE (2020). This diagram shows the distribution chain of products from suppliers to distribution centers (CEDIS), and to Retail Shops in the supermarket industry.

Figure 2 shows how products arrive at retail shops of the modern channel in three ways. First, directly from suppliers, or "truck-side", happens when suppliers have an extensive distribution network. Second, CEDIS are how supermarket chains distribute their products to the retail or main shops. Third, the nearest store in each region is called the "main store", which transfers stock to smaller stores.

4 Theoretical Framework

Considering the composition and dynamics of the industry, it is possible to explain the pricing behaviour of supermarket chains on the basis of three theoretical models. On the one hand, the Hotelling model, which shows the asymmetry of consumer transport costs. On the other hand, the monopolistic competition model. And adding a third one, a search model, which allows explaining the differences in the access to price information in the market.

Considering a retail market where chains establish their shops taking into account socioeconomic level and population, it will be assumed a model with quadratic transport costs, this distinction implies that consumers in poor neighborhoods have different transport costs from those consumers in less poor regions. According to Hotelling (1929), the indifferent consumer in the space is located at:

$$\hat{x} = \frac{1}{2}(l_i + l_j) - \frac{1}{2\tau}\left(\frac{P_i - P_j}{l_j - l_i}\right) \tag{1}$$

The indifferent consumer in the space, $\hat{x}$, determines the demand for competitors. It is composed by two terms: a proportion of the sum of the distance between firms, and a proportion of the ratio between differences of prices and distance, where τ is a disutility for consumer associated to travel. Then, equation (1) shows that benefits of firms are increasing in distance between the stores, and are decreasing in prices. That is to say, the more distant the shops, the greater the demand they capture. In contrast, the more price difference between stores, the lower the demand they have.

Since firms choose those prices such that $P_i \in argmax \;\; \Pi(P_i, \hat{x}(P_i, P_j))$, it is clear to see that prices are increasing with distance. Equation (2) shows how supermarket chains have incentives to establish their stores far away from their potential competitors in the local market looking for the ideal monopolistic competition.

$$\begin{aligned} P_i^* &= \frac{\tau}{3}(l_j - l_i)(2 + l_j + l_i) + c \\ P_j^* &= \frac{\tau}{3}(l_j - l_i)(4 - l_j - l_i) + c \end{aligned} \tag{2}$$

The idea captured in the above equations in that prices are increasing respect to the relative position of their competitors and the distance from the indifferent consumer. Consider the continuous space of consumers $[0, 1]$, if the position difference between stores in 0, $l_i = l_j$, the retail price $P_{i,j}$ equals the marginal cost c – there is a scenario of economic competition. In contrast, if stores are located in the ends, the retail prices equals $\tau + c$, which is the maximum price that can be reached. Hence, it is expected that supermarket chains look for scenarios where their stores are far away from their competitors.

Under the assumption that supermarket chains attempt to achieve monopolistic competition and inelastic demand, the optimal price results from maximizing the following:

$$\max_{P_{sj}} \sum_{s(r),j} (P_{sj} - c_{rj}) Q_{sj}(P_{sj}) - \sum_{s(r)} C_{sj} \quad (3)$$

$$ln(P^*_{sj}) = ln\left(\frac{\eta_{sj}}{1+\eta_{sj}}\right) + ln(c_{rj}) \quad (4)$$

From the above expression, equation (4), it is clear that, in absolute values, the more inelastic the demand the higher the prices will be. In contrast, the prices decrease in the elastic part of demand. Note that this is to be expected since COFECE (2020) argues that the bigger supermarket chains face limited competition and, as a consequence, an inelastic demand in those municipalities with poor socioeconomic conditions. Furthermore, consumers in less developed municipalities can be expected to face higher transport costs due to the proportions of income spent. In this case, it is expected to see more municipalities with few competitors since there are incentives to be far away from other.

In a complementary way, consumer search models are used to explain how retail prices are related to search costs in the economy. According to Stahl II (1989), there are N stores setting prices at the first stage, consumers have a distribution of search costs (c^i), and they adopt strategies to buy $D(p)$ in the store with the lowest price. When search costs are high, then the reserve price of consumers converges to the monopolistic price because of a lack of price information (Rosenthal, 1980; Stahl II, 1989; Diamond, 1971). In contrast, when the search cost are near to zero, retail prices converges to marginal costs because consumers can compare between stores, and competitors internalize this behavior fixing lower retail prices – perfect competition scenario (Stahl II, 1989; Bertrand, 1883).

These theoretical model implies the following: firms have incentives to establish their stores far away from their competitors, the more the distance between them, the higher the prices they charge; firms can charge higher prices in the inelastic part of their demands, markets with more inelastic demand are charged with higher prices. Then, firms attempt to achieve monopolistic competition scenarios because there they can increase their benefits by both less economic competition and higher retail prices. As a consequence, in areas with elastic demand prices charged are lower because disutility consumers is low and they have information about price distribution.

5 An Application to the Mexican Supermarket Industry

In this section, I estimate the uniform pricing that supermarket chains charge across their stores and the degree of price dispersion between urban areas. On the one hand, the objective is to show how Mexican supermarket chains charge nearly uniform prices across stores, despite wide variations in demographics. On the other hand, this section examines price differentials across urban areas considering consumer demographics and competition. The first subsection describes the data used for this work; the second subsection analyzes uniform pricing; and the price differentials analysis is in the last subsection.

5.1 The Data

Primary data sources are the Who is Who on the Prices (QQP)[6] and the National Deprivation Index (NDI)[7]. The first database contains information about the retail prices of products frequently consumed by households – food, beverages, medicine, and electronic devices. Furthermore, this data collect information about stores where products are sold, such as addresses, supermarket chains, etc. The second database contains socioeconomic and demographic information about urban areas based on the Economic Censuses of 2020[8].

5.1.1 Stores

The analysis is focused on food stores, particularly supermarket stores, and the geographic level considered for this work is the basic geostatistical area (AGEB), which represents the smallest urban area that can be studied with the Census 2020. Furthermore, the analysis is limited to seven years, from 2016 to 2022. It allows for interpreting price dynamics before and after the COVID-19 pandemic. Furthermore, the analysis is limited to seven years, from 2016 to 2022. It allows for interpreting price dynamics before and after the COVID-19 pandemic. The QQP and NDP were used to complete the final sample of stores.

National Deprivation Index was used as an auxiliary database to complete information on the supermarket stores about their demographics at the municipality and AGEB levels. The QQP database does not have demographic details of stores, but it has identifiers for entities (*ent_ store*) and municipalities (*mun_ store*). To merge information between both databases, it was define a store as a unique combination of three identifiers: *ent_ store*, *mun_ store*, and *ageb_ store* – identifier of the basic geostatistical area of each store. The last variables is not available in the data from PROFECO, then using the coordinates of stores, an identifier by AGEB was assigned to each store[9]. As a result, every store was identified by a unique combination making possible to match NDI data and QQP data.

[6]Who is Who on the Prices (QQP) is a project run by the Consumers' Federal Prosecutor (PROFECO).

[7]National Deprivation Index database is constructed by the National Population Committee (CONAPO).

[8]Economic Censuses of 2020 collects information and statistics about populations, demographics, and socioeconomic variables of Mexico, this census is managed by the National Institute of Statistics and Demographics (INEGI).

[9]This process resulted from an intersection of layers in QGis.

Since the objective of this work is to analyze prices in the supermarket industry, establishments that are not part of supermarket chains were excluded from the sample. According to COFECE (2020), there are more than 79 supermarket chains, however, I consider six supermarket chains: Walmart[10], Soriana, Chedraui, Comercial Mexicana, Casa LEY, and HEB.

In the final sample there are 189 stores of Walmart, 109 stores of Soriana, Chedraui has 45 stores, Comercial Mexicana represents 12.74% with 54 stores, LEY has 15 stores, and HEB 12 stores with 2.83%. The number of stores in BSHM formats shows that Walmart and Soriana are the two largest supermarket chains, having jointly 57% of the stores; while LEY and HEB are local supermarket chains with 5% of the stores in the sample, approximately.

The retail shops are unevenly distributed across deprivation level in Mexico. Most of the stores are highly concentrated in urban areas with low deprivation index, that is in those AGEB with higher economic development and more population; while only 6 retail stores are located in areas with high deprivation – they represent 0.05% of the stores in the sample.

Table 1: DISTRIBUTION OF STORES BY DEPRIVATION LEVEL

	Lowest No. of Stores	**Low** No. of Stores	**Middle** No. of Stores	**High** No. of Stores	**Highest** No. of Stores
Supermarket Chain					
Walmart	64	30	8	1	1
Aurrera	23	42	19	0	1
Comercial Mexicana	29	18	7	0	0
Soriana	55	41	11	2	0
Chedraui	13	27	4	1	0
LEY	9	6	0	0	0
HEB	6	5	1	0	0

Source: Author's own construction. This table shows the number of stores for each supermarket chain in Mexico, from 2017 to 2022 in the final sample by deprivation level. Although the national deprivation index is calculated as continuous variable, the discrete values are established by the National Population Committee (CONAPO).

Table 1 shows two relevant facts. On the one hand, the shops established in less advantaged areas are owned by the bigger supermarket chains – Walmart, Soriana, and Chedraui. On the other hand, the six supermarket chains established their shops in areas with higher development. These facts are consistent with other papers (Eizenberg et al., 2021; MacDonald and Paul, 1991; Chung and Myers, 1999). Bigger supermarket chains have cost advantages in using their distribution centers because they work as an entry barrier for new entrants in the areas of influence. It can explain why there are few participants in poor municipalities where they face inelastic demand. Contrary, in areas highly populated, there are competing pressures from other agents established, and they face elastic demand.

[10]Although there are six supermarket chains considered for the analysis, Figures and Tables below show Aurrera, which is part of Walmart chain.

5.1.2 Products

I focus on a sample of products that are frequently consumed by households. These goods are registered by PROFECO, and are considered to guarantee comparisons within and between chains. I consider 23 products in different categories – processed and non-processed food, vegetables, and personal care – ranging from canned tunny to milk.

To define the sample of products, two identifiers were used: *key_product* and *key_brand*. The first variable allows to identify the kind of product in a general way; while the second allows to identify specific characteristics of each product by brand. Then, combining both identifiers it is possible to guarantee that every product has a unique key with homogeneous characteristics.

Table 2: DISTRIBUTION OF GOODS

	Percent (%)	**Specification**
Good		
Oil	5.11	Bottle of 975ml
Rice	2.68	Bag of 1kg
Tunny	4.91	Can of 275g
Sugar	4.05	Bag of 1kg
Beef	2.83	Steak 1kg
Onion	4.97	1kg
Jalapeno	4.91	1kg
Pork	4.21	Steak of 1kg
Beans	3.91	Bag of 1kg
Eggs	4.79	12 pieces
Soap	5.09	Bar of 150gr
Tomato	4.94	1kg
Milk	3.63	Bottle of 1lt
Apple	4.84	1kg
Banana	4.94	1kg
Bread	5.11	Bag of 700gr
Potato	4.95	1kg
Toilet Paper	4.53	Pack of 12 pieces
Soup	5.10	Bag of 200gr
Chicken	4.67	Chicken leg 1kg
Pilchard	4.95	Can of 275gr
Carrot	4.86	1kg

Source: Author's own construction. This table shows the distribution of goods and their specifications in the final sample.

Table 2 shows the distribution of goods in the final sample as well as their specifications. Although there are goods that have less observations in the sample, such as beef, rice, and beans, the other goods have an equally distributed data. This allows to have clear comparisons between goods and their prices. The low frequency in beef, rice, and beans could be explained by the way in which the database was created. Since PROFECO collects prices for all sold goods in supermarket stores, the QQP database have prices for goods with specifications that differ from the above, then when I limited the sample of goods some observations were omitted.

5.2 Uniform Pricing

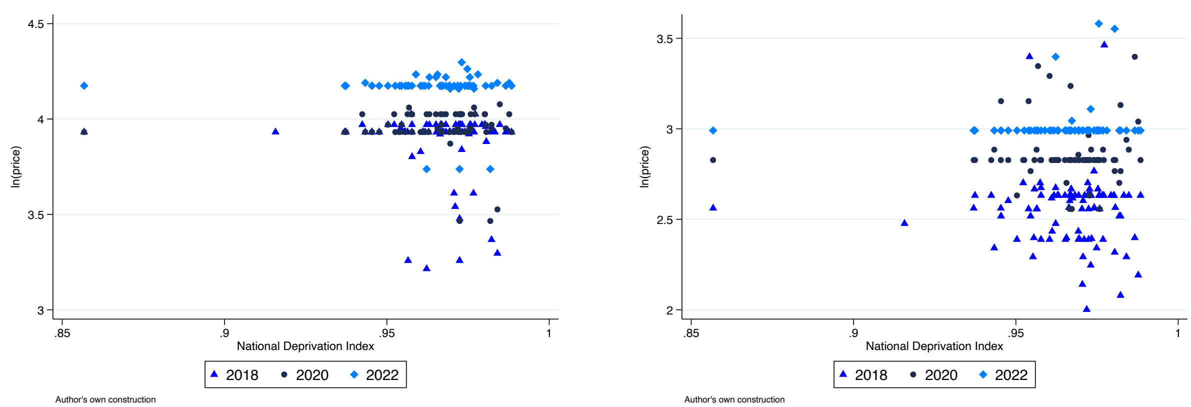
To describe pricing patterns that supermarket chains adopt, extent of pricing similarity is exhibited in two ways: descriptive visualization of pricing by chain, and using a measure of uniform pricing proposed by DellaVigna and Gentzkow (2019). First, descriptive visualization consists of displaying the price of some goods within representative supermarket chains[11] in three years – 2018, 2020, and 2020. Second, the measure of the extent of pricing similarity shows the distribution of price difference for some goods within chains.

Figure 3, Panel (a) shows the log prices of chicken for three different years – 2018, 2020, and 2022. The y-axis corresponds to the prices. The x-axis corresponds to the stores sorted by their level of deprivation from lowest to highest. Solid blue triangles represent prices of chicken in 2018; while the black solid circles are the prices in 2020; and the solid diamonds represent prices in 2022. These year are useful to show pricing behavior before, during, and after the COVID-19 pandemic.

Figure 3: PRICES ACROSS STORES

(a) Price of Chicken

(b) Price of Onion



Notes: Both panels present the $ln(price)$ across deprivation index within Walmart for three different years: 2018, 2020, and 2022. Panel (a) shows price dispersion of chicken. Panel (b) shows price dispersion of onion.

As DellaVigna and Gentzkow (2019), the figure shows clear variation across time, but apparently no variation across stores within chain. Furthermore, price variation is not visibly correlated with deprivation level of the urban area where store is established. Figure 3, Panel (b) shows a similar pricing pattern for onion. Although there is a wider dispersion graphically, this occurs mainly in 2018, and for 2020, and 2021, there is apparently no variation across stores. Next section describes pricing similarity in a better way.

[11] Walmart, Chedraui, HEB, and Casa LEY are considered the representative chains in this work. The first two supermarket chains are the biggest in the Mexican industry, having more than 200 stores in the sample. The last chains are considered to show consistency of results in local markets because these supermarket chains have established stores only in the North of Mexico.

5.2.1 Measure of Pricing Similarity

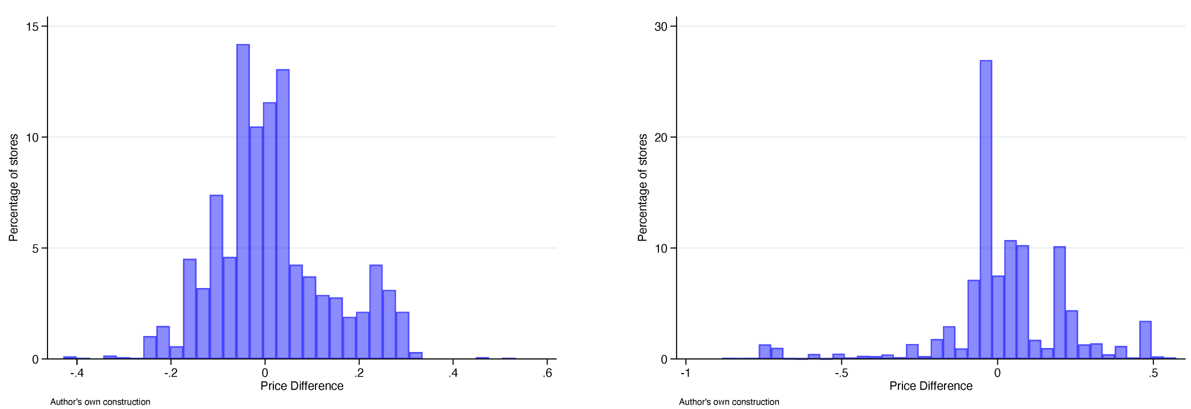
To describe pricing similarity more systematically, based on the measure proposed by DellaVigna and Gentzkow (2019) to evaluate the extent of uniform pricing, for each pair of stores s and s' and product j within the same chain was estimated the log price difference[12]. For each pairs of stores, product j, and month m, I first compute the median of log price of product j in stores s. I then compute the difference between log price and log median price. Finally, I average these monthly differences.

According to DellaVigna and Gentzkow (2019), this measure captures two orthogonal dimensions: cross-sectional differences in log prices, and correlation of price changes over time. That is to say, two stores with with different regular prices would have a high difference, but two stores with similar average prices would have a low difference.

Figure 4: LOG PRICE DIFFERENCE

(a) Price Difference of Beef

(b) Price Difference of Chicken



Notes: Both panels present the distribution of log price difference within Walmart for two different products: beef and chicken. Panel (a) shows the distribution of log price difference of beef. Panel (b) shows the distribution of log price difference of chicken.

Figure 4, Panel (a) shows the distribution of the measure of pricing similarity across stores within supermarket chain. Regular prices of beef for within-chain pairs are similar over time. The log price difference is typically between -0.2 and 0.2 log points from the median. Figure 4, Panel (b) shows a similar log price difference pattern. Regular prices of chicken for within-chain stores are similar, but the difference is typically lower since it is between -0.1 and 0.1 log points.

[12] *Cf.* DellaVigna and Gentzkow (2019). The authors estimate the quarterly absolute log price difference, which is different with the estimation that was made in this work because I considered those stores with prices below the median.

To summarize this measure, Table 3 shows summarizes of the baseline price similarity. Panel (A) shows mean and standard deviation of the log price difference for all store pairs. Panel (B) shows summarizes for store pairs in the lowest deprivation level; while the Panel (C) summarizes log price differences for store pairs in the top deprivation level.

Table 3: PRICING SIMILARITY, WITHIN CHAIN & BETWEEN CHAIN

	Same Chain (1)	**Different Chain** (2)
Panel A: All stores		
Mean	0.003	0.001
Standard Deviation	0.245	0.255
Panel B: Stores in the top NDI		
Mean	0.001	0.012
Standard Deviation	0.232	0.247
Panel C: Stores in the bottom NDI		
Mean	0.004	0.003
Standard Deviation	0.244	0.250

Source: Author's own construction. This table shows the pricing similarity within and between supermarket chain in Mexico, from 2017 to 2022 in the final sample. Panel A shows the mean and standard deviation for all stores; while Panel B and Panel C show the stats for those stores in the top and bottom NDI, respectively.

Table 3 shows in column (1) summarizes of log price difference within chain. Contrary to DellaVigna and Gentzkow (2019), on average, there is a difference in the regular price of 0.003 log points; while in stores of the bottom NDI, the difference increases in 0.001 log point. This difference in results could be explained by the absence of stores from different chains established in the top NDI areas[13]. Table 3, column (2) shows the log price difference between chain. However, on average the difference in prices in lower than in the stores in the bottom NDI.

These results suggest initial evidence against the argument that observed pricing similarity reflects within-chain stores serve homogeneous consumers in terms of demographics. Furthermore, these results suggest that this pricing similarity does not result from market constraints such as price advertising (DellaVigna and Gentzkow, 2019).

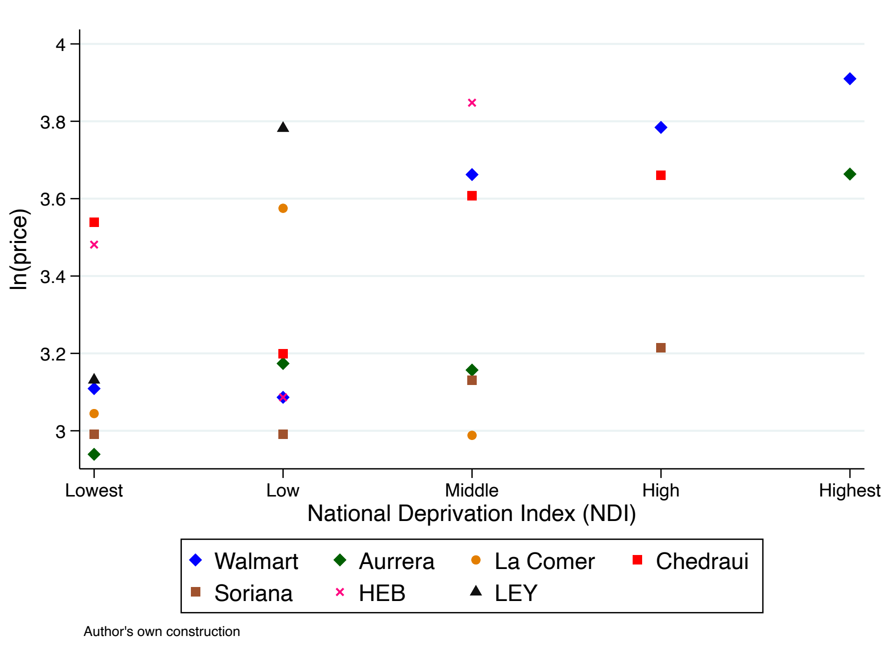
[13]In the final sample, Walmart is the only supermarket chain with stores in the top NDI.

5.3 Price Differentials Across Urban Areas

Given the previous results, apparently regular prices are similar across stores. However, according to COFECE (2020), supermarket chains consider demographics and socioeconomic factors to make decisions about where they should establish new retails stores. Then, the objective of this section is to examine price differentials across urban areas (AGEB).

Using log prices, I first estimate the median log price of every good over chains. The median log price allows to make clean comparisons eliminating outlier biases. Figure 5 plots median log price of chicken against national deprivation index over supermarket chains. Note that Walmart is the only chain with stores established in urban areas with highest deprivation level; while in urban areas with lowest deprivation level all the supermarket chains have presence.

Figure 5: LOG PRICES AGAINST DEPRIVATION LEVEL



This figure shows the median log price of chicken over the six supermarket chains – Aurrera is owned by Walmart.

Figure 5 shows that log prices charged in urban areas with high or highest deprivation level are apparently higher than those prices charged in stores established in lowest or low deprivation level. On average, there is a price difference of 0.4 log points between low and middle deprivation index for Walmart and Chedraui. Nonetheless, the price difference between middle and higher deprivation index is lower, passing from 3.6 to 3.8 for the same supermarket chains.

In order to expand on the previous analysis, considering the two largest supermarket chains – Walmart and Soriana (COFECE, 2020) – as well as two regional chains, such as LEY and HEB, there is a persistence of price patterns in some goods where urban areas with lowest deprivation level paid lower prices than in areas with highest deprivation. Focused on essential food products – chicken and soup – households in areas with low economic development paid more.

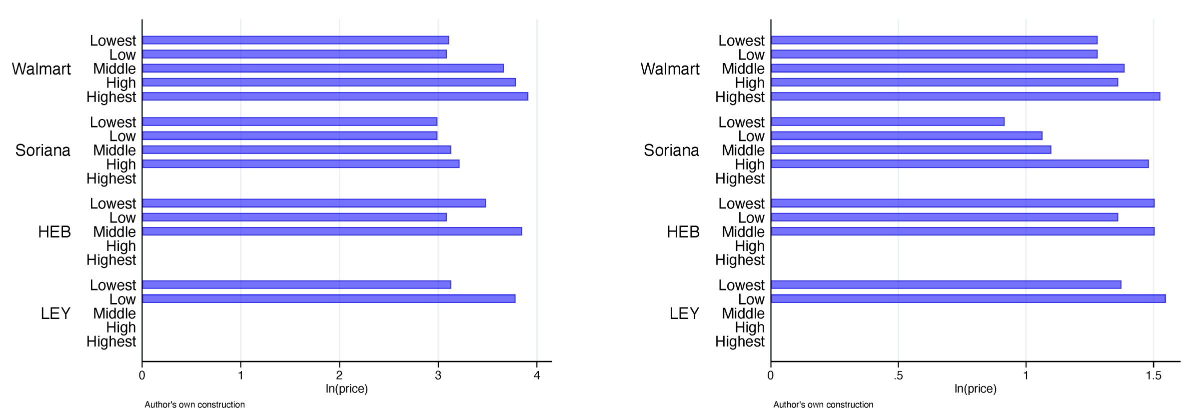
To describe price differentials of these food products across deprivation levels over the main supermarket chains in the industry, I first estimate the minimum log price of chicken and soup. The minimum log price is considered to make clean comparisons excluding differences in the quality of products; that is, the minimum price could guarantee an homogeneous good to compare.

Figure 6 shows the log prices of chicken and soup – two essential food products in a consumption basket. The y-axis corresponds to the deprivation levels over supermarket chains; from the bottom to the top, stores with highest deprivation and stores with lowest deprivation. The x-axis corresponds to the log price of these products. Panel (a) shows the log price of chicken by deprivation level over supermarket chains; while panel (b) shows the same information, but the log price of soup is illustrated.

Figure 6: PRICES ACROSS DEPRIVATION LEVEL OVER SUPERMARKET CHAINS

(a) Price of Chicken

(b) Price of Soup



Notes: Both panels present the $ln(price)$ across deprivation index over supermarket chains. Panel (a) shows prices of chicken. Panel (b) shows prices of soup.

There is a clear price differential across urban areas with different deprivation level. Price differential is in both products chicken and soup. On the one hand, Walmart charges a price of chicken 1 log point higher in the highest deprived areas than in the lowest. In a regional context, the supermarket chain LEY charges a similar price of 1 log point higher in low than in lowest. On the other hand, in the case of soup, the biggest price differential across urban areas occurs within Soriana, which charges in high deprived areas a price of 1.5 log point higher in lowest deprived areas. Although Walmart and LEY present a similar price pattern, the price differential across urban areas is not significant.

Although there is a considerably price difference that can be systematically associated with deprivation level of urban areas where stores are established, this figure does not have explanatory power to make conclusions. Since COFECE (2020) argues that population of the influence area play an important role to fix prices, I estimate the following econometric specifications to estimate the effects of population, deprivation index, and competition on prices of products.

$$ln(P_{sjc}) = \alpha + \beta_1 ln(TP_{as}) + \beta_2 NDI_{as} + \gamma_t + \gamma_m + \epsilon_{sj} \tag{5}$$

where $ln(P_{sjc})$ is the log price of product j in store s owned by supermarket chain c, $ln(TP_a s)$ is the log of total population of the urban area a where store s is established, NDI_{as} is the standardised deprivation index[14] of the urban area a where store s is, and γ_t, and γ_m are time and municipality fixed effects, respectively.

Following the model proposed by Eizenberg et al. (2021), variable of economic competition, NS_a, is added to the above equation. This variable consists in the number of stores established by urban area. Then the second econometric specification is the following:

$$ln(P_{sjc}) = \alpha + \beta_1 ln(TP_{as}) + \beta_2 NDI_{as} + \beta_3 NS_a + \gamma_t + \gamma_m + \epsilon_{sj} \tag{6}$$

where NS_a represents the variable associated with economic competition, and the rest of the variables are the same as in the equation (7).

Both equations were estimated by Ordinary Least Squares (OLS) and by Fixed Effects (FE). The coefficients of interest are β_1, β_2, and β_3; the first represent the effect of percentage changes in population on price level; to measure the effect of changes of deprivation level on prices, β_2; and the competition effects are captured by β_3.

It is expected to have the following relationships between regular prices and the explanatory variables. For the case of log population, it is expected a negative relation. According to COFECE (2020), areas with more population have higher competition pressures. That is, the more population in the urban area, the lower the regular price charged there. Second, for the case of deprivation index, it is expected to have a positive relationship because of there the market structure in the industry. Third, it is expected to have a negative relationship between competition and prices. The more competitors in the market, the lower the prices (Bertrand, cite).

[14] To estimate the standardised deprivation index, I compute the difference with the mean divided by the standard deviation of NDI.

5.3.1 The Estimation

To estimate the econometric specifications, I used Ordinary Least Squares (OLS) and Fixed Effects (FE) methods. First, the model with no economic competition variable was estimated. Secondly, adding the variable NS_a, the estimates by OLS and FE were reported. Using FE model, I controlled by product, time, and municipality fixed effects. Table 4 contains the estimates of log prices considering both OLS, and FE.

Table 4: ESTIMATES LOG PRICES, OLS AND FE

VARIABLES	(1) OLS	(2) FE	(3) OLS	(4) FE
AGEB Population	-0.003**	-0.004	-0.003**	-0.004
	(0.001)	(0.003)	(0.001)	(0.003)
AGEB Deprivation Index	0.018***	0.008***	0.018***	0.008***
	(0.001)	(0.003)	(0.001)	(0.003)
AGEB Stores	-	-	0.001	0.009
			(0.003)	(0.007)
Constant	3.275***	3.281***	3.275***	3.278***
	(0.010)	(0.020)	(0.010)	(0.021)
Observations	413,975	413,975	413,975	413,975
R-squared	0.001	0.909	0.001	0.909
Municipalities		56		56
Product FE		YES		YES
Time FE		YES		YES
Municipality FE		YES		YES

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4, columns (1)-(2) show OLS and FE estimates of equation (6). In the first column, both the AGEB population and AGEB deprivation index have the expected sign and are statistically significant. On the one hand, the increase of 1 percentage point (pp) in population implies a decrease of 0.003 pp of price. On the other hand, an increase of 0.01 standard deviations in the AGEB deprivation index implies an increase in prices. The second column shows the estimates by FE. Sign of the coefficients of interest does not change.

In columns (3)-(4) are shown estimates of equation (7) – number of stores by AGEB is added as explanatory variable. As in the equation (6), both population and deprivation index have the expected sign. However, the sign of AGEB stores is positive suggesting that the more participants in the market, the higher the price of products. This fact could be contradictory, but the coefficient associated is close to zero, and it is not statistically significant.

Finally, columns (3) and (4) show that adding a measure of economic competition – as the number of stores – has no effect of the other explanatory variables. Nevertheless, controlling for product, time, and municipality fixed effects substantially increases the goodness of fit of the regressions from 0.001 to 0.909.

6 Conclusions

This work uses a unique database which combines information on retail prices differentiated by supermarket chains and information on demographics of the urban areas where supermarket retail stores are established. To analyze both the determinants of price differentials across urban areas and the extent of uniform pricing within and between supermarket chains, I used the methodologies proposed by Eizenberg et al. (2021) and DellaVigna and Gentzkow (2019), which consist of a FE model, and log differences.

Using the methodologies to the Mexican supermarket industry, there were analyzed 23 products of regular consumption by households in Mexico, and the main findings are as follows. First, although retail prices differ across time, there is not variation across stores within a supermarket chain. Second, using the measure of pricing similarity, on average the price difference within chains is 0.003 log points, but the difference between chains is 0.001 log points. It is explained by competition pressures, since there is not competition within chains, they can fix higher prices in stores with less population or higher deprivation index. Third, there is a negative relationship between population and retail prices: an increase of 1 percentage point in population implies an decrease of 0.003 percentage points in the price. Furthermore, the relationship between retail prices and deprivation index is positive. However, the number of stores in the same urban area does not reflect a significant effect on regular prices.

Finally, this work can be extended in future works to examine price patterns in other samples of products, beyond regular consumption products. Furthermore, this analysis can be considered as a step toward a better understanding of the role played by demographics and competition in determining retail regular prices, which has important implications for analysis of mergers, and the extent of inequality motivating better policies to alleviate the higher costs in less affluent urban areas.

References

Adams, B., & Williams, K. R. (2019). Zone Pricing in Retail Oligopoly. *American Economic Journal: Microeconomics*, *11*, 124–156.

Aguiar, M., & Hurts, E. (2007). Life-Cycle Prices and Production. *American Economic Review*, *97*(5), 1533–1559.

Alcott, H., Lockwood, B. B., & Taubinsky, D. (2019). Regressive Sin Taxes, with an Application to the Optimal Soda Tax. *Quarterly Journal of Economics*, *134*, 141–156.

Bertrand, J. (1883). Théorie mathématique de la richesse sociale. *Journal des Savants*, 499–508.

Caplovitz, D. (1967). The Poor Pay More. *New York: Free Press*.

Castillo, V., M., M., & Ayala, S. (2018). Innovation in Distribution: New Strategies from Large Commercial Companies.

Cavallo, A. (2017). Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers. *American Economic Review*, *107*, 283–303.

Cawley, J., Frisvold, D., Hill, A., & Jones, D. (2018). The Impact of the Philadelphia Beverage Tax on Prices and Product Availability. *NBER Working Paper 24990*.

Chevalier, J. A., & Kashyap, A. K. (2019). Best Prices: Price Discrimination and Consumer Substitution. *American Economic Journal: Economic Policy*, *11*(1), 126–159.

Chintagunta, P. K., Dubé, J.-P., & Singh, V. (2003). Balancing Profitability and Customer Welfare in a Supermarket Chain. *RAND Journal of Economics*, *1*(1), 111–147.

Cho, S., & Rust, J. (2010). The Flat Rental Puzzle. *Review of Economic Studies*, *77*, 560–594.

Chung, C., & Myers, S. L. J. (1999). Do the Poor Pay More for Food? An Analysis of Grocery Store Availability and Food Price Disparities. *Journal of Consumer Affairs*, *33*(2), 276–296.

COFECE. (2020). Estudio de Competencia en el Canal Moderno del Comercio al Menudeo de Alimentos y Bebidas. *Comisión Federal de Competencia Económica*, *Noviembre*.

Coibion, O., Gorodnichenko, Y., & Hong, H. G. (2015). The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications. *American Economic Review*, *105*, 993–1029.

Davis, D. R., Dingel, J. I., Monras, J., & Morales, E. (2019). How Segregated Is Urban Consumption? *Journal of Political Economy*, *127*(4), 1684–1738.

Davis, P. (2006). Spatial Competition in Retail Markets: Movie Theaters. *RAND Journal of Economics*, *37*(4), 964–982.

DellaVigna, S., & Gentzkow, M. (2019). Uniform Pricing in U.S. Retail Chains. *The Quarterly Journal of Economics*, 2011–2084.

DellaVigna, S., & Malmendier, U. (2004). Contract Design and Self-Control: Theory and Evidence. *The Quarterly Journal of Economics*, *119*, 353–402.

Diamond, P. (1971). A Model of Price Adjustmen. *Journal of Economic Theory*, *3*, 156–168.

Dubois, P., & Jódar-Rosell, S. (2010). Price and Brand Competition between Differentiated Retailers: A Structural Econometric Model. *Ecole d'Économie de Toulouse*, *Working Paper 10-159*(1).

Eizenberg, A., Lach, S., & Oren-Yiftach, M. (2021). Retail Prices in a City. *American Economic Journal: Economic Policy*, *13*(2), 175–206.

Hotelling, H. (1929). Stability in Competition. *Economic Journal*, *39*(153), 41–57.

Leung, J. (2018). Minimum Wage and Real Wage Inequality: Evidence from Pass-Through to Retail Prices. *Working Paper, National University of Singapore Business School*.

MacDonald, J. M., & Paul, E. N. J. (1991). Do the Poor Still Pay More? Food Price Variations in Large Metropolitan Areas. *Journal of Urban Economics*, *30*(3), 344–359.

McMillan, R. S. (2007). Different Flavor, Same Price: The Puzzle of Uniform Pricing for Differentiated Products. *Cauchy Capital*, *Working Paper*.

Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, *5*(1), 65–103.

Nakamura, E. (2008). Pass-Through in Retail and Wholesale. *American Economic Review*, *98*, 430–437.

Orbach, B. Y., & Einav, L. (2007). Uniform Prices for Differentiated Goods: The Case of The Movie-Theater Industry. *International Review of Law and Economics*, *27*, 129–153.

Richburg Hayes, L. (2000). Do the Poor Pay More? An Empirical Investigation of Price Dispersion in Food Retailing. *Princeton Deparment of Economics Industrial Relations*, *Working Paper*(446), 276–296.

Rosenthal, R. W. (1980). A Model in Wich an Increase in the Number of Sellers Leads to a Higher Price. *Econometrica*, *48*, 1575–1579.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *Bell Journal of Economics*, *10*(1), 141–156.

Stahl II, D. O. (1989). Oligopolistic Pricing with Sequential Consumer Search. *The American Economic Review*, *79*(4), 700–712.

List of Figures

List of Tables