Universidad de San Andrés

Departamento de Economía

Doctor en Economía

Essays on Retail Pricing:
Competition, Relative Convergence, and Border Effect

Autor: Leandro Zipitría

Mentor: Walter Alberto Cont

Buenos Aires, Marzo de 2017

Essays on Retail Pricing:
Competition, Relative Convergence, and Border Effect

Leandro Zipitría

A dissertation submitted to the Faculty of the Department of Economics in Candidacy for the Ph.D. in Economics

Universidad de San Andrés, Buenos Aires, Argentina

Contents

Abstract:

This dissertation consists of three essays on price setting in retail markets using a detailed price database for Uruguay. The first essay examines the effects of local competition on price dispersion. While the literature has concentrated on the effects of borders on relative prices between countries – the distance dimension –, less is known about how the availability of different local products affect relative prices and the estimation of the border effect. I offer a model that incorporates substitution on the bases of distance between stores and different qualities – local competition – within stores to study the effect on relative prices between different locations. The model shows that: (i) relative prices will differs even if no border is in place when local competition differ between stores; and (ii) the border estimation will be biased if local competition is not controlled for. I test the predictions of the model for three markets in Uruguay. The empirical analysis shows that when borders are relevant in explaining relative prices between stores, controlling for local competition enables one to adjust their effects.

The second paper is co-authored with Fernando Borraz, Alberto Cavallo, and Roberto Rigobon. It also explores the relative convergence of prices between stores. The paper proposes a new methodology to estimate price dispersion, by estimating distance and border parameters using the upper quantile of the distribution of price differences. When the average distribution of price differences is used in the estimations, the border effect triples the distance effect. Conversely, when the upper quantile of the distribution of price differences is used in the estimations, the border effect is nil, while the effect – and implicit cost – of distance increases substantially.

The third paper is co-authored with Fernando Borraz. The paper explores the macroeconomic effects of retail price setting. In particular, we explore whether retail prices are consistent with state-dependent or time-dependent models in Uruguay. We find that prices change about five times a year with no seasonal pattern. Prices changes are highly synchronized and are concentrated on the first day of the month. Our paper is the first to present evidence of high synchronization of prices, which in turn could be explained mainly by the data periodicity. We found that, overall the analysis seems to be consistent with state-dependent pricing models, although we found some interesting features relating to prices that could not be explained by these models.

Acknowledgments:

Many people have contributed toward this project coming to fruition. First and foremost, I am extremely grateful to my advisor Walter Cont for his dedication and support throughout these years. He was always available to read and discuss the ideas underlying this dissertation. His comments, insightful criticisms, and suggestions not only enriched the papers, but also helped me understand the underlying economic forces behind each paper. I am also grateful to Martín Rossi and Christian Ruzzier of Universidad de San Andrés for their encouragement and support throughout my Ph.D. research.

I have benefited from very useful comments that enriched the papers. I would like to thank Andrew Bernard, Ariel Burstein, Lorenzo Caliendo, Alberto Cavallo, Juan Dubra, Alvaro Forteza, Francisco Gallego, Juan Carlos Hallak, Daniel Heymann, Ignacio Presno, Roberto Rigobon, Martín Rossi, Christian Ruzzier, Mariano Tommasi, Marcel Vaillant, and the seminar participants at UdeSA, dECON, LACEA, LAMES, RIDGE, and the Central Bank of Uruguay for very useful comments.

The General Directorate of Commerce of the Ministry of Economy and Finance of Uruguay kindly provided me with the database used for the empirical analysis in the dissertation. I am particularly indebted to Fernando Antía and Bruno Delgado for their support in building the database.

Many colleagues and friends have encouraged me during the dissertation process. My colleagues at dECON – Álvaro, Marcel, Patricia, and Rosario – provided constructive support and an appropriate environment that allowed me to concentrate on completing my PhD. I am also grateful to my co-authors Mario Bergara, Marcelo Caffera, Alberto Cavallo, Juan Dubra, Jorge Ponce, and Roberto Rigobon for their generosity in sharing their knowledge, which helped me to better understand economics. They made all the work enjoyable and fun. Finally, I have a special debt of gratitude to Fernando Borraz who has been a co-author for many years now. He has been an acute critic of many of the ideas we discussed from the inception, but he is also very generous with his knowledge and insights. I am glad to now call him a friend.

This dissertation would not have been possible without the support of my family, to whom I owe my endless gratitude. My wife Elizabeth has been – and still is – a source of constant love and support in helping me find the necessary balance in life to finish this dissertation. I have no words to express my love and gratitude to her. My children Sofia, Emilio, and Alicia have being a source of joy and happiness that has reminded me of the important things in life. My hope is for them to some-day look back and see this moment as a motivation for their own lives, whichever path they choose to follow. My brother Felipe has been a lovely support along these years. He also helped and advised me countless times with the script programming and gave me access to the servers to run them. Lastly, let this dissertation be a humble tribute to the memory of my sister Amalia and my parents Pipo and Graciela. I still remember Amalia's happiness and joy for life, the conversations with my father and his points of view about life, as well as my mother's love for learning and her curiosity for understanding the world around us, all of which have had a lasting effect on me. I am grateful for having had their love and company in my life. They will live forever in my heart.

Para mi esposa Eli y mis hijos Sofia, Emilio y Alicia.
En recuerdo de mi hermana Amalia y de mis padres Pipo y Graciela.

Chapter 1

Quality and Non-Convergence of Prices in Trade

Abstract

I propose a decomposition of the border effect in international trade by controlling for differences in competition at the local level. An extension of the Hotelling (1929) model shows that the availability of local substitutes increases price dispersion and biases the estimation of the border effect. I test these predictions for three markets using a detailed price database at the supermarket level for Uruguay. This stylized setting makes it possible to control for other potential explanations of the border effect (i.e., exchange rates, taxes, or transport costs). The results of the empirical methodology highlight the relevance of local competition in explaining the border effect: it increases for those goods not subject to local competition while decreasing for those exposed to local competitors. As the literature suggests, results should be even larger for different countries than for different cities. The methodology developed in the paper allows a finer explanation for understanding the relevance of borders in price dispersion.

JEL CODE: F14; F15; L13.

Keywords: border effect, relative prices, competition

1.1 Introduction

The impact of political borders on relative prices was empirically documented in a seminal paper by Engel and Rogers (1996). Using CPI data, the authors showed that the US–Canadian border had an effect on price dispersion equivalent to adding a distance of at least 1,780 miles between locations (approximately the distance between Miami and Quebec). A border is said to exist if, controlling for distance, the relative prices of the same good differs if the stores are in different geographical locations (either cities, counties, states, or countries). Their work spurred a large stream of literature that found similarly large "border effects" across countries, states, and even cities.[1] These results have been heavily debated over the years. The emphasis

[1]For example, see Parsley and Wei (2001) for results between the US and Japan and Ceglowski (2003) for the effects of provincial borders in Canada.

of the debate has been on the bias in the estimation of the border estimation due to different measurement and methodological issues.

In the debate over measurement in border estimation, it was argued that the distances between cities have been mis-measured (see Head and Mayer, 2002), and that regressions suffer from aggregation bias (see Evans, 2003 and Broda and Weinstein, 2008). The main methodological criticism was issue by Gorodnichenko and Tesar (2009) and established that differences in price dispersion within countries may bias the estimation of the border (i.e, price dispersion between countries), which they called the country heterogeneity effect. Borraz, Cavallo, Rigobon, and Zipitría (2016) also pointed to measurement bias in the estimation of the border effect due to the need to use maximum price distance (i.e., the upper quantile of price differences) to estimate transport costs. Previous papers found an upward bias in the estimation of the border effect, although a few found the border equal to zero after correcting for potential biases (see Broda and Weinstein 2008 and Borraz, Cavallo, Rigobon, and Zipitría, 2016).

One of the main debated issues that underlies most papers involves the differences in the implicit markups of prices between locations (see Coşar, Grieco, and Tintelnot, 2015a and Gopinath, Gourinchas, Hsieh, and Li, 2011). Building from insights in industrial organization, and using detailed cost information, some papers have overcome such limitations to estimate the impact of borders between countries. Gopinath, Gourinchas, Hsieh, and Li (2011) found a median discontinuity in relative prices of 24 percent between US and Canada, after controlling for costs and markups. Coşar, Grieco, and Tintelnot (2015b) found that borders in the wind turbine industry explain up to half of the differences in producer market shares between their home country and neighboring ones.

The literature on the estimation of the border effect has also moved from cities to stores, and from aggregate goods to precisely defined ones –mainly at the UPC code. Therefore, a typical analysis estimates the distance between two stores, either in the same city or across cities. Then, it usually identifies the exact same item in both stores (i.e., regular Coke sold in cans) and compares both prices in the same monetary unit.[2] As a result, products not sold in both geographical locations under analysis are discarded.

My paper is motivated by the fact that the previous analyses could be missing useful information about the border effect. The availability of local goods, i.e., goods sold only in one store but not in another, should distort the relative price in different countries.[3] However, this distortion is independent of the border, at least for those goods available in both locations. As a result, the literature has concentrated on one dimension of product arbitrage –substitution–: geographical distance. But, another arbitrage is possible for the consumer: to substitute for similar goods at the same location. Local competition will also influence markups, nor just geographical substitution.

Take the case of carbonated soft drinks as an example. When shopping at a store, the price consumers are charged for a given product, presenting them with a trade-off between moving to

[2]Gopinath, Gourinchas, Hsieh, and Li (2011) established, "Our first task consist in restricting the initial sample of 125,048 unique products to a set of products that appears on both sides on the border..." (page 2455). Nevertheless, Broda and Weinstein (2008) used the whole sample of products; see tables 3 and 4 in their Appendix.

[3]This was also established by Gopinath, Gourinchas, Hsieh, and Li (2011); see page 2451.

the next store –and buying the same preferred product– or purchasing a different good at the same store. Suppose a consumer is at a store to buy a Coke and she realizes its price is higher than the price charged at the next proximate store. She could either buy the Coke anyway -not moving to the next store- or she could move to the next store to buy the less expensive Coke. This is the classical analysis implied in the border literature. However, she could also buy Pepsi at the store she is currently in rather than buying Coke. Previous literature does not control for this dimension of substitution.

I analyze the border effect within a country. This methodology is adequate for avoiding the problems associated with exchange rates, taxes, language, non-price tariffs, factor market rigidities, and other restrictions that could affect the estimation of prices. Moreover, it also avoids the problems associated with transportation costs (see Gorodnichenko and Tesar, 2009). Uruguay is a small homogeneous country. People speak the same language, taxes are homogeneous at the country level, movements of goods and factors are free, and the maximum distance between stores in the sample is just 526 kilometers. No barriers between cities or states should be expected but rather a homogeneous convergence of prices. A similar analysis for different cities was made by Parsley and Wei (1996) and Yazgan and Yilmazkuday (2011) for the US, Crucini, Shintani, and Tsuruga (2010) for Japan, and Ceglowski (2003) for Canada.[4] These papers found a milder effect of intra-national borders for price convergence in relation to national borders.

The empirical approach and the nature of the data also address four additional sources of concerns that have been raised since the original Engel–Rogers regressions. First, I use product-level data with identical goods across locations. As suggested by Goldberg and Knetter (1997), product-level data is crucial to understanding deviations from the Law of One Price (LOP). Indeed, Evans (2003) and Broda and Weinstein (2008) argued that a significant problem in the border effect literature is the aggregation bias induced by price indexes. Second, the database has information on the exact location of each store. As pointed out by Head and Mayer (2002), using approximate distances (such as from one country capital to another) can greatly overestimate the border effect. Finally, the database has information for different supermarket chains that sold the same basket of products. This make it possible to control for competition between stores, as the sample of competing stores is larger than in previous papers.[5]

My paper is related to the work of Evans (2003), who addressed the problem of the relative substitution of similar goods across countries, Gorodnichenko and Tesar (2009), who established how differences in baskets of goods are a source of bias in the estimation of border effects, and Gopinath, Gourinchas, Hsieh, and Li (2011), who accounted for differences in markups to estimate borders. Nevertheless, it differ on several grounds. First, the paper explicitly introduces the substitution of goods within stores and relates it to the substitution across stores. Second, the theoretical model makes it possible to explain the convergence of prices of the same goods in different locations, taking into account both distance and local competition dimensions. Third, as I analyze the convergence of prices within a small country, I can isolate problems associated with exchange rate, language, and tariff barriers, which usually make

[4] Papers for trade within countries include Hillberry and Hummels (2003) and Wolf (2000).

[5] Gopinath, Gourinchas, Hsieh, and Li (2011) provided information on just one chain store.

the comparison of prices difficult. The problem is reduced to one of distance, local product substitutes, and the characteristics of stores or cities. Fourth, the analysis is based on a database that comprises nearly all the supermarkets in Uruguay. This makes it possible to capture the influence of local competitors that affects the price setting by each store. Fifth, I provide a simple technique for unfolding trade –border– costs from local product competition conditions. This make it possible to disentangle the relative importance of border costs –and local competition– on relative prices.

The model developed here shows that relative prices will differ if local competition is different between geographical locations. It also shows that not accounting for these different competitive conditions could bias the border estimation, although the direction of the bias should be empirically estimated.

The empirical section proposes a simple methodology for estimating the effect of local competition on the border estimation for three markets in Uruguay. It shows that more competition at the local level has a sizable effect on the border estimation. The border increases substantially for those goods not subject to local competition, while shrinking for those goods that do have local competitors at the store level. This shows that the border parameter is highly sensitive to local competitive conditions in each market, and this effect has a relevant impact in its determination. It also shows that the size of the border depends on local competition conditions.

The paper is organized as follows. The next section introduces the model and specifies the conditions that allows the prices of goods sold in different places to converge, when substitutes are available. Section 1.3 describes the dataset used to estimate the effect of the availability of substitutes on the estimation of the border effect. Section 1.4 introduces the equation to be estimated, the econometric results, and the robustness test to check the main results. Finally, Section 1.5 presents the conclusions of the analysis.

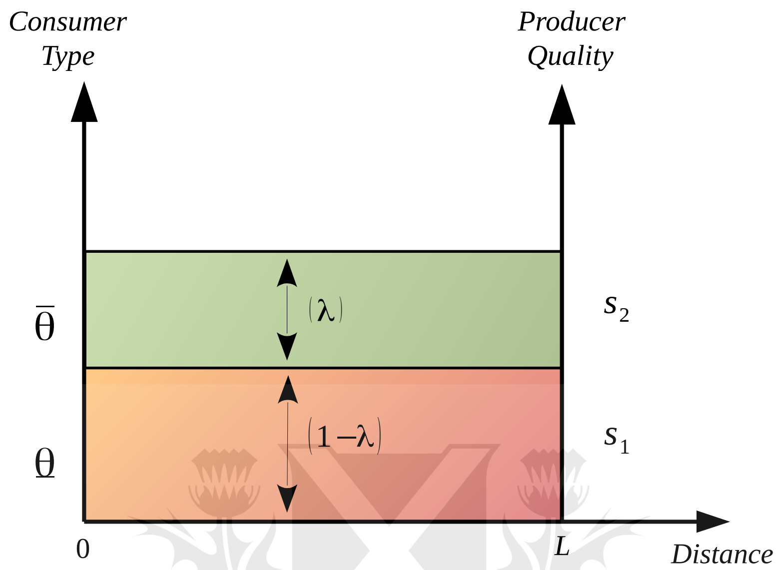
1.2 A Simple Model of Quality and Distance

I propose a simple extension of the Hotelling (1929) model, which incorporates vertical and horizontal differentiation in order to capture the essential claim. I add both dimensions to develop a model in which each one is a special case. In order to keep the model tractable and focus on the main message, I develop a discrete version of the quality dimension and assume that there is perfect information about consumers' characteristics. The model is related to the new verti-zontal models, as found in Di Comite, Thisse, and Vandenbussche (2014) or Gabszewicz and Wauthy (2012). This is a two-dimensional variant of the Hotelling (1929) and Salop (1979) models used in Gopinath, Gourinchas, Hsieh, and Li (2011).

As in the Hotelling (1929) model, consumers are uniformly located along a line of distance L. At each point of the line, there are two types of consumers that differ in their valuation of quality $\theta_i = \left\{\underline{\theta}, \overline{\theta}\right\}$, with $\overline{\theta} > \underline{\theta}$. At each point in the line there is a population of λ consumers that has a valuation for quality equal to $\overline{\theta}$ and $1 - \lambda$ consumers with a quality valuation of $\underline{\theta}$. The model could be thought of as being two lines of distance L, one on top of the other. The first line is for consumers of valuation $\overline{\theta}$, its thickness is λ, and the total mass of consumers is

$L \times \lambda$. The second line is for consumers with valuation $\underline{\theta}$, its thickness is $(1 - \lambda)$, and there is a total mass of consumers of $L \times (1 - \lambda)$. The next figure depicts the concept of the model.

Figure 1.1: The model with two type of consumers.



Products have a physical –distance– identification (d) but also a quality identification (s). Producers are located at one point in distance, and there could be different qualities of goods in a store. Consumers have an -indirect- utility function:

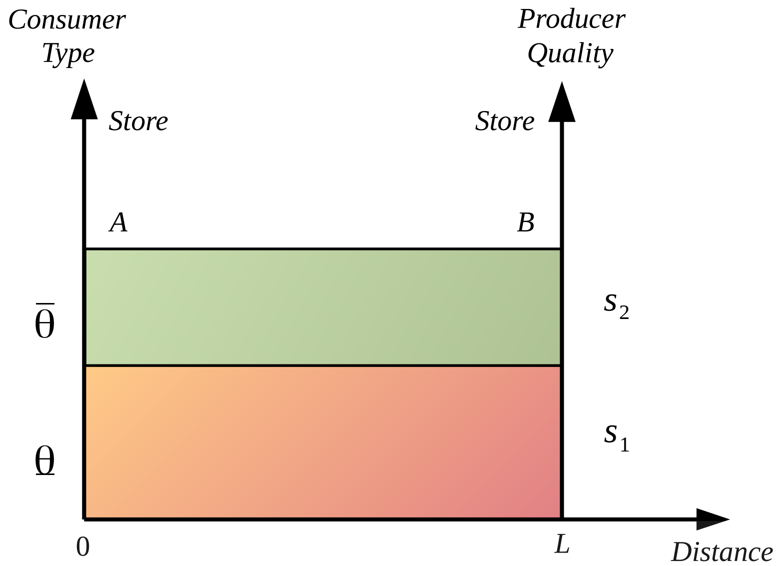
$$U_{ij} = r_{ij} + \theta_i s_q - t\,|x_j - x_d| - p_{qd},$$

where r_{ij} is the reservation utility of the consumer, i indicates the quality characteristic of the consumer $\left(ie.\,\theta_i = \left\{\underline{\theta},\, \overline{\theta}\right\}\right)$, j is the location of the consumer on the street line, s_q is the quality of the good, and d is the point of the store in the street line. For simplicity, I assume that the production costs of firms is equal to zero. Also, I assume the number of goods and their location are both exogenously given, and that production costs are zero. First, I derive the equilibrium conditions for two goods of the same quality (i,e., the traditional Hotelling problem), then I add a third good that differs in quality and derive the pricing equilibrium conditions. Finally, I will assume that each good is sold by a different producer/store.

Suppose there are two stores that sell the same quality $\underline{\theta}$ of the good. The stores are located in opposite places on the street. The first store is located at 0 and the second store at L, so L is also the distance between the stores.[6] The locations are indexed from the beginning of the street, either for consumers or stores (i.e., the consumer/store located at 0 is at the beginning of the street). The situation is depicted in Figure 1.2.

[6]Fixing the location of the stores eliminates one variable in the analysis (i.e., distance). I fix the store location to concentrate on the effects of quality.

Figure 1.2: The model with two stores.



This is the traditional Hotelling (1929) model with two stores, were A is the store located on the left and B is the one located on the right of the line. In order to find the price equilibrium, as I have assumed that the locations of both stores are exogenously given, the indifferent consumers must be found in order to establish the demand.[7] Fix quality at s_1 and the indifferent consumer is located between both stores:[8]

$$r_{ij} + \underline{\theta} s_1 - t\widehat{x} - p_A = r_{ij} + \underline{\theta} s_1 - t\left|L - \widehat{x}\right| - p_B \tag{1.1}$$

and solving for $\widehat{x}$ we obtain:

$$\widehat{x} = \frac{p_B - p_A + tL}{2t}. \tag{1.2}$$

The demand for store A is $\widehat{x}$: $D_A = \widehat{x} = \frac{p_B - p_A + tL}{2t}$, as consumers at the left of $\widehat{x}$ bought at that store regardless of their valuation of quality, and the mass of consumers in each point is 1 (i.e., λ consumers of quality $\overline{\theta}$ and $1-\lambda$ consumers of quality $\underline{\theta}$) and for store B: $D_B = L - \widehat{x} = \frac{p_A - p_B + tL}{2t}$.

Then, profits are $\Pi_A = p_A \times D_A$ and $\Pi_B = p_B \times D_B$, as I have assumed that cost are zero. Maximizing profits we find the reaction functions in prices, $p_A = \frac{p_B + tL}{2}$ and $p_B = \frac{p_A + tL}{2}$, and solving for the reaction functions in prices, we find:

$$p_A = p_B = tL,$$

and prices of both firms converge. This result holds as both firms have the same costs (zero in this case) and the same demand –in this case, $L/2$–.

[7] I assume that the minimum valuation of quality is large enough such that all consumers on the street buy the good; i.e., that $r_{ij} + \underline{\theta} s_1 - tx - p_A \geq 0$ or $r_{ij} + \underline{\theta} s_1 - t\left|L - x\right| - p_B \geq 0$ or both, $\forall x \in [0,\, L]$.

[8] Note that the same reasoning applies for the $\overline{\theta}$ consumer.

1.2.1 Quality

Now I assume that at location 0 there is another good of quality s_2 available for consumers, which I will call brand C. This brand also has zero production cost. As the model is continuous in the distance dimension but not on the quality dimension, I need to introduce additional assumptions in order to consumers buying product C. I will assume that, at 0, consumers of quality $\underline{\theta}$ will prefer to buy the low-quality good (that is to buy A), but consumers of quality $\overline{\theta}$ will prefer to buy the high-quality good (to buy C). This guarantees consumption for both goods, or entry of the new brand.

These assumptions add two restrictions to the model. Low-quality consumers located at 0 will prefer to buy A rather than C if $\underline{\theta}s_1 - p_A > \underline{\theta}s_2 - p_C \iff p_C - p_A > \underline{\theta}(s_2 - s_1)$. High-quality consumers located at 0 will prefer to buy C rather than A if $\overline{\theta}s_2 - p_C > \overline{\theta}s_1 - p_A \iff p_C - p_A < \overline{\theta}(s_2 - s_1)$. Both inequalities establish upper and lower bounds for the prices of stores A and C:

$$\underline{\theta}(s_2 - s_1) < p_C - p_A < \overline{\theta}(s_2 - s_1). \tag{1.3}$$

The upper bound, those prices that meet equality $p_C - p_A = \overline{\theta}(s_2 - s_1)$, allow us to establish the demand for stores B and C. If this upper bound binds, then high valuation consumers are indifferent in regard to buying at C or A, but low valuation consumers strictly prefer to buy at A.

Now I find the consumers who are indifferent about buying from stores B and C. Take the case of a consumer located on the high quality segment:

$$r_{ij} + \overline{\theta}s_2 - t\widetilde{x} - p_C = r_{ij} + \overline{\theta}s_1 - t\left|L - \widetilde{x}\right| - p_B, \tag{1.4}$$

and

$$\widetilde{x} = \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + tL}{2t}. \tag{1.5}$$

A comparison of equations 1.2 and 1.5 shows that $\widetilde{x} > \widehat{x} \iff p_C - p_A < \overline{\theta}(s_2 - s_1)$, with equality if $p_C - p_A = \overline{\theta}(s_2 - s_1)$. Appendix 1.6 shows that the model does not have a solution if $\widetilde{x} = \widehat{x}$; that is, if there is a corner solution $p_C - p_A = \overline{\theta}(s_2 - s_1)$.

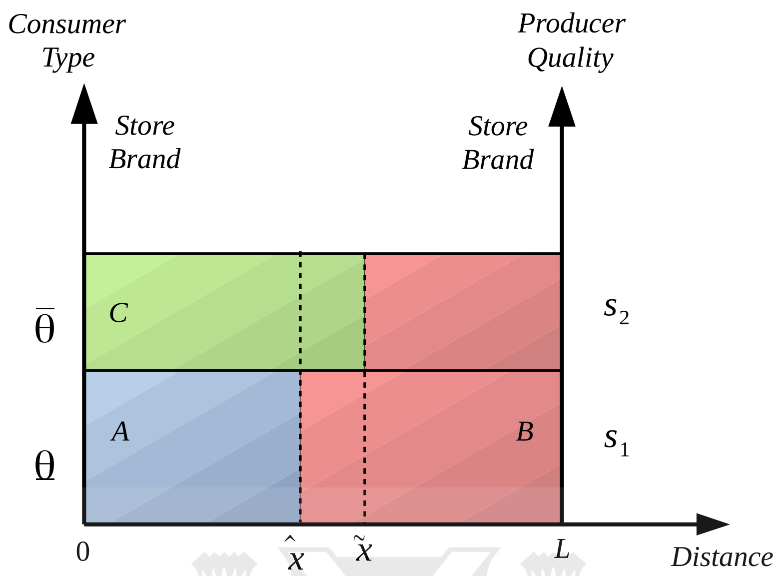
An interior solution to the model implies that the indifferent consumer for the quality segment should be at the right of $\widehat{x}$ if the previous equation does not hold with equality, but it could not be at the left. Figure 1.3 depicts the possible location of $\widetilde{x}$ for a given location of $\widehat{x}$ and the demand for each store.

Now I proceed to find the demand for each brand, taking into account the previous results. Demand for firm A is: $D_A = (1-\lambda)\widehat{x} = (1-\lambda)\frac{p_B - p_A + tL}{2t}$. Profits are $\Pi_A = p_A \times D_A$. The first order constraint of the problem is $\frac{\partial \Pi_A}{\partial p_A} = 0 = \frac{(1-\lambda)}{2t}\left[p_B - 2p_A + tL\right]$, so the reaction function is

$$p_A = \frac{p_B + tL}{2}. \tag{1.6}$$

Note that the reaction function of firm A depends –increasingly– only on the price of firm B, but not on the price of brand C. This result holds by construction of the model and the

Figure 1.3: Possible equilibrium values of $\tilde{x}$ and $\widehat{x}$. Demand for store A is depicted in blue, demand for store B in red, and demand for store C in green.



discrete nature of the quality decision.

For firm B, as $\tilde{x} > \widehat{x}$, its demand is affected by the entry of firm C, that is, $D_B = \underbrace{(1-\lambda) \times (L - \widehat{x})}_{low\, quality} + \underbrace{\lambda \times (L - \tilde{x})}_{high\, quality} = (L - \widehat{x}) - \lambda (\tilde{x} - \widehat{x})$.

The profit function is: $\Pi_B = p_B \left[\left(\frac{p_A - p_B + tL}{2t} \right) - \lambda \left(\frac{p_A - p_C + \overline{\theta}(s_2 - s_1)}{2t} \right) \right] = p_B \left(\frac{(1-\lambda)p_A - p_B + \lambda p_C - \lambda \overline{\theta}(s_2 - s_1) + tL}{2t} \right)$. From the FOC we obtain:

$$p_B = \frac{(1-\lambda) p_A + \lambda p_C - \lambda \overline{\theta} (s_2 - s_1) + tL}{2}. \tag{1.7}$$

The reaction function of firm B is increasing in p_A and p_C as they are both substitutes.

Lastly, the demand for firm C is $D_C = \lambda \tilde{x} = \lambda \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + tL}{2t}$. Profits are $\Pi_C = p_C \times \left[\lambda \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + tL}{2t} \right]$. The first order constraint is $\frac{\partial \Pi_C}{\partial p_C} = 0 = \frac{\lambda}{2t} \left(p_B - 2p_C + \overline{\theta} (s_2 - s_1) + tL \right)$. The reaction function for firm C is

$$p_C = \frac{p_B + \overline{\theta} (s_2 - s_1) + tL}{2}. \tag{1.8}$$

The solution to the three equations system is:

$$p_A^{'} = tL - \frac{\lambda \overline{\theta} (s_2 - s_1)}{6}, \tag{1.9}$$

$$p_B^{'} = tL - \frac{\lambda \overline{\theta} (s_2 - s_1)}{3}, \tag{1.10}$$

$$p_C^{'} = tL + \frac{(3-\lambda) \overline{\theta} (s_2 - s_1)}{6}. \tag{1.11}$$

We need to check if this solution is interior, in the sense that $p_C - p_A$ is a solution to inequality 1.3. The difference between p_C and p_A is $\frac{1}{2}\overline{\theta}\left(s_2 - s_1\right)$, so $\underline{\theta}\left(s_2 - s_1\right) < \frac{1}{2}\overline{\theta}\left(s_2 - s_1\right) < \overline{\theta}\left(s_2 - s_1\right)$ is satisfied only if $\frac{\overline{\theta}}{\underline{\theta}} < 2$. The results show that the prices of goods A and B are lower than if good C is not in place. As competition increase, prices decrease. Also, in this model, the effect of quality is independent of the effect of distance. The next Proposition summarizes the effect of quality on pricing.

Proposition 1.1. *Introducing quality into the model:*

1. *Decreases the price of low-quality goods;*
2. *Makes prices more volatile (i.e., price convergence less likely to hold)*

Proof. For 1, it is sufficient to note that $p_A^{'} = p_A - \frac{\lambda\overline{\theta}(s_2 - s_1)}{6}$ while $p_B^{'} = p_B - \frac{\lambda\overline{\theta}(s_2 - s_1)}{3}$. For 2, $p_A^{'} = p_B^{'} \iff \lambda = 0$ or $s_2 = s_1$, that is, if there is not other quality in the market. $\square$

Although the price of brand A does not depend on the price of good C, it has an effect through the price of brand B. As brand C induces the price of brand B to decrease, this affects the price of brand A in equilibrium. The effect of competition is more intense for good A.

The results do not change if we assume that there are two high-quality brands and one low-quality brand, as shown in the Appendix 1.7. In the next section, a border is added between the stores, and its effect on price convergence is evaluated.

1.2.2 Border

I modify the previous analysis and introduce a cost for the consumer to cross a hypothetical border between stores. This border cost could be language, the use of different paper money, paying a tax, etc. I assume that any of these factors imposes a cost on the utility of consumers, which they avoid by not crossing the border. I also assume that the border is between both stores, at distance ℓ_b. The border imposes a cost b for consumers that cross it in order to buy from a store located on the other side. Formally:

$$U_{ij} = r_{ij} + \theta_i s_q - \delta \times d - t\left|x_j - x_d\right| - p_{qd},$$

and δ equals 1 if the consumer located at j needs to cross the border to buy at a store located at d, and 0 otherwise. To understand the effect of the border, I return for a moment to the model with just one quality. Assume in that model a border located at point $\widehat{x}$, that is, where consumers are indifferent about which store they will buy from. Imposing a border implies that there is not one indifferent consumer but two: one located at the left and the other at its right. In turn, this implies that the border does not play any role if it is located where the indifferent consumer is.

Lemma 1.2. *If the border is located at the point where the indifferent consumer is, then the border cost is not relevant in the analysis.*

Proof. Assume two consumers, each one located at ε of the border $\widehat{x}$. For the consumer at the left, his utility for buying in stores A and B is

$$r_{ij} + \underline{\theta} s_1 - t\,(\widehat{x} - \varepsilon) - p_A > r_{ij} + \underline{\theta} s_1 - t\,[L - (\widehat{x} - \varepsilon)] - p_B + d,$$

and solving for $(\widehat{x} - \varepsilon)$ we obtain $(\widehat{x} - \varepsilon) > \frac{p_B - p_A + tL}{2t} - \frac{d}{2t}$. For the consumer located at the right, his utility is

$$r_{ij} + \underline{\theta} s_1 - t\,(\widehat{x} + \varepsilon) - p_A + d < r_{ij} + \underline{\theta} s_1 - t\,[L - (\widehat{x} - \varepsilon)] - p_B,$$

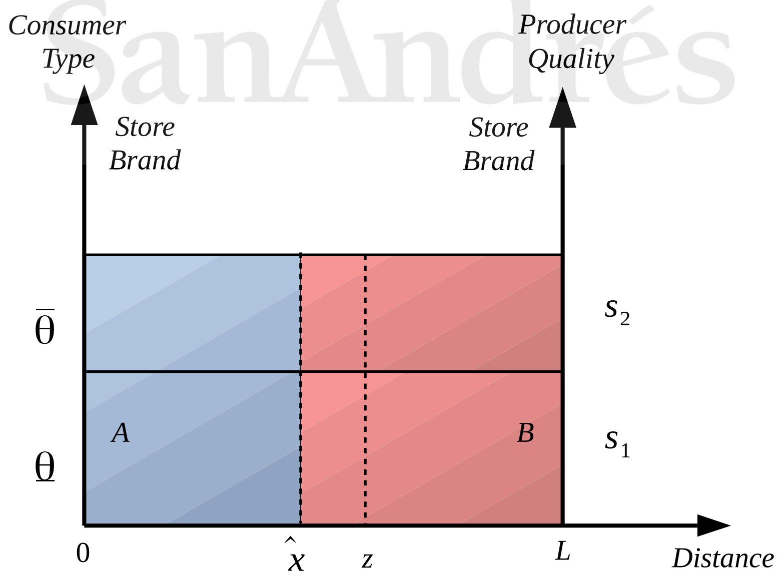
and solving for $(\widehat{x} + \varepsilon)$ we obtain $(\widehat{x} + \varepsilon) < \frac{p_B - p_A + tL}{2t} + \frac{d}{2t}$. As $\varepsilon \to 0$, we obtain $\frac{p_B - p_A + tL}{2t} - \frac{d}{2t} < \widehat{x} < \frac{p_B - p_A + tL}{2t} + \frac{d}{2t}$. Then, $\widehat{x} = \frac{p_B - p_A + tL}{2t}$. □

Lemma 1.2 says that the border is relevant only if it shifts consumers from buying in one store to buying in the other store. If consumer choice is not affected by the border –i.e., if demands do not change due to the border– then the border is irrelevant. As the border moves the indifferent consumer, this movement reaches a bound equal to the border itself.

1.2.2.1 Border with One Quality

Assume that there is only one quality and a border between stores. Assume that the border is at z to the right of $\widehat{x}$, as the next figure shows.

Figure 1.4: A border at the right of $\widehat{x}$.



For every positive border cost, the indifferent consumer should move from $\widehat{x}$ through z. The new indifferent consumer $\widehat{x}'$ should be equal to $\widehat{x} + b$, as the utility is lineal in cost. As a result, $\widehat{x}' = \widehat{x} + b = \frac{p_B - p_A + tL}{2t} + b$, where $b \in [0, (z - \widehat{x})]$. If b is bigger than $(z - \widehat{x})$, then Lemma

1.2 establishes that the demand for brand A should be z. Now $D_A = \frac{p_B - p_A + tL + 2tb}{2t}$, and the new reaction function is $p_A = \frac{p_B + Lt + 2tb}{2}$. Demand for store B is $D_B = \frac{p_A - p_B + tL - 2tb}{2t}$, and the reaction function for price B is $p_B = \frac{p_A + Lt - 2tb}{2}$. The new equilibrium prices are:

$$p_A = tL + \frac{2tb}{3}, \tag{1.12}$$

$$p_B = tL - \frac{2tb}{3}, \tag{1.13}$$

Lemma 1.3. *Borders make price convergence more difficult.*

Proof. Now $p_A - p_B = \frac{4}{3}tb$. □

If z is at the left of $\widehat{x}$, then the sign of the border coefficients in equations 1.12 and 1.13 reverse, but the Lemma remains unchanged by simply reversing the price difference. I now compute the size of the border by substituting p_A and p_B in $\widehat{x}' = \widehat{x} + b = \frac{5}{3}b + \frac{L}{2}$. As $\widehat{x}' \in \left[\frac{L}{2},\, L\right]$, then $b \in \left[0,\, \frac{3}{10}L\right]$.

Borders shift demand, so prices change with borders and price convergence becomes more difficult. This is the standard result found in the literature, where borders increase price variability in relation to the volatility of prices within countries.

1.2.2.2 Quality and Border

Now I extend the analysis of the effect of borders in a setting with different qualities. I will analyze the case where the border z is at the right of $\widetilde{x}$, and show the results for the case where the border z is at the left of $\widehat{x}$.[9] As $\widehat{x} \neq \widetilde{x}$, the effect of the border will be different for the high-quality and low-quality consumers. The next figure shows the case.

The new indifferent consumers will be

$$\widehat{x}' = \widehat{x} + \widehat{b} = \frac{p_B - p_A + tL}{2t} + \widehat{b}, \tag{1.14}$$

$$\widetilde{x}' = \widetilde{x} + \widetilde{b} = \frac{p_B - p_C + tL + \overline{\theta}\,(s_2 - s_1)}{2t} + \widetilde{b}, \tag{1.15}$$

where $\widehat{b} \in [0,\, (z - \widehat{x})]$ and $\widetilde{b} \in [0,\, (z - \widetilde{x})]$ and $\widetilde{b} \leq \widehat{b}$.[10] The border coefficient will be subtracted if the border z is at the left of $\widehat{x}$. The reaction function for brand A is the same as in the previous subsection: $p_A = \frac{p_B + Lt + 2t\widehat{b}}{2}$. Demand for firm B will now be $D_B = (1-\lambda) \times \left(L - \widehat{x}'\right) + \lambda \times \left(L - \widetilde{x}'\right)$ and substituting equations 1.14 and 1.15 and rearranging terms we obtain $D_B = \frac{(1-\lambda)p_A - p_B + \lambda p_C + Lt - \lambda\overline{\theta}(s_2 - s_1) - 2t\left[\widehat{b} - \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{2t}$.[11] Now the reaction function for firm B is

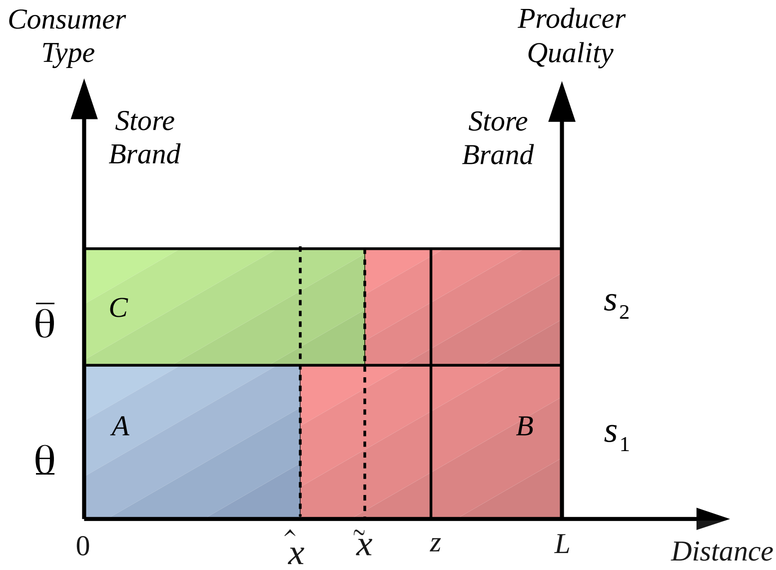
$$p_B = \frac{(1-\lambda)\,p_A + \lambda p_C + Lt - \lambda\overline{\theta}\,(s_2 - s_1) - 2t\left[\widehat{b} + \lambda\left(\widetilde{b} - \widehat{b}\right)\right]}{2}.$$

[9] The case where z is between both $\widehat{x}$ and $\widetilde{x}$ cancel out, as the analysis below shows.

[10] The inequality is reversed if the border z is at the left of $\widehat{x}$.

[11] If border z is at the left of $\widehat{x}$, then the border coefficients will be subtracting. Thus, we obtain $D_B = \frac{(1-\lambda)p_A - p_B + \lambda p_C + Lt - \lambda\overline{\theta}(s_2 - s_1) + 2t\left[\widehat{b} - \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{2t}$

Figure 1.5: A border at the right of $\widehat{x}$ and $\widetilde{x}$ when there are two qualities.



Demand for firm C is $D_C = \lambda \widetilde{x}' = \lambda \left[\frac{p_B - p_C + tL + \overline{\theta}(s_2 - s_1) + 2t\widetilde{b}}{2t}\right]$,[12] and the new reaction function is

$$p_C = \frac{p_B + tL + \overline{\theta}\left(s_2 - s_1\right) + 2t\widetilde{b}}{2}.$$

Substituting reaction functions we obtain:

$$p_A^{''} = tL - \frac{\lambda\overline{\theta}\left(s_2 - s_1\right)}{6} + \frac{t\left[2\widehat{b} + \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{3},$$

$$p_B^{''} = tL - \frac{\lambda\overline{\theta}\left(s_2 - s_1\right)}{3} - \frac{2t\left[\widehat{b} - \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{3},$$

$$p_C^{''} = Lt + \frac{\left(3 - \lambda\right)\overline{\theta}\left(s_2 - s_1\right)}{6} + \frac{t\left[\left(3 - \lambda\right)\widetilde{b} - \left(1 - \lambda\right)\widehat{b}\right]}{3}.$$

If the border z is at the left of $\widehat{x}$, the last coefficient in the three price equation is reversed. This implies that the border coefficient could either be positive or negative, dependent upon where the border is displaced. As a result, the border effect could either reinforce or hinder the quality effect.

Lemma 1.4. *The border could diminish or augment the quality effect.*

Proof. Price difference $p_A^b - p_B^b = \frac{\lambda\overline{\theta}(s_2 - s_1)}{6} + \frac{t\left[4\widehat{b} - \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{3}$ if the border z is at the right of $\widetilde{x}$. For the second case, if the border z is at the left of $\widehat{x}$, we have $p_A^b - p_B^b = \frac{\lambda\overline{\theta}(s_2 - s_1)}{6} - \frac{t\left[4\widehat{b} - \lambda\left(\widehat{b} - \widetilde{b}\right)\right]}{3}$. $\square$

When there are quality differences, the border effect always reinforces the quality effect. The main point of this section is twofold. First, the border coefficient changes when there is

[12] Accordingly, $D_C = \lambda \left[\frac{p_B - p_C + tL + \overline{\theta}(s_2 - s_1) - 2t\widetilde{b}}{2t}\right]$ if the border z is at the left of $\widehat{x}$.

a competition effect. A comparison between price differences in Lemmas 1.3 and 1.4 shows that border coefficients change due to the border effect. In Lemma 1.3, the border coefficient is $\frac{4}{3}b$ while in Lemma 1.4 it is $\frac{\left[4\widehat{b}-\lambda\left(\widehat{b}+3\widetilde{b}\right)\right]}{3}$ in absolute terms. Second, there is a quality effect in Lemma 1.4 that, if not accounted for, could bias the estimation of the border coefficient. In addition to the border coefficient, the term $\frac{\lambda\overline{\theta}(s_2-s_1)}{6}$ in Lemma 1.4 will be added to the border if not accounted for in the estimation. These results are shown in the paper's main proposition.

Proposition 1.5. *The availability of competitive substitutes bias the estimation of the border effect through two channels*

1. *A direct effect bias* $\left(e.g.,.\ \frac{4}{3}b \text{ vs. } \frac{t\left[4\widehat{b}-\lambda\left(\widehat{b}-\widetilde{b}\right)\right]}{3}\right)$
2. *An indirect effect bias* $\left(\frac{\lambda\overline{\theta}(s_2-s_1)}{6}\right)$ *due to the availability of different qualities*

The following table offers a summary of the results of the section.

Table 1.1: Results of the theoretical model.

	Equilibrium	Price diff.: $p_A - p_B$
Horizontal Price Dispersion	$p_A = p_B = tL$	0
Vertical Price Dispersion	$p_A = tL - \frac{\lambda\overline{\theta}(s_2-s_1)}{6}$; $p_B = tL - \frac{\lambda\overline{\theta}(s_2-s_1)}{3}$	$\frac{\lambda\overline{\theta}(s_2-s_1)}{6}$
Horizontal Price Dispersion and Border Effect	$p_A = tL + \frac{2tb}{3}$; $p_B = tL - \frac{2tb}{3}$	$\pm\frac{4}{3}tb$
Horizontal and Vertical Price Dispersion and Border Effect	$p_A = tL - \frac{\lambda\overline{\theta}(s_2-s_1)}{6} + \frac{t\left[2\widehat{b}+\lambda\left(\widehat{b}-\widetilde{b}\right)\right]}{3}$ $p_B = tL - \frac{\lambda\overline{\theta}(s_2-s_1)}{3} - \frac{2t\left[\widehat{b}-\lambda\left(\widehat{b}-\widetilde{b}\right)\right]}{3}$	$p^b_A - p^b_B = \frac{\lambda\overline{\theta}(s_2-s_1)}{6} \pm \frac{t\left[4\widehat{b}-\lambda\left(\widehat{b}-\widetilde{b}\right)\right]}{3}$

1.3 Data

I use a good-level database of daily prices compiled by The General Directorate of Commerce (DGC) in Uruguay, which comprises grocery stores all over the country.[13] The DGC is the authority responsible for the enforcement of the Consumer Protection Law at the Uruguayan Ministry of Economy and Finance.

In 2006 a new tax law was passed by the Uruguayan legislature that changed the tax base and rates of the value added tax (VAT). The Ministry of Economy and Finance was concerned about incomplete pass-through from tax reductions to consumer prices and hence decided to collect and publish a dataset of prices in different grocery stores and supermarkets across the country. The DGC issued Resolution Number 061/006, which mandates that grocery stores and supermarkets report their daily prices for a list of products if they meet the following two conditions: i) they sell more than 70% of the products listed, and ii) either have more than four grocery stores under the same brand name or have more than three cashiers in a store. The information sent by each retailer is a sworn statement, and there are penalties in case of

[13]This is an updated database from Borraz and Zipitría (2012) and Borraz, Cavallo, Rigobon, and Zipitría (2016).

misreporting it. The objective of the DGC is to ensure that prices posted on the DGC website reflect the real posted prices of the stores. In this regard, stores are free to set the prices they optimally choose, but they face a penalty if they try to misreport them to the DGC in an attempt to mislead costumers.

The data include daily prices from March 1st of 2007 to September 30th of 2014 for 212 universal product code (UPC) item.[14] The products in the sample represent 16.34% of the goods and services in the CPI basket. The DGC requires large retailers to report their daily prices once a month using an electronic survey. The three best-selling brands are reported for each product category, disregarding the supermarket's own brands. Most items have to be homogenized in order to be comparable, and each supermarket must always report the same item. For example, sparkling water of the local brand "Salus" is reported in its 2.25 liter variety by all stores. If this specific variety is not available at a store, then no price is reported. The data are then used on a public web site that allows consumers to check prices in different stores or cities and to compute the cost of different baskets of goods across locations.[15]

The database has information on 50 markets defined at the product category level. For many of them, the information does not allow the identification of the goods at the UPC level or determination of whether there are local competitors. However, in the meat and bread markets, products do not have brand. While the database has information for the three best-selling brands of toothpaste, toilet paper, beer, and several other items, there is only one firm that sells all three brands. For the remaining markets, I chose the three that came closer to the international setting I am trying to mimic. I pick between those that have either –at least– an international brand and a local brand (as in the mayonnaise and soft drink markets), or an international firm that owns a local brand and has other local competitors (as in the sparkling water market). I identify three categories for the analysis: Sparkling Water, Soft Drinks, and Mayonnaise.[16]

The Soft Drinks market has international brands –Coke and Pepsi– but also a local brand called Nix. In the Sparkling Water sector, nearly all firms are local. The products "Salus" and "Matutina" are sold by the same firm (Salus is owned by the international group Danone), while "Nativa" is produced by the same firm that produces "Nix" in the Soft Drink market. Finally, the Mayonnaise market includes information for two brands sold by Unilever (Hellmans and Fanacoa) and one local brand (Uruguay, sold by a local oil producer COUSA).

The database has a larger number of supermarket chains than in Gopinath, Gourinchas, Hsieh, and Li (2011), who provide information for only one supermarket chain, although they also had daily prices. Nevertheless, the database has information for the three best-selling goods in each market. Some small brands and supermarket own brands are not available in the database. The next table shows the information for each product in the database.

[14] The database contain nine prices for October, November, and December of 2014 for two supermarkets and for one product. I keep those prices, as when price differences are constructed the results do not change.

[15] See http://www.precios.uy/servicios/ciudadanos.html and Borraz and Zipitría (2012) for a detailed description of the database and an analysis on its price stickiness.

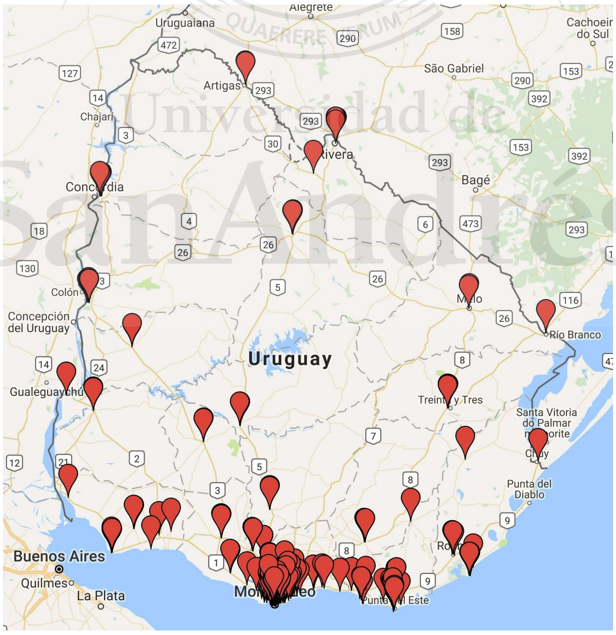
[16] A detailed description of the database will be available in future versions.

Table 1.2: Products in the database.

Sparkling water		Soft Drinks		Mayonnaise	
Brand	*Presentation*	*Brand*	*Presentation*	*Brand*	*Presentation*
Salus	2.25 liters	Coke	1.5 liters	Hellmans	0.5 kilos
Matutina	2 liters	Pepsi	1.5 liters	Fanacoa	0.5 kilos
Nativa	2 liters	Nix	1.5 liters	Uruguay	0.5 kilos

For each supermarket, I have detailed information about the exact location given by its Universal Transverse Mercator (UTM), its size –measured by the number of cashiers–, and if it belongs to a chain. Uruguay is divided into nineteen political states called "departamentos." The database has information for supermarkets across all nineteen political zones, comprising 54 cities. There are up to 386 supermarkets that report Sparkling Water, Soft Drinks, and Mayonnaise in the sample. A detailed description of each supermarket chain that includes its location and the availability of each good in the sample is available in Table 1.8 at the Annex. Montevideo, the capital city of Uruguay, is also the largest city, with nearly forty percent of the Uruguayan population.[17] The following figure shows the cities in the database and the supermarket distribution for Montevideo, which accounts for 54% of all supermarkets in the sample.

Figure 1.6: Cities covered in the sample and distribution of supermarkets.



Note: Each dot represents a store location across the 19 Uruguayan states.

For each brand and store, I chose the mode of the monthly prices to reduce the database dimension, although I tested the robustness of the results using the median, average, and observation on the first day of the month. According to Borraz and Zipitría (2012), prices change on the first day of the month 10 times more frequently than on any other day. As a

[17]More information is available at http://www.ine.gub.uy/uruguay-en-cifras.

result, the first observation will reasonably capture the main price changes in the dataset. This reduction in the dimension of the database is crucial because of the calculations that must be performed to obtain the results.

I check for outliers in the sample by filtering each series to exclude those observations above three times (or a third of) the monthly median price.[18] However, deleted prices only account for a small 0.0045% of the whole database (8 observations in 177,987). Also, information for 853 observations had to be eliminated due to a lack of information about two supermarkets. There are 177,126 observations in the database.

1.3.1 Descriptive Statistics

I first show some statistics for products and then for supermarkets. The following table shows sample information for each product in the as well as descriptive statistics for the mode price of the month.

Table 1.3: Sample information and price descriptive statistics for each product.

	Sparkling Water (Brands)			**Soft Drinks (Brands)**			**Mayonnaise (Brands)**		
	Salus	**Matutina**	**Nativa**	**Coke**	**Pepsi**	**Nix**	**Hellmans**	**Fanacoa**	**Uruguay**
Sample	2007/03 –2014/09	2007/03 – 2014/09	2007/03 – 2014/09	2007/03 – 2014/09	2011/11 – 2014/09	2007/05 – 2014/09	2007/03 – 2014/09	2007/03 – 2014/09	2007/07 – 2014/09
Panel A: Price database									
# observations	27,003	24,880	17,657	27,299	13,102	6,366	26,622	21,543	12,654
% in market	39	36	25	58	28	14	44	35	21
% of Montevideo	57	58	64	57	56	74	58	59	47
% supermarkets	100	97	77	99	96	37	99	96	56
Panel B: Price difference database									
# observations	4,063,790	3,488,074	1,745,884	4,144,790	1,849,121	230,171	3,941,040	2,646,392	934,171
% of market	44	37	19	66	30	4	52	35	13

Source: author's calculation.

	Sparkling Water (Brands)			**Soft Drinks (Brands)**			**Mayonnaise (Brands)**		
	Salus	Matutina	Nativa	Coke	Pepsi	Nix	Hellmans	Fanacoa	Uruguay
Panel A: Price database									
Minimum	14.5	12.9	13.0	16.1	29.9	15.7	19.9	14.5	9.9
Median	26.0	19.0	23.0	42.0	52.0	30.0	52.6	32.9	31.0
Maximum	37.0	32.5	33.0	68.0	70.0	45.0	89.0	67.0	52.0
SD	4.1	3.1	3.3	9.3	6.1	3.4	11.1	6.9	5.4
Panel B: Price difference database (percentage of price difference)									
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	0.0	4.4	3.5	2.5	4.1	5.1	6.2	7.0	7.5
Maximum	70.2	91.6	55.7	93.0	82.2	93.5	91.2	107.4	110.5
SD	3.8	5.5	5.8	6.0	5.6	10.6	6.5	7.3	7.8

Source: author's calculation.

[18]This is similar to Borraz, Cavallo, Rigobon, and Zipitría (2016) and more stringent than Klenow and Kryvtsov (2008), who excluded those prices 10 times larger (see page 867).

The brand "Pepsi" in its 1.5 liters presentation starts on November 2011 in the sample. Previously, supermarkets reported the 2.0 liters presentation. I pick the 1.5 liters presentation as I want to compare products with the most similar presentations. The following figures plot the price differences –in percentage– for each market and good.

Figure 1.7: Price difference distribution in the Sparkling Water market.

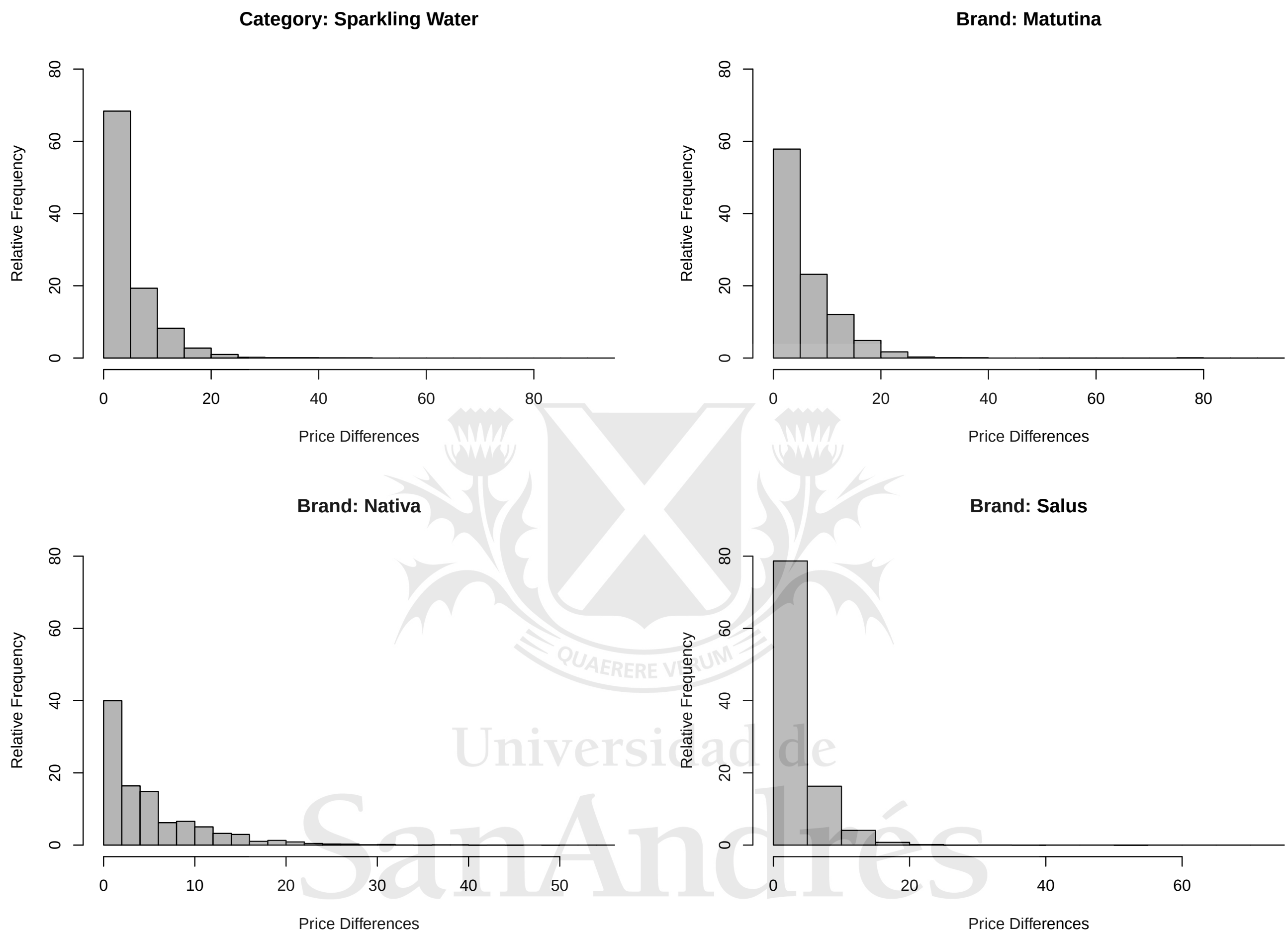


Figure 1.8: Boxplot of price difference in the Sparkling Water market.

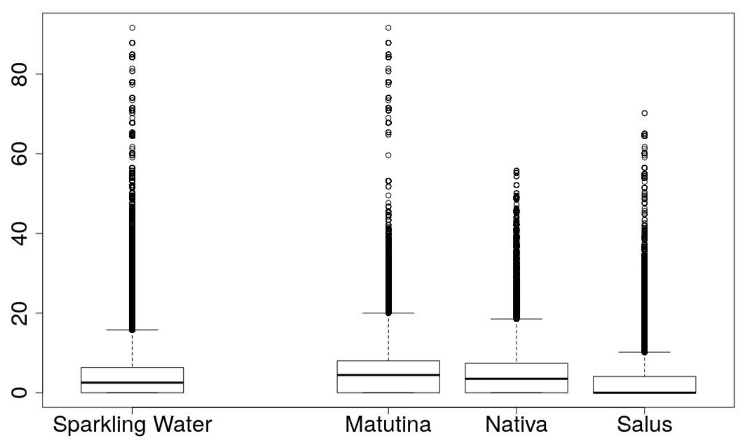


Figure 1.9: Price difference distribution in the Soft Drinks market.

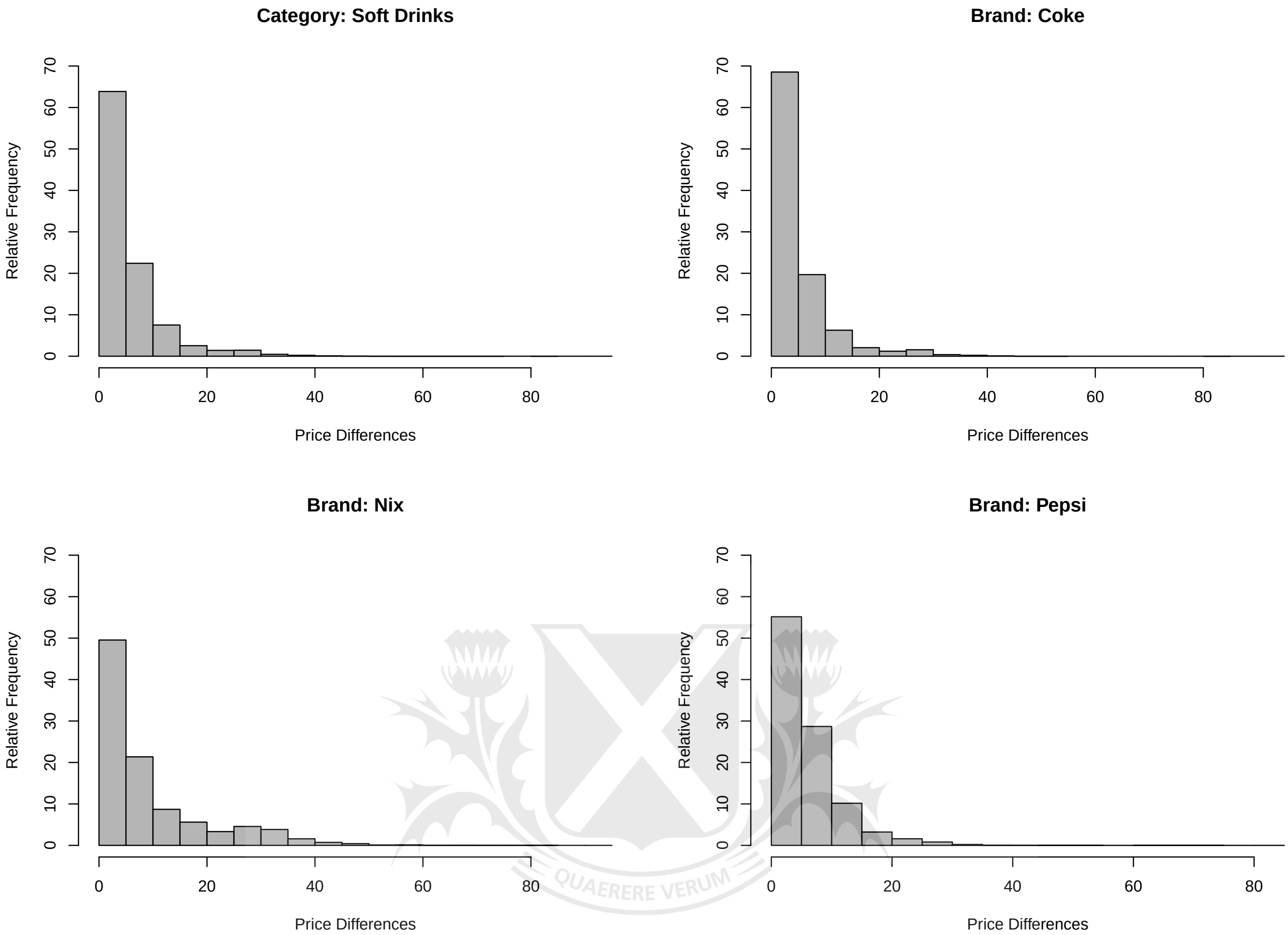


Figure 1.10: Boxplot of price difference in the Soft Drinks market.

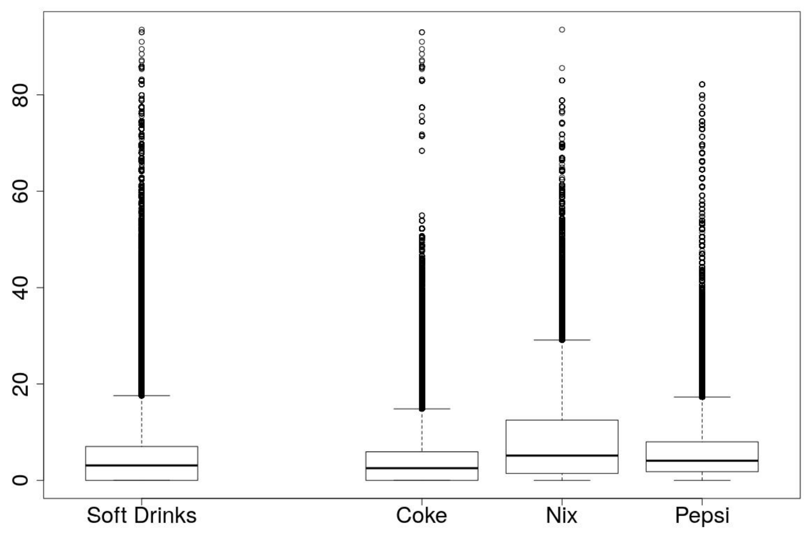


Figure 1.11: Price difference distribution in the Mayonnaise market.

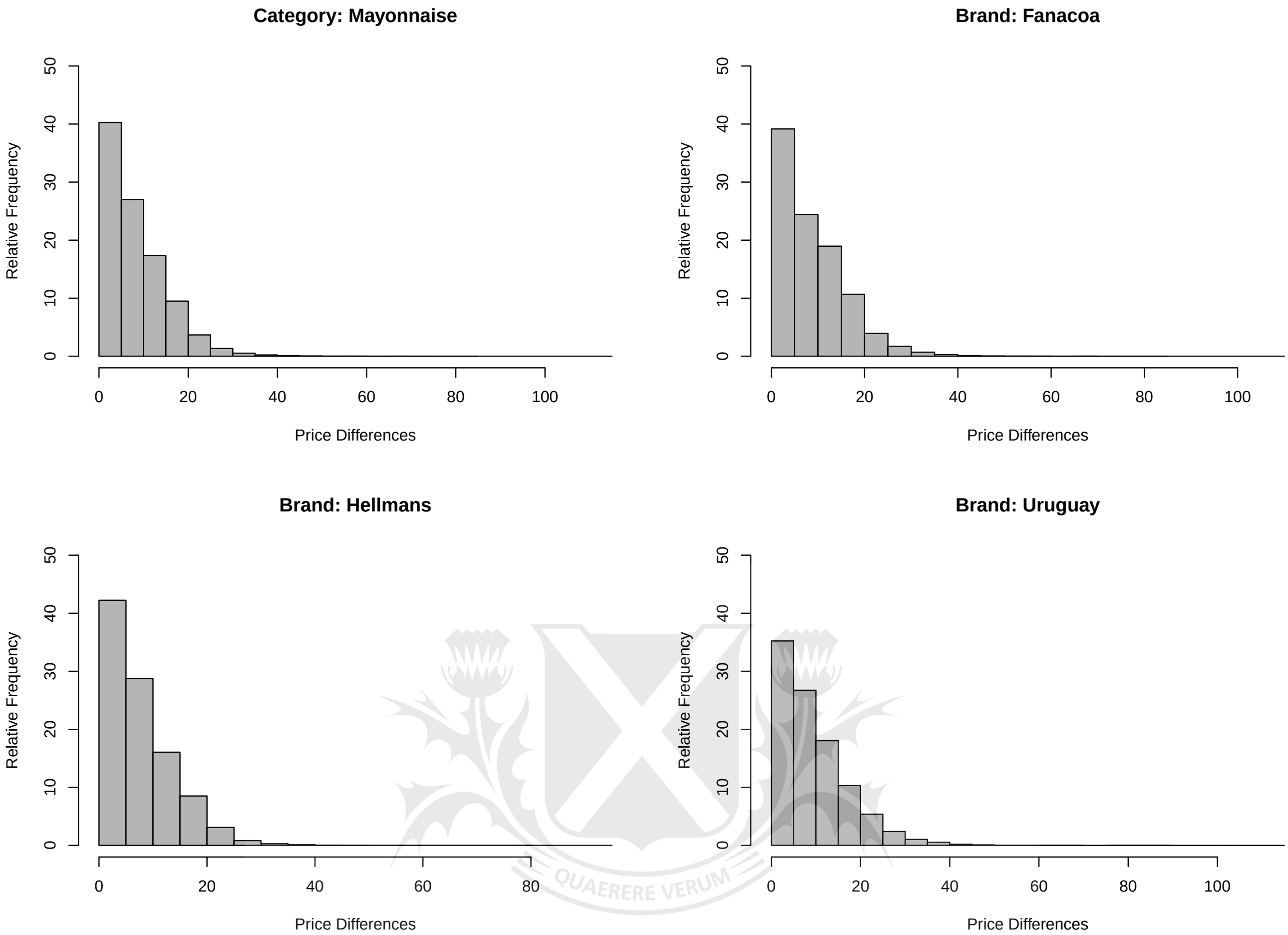
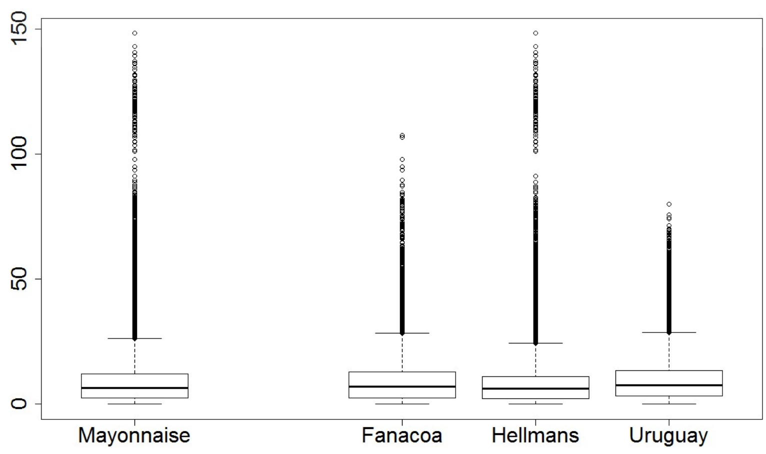


Figure 1.12: Boxplot of price difference in the Mayonnaise market.



The previous figures show that there is much more price variation for local brands, that is, for those brands that are less available. Because these figures show the histogram of the absolute difference of the log prices they are not useful for estimating the proportion of exact equal prices. The next table shows the proportion of zeros –i.e., equal prices– in the database.

Table 1.4: Percentage of zero price difference for each market and brand.

	Total	**Within Cities**	**Between Cities**
Water	41	52	35
Matutina	32	39	28
Nativa	36	48	27
Salus	51	66	44
Soft Drinks	28	48	18
Coke	32	56	21
Nix	20	24	15
Pepsi	19	34	12
Mayonnaise	11	16	16
Hellmans	11	16	16
Fanacoa	12	16	16
Uruguay	6	11	11

Source: author's calculation.

The chance of obtaining equal prices seems to vary according to markets and brands. The Mayonnaise market has much more volatile prices than the Soft Drinks and Sparkling Water markets. Also, small local brands seem to have more price volatility, with the exception of Nativa in the Sparkling Water market.

The distance between pairs of stores varies a lot, taking into account whether the stores are within or between cities. The next table shows statistics for the distance between supermarkets pairs.

Table 1.5: Descriptive statistics for distance between supermarkets (in kilometers).

	Total	**Within City**	**Between Cities**
Minimum	0.0	0.0	0.4
Median	78	6	119
Maximum	526	29	526

Source: author's calculation.

The following figures show the distribution of prices by distance in the sample. Figure 1.13 shows the distribution of observations for the whole sample, while figures 1.14 and 1.15 show the distributions between and within cities for each market.

Figure 1.13: Observations by distance in the sample for each market.

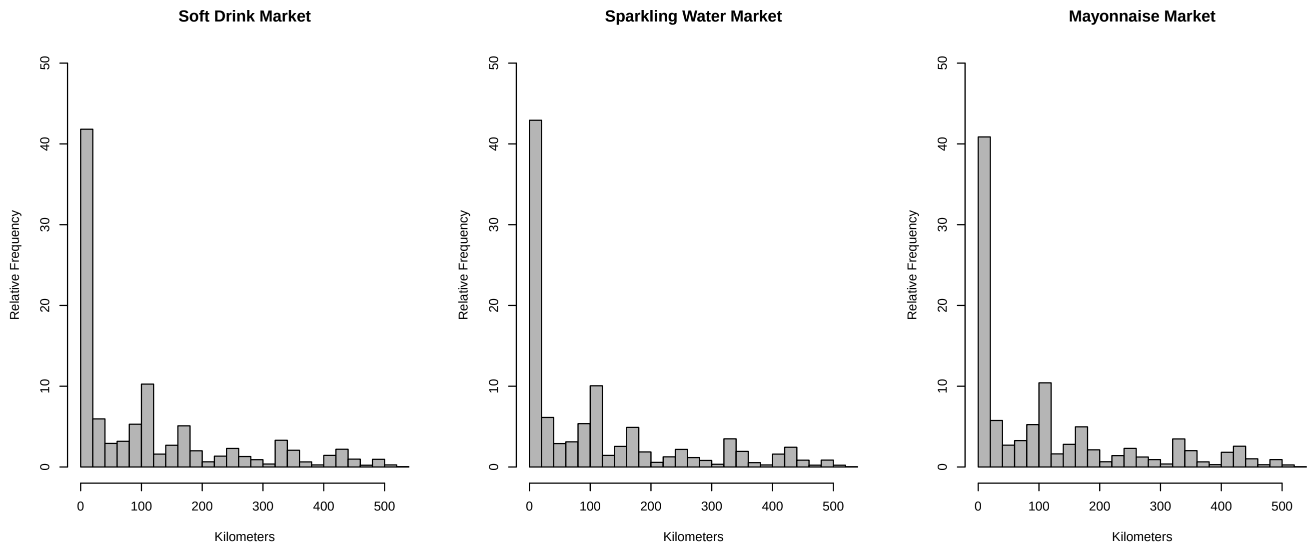


Figure 1.14: Observations by distance between cities in the sample for each market.

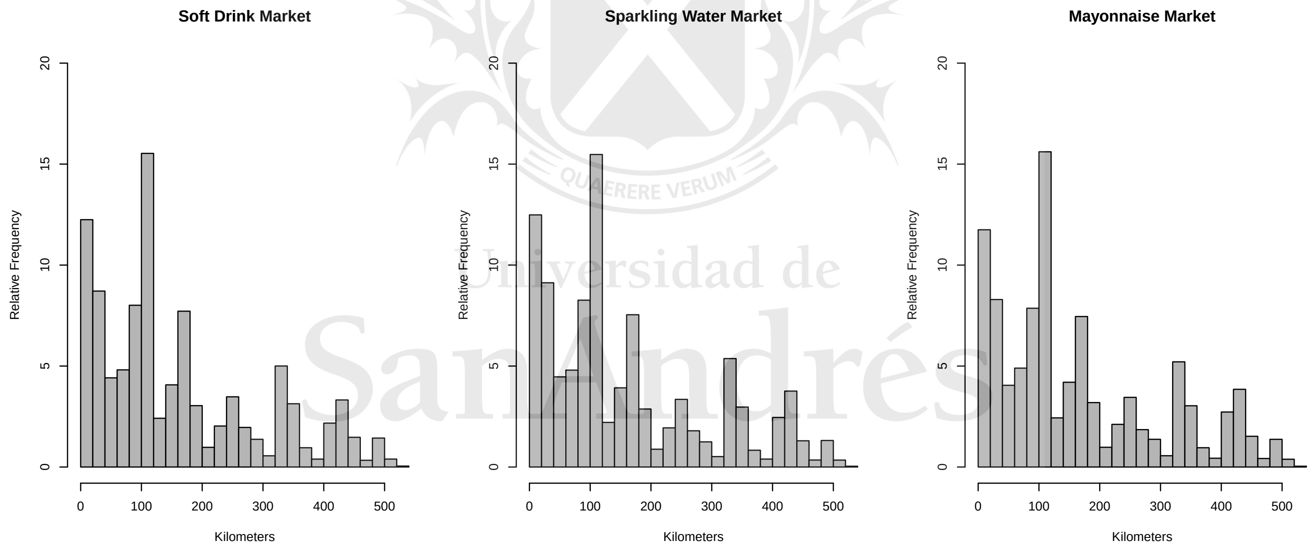
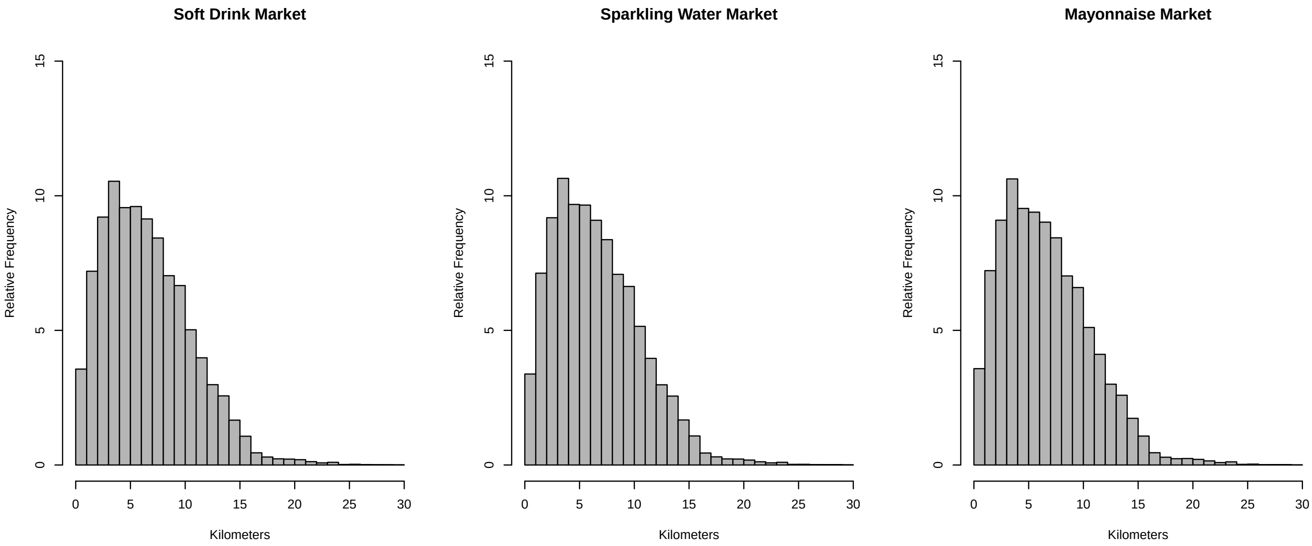


Figure 1.15: Observations by distance within cities in the sample for each market.



There do not seem to be differences in the distribution of prices by market. Nevertheless, there are differences in the distribution of prices based on whether a pair of stores are in the same or different cities. Within cities, nearly 90% of price observations are between 0 and 15 kilometers. Between cities there is not a clear pattern for the distribution of prices.

1.4 Estimation Strategy

The methodology for estimating the border effect and transport costs is standard in the literature. Engel and Rogers (1996) estimated the following equation:

$$|p_{ist} - p_{irt}| = \alpha_i + \alpha_{ch} + \alpha_t + \beta_1 \times Dist_{sr} + \beta_2 \times City_{sr} + \varepsilon_{isrt}, \tag{1.16}$$

where i is the indexed product and $i \in I$ is the product space; s, r are two stores, where s, $r \in S$ is the store's space in the sample and $s \neq r$; $|p_{ist} - p_{irt}|$ is the (absolute) difference of the logs of the price of good i between stores s, r at moment t;[19] α_i is a dummy variable for product i; α_{ch} is a dummy variable that takes the value one if stores s, r belong to the same chain; α_t is a time dummy; $Dist_{sr}$ measures the actual distance in (logs of) kilometers between stores s, r;[20] $City_{sr}$ is a dummy variable that takes the value one if stores s, r are located in different cities; and ε_{isrt} is a stochastic error term. In a second estimation, I add an interaction term for distance and border to the previous equation in order to control for nonlinear effects of the border parameter (see Borraz, Cavallo, Rigobon, and Zipitría (2016) for details):

$$|p_{ist} - p_{irt}| = \alpha_i + \alpha_{ch} + \alpha_t + \beta_1 \times Dist_{sr} + \beta_2 \times City_{sr} + \beta_3 \times Dist_{sr} \times City_{sr} + \varepsilon_{isrt}, \tag{1.17}$$

[19]The literature also studies the standard deviation of the price difference.

[20]As some distance are less than one kilometer and I want to avoid negative distance, I actually add 1 to the distance in kilometers.

where the interaction term between distance and border ($Dist_{sr} \times City_{sr}$) is due to the fact that, according to table 1.5, the median distances between and within cities are very different.

My analysis proposes a simple modification of equation 1.17. The database has data for each good sold in each store for each month, so I compute a binary variable that takes the value one if a local -smaller- competitor is present at one or both stores. In the analysis, the local brand will be Nix in the Soft Drink market, Nativa in the Sparkling Water market, and Uruguay in the Mayonnaise market. This simple strategy makes it possible to introduce the competitive effect previously established in Section 1.2. Now equation 1.17 is:

$$|p_{ist} - p_{irt}| = \alpha_i + \alpha_{ch} + \alpha_t + \beta_1 \times Dist_{sr} + \beta_2 \times City_{sr} + \beta_3 \times Dist_{sr} \times City_{sr} +$$

$$\underbrace{\alpha_1 \times OneLocal + \beta_4 \times OneLocal \times City_{sr}}_{\text{One store has a local competitor}} + \underbrace{\alpha_2 \times BothLocal + \beta_5 \times BothLocal \times City_{sr}}_{\text{Both stores have a local competitor}} + \varepsilon_{isrt}\,, \tag{1.18}$$

where $OneLocal$ takes the value one –at time t– if either store (s, t) sold the local brand, and $BothLocal$ takes the value of one if –at time t– both stores (s, t) sold the local brand. As Table 1.5 shows, the median distances between and within cities are very different. Therefore, in the empirical estimation I will interact the border dummy and each local dummy with the border parameter.

Table 1.6 shows the results for the estimation of equations 1.16, 1.17, and 1.18 for each of the three markets.[21] I exclude the local goods in the estimation, as they would be dropped if the analysis were to be made between different countries. The brands in the estimations include Coke and Pepsi for the Soft Drink market, Salus and Matutina for the Sparkling Water market, and Hellmans and Fanacoa for the Mayonnaise market.

In equations 1.17 and 1.18 the border parameter is interacted with distance, so a benchmark must be established to calculate the –distance equivalent– size of the border.[22] In the analysis that follows, I set 29 kilometers –the maximum distance between two stores within a city, see Table 1.5– as the benchmark for calculating the border size. To estimate it –and the local competition effect size– the following calculations are performed. For equation 1.17 the size of the border is calculated as $\beta_1 \times \ln(x+1) = \beta_2 + (\beta_3 - \beta_1) \times \ln(29+1) \Rightarrow x = distance = \exp\left(\frac{\beta_2 + (\beta_3-\beta_1)\times\ln(29+1)}{\beta_1}\right) - 1$. For equation 1.18 I perform several calculations to calculate border and local competition effects. First, the size of the local competition effect when there is a local competitor in one store is $\beta_1 \times \ln(x+1) = \alpha_1 \Rightarrow x = distance = \exp\left(\frac{\alpha_1}{\beta_1}\right) - 1$, while if there are local competitors at both stores the size of the effect is $\beta_1 \times \ln(x+1) = \alpha_2 \Rightarrow x = distance = \exp\left(\frac{\alpha_2}{\beta_1}\right) - 1$. Second, I calculate the three border sizes:

- If there is no local competition: $\beta_1 \times \ln(x+1) = \beta_2 + (\beta_3 - \beta_1) \times \ln(29+1) \Rightarrow x = distance = \exp\left(\frac{\beta_2 + (\beta_3-\beta_1)\times\ln(29+1)}{\beta_1}\right) - 1$

[21] Price differences are multiplied by 100. Intercept and same chain dummies are omitted in all equations.

[22] Distance equivalent measures, either of the border or the local competition effect, will be referred to as the size of the variable.

Table 1.6: Estimation of distance and border effect for each market.

	Sparkling Water			Soft Drinks			Mayonnaise		
	Eq. 1.16	Eq. 1.17	Eq. 1.18	Eq. 1.16	Eq. 1.17	Eq. 1.18	Eq. 1.16	Eq. 1.17	Eq. 1.18
Distance	0.106***	0.388***	0.386***	2.267***	0.070***	0.044***	-0.103***	0.179***	0.035***
	(0.002)	(0.005)	(0.004)	(0.002)	(0.004)	(0.006)	(0.003)	(0.008)	(0.008)
Border	0.880***	1.608***	0.800***	-2.455***	-8.101***	-8.019***	0.456***	1.221***	1.450***
	(0.006)	(0.013)	(0.016)	(0.007)	(0.015)	(0.016)	(0.010)	(0.021)	(0.022)
Distance×Border		-0.324***	-0.269***		2.525***	2.562***		-0.326***	-0.193***
		(0.005)	(0.005)		(0.005)	(0.005)		(0.008)	(0.008)
One Local			0.193***			0.508***			1.922***
			(0.010)			(0.005)			(0.010)
One Local ×Border			0.305***			-0.482***			-0.847***
			(0.011)			(0.009)			(0.012)
Both Local			0.080***			-0.007			2.194***
			(0.010)			(0.009)			(0.014)
Both Local ×Border			1.106***			0.316***			-1.207***
			(0.012)			(0.015)			(0.017)
# Observations	7,551,864	7,551,864	7,551,864	5,993,911	5,993,911	5,993,911	6,587,432	6,587,432	6,587,432
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Product dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Same Chain Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R square	0.09	0.09	0.10	0.24	0.25	0.25	0.04	0.04	0.05

*** $p < 0.01$. Heteroscedastic standard errors in parentheses.

- If there is one local competitor at any store: $\beta_1 \times \ln(x+1) = \beta_2 + \beta_4 + (\beta_3 - \beta_1) \times \ln(29+1) \Rightarrow x = distance = \exp\left(\frac{\beta_2+\beta_4+(\beta_3-\beta_1)\times\ln(29+1)}{\beta_1}\right) - 1$
- If both stores have a local competitor: $\beta_1 \times \ln(x+1) = \beta_2 + \beta_5 + (\beta_3 - \beta_1) \times \ln(29+1) \Rightarrow x = distance = \exp\left(\frac{\beta_2+\beta_5+(\beta_3-\beta_1)\times\ln(29+1)}{\beta_1}\right) - 1$

The results of performing the previous calculations show differences in the estimation of the size of the border for each market. Nevertheless, in line with the theoretical model in Section 1.2, controlling for local competition results in corrections of the estimated size of the border effect. Border coefficients are statistically different from zero in all equations (positive for the Sparkling Water market and negative for the other two markets). At first glance, the data seem to show that some segmentation of geographical market exists in Uruguay for the markets under analysis. The Distance coefficient changes substantially in the three markets. In the Sparkling Water market, the Distance coefficient triples when adding controls to the Border coefficient either by nonlinearity in the border –equation 1.17– or by local competition –equation 1.18–. In the Soft Drink market, the Distance coefficient shrinks substantially from estimations of equations 1.16 to 1.17 and again from equation 1.18. In this market, the role of distance between stores is nearly zero in explaining price dispersion.[23] In the Mayonnaise market, the Distance coefficient has no economic meaning when estimated by equation 1.16, as it is negative, became positive when estimated by equation 1.17, and then decreases to a fifth when local competition is controlled –equation 1.18–. As in the Soft Drink market, it seems that distance between stores becomes less relevant in explaining price variation when local competition is controlled for. Now I proceed to a detailed analysis of the border and local competition effect in each market.

In the Sparkling Water market neither border nor local competition seems to play a role. Although the Border coefficient seems to be positive and significant in all estimations, once interacted with distance it becomes negative. The size of the border implied in the estimation of equation 1.17 is to decrease distance by one kilometer (i.e., if stores were located at 28 kilometers). The same result is found for the border size when controlling for local competitive conditions, as established in equation 1.18. The size of the border decreases from -0.88 to -1 kilometers when controlling for local competition –equation 1.18– if no local competitors are present, while if there is one competitor at a store (both stores) it decreases to -0.95 (increase to -0.57) kilometers. The effect of local competition is to add one kilometer (none) to price variation if at one store (both stores). As a result, neither borders nor local competition seems to be explaining price variations in the Sparkling Water market. The main source of price variation is the distance between stores.

The Soft Drinks market shows different results. As previously noted, the Distance parameter decreases substantially when adding controls to equation 1.16. This result translates into a larger size of the border estimation: 34 kilometers implied in the estimation of equation 1.17; while in the estimation of equation 1.18 it is 240 thousand kilometers if no local competition is available, and just 3 kilometers (316 millions) if one competitor is available at one (both)

[23]The maximum distance, 526 kilometers, adds a price variation of 0.44% in column (2) and just 0.27% in column (3).

stores. In this market, the border seems to play a relevant role, and its estimation is affected by local competitive conditions. Lastly, the size of the local competition effect is about one hundred thousand kilometers if one store has a local competitor, but zero if there are local competitors at both stores. This last result is due to the fact that the α_2 coefficient is zero in the estimation of equation 1.18.

The Mayonnaise market shows an even larger impact of borders and local competition on price dispersion. The border estimation implied in the estimation of equation 1.17 decreases the distance by one kilometer, while the estimation of equation 1.18 yields 234 million kilometers if no local competition is available, and the distance decrease by one kilometer if one or both stores sold a local brand. The size of the local competition effect is larger than 10 zeros of magnitude, so they are not reported here. While initially the border does not seems to play any role in the market, once the effect of local competition is controlled for, the conclusions of the analysis change substantially. If there are no competitors at the store, then the border effect drives price dispersion. If there are local competitors, they are the main explanation of price dispersion in the market, while the border effect is negligible.

The empirical analysis for the three markets shows support for the impact of local competition in explaining price dispersion, as in the model developed in Section 1.2. In turn, this effect is expressed in different impacts of the border effect on relative prices. The analysis also highlights differences in each market in the impact of each source of price dispersion: distance, local competition, and borders. In the Sparkling Water market, distance is the relevant source of dispersion. Neither local competition nor border are important to explaining relative prices. In the Soft Drink market, controlling for local competition increases the impact of the border in explaining price dispersion only when there are local competitors at both stores or no local competitors in any store. In the case when there is a local competitor at one store, the primary explanation for price dispersion is local competition. Finally, in the Mayonnaise market price dispersion is explained by the segmentation introduced by borders and by the effect of local competitors at the store.

When borders are relevant in explaining relative prices between stores, controlling for local competition makes possible to adjust its effects. The next section analyzes several robustness tests for the previous results.

1.4.1 Robustness

This section shows the results of several robustness tests for the main results. All tables are available upon request, and the results are summarized in Table 1.7. First, I estimate equations 1.16, 1.17, and 1.18 using other central measures (e.g., monthly average and median price) and the first day of the month. When summary measures are used, price differences could be the result of contrasting prices in different days of the month. I pick the first day to calculate price differences, as the probability of price change on that day is nine times higher than on any other day of the month (see Borraz and Zipitría (2012)).

In the Sparkling Water market, the results do not change when using alternative central measures. Borders do not seem to play a role in the market. In the Soft Drinks market, the results are similar to previously results, except that the border estimation when there is one local brand increases substantially. This result is mainly driven by a sharp decrease in the estimation of the Distance parameter, which in turn increases the size of the border and local competition effects. In the Mayonnaise market, the results found in the previous section also hold, with the exception of the estimation using the Average price. In this case, the Distance parameter is not statistically different from zero, so the size of the border and of the local competition effect could not be calculated.

Second, the results could be driven by some omitted variable underlying the analysis. In particular, Uruguay has a temperate climate. In summer months the demand for liquids rises with temperature. Also, in summer some cities in the country receive tourism either from neighboring countries or from the capital city Montevideo. In summer months prices are expected to increase in those cities that experience increases in demand. I account for this seasonality by splitting the sample in two: summer months (December, January, February, and March) and excluding summer months. The results for both sub samples, with the exception of the Soft Drink market and in those cases where the Distance parameter is statistically different from zero, remain similar to those previously discussed. For the Soft Drink market, the size of the border decreases in all cases if each store has a local.

Third, as shown in table 1.3, Montevideo (the capital city of Uruguay) accounts for nearly half of the supermarkets and observations in the sample. Thus, I run regressions 1.16, 1.17, and 1.18, adding a dummy that takes the value one if any supermarket is located in Montevideo. Now the estimations reported in Table 1.7 include border and a local competition effect except for Montevideo city. With the exception of the Soft Drink market, the results remain similar to those previously found. For the Soft Drink market, the size of the border decreases in all cases unless there is a local competitor in both stores.

This section has proposed several robustness tests for the results previously reported for the three markets studied. The results of all tests, when the Distance parameter was statistically different from zero, are in line with the results of the previous section. The following table sums up the results obtained in the robustness estimations.

Table 1.7: Distance and local competition estimations (in kilometers).

Sparkling Water Market

	Border (eq. 2)	**Border (eq. 3)**	**Border (one local)**	**Border (both local)**	**One Local**	**Both Local**
Main regression	-1	-1	-1	-1	1	0
Average price	-1	-1	-1	-1	1	0
Median price	-1	-1	-1	-1	1	0
Day 1	-1	-1	-1	-1	1	1
Excluding summer month	-1	-1	-1	-1	1	0
Summer month	-1	-1	-1	-1	0	0
Controlling for Montevideo	-1	-1	-1	-1	1	0

Soft Drinks Market

	Border (eq. 2)	**Border (eq. 3)**	**Border (one local)**	**Border (both local)**	**One Local**	**Both Local**
Main regression	38	large	3	large	large	0
Average price	27	large	large	large	large	0
Median price	33	large	large	large	large	0
Day 1	24	large	large	large	large	0
Excluding summer month	0	51	-1	large	large	0
Summer month	nd	nd	nd	nd	nd	nd
Controlling for Montevideo	0	68	-1	large	large	0

Mayonnaise Market

	Border (eq. 2)	**Border (eq. 3)**	**Border (one local)**	**Border (both local)**	**One Local**	**Both Local**
Main regression	-1	large	-1	-1	large	large
Average price	-1	nd	nd	nd	nd	nd
Median price	-1	large	-1	-1	large	large
Day 1	-1	large	-1	-1	large	large
Excluding summer month	-1	nd	nd	nd	nd	nd
Summer month	-1	143	-1	-1	large	large
Controlling for Montevideo	-1	nd	nd	nd	nd	nd

Note: *nd*, not defined (distance not significant); *large*, if distance is larger than 526 kilometers.

1.5 Conclusion

The literature has found that borders affect relative prices between countries, but also between states and even cities within a country. This paper claims that local competition affects the estimation of the border effect. I develop a stylized model that shows that the availability of local competitors not only affects the relative prices of products but also the estimation of the border parameter. The model allows a richer setting than the linear city model traditionally used in the literature on the border effect (see Gopinath, Gourinchas, Hsieh, and Li, 2011). The model shows that borders could either increase or decrease the positive effect of distance on relative prices. The sign of the effect depends on the direction of the demand shift due to the border. The introduction of local competition in the model, if not accounted for, induces a bias in the border estimation. As a result, the model shows that the interrelation between distance, local competition, and border is much richer than previously found.

Next, I develop a simple methodology to account for local competitive conditions in the estimation of the border effect. I test the predictions of the model using a detailed price database for all supermarkets in Uruguay and for three markets. No influence of the border was found for the Sparkling Water market. In the Soft Drink market, controlling for local competition increases the impact of the border in explaining price dispersion only when there are local competitors at both stores or no local competitors in any store. In the case when there is a local competitor at one store, the primary explanation for price dispersion is local competition. Finally, in the Mayonnaise market, price dispersion is explained by the segmentation introduced by borders and by the effect of local competitors at the store. The empirical analysis shows that when borders are relevant to explain relative prices between stores, controlling for local competition makes it possible to adjust their effects. The results are robust to different specifications of the variables (median, average, first day of the month), to controls for Montevideo city, or by splitting the sample to include or exclude summer months.

Bibliography

Borraz, F., A. Cavallo, R. Rigobon, and L. Zipitría (2016): "Distance and Political Boundaries: Estimating Border Effects under Inequality Constraints," *International Journal of Finance & Economics*, 21(1), 3–35.

Borraz, F., and L. Zipitría (2012): "Retail Price Setting in Uruguay," *Economia*, 12(2), 77–109.

Broda, C., and D. E. Weinstein (2008): "Understanding International Price Differences Using Barcode Data," *NBER Working Paper 14017*.

Ceglowski, J. (2003): "The Law of One Price: Intranational Evidence for Canada," *The Canadian Journal of Economics*, 36(2), 373.

Coşar, A. K., P. L. Grieco, and F. Tintelnot (2015a): "Bias in estimating border- and distance-related trade costs: Insights from an oligopoly model," *Economics Letters*, 126, 147 – 149.

——— (2015b): "Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry," *The Review of Economics and Statistics*, 97(3), 623–637.

Crucini, M. J., M. Shintani, and T. Tsuruga (2010): "The Law of One Price without the Border: The Role of Distance versus Sticky Prices," *The Economic Journal*, 120(544), 462–480.

Di Comite, F., J.-F. Thisse, and H. Vandenbussche (2014): "Verti-zontal differentiation in export markets," *Journal of International Economics*, 93(1), 50 – 66.

Engel, C., and J. H. Rogers (1996): "How Wide is the Border?," *American Economic Review*, 86(5), 1112–1125.

Evans, C. L. (2003): "The Economic Significance of National Border Effects," *American Economic Review*, 93(4), 1291–1312.

Gabszewicz, J. J., and X. Y. Wauthy (2012): "Nesting horizontal and vertical differentiation," *Regional Science and Urban Economics*, 42(6), 998 – 1002.

Goldberg, P. K., and M. M. Knetter (1997): "Goods Prices and Exchange Rates: What Have We Learned?," *Journal of Economic Literature*, 35(3), 1243–1272.

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li (2011): "International Prices, Costs, and Markup Differences," *American Economic Review*, 101(6), 2450–86.

Gorodnichenko, Y., and L. Tesar (2009): "Border Effect or Country Effect? Seattle may not be so far from Vancouver after all," *American Economic Journal – Macroeconomics*, 1(1), 219–241.

Head, K., and T. Mayer (2002): "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade," *CEPII Research Center, Working Paper 2002-01.*

Hillberry, R., and D. Hummels (2003): "Intranational Home Bias: Some Explanations," *The Review of Economics and Statistics*, 85(4), 1089–1092.

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

Klenow, P. J., and O. Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?," *The Quarterly Journal of Economics*, 123(3), 863–904.

Parsley, D. C., and S.-J. Wei (1996): "Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuations," *The Quarterly Journal of Economics*, 111(4), 1211–1236.

——— (2001): "Explaining The Border Effect: The Role of Exchange Ratee Variability, Shipping Costs, and Geography," *Journal of International Economics*, 55(1), 87–105.

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

Wolf, H. C. (2000): "Intranational Home Bias in Trade," *The Review of Economics and Statistics*, 82(4), 555–563.

Yazgan, M. E., and H. Yilmazkuday (2011): "Price-level convergence: New evidence from U.S. cities," *Economics Letters*, 110(2), 76 – 78.

Appendix

1.6 Non Equilibrium in Model of Section 1.2

From now on, I will assume that $\widetilde{x} = \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + t(\ell_1 + \ell_2)}{2t} = \widehat{x} = \frac{p_B - p_A + t(\ell_1 + \ell_2)}{2t}$, which implies that

$$p_C - p_A = \overline{\theta}\left(s_2 - s_1\right). \tag{1.19}$$

I proceed to find the demand for each brand, taking into account the previous results. Demand for firm A is: $D_A = (1-\lambda)\,\widehat{x} = (1-\lambda)\,\frac{p_B - p_A + t(\ell_1 + \ell_2)}{2t}$. Profits are $\Pi_A = p_A \times D_A$. The first order constraint of the problem is $\frac{\partial \Pi_A}{\partial p_A} = 0 = \frac{(1-\lambda)}{2t}\left[p_B - 2p_A + t\left(\ell_1 + \ell_2\right)\right]$, so the reaction function is

$$p_A = \frac{p_B + t\left(\ell_1 + \ell_2\right)}{2}. \tag{1.20}$$

Note that the reaction function of firm A depends –increasingly– only on the price of firm B. This result holds by construction, because implicit in the model is that $p_A = p_C - \overline{\theta}\left(s_2 - s_1\right)$.

For firm B, as I have assumed that $\widetilde{x} = \widehat{x}$, the demand for the firm is the same as if firm C was not on the market, that is, $D_B = L - \widehat{x} = \frac{p_A - p_B + t(2L - \ell_1 - \ell_2)}{2t}$. The reaction function is the same as before:

$$p_B = \frac{p_A + t\left(2L - \ell_1 - \ell_2\right)}{2}. \tag{1.21}$$

Again, the reaction function does not depend on the price of firm C explicitly, but will impact indirectly through the price of firm A.

Lastly, the demand for firm C is $D_C = \lambda \widetilde{x} = \lambda \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + t(\ell_1 + \ell_2)}{2t}$. Profits are $\Pi_C = p_C \times \left[\lambda \frac{p_B - p_C + \overline{\theta}(s_2 - s_1) + t(\ell_1 + \ell_2)}{2t}\right]$. The first order constraint is $\frac{\partial \Pi_C}{\partial p_C} = 0 = \frac{\lambda}{2t}\left(p_B - 2p_C + \overline{\theta}\left(s_2 - s_1\right) + t\left(\ell_1 + \ell_2\right)\right)$. The reaction function for firm C is

$$p_C = \frac{p_B + \overline{\theta}\left(s_2 - s_1\right) + t\left(\ell_1 + \ell_2\right)}{2}. \tag{1.22}$$

I now have a system of four equations (1.19, 1.20, 1.21, 1.22) with three unknowns (p_A, p_B, p_C). The model does not have an equilibrium.

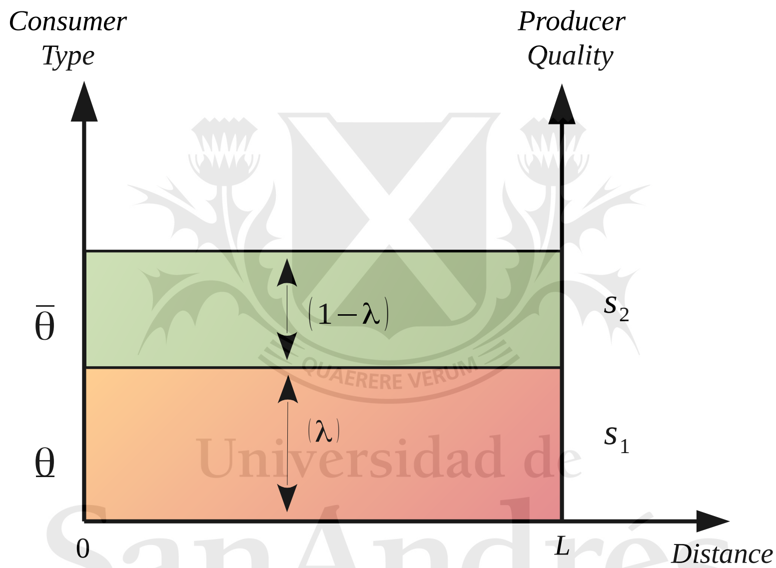
Proposition 1.6. *The model do not have an equilibrium.*

Proof. As equations 1.20 and 1.21 are the same as if product C does not exist, then the solution $p_A = \frac{t(2L + \ell_1 + \ell_2)}{3}$ and $p_B = \frac{t(4L - \ell_1 - \ell_2)}{3}$ is also a solution for those equations. Now we substitute $p_B = \frac{t(4L - \ell_1 - \ell_2)}{3}$ in equation 1.22 and find $p_C = \frac{2t(2L - \ell_1 - \ell_2) + 3\overline{\theta}(s_2 - s_1)}{6}$. Substituting p_A in equation 1.19 we find $p_C^{'} = \frac{t(2L - \ell_1 - \ell_2) + 3\overline{\theta}(s_2 - s_1)}{3}$. Then $p_C^{'} \neq p_C$. □

1.7 The Model with Two High-Quality and One Low-Quality Brands

Assume that the setup is reversed, and two high qualities compete with one low quality. This section shows that the results of Section 1.2.1 remain unchanged. I made just one change in order to make the analysis easier. Now I assume that there are two high-quality brands (A located at 0, and B located at L). The proportion of high-quality consumers is $(1-\lambda)$ instead of λ. This assumption guarantees that the demands remain the same as in Section 1.2.1. The next figure depicts the new setup.

Figure 1.16: The model with two type of consumers.



First, note that the equilibrium with two identical quality brands remains unchanged. Equation 1.17 is now $r_{ij}+\overline{\theta}s_2-t\widehat{x}-p_A=r_{ij}+\overline{\theta}s_2-t\left|L-\widehat{x}\right|-p_B$, and solving for $\widehat{x}$ we obtain $\widehat{x}=\frac{p_B-p_A+tL}{2t}$, as before. So, the equilibrium prices do not change: $p_A=p_B=tL$.

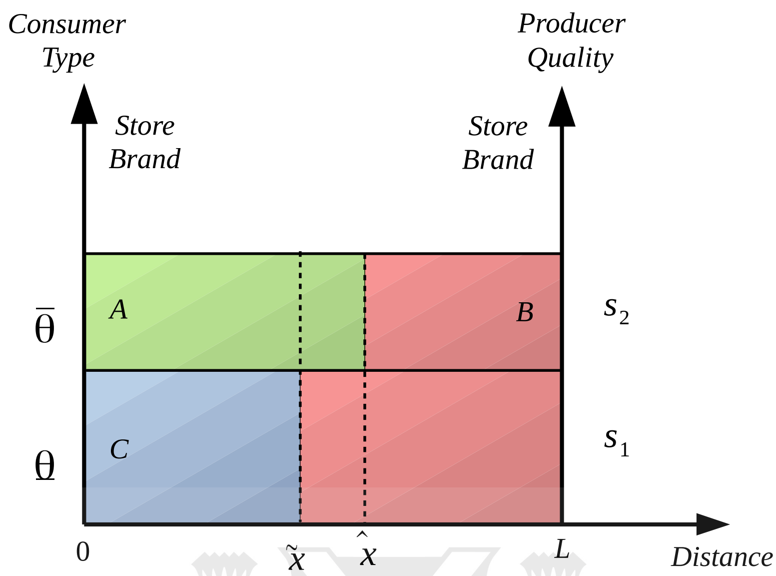
Now I add a low-quality brand at 0. The following figure shows the new setting.

The analysis is similar, but now $\widehat{x}>\widetilde{x}$. The indifference between low-quality brand C and high-quality brand B is now $r_{ij}+\underline{\theta}s_1-t\widetilde{x}-p_C=r_{ij}+\underline{\theta}s_2-t\left|L-\widetilde{x}\right|-p_B$ and $\widetilde{x}=\frac{p_B-p_C-\underline{\theta}(s_2-s_1)+tL}{2t}$. Now $\widehat{x}>\widetilde{x}$ implies that $p_A-p_C<\underline{\theta}\left(s_2-s_1\right)$, and the same analysis of Appendix 1.6 holds.

As I reverse λ, the profits remain unchanged; e.g., $D_A=(1-\lambda)\,\widehat{x}$ and the reaction function $p_A=\frac{p_B+tl}{2}$; $D_B=(1-\lambda)\left(L-\widehat{x}\right)+\lambda\left(L-\widetilde{x}\right)$ and the reaction function is now $p_B=\frac{(1-\lambda)p_A+\lambda p_C+\lambda\underline{\theta}(s_2-s_1)+tL}{2}$; and $D_C=\lambda\widetilde{x}$, with reaction function $p_C=\frac{p_B-\underline{\theta}(s_2-s_1)+tL}{2}$. The equilibrium prices are now:

$$p_A=tL-\frac{\lambda\underline{\theta}\left(s_2-s_1\right)}{6},$$

Figure 1.17: Possible equilibrium values of $\tilde{x}$ and $\hat{x}$. The demand for store A is depicted in blue, the demand for store B in red, and the demand for store C in green.



$$p_B = tL - \frac{\lambda \underline{\theta}\,(s_2 - s_1)}{3},$$

$$p_C = tl - \frac{(3+\lambda)\,\underline{\theta}\,(s_2 - s_1)}{6}.$$

Price differences remains unchanged: $p_A - p_B = \frac{\lambda \underline{\theta}(s_2 - s_1)}{6}$. The border analysis remains unchanged, with the previous adjustment.

1.8 Chain Description

Table 1.8: Chain description.

Chain	Montevideo	# Cities	Soft Drinks			Sparkling Water			Mayonnaise		
			Coke	Pepsi	Nix	Salus	Matutina	Nativa	Hellman's	Fanacoa	Uruguay
Devoto	Yes	6	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Disco	Yes	5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
El Clon	Yes	5	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
El Dorado	No	20	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Frigo	Yes	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Géant	Yes	2	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Iberpark	Yes	2	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
La Colonial	Yes	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Los Jardines	Yes	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macromercado	Yes	3	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Micro Macro	Yes	4	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
MultiAhorro	Yes	8	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
None	Yes	27	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Red Market	Yes	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Super XXI	No	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Super Star	No	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TATA	Yes	25	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tienda Inglesa	Yes	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Ubesur	Yes	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: author's calculations based on data from the General Directorate of Commerce.

Chapter 2

Distance and Political Boundaries: Estimating Border Effects under Inequality Constraints

Abstract

The "border effects" literature finds that political boundaries have a large impact on relative prices across locations. In this paper we show that the standard empirical specification suffers from selection bias, and propose a new methodology based on binned-quantile regressions. We use a novel micro-price dataset from Uruguay and focus on city borders. We find that when the standard methodology is used, two supermarkets separated by 10 kilometers across two different cities have the same price dispersion as two supermarkets separated by 30 kilometers within the same city, implying that crossing a city border is equivalent to tripling the distance. By contrast, when upper quantiles are used the city border effect disappears. These findings imply that transport cost have been systematically underestimated by the previous literature. Our methodology can be applied to measure any kind of border effect. [1]

2.1 Introduction

Political borders can have a significant impact on relative prices. The degree of price segmentation caused by such boundaries was empirically documented in a seminal paper by Engel and Rogers (1996), who showed with CPI data that the US–Canadian border had an effect on price dispersion equivalent to adding a distance of at least 1,780 miles between locations (approximately the distance between Miami and Quebec). Their work spurred a large literature that found similarly large "border effects" across countries, states, and even cities.[2] These results have been heavily debated over the years. Some papers have argued that (i) the distances have been mis-measured (see Head and Mayer (2002)), (ii) the regressions suffer from aggregation bias (see Evans (2003) and Broda and Weinstein, 2008), (iii)

[1] This chapter is a revised version of the paper written with Fernando Borraz, Alberto Cavallo and Roberto Rigobon and published as "Distance and Political Boundaries: Estimating Border Effects under Inequality Constraints", International Journal of Finance and Economics 21(1) (2016), pp. 3-35.

[2] For example, see Parsley and Wei (2001) for results between the US and Japan and Ceglowski (2003) for the effects of provincial borders in Canada.

the gravity equation implied in the standard specification has been misspecified (see Anderson and van Wincoop, 2003 and Hillberry and Hummels, 2003), and that (iv) the regressions do not have a proper benchmark due to the fact that country distributions of prices are very different across countries (see Gorodnichenko and Tesar, 2009). Despite all this work, the magnitude and reasons behind the segmentation introduced by political borders is still an open question in the literature.

In this paper we propose a simple method to estimate the size of the "border effect" based on Samuelson's iceberg cost model. This methodology imply that largest price differences observed between locations are relevant for transport cost estimation. We first argue that the standard regression is based on an arbitrage inequality constraint, and that using all price observations creates a selection bias that affects both the distance and border coefficients (and therefore the estimates of the "border effect"). We propose an alternative approach based on quantile regressions that corrects for the selection bias while simultaneously controlling for potential measurement errors.

Our method can be explained using a very simple framework along the lines of Engel and Rogers (1996). Consider the problem of a firm that sets a price bounded by the existence of an arbitrage constraint. If the arbitrage cost between two establishments (i and j) is τ, and p denotes the log price in each location, then the arbitrage constraint can be expressed as a simple inequality:

$$|p_i - p_j| \leq \tau \tag{2.1}$$

where τ is a function of distance, political boundaries, and other regional and product characteristics. The literature typically estimates τ and the border effects by running the following regression on price dispersion:[3]

$$|p_{i,t} - p_{j,t}| = \alpha + \beta D_{i,j} + \gamma B_{i,j} + \delta X_{i,j,t} + \epsilon_{i,j,t} \tag{2.2}$$

where $p_{i,t} - p_{j,t}$ is the log price difference between locations i and j at time t. The locations can be countries, provinces, cities or establishments. $D_{i,j}$ is the distance between the two locations, $B_{i,j}$ is a dummy that takes value 1 if a border exists between locations i and j, and $X_{i,j,t}$ is a series of additional controls. In this context, the "border effect" is the equivalent number of miles that would produce the same dispersion as the estimated border dummy coefficient γ. In its simplest form, it is the ratio γ/β, which means that a bias in either (or both) of these coefficients will have an impact on the estimate of the border effect.

We argue that τ and its determinants cannot be estimated through a simple OLS regression because prices in the two locations are an optimal choice subject to an inequality constraint that is not necessarily binding. If the optimal prices of the two stores lie within the constraint, then their difference is smaller than τ and these observations are not relevant to estimate the arbitrage costs. To illustrate this, consider two markets that are highly segmented but have identical supply and demand characteristics. Goods will have the same price across the two locations, but this price gap tells us nothing about the arbitrage costs or the degree of segmentation between the markets. In fact, all observations within the no-arbitrage range suffer from selection bias, and estimates that use the mean or the standard deviation of $|p_1 - p_2|$ will be biased downward as well.

The arbitrage cost τ is better estimated when we use only the largest observed price differences

[3] A common alternative specification used by papers such as Engel and Rogers (1996) has the standard deviation $\sigma(p_{i,t} - p_{j,t})$ instead. In both cases, the objective is to measure the effect of the righ-hand side variables on price dispersion, which can be done either through the mean of the absolute value or the standard deviation of the price differences. Our results do not change if we use the standard deviation. See Broda and Weinstein (2008) for an overview of the papers that use these two regressions in the literature.

between locations. Those are the observations that provide information about the limit that arbitrage imposes of the magnitude of price dispersion.[4] Ideally we would like to use the maximum observed price gap between locations, but it is potentially sensitive to measurement errors.[5] Instead, we estimate a series of binned-quantile regressions that allow us to measure the sensitivity of our estimates to the errors-in-variables. We start with the mean price gap between locations (equivalent to the method typically used in the literature), and then use only the observations in the 80th, 90th, 95th, 99th percentiles, and the maximum observed price difference.

We apply this method to study the impact of city borders on price dispersion in Uruguay. We use a novel good-level dataset composed by daily prices from 202 UPC-level products sold in 333 supermarkets across 47 cities collected between 2007 and 2010. When we first estimate the border effect using standard methods, we find that the city border between two stores separated by 10 kilometers is larger than 20 kilometers wide, and statistically different from zero. This implies that the border triples the distance of stores across the city borders. However when we re-estimate using distance-binned quantile regressions, the border declines until it is not significantly different from zero. As expected from our discussion both the distance and border dummy coefficients are downward biased in the standard regression, but the bias is largest on the distance parameter.[6] As a result, the net impact is that the implied border effect (in kilometers) falls.

We perform robustness tests to correct for outliers, product mix, and we change the specification to include non-linearity and interaction terms. In all of them, the city-border effect measured in kilometers tends to disappear when higher percentiles are used. Furthermore, the results are similar at the 99th, 99.5th, 99.9th percentile, and the maximum, suggesting that the estimates are not significantly affected by potential errors in the data.

Our approach and the nature of the data address four additional sources of concerns that have been raised since the original Engel-Rogers regressions. First, we use product-level data with identical goods across all locations. As suggested by Goldberg and Knetter (1997), product-level data is crucial to understand deviations from the Law of One Price (LOP). Indeed, Evans (2003) and Broda and Weinstein (2008) argue that a significant problem in the border effect literature is the aggregation bias induced by price indexes. Second, we use retail prices. Hillberry and Hummels (2003) have argued that business-to-business data tends to overestimate trade flows and to underestimate price differences within countries. Third, we know the exact location of each store. As pointed out by Head and Mayer (2002), using approximate distances (such as from one country capital to another) can greatly overestimate the border effect. Finally, all the stores in our sample sell the same set of products. As Evans (2003) points out, the mix of products sold across borders can lead to a bias in the standard regressions.

Compared to recent papers in the literature, our results are consistent with Gorodnichenko and Tesar (2009), who argue that with "cross-country heterogeneity in the distribution of within-country

[4]The estimation problem is analogous to estimating using inequality moments as opposed to equality moments. This area has received significant attention recently. See for example Andrews, Berry, and Jia (2004), Andrews and Guggenberger (2009), Andrews and Soares (2010), Andrews and Shi (2014), Ponomareva and Tamer (2011), and Rosen (2008).

[5]A related idea in the context of trade can be found in Eaton and Kortum (2002) who propose estimating trade friction using the maximum price difference. Simonovska and Waugh (2014) criticizes the use of the maximum price difference in the estimating strategies, based on the possibility of bias of the estimator on finite samples.

[6]The reason is that price gaps within the arbitrage constraint are less common for observations across cities, and therefore the border coefficient is less affected by the selection bias. Within cities, by contrast, small price gaps are very frequent and can greatly bias the distance coefficient.

price differentials there is no clear benchmark from which to gauge the effect of the border". We agree with this statement, but we show that even in the absence of a structural model it is still possible to obtain a simple and reliable estimate for the magnitude of the border effect using quantile regressions. Our paper is also complementary to the work of Gopinath, Gourinchas, Hsieh, and Li (2011) who estimate the border effect by studying the response of average prices in one market to cost shocks in another market. An advantage of our approach is that it does not require any cost data.

2.2 Methodology

In this section we present a simple model of price-setting across locations that provides the inequality we use to estimate arbitrage costs and the border effect. In particular, we propose a model where the firms' pricing decision is constrained by the ability of the consumer to arbitrage the price gap between two locations. Standard OLS regressions used in the literature consider all pairs of prices, including those that lie within the arbitrage constraint, which introduces a bias in the estimates of factors that affect the cost to arbitrage, such as distance and political borders. We propose an alternative methodology that focuses on the largest observed price differences between locations using binned-quantile regressions.

2.2.1 A simple model of price-setting with arbitrage

2.2.1.1 Consumers

Consider an economy with a mass of consumers uniformly distributed along a line. This line encompasses two cities (A, B) of equal distance. There are J stores in the economy, J_A stores in city A and J_B stores in city B. There is also a "border" between A and B, in the sense that consumers pay a cost whenever they cross to another city. This border cost may arise due to differences in taxes, convenience in shopping hours, and other characteristics associated with the city but not driven by distance. A consumer located on point ℓ and buying in store i has an indirect utility function represented by

$$u_\ell(i) = v - \theta p_i - \tilde{\beta} \mid \ell - \ell_i \mid - \tilde{\gamma} b_i - \tilde{\delta} I_i \tag{2.3}$$

where v is the reservation price of the consumer, and θ captures how sensitive the consumer is to prices. The rest of the parameters measure transaction costs: $\tilde{\beta}$ measures unit transportation costs, multiplied by the distance between the consumer location (ℓ) and the store position (ℓ_i) (including information costs about the store, such as knowing the distance to the store or its prices); b_i is a dummy that takes value 1 if the consumer and the store are in different cities; and I_i measures additional store-specific costs not related with distance, such as learning the layout and sale events of a given store.

Since the consumer buys the one item that maximizes his utility, we can compare the price each consumer pays across all possible pairs of stores.[7] The consumer ℓ weakly prefers store i to store j if $u_\ell(i) \geq u_\ell(j)$, for each i, $j \in J = (J_A + J_B)$, $i \neq j$. For simplicity assume the price elasticity and the transportation cost are symmetric in all locations. This implies that:

$$v - \theta p_i - \tilde{\beta} \mid \ell - \ell_i \mid - \tilde{\gamma} b_i - \tilde{\delta} I_i \geq v - \theta p_j - \tilde{\beta} \mid \ell - \ell_j \mid - \tilde{\gamma} b_j - \tilde{\delta} I_j \tag{2.4}$$

[7] We require that v is large enough so that is $u_\ell(i)$ is positive in at least one store.

Rearranging terms we obtain:

$$p_i - p_j \leq \frac{\tilde{\beta}}{\theta}(\mid \ell - \ell_j \mid - \mid \ell - \ell_i \mid) + \frac{\tilde{\gamma}}{\theta} \vartriangle b_{i,j} + \frac{\tilde{\delta}}{\theta} \vartriangle I_{i,j} \tag{2.5}$$

where $\vartriangle b_{ij}$ is equal to 1 if both stores are located in different cities and 0 otherwise, and $\vartriangle I_{i,j}$ measures the incremental information cost incurred by changing the store. Thus for each pair of stores the consumption decision can be expressed as the result of inequality (2.6). The value of the distance terms depends on which store is further away from the consumer. If the difference between $\mid \ell - \ell_j \mid - \mid \ell - \ell_i \mid$ is negative, the price difference could simply be defined as $(p_j - p_i)$. Therefore, the expression is simplified to the absolute difference of the location between stores:

$$\mid p_i - p_j \mid \leq \frac{\tilde{\beta}}{\theta} |\ell_i - \ell_j| + \frac{\tilde{\gamma}}{\theta} \vartriangle b_{i,j} + \frac{\tilde{\delta}}{\theta} \vartriangle I_{i,j} \tag{2.6}$$

Comparing across different pairs, if the distance increases, a border exists between the stores, or there is a positive cost of switching stores, the level of price dispersion rises. The opposite occurs if consumers are more sensitive to prices. It can be shown that the price space is not empty and that the inequality constraint is only binding for the marginal consumer.[8] Where the marginal consumer is traditionally defined as indifferent between buying in two different stores. This implies that the rest of the consumers are not indifferent between two stores and always prefer to buy on a particular one. In the end, the marginal consumer is the one for which the inequality is binding, and defines the demand for each store.

2.2.1.2 Producers

Assume there are J_A and J_B identical producers (or stores) in each city that sell an identical good at price p_j, where $j \in J = (J_A + J_B)$. Each producer maximizes profits, given the prices of the other stores and subject to the participation constraint of the consumers. Suppose all producers, except for j, are in equilibrium. Then firm j sets its price subject to the participation constraint of consumer ℓ:

$$\underset{p_j}{max} \prod_j (p_j/p_{-j})$$

$$st \quad p_j \in \Re_+$$

and to the other $J-1$ consumer constraints

$$\mid p_i - p_j \mid \leq \frac{\tilde{\beta}}{\theta} |\ell_i - \ell_j| + \frac{\tilde{\gamma}}{\theta} \vartriangle b_{ij} + \frac{\tilde{\delta}}{\theta} \vartriangle I_{i,j},\ \forall i \in J,\, i \neq j$$

where this condition applies to all the J firms in the sample.

Firms maximize profits when setting the maximum price for the marginal consumer, as shown in Appendix 2.7. This in turn implies that the inequality will not be binding for the rest of the consumers. Therefore comparing pairs of prices for all stores to estimate equation 2.6 will not result in the correct measure of the consumers' parameters, as only those where the restriction is binding are valid.

[8]See the proof in Appendix 2.7

2.2.2 Binned-Quantile Regressions

This model of inequality constraints provides an equation that can be estimated as any other regression in the literature of border effects.[9] In this case, the specification can be defined as follows

$$| p_{i,t} - p_{j,t} | \leq \beta D_{i,j} + \gamma B_{i,j} + \delta X_{i,j,t} + \varepsilon_{i,j,t} \tag{2.7}$$

where $\beta D_{i,j} \equiv \frac{\tilde{\beta}}{\theta} |\ell_i - \ell_j|$, $\gamma B_{i,j} \equiv \frac{\tilde{\gamma}}{\theta} \triangle b_{i,j}$, and $\delta X_{i,j,t} \equiv \frac{\tilde{\delta}}{\theta} \triangle I_{i,j}$.

Notice that this inequality implies that all the residuals ($\epsilon_{i,j,t}$) in equation 2.7 are either zero or negative, in which case $E\left[\epsilon_{i,j,t}\right] \leq 0$. When this happens the estimation by OLS is expected to produce biased estimates due to the failure of the orthogonality conditions, where the bias is downward. There is only one case in which the estimates remain unbiased, and is if the price deviations are exactly equal to the arbitrage cost –i.e. the constraint is always binding. The residuals are identical to zero and OLS produces unbiased estimates. Intuitively prices are assumed free of errors-in-variables, so that the extreme in the distribution of price differences is the closest estimator to the arbitrage costs. It is important to mention that if all prices are optimally chosen to lie within the no-arbitrage region then even the estimation using the extreme of the price distribution will produce downward biased estimates. However by construction the biases will be smaller. Formally, the expected value of the errors at the medium are more negative than the errors at the 95th quantile: $E\left[\epsilon_{i,j,t}|50th\right] \leq E\left[\epsilon_{i,j,t}|95th\right] \leq 0$. Nevertheless, as pointed out by Simonovska and Waugh (2014) the maximum price difference could introduce additional biases, so we use instead the upper quintiles for our estimation.[10]

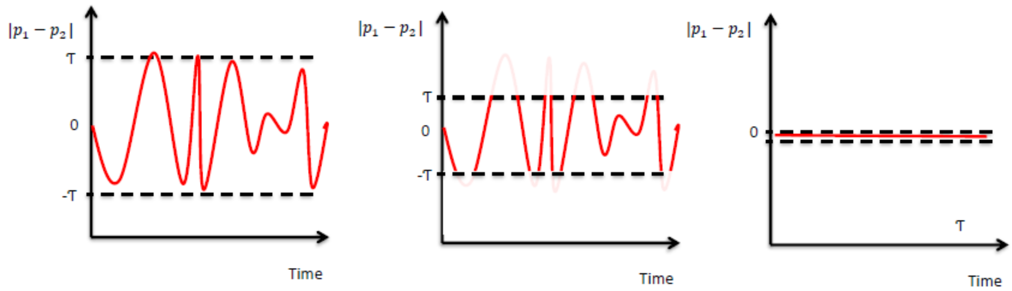
Figure 2.1 illustrates how the observed price dispersion may not be informative of the arbitrage cost. Panel (a) in Figure 2.1 shows hypothetical prices over time of one product in two locations. If the no-arbitrage condition is binding, then as the arbitrage cost τ decreases, so does the price dispersion. However if the condition is not binding, as shown in Panel (b), distributional statistics such as the mean or the standard deviation will not be associated with the estimate of τ. In both cases, however, we would be able to obtain good estimates of τ by using only the maximum observed price difference.

[9]The results from our model are also related to Samuelson (1954). See Appendix 2.7.4.

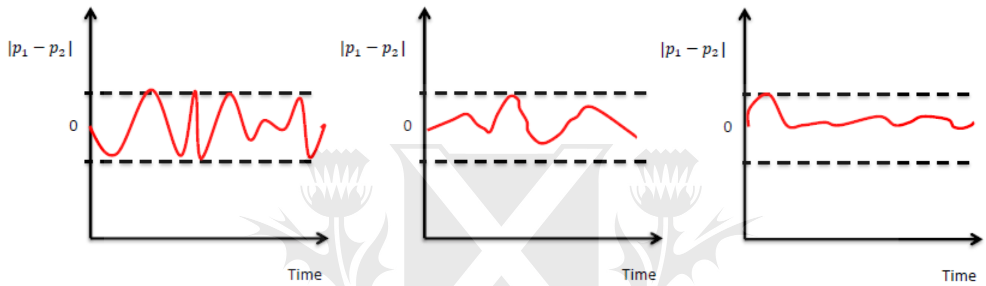
[10]Note that Simonovska and Waugh (2014) argues that the estimation of the transport costs can be downward biased if the maximum price difference is used, but we find a monotonic increase in this parameter as we use move from lower quintiles to the maximum price difference.

Figure 2.1: Price dispersion and the arbitrage cost

(a) Binding no-arbitrage condition



(b) Non binding no-arbitrage condition



Note: Panel (a) shows cases where the no-arbitrage condition constrains the price dispersion. Panel (b) cases where the price dispersion is not correlated with the arbitrage cost.

Figure 2.2 makes a similar point with real data. We plot the distribution of price differences for all goods between two locations of a given retailer, and compare the results for stores located at 1 km, 10 km (same city), 10 km (different cities), and 20 km (different cities) of each other. As expected, as the distance increases the share of price gaps at 0% falls (see table), and the mass between 1% and 20% increases. Interestingly, when we compare the two pairs of stores located at 10 km from each other (one of which is for stores in different cities), we find that crossing the city border has an effect on the mean and 90th percentile. The two pairs, however, have exactly the same gaps at the 99th and 99.9th percentiles of the distribution. This last result is consistent with the idea that city borders should not affect the cost to arbitrage across locations. Using the mean and lower percentiles of the price gaps that lie within the arbitrage constraint can therefore lead to biased results.

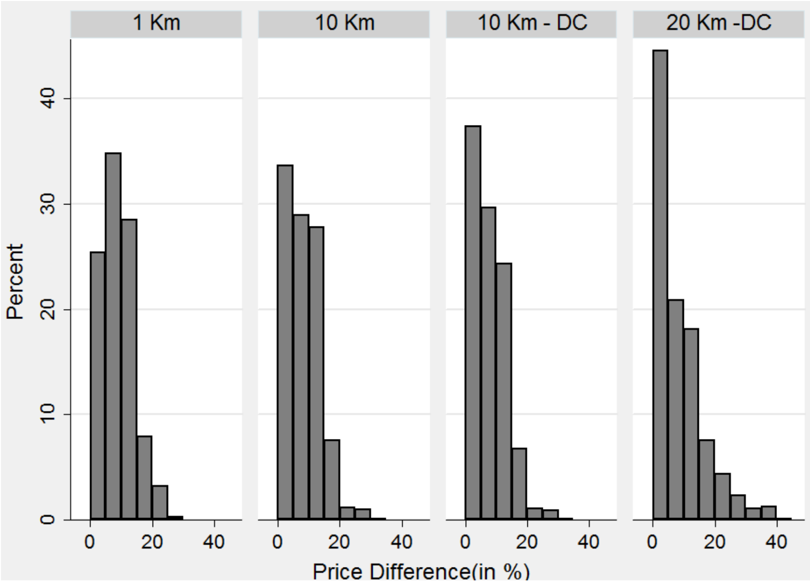


Figure 2.2: Example of Price Gaps in Different Store Pairs

Distance (Km)	Same City	Share at 0% (in %)	Price Gaps (in %) Mean	90th	99th	99.9th
1	yes	96.5	0.3	0	11.1	20.5
10	yes	76.6	1.9	8.64	16.9	25.1
10	no	73.6	2.0	9.46	16.9	25.1
20	no	69.3	2.6	10.6	26.9	39.3

Note: We calculate the price gaps (in absolute value) for all goods sold in a single retailer across two locations. We picked a random store from the largest retailer in the country and compared its prices to those of other stores from the same retailer located at 1 km, 10 km (same city), 10 km (different city), and 20 km (different city). Where "DC" denotes pairs in different cities. The graph excludes the mass at 0% to facilitate the comparison of positive gaps. The table shows the distributional statistics for all price gaps, including those at 0% (identical prices).

In order to address this selection bias, we propose a new method to estimate border effects using distance-binned quantile regressions. The methodology can be described as follows:

First, compute the absolute price difference for all possible location (stores) pairs. Repeat this exercise across time and all goods, and pool all observations.

Second, define distance-border-bins according to a discrete spacing criteria that depends on the unit of observation (city vs countries) and the availability of enough observations within each bin. In the case of the city effect, stores are assigned to bins of a few miles apart. If the unit of analysis is countries, bins should be larger to contain stores that are separated by larger distances. The distance in each bin does not have to be set in linear increments. For simplicity denote each bin as n, where $n = \{1, \ldots, N\}$, and N is the number of bins. Each bin is defined by a distance D_n, a dummy $B_n = 1$ if there is a border between the two stores, and additional controls X_n. In our case, X_n includes a chain dummy and an interaction term between distance and city dummy.

Third, compute the relevant quantile statistic of the absolute price differences for each bin. Denote the statistic as $Q_n\left(|p_{i,t} - p_{j,t}|, q\right)$ for the q^{th} percentile of bin n.

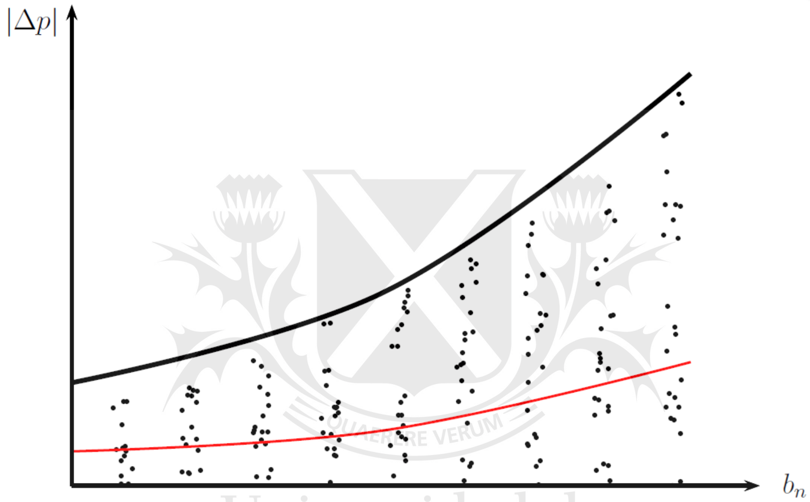
Finally, estimate the following equation:

$$Q_n\left(|p_{i,t} - p_{j,t}|, q\right) = \alpha + \beta D_n + \gamma B_n + \delta X_n + \epsilon_n \tag{2.8}$$

In Figure 2.3 we depict the impact of the bias and the intuition behind our methodology. The horizontal axis shows the bins for a range of distances, and the vertical axis is the absolute price difference. The dots mark the absolute differences in the data for each of the selected "bins". The thick black line reflects the price difference implied by the no-arbitrage constraint. Because all the observed price differences are less than or equal to the thick line, the estimation in the standard regression –which implicitly uses the mean within each bin– is downward biased as denoted by the red line.

In small samples, the true maximum per bin might not be observed, and therefore estimating via the sample maximum will also be downward biased. Still, in this case the bias is smaller than using the mean. In other words, it is possible that there is no realization on the black line, but using the maximum within each bin gets closer to the "true" line. This explains why we interpret our results as a lower bound estimate of the degree of segmentation.

Figure 2.3: Bias in Standard Regressions



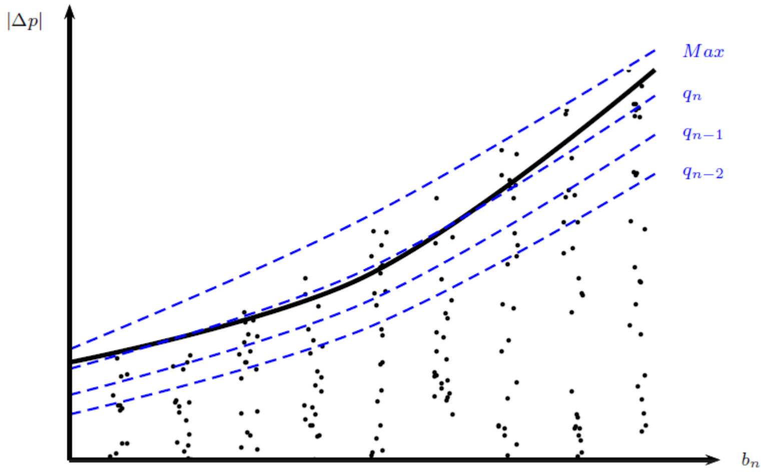
Note: This figure illustrates the source of the selection bias. The horizontal axis shows the bins for a range of distances. The vertical axis is the absolute price difference across locations. For each bin, all the absolute differences from the data are shown as the black dots. The thick black line reflects the price difference implied by the no-arbitrage constraint. Because all the observed price differences are less or equal to the thick line, the estimation in the standard regression which implicitly uses the mean within each bin (red line) is downward biased.

2.2.3 Dealing with errors-in-variables

One of the reasons we use quantile regressions to estimate arbitrage costs, instead of only the maximum, is to relax the assumption of no errors-in-variables (EIV). In particular, the maximum of the price difference distribution within each bin can be significantly affected if prices are mis-measured. These errors can arise either because prices are observed and/or reported with errors, or because stores make mistakes and post prices outside the no-arbitrage range. When we describe our data in Section 2.3 it will become clear that the errors from misreporting are very small, because of the way the data is collected. However, there is still the possibility that the prices are incorrectly reported, and thus concentrating the estimates on the maximum within each bin would exacerbate the impact of any errors-in-variables.

This case is depicted in Figure 2.4. The black thick line is still the "true" upper bound of the no-arbitrage band, that is the true degree of segmentation. However due to EIV, some price differences might even be above the no-arbitrage range. In this case, using the maximum within each bin also produces a bias in the estimation.

Figure 2.4: Bias in Standard Regression in the presence of EIV



Note: The black thick line is still the "true" upper bound of the no-arbitrage band, i.e. the true degree of segmentation. However due to EIV, some price differences might even be above the no-arbitrage range. In this case, using the maximum within each bin also produces a bias in the estimation. For this reason we use a series of quantile regressions instead.

We address errors-in-variables in two ways. One is to eliminate outliers from the distribution. As we discuss below, the type of errors that are likely to be present in our data are misplacement of the decimal point or flipping digits, both of which are likely to produce large price changes at the item level that we can identify. This approach, however, does not provide a definite answer. For example, if the estimates change little then it is not clear whether the EIV had a small impact, or not enough observations were eliminated to remove the bias. The alternative we propose it to estimate the regression using different quantiles. Within each bin we compute several quantiles –the median, 80th, 90th, 95th, and 99th percentiles.[11] The 50th and 80th percentiles are clearly less affected by the EIV than the maximum, but those estimates will be affected by the sample selection of prices within the no-arbitrage range. As we move towards higher percentiles, the estimates are less affected by the sample selection, and more affected by the EIV. If the EIV is small, it should be the case that the estimates are monotonically increasing. We evaluate the robustness and sensitivity of our estimates in Section 2.4.1.

2.3 Data

We use a good-level dataset of daily prices compiled by The General Directorate of Commerce (DGC) which comprises grocery stores all over the country.[12] The DGC is the authority responsible for the enforcement of the Consumer Protection Law at the Ministry of Economy and Finance.

In 2006 a new tax law was passed by the Uruguayan legislature which changed the tax base and rates of the value added tax (VAT). The Ministry of Economy and Finance was concerned about incomplete pass-through from tax reductions to consumer prices, and hence decided to collect and

[11]We also evaluate the robustness of our estimates to the elimination of price change outliers.
[12]The same dataset is used in Borraz and Zipitría (2012).

publish a dataset of prices in different grocery stores and supermarkets across the country. The DGC issued Resolution Number 061/006 which mandates grocery stores and supermarkets to report its daily prices for a list of products if they meet the following two conditions: i) they sell more than 70% of the products listed, and ii) either have more than four grocery stores under the same name, or have more than three cashiers in a store. The information sent by each retailer is a sworn statement, and they are subject to penalties in case of misreporting. The objective of the DGC is to ensure that prices posted reflect real posted prices by stores. In this regard, stores are free to set the prices they optimally choose, but they face a penalty only if they try to misreport them

The data includes daily prices from April 1st of 2007 to December 31th of 2010 for 202 items corresponding to 61 product categories, where each item is defined by its universal product code (UPC).[13]

The three highest-selling brands are reported for each product category. Most items had to be homogenized in order to be comparable, and each supermarket must always report the same item. For example, sparkling water of the "Salus" brand is reported in its 2.25 liter variety by all stores. If this specific variety is not available at a store, then no price is reported. Whenever prices are 50% greater (or less) than the average price, the retailer is contacted to confirm whether the submitted price is correct. The data are then used in a public web site that allows consumers to check prices in different stores or cities and to compute the cost of different baskets of goods across locations.[14]

After the exclusion of observations labeled as "preliminary" as well as wrongly categorized or unidentified data (e.g. products that share the same UPC), our final dataset is composed by 202 products at the UPC level in 333 grocery stores from 47 cities. Table 2.3 describes the summary statistics of the coverage in the data, and Appendix 2.9 provides a detailed list of the products. In addition, see Figure 2.9 for a map with the cities covered in the dataset. These cities represent more than 80% of the total population of Uruguay. Montevideo, with 45% of the population, accounts for 58% of the supermarkets in the sample. The maximum distance between two supermarkets is 526 kilometers.[15]

We consider two datasets separately to account for outliers that may have a greater impact on the largest price differences between one good. A baseline case with the complete sample, and a second case in which we exclude those prices higher than 3 times (or less than a third) of the median daily price. However, deleted prices only account for a small 0.034% of the whole database.

In order to compute the linear distance between each pair of stores in our sample, we use information on the exact geographical location of each supermarket as provided by Ciudata, an industry organization. We then construct distance bins using a geometric sequence starting from 0.1 kilometers, and incrementing by $((526/0.1)^{1/N})$%. Our baseline estimation uses N=500 bins, but we re-estimated our results using 50, 100, and 1,000 bins as well. We then calculate the distance between all supermarkets in the sample (333) and assign each pair of supermarkets (55,278) to its proper bin according to their distance range.

Finally, we define the following specification:

$$Q_n\left(|p_{i,t}-p_{j,t}|,q\right) \;=\; \alpha+\beta D_n+\gamma B_n+\delta_1 B_n\times D_n+\delta_2 Firm_n+\epsilon_n \tag{2.9}$$

[13] The only exceptions are meat, eggs, ham, some types of cheese, and bread. However, as we later show, the exclusion of these goods which could potentially be affected by an imperfect matching, does not modify the results.

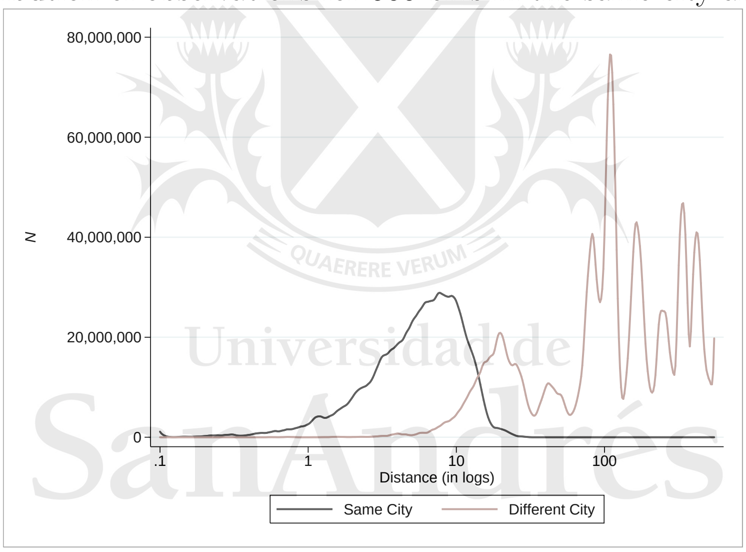
[14] See http://www.precios.uy/servicios/ciudadanos.html.

[15] See Borraz and Zipitría (2012) for a detailed description of the database and an analysis on its price stickiness.

where $Q_n(|p_{i,t} - p_{j,t}|, q)$ is the q^{th} quantile of the absolute price differences for all store pairs i and j whose distance belongs to bin n; D_n measures the distance between stores that belong to bin n; B_n is a dummy that takes the value 1 if the supermarkets are in different cities; $Firm_n$ is a dummy variable that takes the value 1 if the price difference in bin n comes from the same supermarket chain. We also add an interaction dummy between distance and the city border dummy, and fixed effects for each good.

Notice that this regression requires that we have both observations within and across cities that overlap in distance bins. Figure 2.5 shows the distribution of observations for each of the 500 bins for the same city pairs and the different city pairs. The horizontal axis is the log distance starting at 100 meters to a maximum of 526 km. The black line is the number of observations per bin for the stores within the same city boundaries, while the gray line are the observations for the stores in different cities. There is a non-trivial range in which stores are separated exactly by the same distance within cities and across cities –although almost all of them between 5 to 15 kilometers. This is the source of the variation where the city-border effect is actually estimated.

Figure 2.5: Distribution of observations for 500 bins in the same city and between cities



Note: The black line shows the distribution of bilateral observations for each of the 500 bins within cities, while the gray line (extending to the right, with multiple peaks) shows the distribution across cities. Lines are smoothed for better visualization.

2.4 Results

As described in Section 2.2.2, we pool all the prices into each corresponding bin and estimate the distribution of price differences. We select the mean, median, 80, 85, 90, 95, 97.5, 99, 99.5 and 99.9th percentiles. For each of these we estimate equation (2.9) by weighted least squares to account for differences in the number of observations inside each bin. Price differences are expressed in percentage terms, while distance is measured in hundreds of kilometers.

The results are presented in Table 2.1. The first coefficient is the segmentation generated by distance. The second and third estimate the effect of the city boundaries (border dummy) and the interaction term (how the effect of distance changes once the stores are in different cities), respectively. The fourth coefficient is the impact of both stores belonging to the same retailer, and the last one is

Table 2.1: Weighted Least Square Estimation of Price Differential using the complete database and for 500 bins
Linear Specification

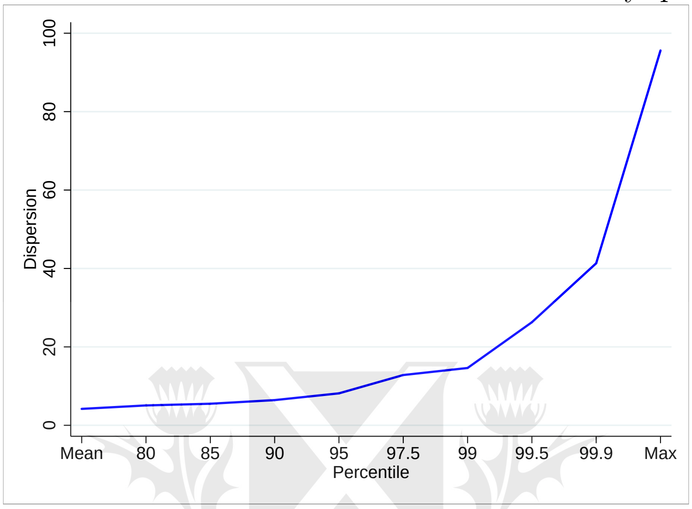
	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	4.188***	3.528***	5.073***	5.504***	6.428***	8.153***	12.822***	14.618***	26.287***	41.329***	95.596***
	(0.186)	(0.202)	(0.297)	(0.323)	(0.368)	(0.510)	(0.733)	(0.967)	(1.414)	(2.601)	(4.677)
City	1.260***	1.243***	1.691***	1.738***	1.880***	1.926***	1.890***	2.478***	2.794***	2.889***	5.105***
	(0.017)	(0.018)	(0.026)	(0.029)	(0.033)	(0.045)	(0.065)	(0.086)	(0.126)	(0.232)	(0.417)
City*Dist	-4.049***	-3.350***	-4.930***	-5.364***	-6.323***	-8.083***	-12.880***	-14.670***	-26.460***	-41.833***	-92.579***
	(0.186)	(0.202)	(0.297)	(0.323)	(0.368)	(0.510)	(0.733)	(0.967)	(1.414)	(2.602)	(4.678)
Chain	-6.012***	-5.196***	-9.652***	-10.738***	-12.101***	-14.642***	-18.086***	-22.188***	-25.305***	-38.565***	-68.955***
	(0.020)	(0.022)	(0.032)	(0.035)	(0.040)	(0.056)	(0.080)	(0.106)	(0.155)	(0.285)	(0.508)
Const	5.081***	3.782***	8.541***	9.956***	11.745***	14.832***	18.150***	22.106***	25.807***	41.789***	130.155***
	(0.038)	(0.041)	(0.061)	(0.066)	(0.075)	(0.104)	(0.150)	(0.198)	(0.289)	(0.532)	(0.975)
N	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	184,328
R2	0.752	0.645	0.749	0.761	0.762	0.725	0.744	0.766	0.68	0.58	0.492

Note: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors in parenthesis.

the constant term. Each column reflects a different regression. The first one uses the mean within each bin, which replicates the standard regressions in the literature. After that we present the results for the quantiles moving from the 50th until 99.9th percentiles and finally the maximum.

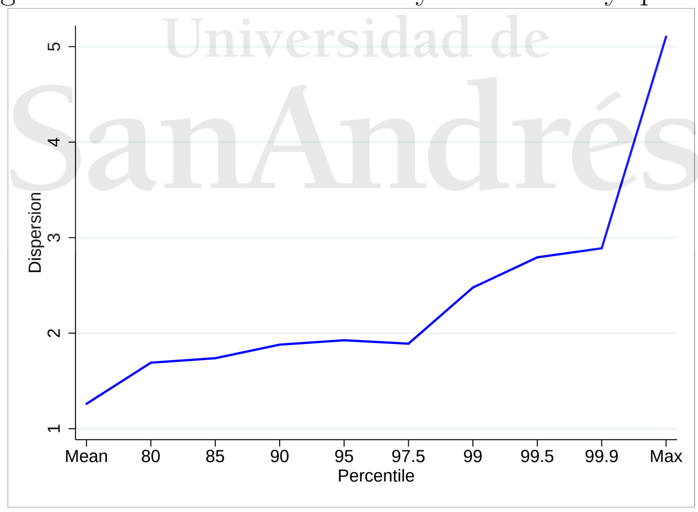
Notice that as we increase the percentile, all individual coefficients increase –in line with the intuition we discussed before. This pattern can be easily appreciated in Figures 2.6 and 2.7, which show the coefficient on distance and city dummy, respectively, as a function of the percentile.

Figure 2.6: Estimation of the distance coefficient by quantile



Note: Estimated distance coefficient when different quantiles are used for the baseline regression.

Figure 2.7: Estimation of the city coefficient by quantile



Note: Estimated city dummy coefficient when different quantiles are used for the baseline regression.

There are two alternatives to compute the border effect. One way is to base the effect upon a specific distance. First, we calculate the degree of price dispersion when the two stores are located in different cities. Then we solve for the distance that would be needed for two stores within the same city to have the same degree of price dispersion. The following example clarifies the analysis.[16] Using the

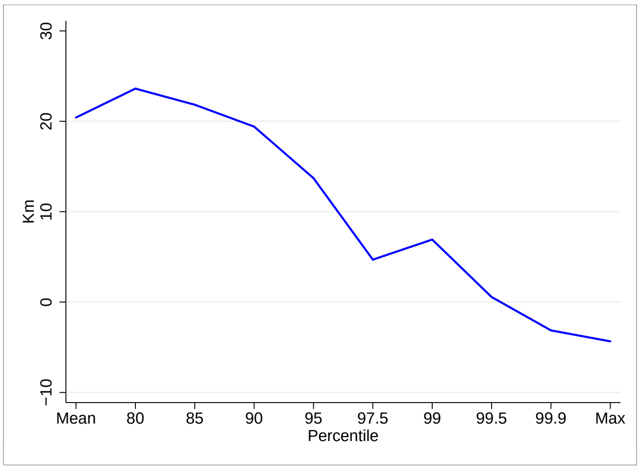
[16]We show the results for 10 kilometers but results remain qualitatively the same for stores 15 and 20 kilometers apart. Given the characteristics of our data, it makes no sense to go beyond that distance because in the city of Montevideo there are very few observations with stores more than 20 kilometers apart.

results in the first column (average) in Table 2.1, we compute the price dispersion of two cities across the border that are 10 km apart. The price dispersion is $5.081 + 4.188 * 0.1 + 1.260 - 4.049 * 0.1 = 6.355$. Two stores in the same city exhibit a segmentation equal to $5.081 + 4.188 * X$. Solving for X to make the within city segmentation equal to 6.355 yields 30.5 km. Therefore the border adds 20 kilometers to two stores 10 kilometers apart –that is, the city border triples the distance. Although the literature simply uses the ratio of the two coefficients to compute the border effect, our specification also allows for non-linearities. Therefore the implied border effect needs to be estimated conditional to a given distance.

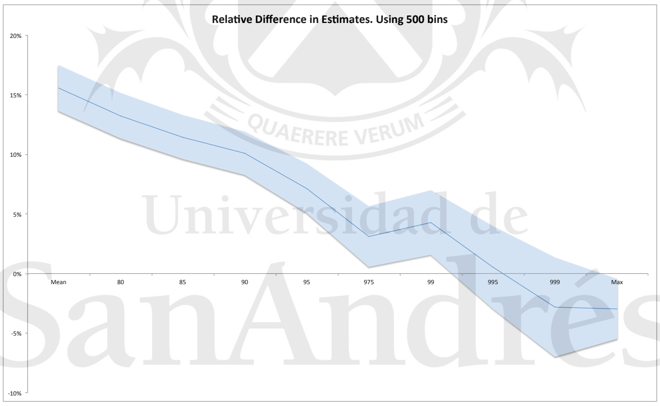
In Panel (a) of Figure 2.8 we compute this implied additional distance for two stores 10 km apart for each of the quantiles. The border effect, as measured in kilometers, collapses towards zero around the 99.5th percentile. Interestingly, the effect is even found negative at the highest percentiles. In addition, notice the (almost) monotonic decrease in the estimates. This is encouraging from an errors-in-variables point of view. If the maximum of the distribution were the result of large errors-in-variables, there is no reason to expect the estimates and the impact of the border effect to remain similar to the upper percentiles.

Figure 2.8: Estimation of the city border effect using all data

(a) Implied Kilometers
Additional Km implied by City Border Effect for Stores 10 Km Apart



(b) Relative Increase in Price Dispersion
of City Borders for Stores 10 Km Apart.



Note: Panel (a) shows the implied additional km for the linear specification using all data and 500 bins. Panel (b) shows the relative increase in the degree of segmentation for the baseline linear specification, with its 95th percent confidence band.

The second way to compute the border effect is to focus on the relative price dispersion for a given distance. In other words, we compute how large is the implied degree of segmentation for a pair of stores 10 km apart across two cities, relative to another pair of stores 10 km apart within the same city. In both cases we consider all stores that do not belong to the same retailer. For instance, in the average case (column 1 in Table 2.1) the price dispersion for $D_n = 0.1$ and $B_n = 1$ is, as before, $5.081 + 4.188 * 0.1 + 1.260 - 4.049 * 0.1 = 6.355$. The price dispersion when $B_n = 0$ is $5.081 + 4.188 * 0.1 = 5.499$. Thus the border implies a 15.57 percent higher degree of segmentation. However, this relative effect becomes small and insignificant using higher quantiles. In Panel (b) in Figure 2.8 we present the decreasing pattern in the relative degree of segmentation, together with its 95th confidence band.

Panels (a) and (b) show that the degree of segmentation is overestimated –and the impact of distance is underestimated– when the average price deviations are used. By contrast, the effect of the border becomes not significantly different from zero when the upper quantiles of the distribution within each bin are used. Notice that the reduction in the border effect is not a mechanical consequence of the methodology. The estimation using upper quantiles should in fact increase the absolute value of all coefficients –simply because there is less sample selection. The decrease in the final border effect, however, comes from the fact that the bias in the distance coefficient is larger than in the border coefficient.

2.4.1 Robustness

In this section we test the sensibility of the baseline estimates to changing the specification of the regression, to different subsamples of product mix, elimination of outliers, and to different number of bins. In all cases we find that the results are qualitatively similar. That is, the traditional regression (average price dispersion) estimates a large and significant city border effect, whereas quantile regressions show that the city border becomes insignificant using upper quantiles of the distribution. Furthermore, the results are similar at the highest percentiles and the maximum, suggesting that the estimates are not significantly affected by measurement errors.

First, we modify the equation to the following non-linear specification:

$$\begin{aligned} Q_n\left(|p_{i,t} - p_{j,t}|, q\right) &= \alpha + \beta D_n + \gamma B_n + \delta_1 B_n \times D_n + \delta_2 D_n^2 + \delta_3 D_n^3 \\ &\quad + \delta_4 B_n \times D_n^2 + \delta_5 B_n \times D_n^3 + \delta_6 Firm_n + \epsilon_n \end{aligned} \tag{2.10}$$

where the variables are defined as in equation 2.9.

The results, presented in Table 2.2, yield the same patterns we described above for the baseline estimation. In absolute value, all point estimates increase as the estimation is performed over the higher quantiles. Furthermore, if one computes the implied additional distance, the results remain qualitatively the same as those in Panel (a) in Figure 2.8. The border effect, as measured in kilometers, is close to 25 km using the traditional regression (average). However it decreases with higher percentiles, until it becomes small and insignificant at the 97.5th percentile.

In addition, we perform three robustness tests using different subsamples. Results are presented in Appendix 2.8, for both the linear and non-linear specifications. First, we eliminate products in which the matching across stores is not perfect. In particular, we exclude meat, bread, among other categories. Quantile regressions yield identical patterns as when using the complete dataset. Second, we use all products but eliminate the outliers, defined here as those whose price is above three times (or a third below) the median price. This approach is more conservative that the one typically used in the literature. For example, Gopinath and Rigobon (2008) and Klenow and Kryvtsov (2008) eliminate prices that are more than 10 times higher or less that a tenth of the median price. Still, our rule only excludes 11.2 thousand in 32.8 million, or just 0.034% of the observations. Once again, the patterns are almost identical to the ones obtained using the complete number of observations. The only minor difference is that, for a given percentile, the border effects are smaller in absolute terms. In other words, the estimated implied distances are smaller than those in Panel (a) in Figure 2.8.

Third, we further combine the two cases above and estimate quantile regressions excluding goods with imperfect matching as well as those defined as outliers. The results do not yield significant differences with respect to our baseline estimation.

Lastly, we tested the linear and non-linear specifications to the sensibility of the number of distance bins. Instead of 500 bins, we re-estimated using 50, 100, and 1,000 bins. Notice the trade-off in the

Table 2.2: Weighted Least Square Estimation of Price Differential using the complete database and for 500 bins

	Non Linear Specification										
	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	-1.396**	-4.359***	1.854**	3.213***	3.132***	4.005***	24.295***	43.301***	94.219***	221.741***	946.493***
	(0.542)	(0.589)	(0.870)	(0.947)	(1.081)	(1.498)	(2.153)	(2.840)	(4.152)	(7.636)	(13.542)
City	0.610***	0.486***	1.042***	1.091***	1.132***	0.982***	1.355***	2.276***	3.575***	7.527***	18.414***
	(0.025)	(0.028)	(0.041)	(0.044)	(0.051)	(0.070)	(0.101)	(0.133)	(0.195)	(0.358)	(0.640)
City*Dist	2.467***	5.559***	-0.669	-1.974**	-1.780*	-2.262	-22.579***	-41.457***	-92.334***	-221.689***	-919.021***
	(0.543)	(0.590)	(0.872)	(0.948)	(1.082)	(1.500)	(2.156)	(2.844)	(4.158)	(7.647)	(13.564)
Chain	-6.001***	-5.186***	-9.637***	-10.721***	-12.083***	-14.618***	-18.044***	-22.122***	-25.194***	-38.355***	-67.614***
	(0.020)	(0.022)	(0.032)	(0.035)	(0.040)	(0.056)	(0.080)	(0.106)	(0.155)	(0.284)	(0.500)
$Dist^2$	34.149***	48.223***	19.697***	14.030***	20.175***	25.390***	-70.068***	-175.224***	-415.058***	-1102.431***	-5203.545***
	(3.117)	(3.387)	(5.002)	(5.440)	(6.210)	(8.607)	(12.373)	(16.321)	(23.862)	(43.883)	(77.900)
$Dist^3$	0.026***	0.031***	0.025***	0.028***	0.034***	0.061***	0.062***	0.032***	0.008	-0.112***	0.870***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.008)	(0.010)	(0.015)	(0.027)	(0.060)
City*$Dist^2$	-34.466***	-48.581***	-20.037***	-14.394***	-20.597***	-26.020***	69.408***	174.667***	414.569***	1102.788***	5194.370***
	(3.117)	(3.387)	(5.002)	(5.440)	(6.210)	(8.607)	(12.373)	(16.321)	(23.862)	(43.883)	(77.901)
Const	5.241***	4.010***	8.630***	10.017***	11.835***	14.946***	17.800***	21.239***	23.768***	36.406***	104.760***
	(0.041)	(0.044)	(0.065)	(0.071)	(0.081)	(0.112)	(0.161)	(0.213)	(0.311)	(0.572)	(1.031)
N	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	179,215	184,328
R2	0.755	0.649	0.751	0.763	0.763	0.726	0.745	0.767	0.681	0.581	0.509

Note: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors in parenthesis.

selection of bins. The advantage of a larger number of bins is that each pair of stores is allocated to a very specific distance bin and the distance representing the bin is closer to the real distance across the stores. The disadvantage is that the number of observations within each bin decreases. In the limit, if the bins are so narrow that each store pair belongs to a single bin, then the problem is that the estimation at the 99.9 percentile becomes very noisy.[17] The results are qualitatively the same to the baseline estimation. The city-border effect measured in kilometers falls and becomes insignificant when the upper quantiles are used in the estimation.

2.5 Conclusions

The extensive literature on the degree of segmentation resulting from political borders has reported extremely large transaction costs introduced by country, province, and even city borders. In this paper we argue that some of those estimates have been overstated because the empirical strategy has not taken into account the selection problem in posted prices: when a firm faces the possibility of arbitrage due to the existence of a transaction cost, the firm sets prices subject to a no-arbitrage constraint. However, if the optimal price falls within the no-arbitrage range, the dispersion in prices is not informative of the tightness of the constraint. A firm may set the same price in two locations, but it does not mean that the arbitrage cost for the consumer is zero. This implies that the estimation using average absolute price differences or standard deviations of price differences will not capture the size of the trade or arbitrage cost.

This paper builds on the existing literature with two main contributions. In the first place, it offers an alternative methodology to estimate transactions costs –which not only can be applied in international trade, but also in other areas as in empirical finance, measurement of liquidity, or the cost of regulatory restrictions. In the second place, we show that city borders matter little for price dispersion within a country. Although the border effect of a city should be small from an intuitive point of view, the traditional methods still estimated a very wide border effect (20 additional kilometers to two stores separated 10 km apart, that is, the border triples the distance). This is particularly large in a country where the largest city is less than 40 kilometers wide and there are no significant differences between cities in terms of taxes, language and the like. By contrast, the border becomes insignificant once we estimate using our method of distance-binned quantile regressions.

Finally, we believe further research should advance in at least two dimensions. From a methodological point of view, it is important to further examine the definition of optimal bandwidths. Although in our paper we used different bin sizes and results remained consistent across all specifications, this may not be the case in other applications in economics. And second, similar micro-level data needs to be collected across several countries to shed light on the actual width of international borders.

[17]Future research should formally address the optimal bandwidth selection. For the moment we compare the results across different specifications, and do not explore the issue further because the results remain essentially identical. It is possible that if the estimation is done using less frequent data such as month by month, or using a much smaller dataset, then the issue of bandwidth selection becomes more important. This was not the case in our application.

Bibliography

——— (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–192.

Andrews, D. W. K., S. Berry, and P. Jia (2004): "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location," *Working Paper, Yale University.*

Andrews, D. W. K., and P. Guggenberger (2009): "Validity of Subsampling and "Plug-in Asymptotic" Inference for Parameters defined by Moment Inequalities," *Economic Theory*, 25(03), 1–41.

Andrews, D. W. K., and X. Shi (2014): "Nonparametric Inference based on Conditional Moment Inequalities," *Journal of Econometrics*, 179(1), 31–45.

Andrews, D. W. K., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78(1), 119–157.

Borraz, F., and L. Zipitría (2012): "Retail Price Setting in Uruguay," *Economia*, 12(2), 77–109.

Broda, C., and D. E. Weinstein (2008): "Understanding International Price Differences Using Barcode Data," *NBER Working Paper 14017.*

Ceglowski, J. (2003): "The Law of One Price: Intranational Evidence for Canada," *The Canadian Journal of Economics*, 36(2), 373.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Engel, C., and J. H. Rogers (1996): "How Wide is the Border?," *American Economic Review*, 86(5), 1112–1125.

Evans, C. L. (2003): "The Economic Significance of National Border Effects," *American Economic Review*, 93(4), 1291–1312.

Goldberg, P. K., and M. M. Knetter (1997): "Goods Prices and Exchange Rates: What Have We Learned?," *Journal of Economic Literature*, 35(3), 1243–1272.

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li (2011): "International Prices, Costs, and Markup Differences," *American Economic Review*, 101(6), 2450–86.

Gopinath, G., and R. Rigobon (2008): "Sticky Borders," *The Quarterly Journal of Economics*, 123(2), 531–575.

Gorodnichenko, Y., and L. Tesar (2009): "Border Effect or Country Effect? Seattle may not be so far from Vancouver after all," *American Economic Journal – Macroeconomics*, 1(1), 219–241.

Head, K., and T. Mayer (2002): "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade," *CEPII Research Center, Working Paper 2002-01.*

Hillberry, R., and D. Hummels (2003): "Intranational Home Bias: Some Explanations," *The Review of Economics and Statistics*, 85(4), 1089–1092.

Klenow, P. J., and O. Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?," *The Quarterly Journal of Economics*, 123(3), 863–904.

——— (2001): "Explaining The Border Effect: The Role of Exchange Ratee Variability, Shipping Costs, and Geography," *Journal of International Economics*, 55(1), 87–105.

Ponomareva, M., and E. Tamer (2011): "Misspecification in Moment Inequality Models: Back to Moment Equalities?," *The Econometrics Journal*, 14(2), 186–203.

Rosen, A. M. (2008): "Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities," *Journal of Econometrics*, 146(1), 107–117.

Samuelson, P. A. (1954): "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments," *The Economic Journal*, 64(254), 264–289.

Simonovska, I., and M. E. Waugh (2014): "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 92(1), 34–50.

Appendix

2.6 Figures and Tables

Figure 2.9: Cities covered in the sample

Note: Each dot represents a store location across the 19 Uruguayan departments.

Table 2.3: Product, time, and regional coverage in the data

		All Stores
(i)	Retailers	136
(ii)	Stores	333
(iii)	Products	202
(iv)	Categories	61
(v)	Country	Uruguay
(vi)	Cities	47
(vii)	Departments	19
(viii)	Time Period	April 1st 2007 to December 31st 2010
(ix)	Days	1,154
(x)	Observations (bins)	179,215
(xi)	Observations (pairs)	32,159,865

Note: Summary statistics of the data compiled by The General Directorate of Commerce (DGC).

2.7 Proofs to the model of price-setting with arbitrage

2.7.1 Consumers

Lemma 2.1. *The price space is non empty.*

Proof. Given prices p_i and p_j the right hand side of the inequality is non negative for the consumer located on store ℓ_i. In this case, equation 2.6 is $\mid p_i - p_j \mid \leq \frac{\tilde{\beta}}{\theta} |\ell - \ell_j| + \frac{\tilde{\gamma}}{\theta} b_j + \frac{\tilde{\delta}}{\theta} I_j$, as $\ell = \ell_i$. This implies that there must exist at least one price in order for the consumer to compare its bundles. Thus both the right and left hand side of the inequality are positive.

□

Lemma 2.2. *The inequality constraint is binding only for the marginal consumer.*

Proof. The marginal consumer can be defined as the one obtaining the same utility from buying in both stores, that is $u_\ell(i) = u_\ell(j)$. This in turn implies that $v - \theta p_i - \tilde{\beta} |\ell - \ell_i| - \tilde{\gamma} b_i - \tilde{\delta} I_i = v - \theta p_j - \tilde{\beta} \mid \ell - \ell_j \mid - \tilde{\gamma} b_j - \tilde{\delta} I_j$. Rearranging terms we obtain that $\mid p_i - p_j \mid = \frac{\tilde{\beta}}{\theta} |\ell_j - \ell_i| + \frac{\tilde{\gamma}}{\theta} \bigtriangleup b_{i,j} + \frac{\tilde{\delta}}{\theta} \bigtriangleup I_{i,j}$.

□

2.7.2 Producers

Lemma 2.3. *Firms maximize profits by setting the price that binds the participation constraint for consumer i.*

Proof. Given the prices of all stores except for j, and given the right hand side of the equation in terms of p_j, Kuhn Tucker conditions determine that the price difference should be maximal. That is, when the consumer restriction is binding. At the same time, notice that the marginal consumer for firm j determines the demand for its products.

□

Proposition 2.4. *The consumer that maximizes profits is the marginal consumer.*

Proof. From Lemma 2.2 and Lemma 2.3, the firm sets the maximum price for the marginal consumer.

□

Proposition 2.5. *For any given two stores (locations), the parameters can only be estimated for the marginal consumer, i.e. where the inequality is binding.*

Proof. From Lemma 2.2 and Proposition 2.4, the firm sets its price for the marginal consumer such that equation 2.6 is binding. Therefore the price differences will be maximum given the store location and other exogenous variables for the consumer.

□

Lemma 2.6. *If transportation costs increase (beta) or a border exists between two stores, or if the sensitivity of the consumer to price changes decrease, then the price dispersion increase.*

Proof. Take partial derivatives of each coefficient on the last equation of Lemma 2.2.

□

2.7.3 Consumer heterogeneity: Discussion

So far we have assumed that consumers only differ in their location on the line. However consumers can also differ in their valuation of the good. This feature can be introduced to the original model in either two ways. First, consumers can differ in their maximum valuation of the good, in which case $v \in [\underline{v}, \overline{v}]$. In this case, previous results are easily maintained as well, although now satisfying two conditions for the marginal consumer: indifference in distance and in valuation. Recall that previous results are for the medium consumer. Second, consumers can differ in their disposition to pay for the good, i.e. $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$. Similarly as before, there are two constraints to estimate the demand for each store: the distance constraint and the valuation constraint.

Therefore adding heterogeneity to consumers' taste does not change the estimation procedure. In order to estimate the demand for each store, we must still solve the model for the marginal consumer. As previously shown, the price inequality should only be binding for this consumer, and slack for non-marginal consumers.

2.7.4 Samuelson's Iceberg Costs

The results from our model of product arbitrage is also related to the no-arbitrage pricing region generated in Samuelson's Iceberg costs.[18] Assume that there is an arbitrage cost between two locations that can be described as follows:

$$\tau_{i,j,t} = \alpha + \beta D_{i,j} + \gamma B_{i,j} + \delta X_{i,j,t} \tag{2.11}$$

where the variables are defined as before. This arbitrage cost τ represents the proportion of the item that is lost when a customer transports one unit from i to j.[19] Under this form of arbitrage costs, prices need to lie within the range $|p_i - p_j| \leq \tau_{i,j,t}$ to avoid the possibility that a customer arbitrates across locations. In particular, assume that p_i is set. The second store, when deciding its price, maximizes profits subject to the no-arbitrage constraint. If the optimal price is such that the difference between p_i and p_j is smaller than τ then the constraint is not binding and the price difference is a biased estimate of τ. But if the difference is larger, then the store sets the price at the corner solution and the constraint is binding. This simple behavior implies that the absolute difference of log prices satisfies inequality 2.1, which can be rewritten as $|p_i - p_j| \leq \tau_{i,j,t} = \alpha + \beta D_{i,j} + \gamma B + \delta X_{i,j,t}$.

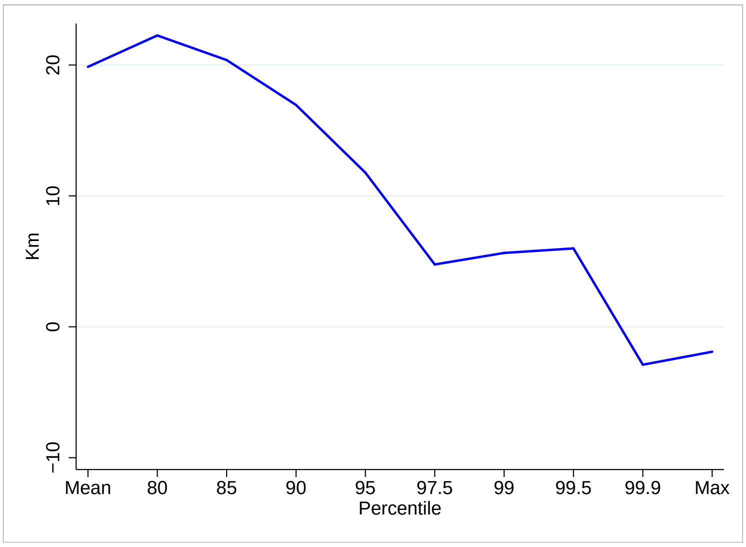
[18] See Samuelson (1954).

[19] For simplicity in the exposition it is assumed that the arbitrageur is the customer itself. Thus the arbitrage cost can be interpreted not only as the loss of physical items, but also the loss in terms of utility that the customer would experience if were forced to travel from one location to another.

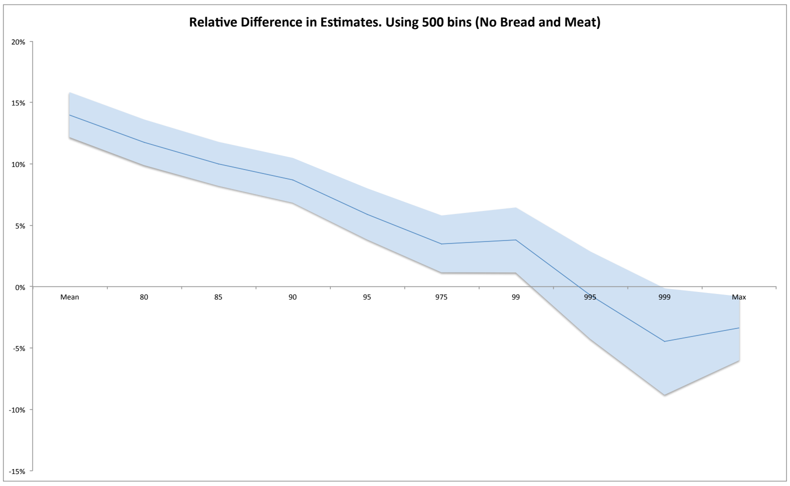
2.8 Alternative Specifications

Figure 2.10: Estimation of the city border effect excluding meat and bread

(a) Implied Kilometers
Additional Km implied by City Border Effect for Stores 10 Km Apart



(b) Relative Increase in Price Dispersion
of City Borders for Stores 10 Km Apart



Note: Panel (a) shows the additional km implied by the city border effect for the linear specification, excluding meat and bread, and using 500 bins. Panel (b) shows the relative increase in the degree of segmentation, with its 95th percent confidence band, for the same specification.

Table 2.4: Price Differential (excluding Meat and Bread, 500 bins, Weighted Least Square)

Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	3.892***	3.118***	4.602***	5.029***	6.207***	7.902***	13.154***	15.652***	28.285***	45.346***	104.739***
	(0.182)	(0.199)	(0.291)	(0.317)	(0.365)	(0.513)	(0.684)	(0.968)	(1.466)	(2.724)	(4.906)
City	1.156***	1.136***	1.523***	1.554***	1.695***	1.718***	1.999***	2.472***	2.650***	2.513***	5.423***
	(0.016)	(0.018)	(0.026)	(0.028)	(0.033)	(0.046)	(0.061)	(0.086)	(0.131)	(0.242)	(0.437)
City*Dist	-3.773***	-2.966***	-4.484***	-4.913***	-6.119***	-7.846***	-13.222***	-15.725***	-28.474***	-45.858***	-101.768***
	(0.182)	(0.199)	(0.291)	(0.317)	(0.366)	(0.513)	(0.684)	(0.968)	(1.466)	(2.724)	(4.907)
Chain	-5.831***	-5.020***	-9.383***	-10.459***	-11.807***	-14.329***	-17.488***	-21.293***	-24.357***	-37.604***	-68.527***
	(0.020)	(0.022)	(0.032)	(0.035)	(0.040)	(0.056)	(0.075)	(0.106)	(0.161)	(0.298)	(0.534)
Const	5.170***	3.884***	8.680***	10.100***	11.878***	14.975***	18.057***	22.072***	25.835***	41.914***	129.772***
	(0.036)	(0.039)	(0.058)	(0.063)	(0.073)	(0.102)	(0.136)	(0.192)	(0.291)	(0.541)	(0.995)
N	159566	159566	159566	159566	159566	159566	159566	159566	159566	159566	165297
R2	0.705	0.593	0.712	0.726	0.724	0.684	0.71	0.679	0.576	0.523	0.479

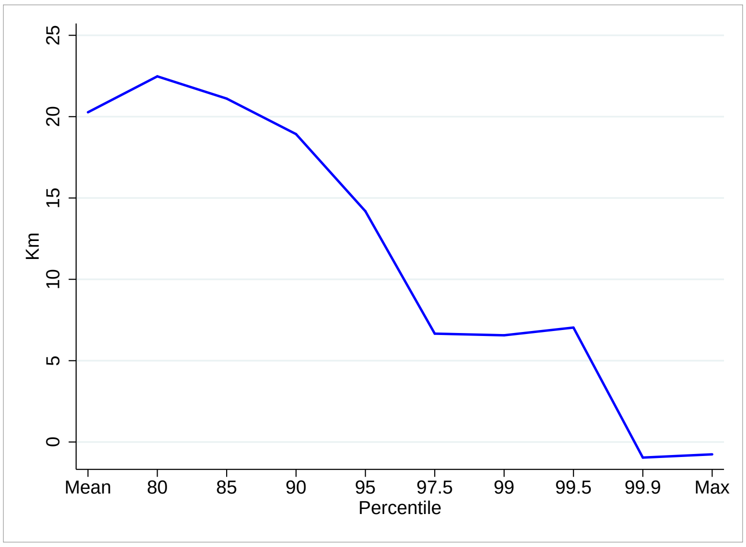
Non Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	-1.721***	-4.583***	0.963	2.260**	1.818*	1.948	19.575***	38.915***	91.935***	220.458***	960.037***
	(0.534)	(0.582)	(0.853)	(0.931)	(1.073)	(1.508)	(2.011)	(2.845)	(4.309)	(8.002)	(14.213)
City	0.521***	0.389***	0.878***	0.912***	0.949***	0.775***	1.299***	2.045***	3.363***	7.466***	19.508***
	(0.025)	(0.027)	(0.040)	(0.044)	(0.050)	(0.071)	(0.094)	(0.133)	(0.202)	(0.375)	(0.672)
City*Dist	2.776***	5.783***	0.214	-1.029	-0.501	-0.27	-17.751***	-36.835***	-90.069***	-221.470***	-934.595***
	(0.534)	(0.583)	(0.855)	(0.932)	(1.075)	(1.510)	(2.013)	(2.849)	(4.315)	(8.013)	(14.236)
Chain	-5.821***	-5.010***	-9.369***	-10.443***	-11.791***	-14.309***	-17.450***	-21.232***	-24.252***	-37.419***	-67.215***
	(0.020)	(0.022)	(0.032)	(0.035)	(0.040)	(0.056)	(0.075)	(0.106)	(0.160)	(0.298)	(0.525)
$Dist^2$	34.381***	47.169***	22.303***	16.981***	26.898***	36.483***	-39.254***	-142.333***	-389.547***	-1071.870***	-5238.861***
	(3.072)	(3.351)	(4.913)	(5.358)	(6.178)	(8.679)	(11.575)	(16.379)	(24.805)	(46.066)	(81.902)
$Dist^3$	0.031***	0.037***	0.034***	0.037***	0.041***	0.068***	0.080***	0.061***	0.027*	-0.159***	0.725***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.007)	(0.010)	(0.015)	(0.029)	(0.063)
City*$Dist^2$	-34.723***	-47.560***	-22.682***	-17.386***	-27.347***	-37.132***	38.493***	141.596***	388.977***	1072.661***	5230.745***
	(3.072)	(3.351)	(4.913)	(5.358)	(6.178)	(8.679)	(11.575)	(16.379)	(24.805)	(46.066)	(81.903)
Const	5.332***	4.107***	8.782***	10.177***	12.001***	15.144***	17.858***	21.370***	23.931***	36.707***	104.313***
	(0.039)	(0.042)	(0.062)	(0.068)	(0.078)	(0.110)	(0.147)	(0.208)	(0.314)	(0.584)	(1.055)
N	159566	159566	159566	159566	159566	159566	159566	159566	159566	159566	165297
R2	0.708	0.599	0.714	0.728	0.726	0.685	0.711	0.68	0.577	0.525	0.497

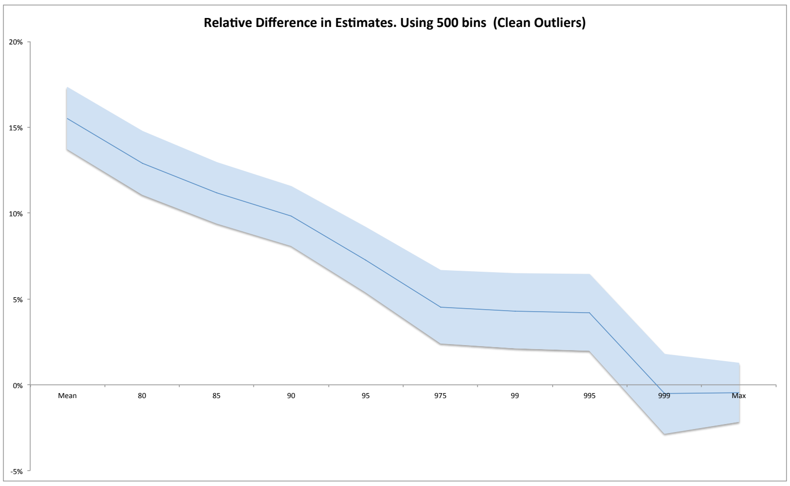
Note: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors in parenthesis.

Figure 2.11: Estimation of the city border effect using all data and excluding outliers

(a) Implied Kilometers
Additional Km implied by City Border Effect for Stores 10 Km Apart



(b) Relative Increase in Price Dispersion
of City Borders for Stores 10 Km Apart



Note: Panel (a) shows the additional km implied by the city border effect for the linear specification, excluding outliers, and using 500 bins. Panel (b) shows the relative increase in the degree of segmentation, with its 95th percent confidence band, for the same specification.

Table 2.5: Price Differential (excluding Outliers, 500 bins, Weighted Least Square)

Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	4.188***	3.684***	5.205***	5.564***	6.442***	8.025***	13.066***	15.412***	16.448***	21.987***	43.223***
	-0.173	-0.189	-0.285	-0.31	-0.35	-0.469	-0.617	-0.775	-0.925	-1.313	(1.719)
City	1.252***	1.242***	1.674***	1.715***	1.851***	1.933***	2.178***	2.554***	2.814***	2.036***	3.885***
	-0.015	-0.017	-0.025	-0.028	-0.031	-0.042	-0.055	-0.069	-0.083	-0.117	(0.153)
City*Dist	-4.031***	-3.489***	-5.040***	-5.404***	-6.316***	-7.945***	-13.075***	-15.427***	-16.567***	-22.461***	-42.125***
	-0.173	-0.189	-0.285	-0.31	-0.35	-0.469	-0.617	-0.775	-0.925	-1.313	(1.720)
Chain	-6.114***	-5.327***	-9.843***	-10.971***	-12.380***	-14.943***	-18.049***	-21.893***	-24.775***	-34.202***	-48.398***
	-0.02	-0.021	-0.032	-0.035	-0.039	-0.053	-0.069	-0.087	-0.104	-0.148	(0.187)
Const	5.046***	3.773***	8.537***	9.959***	11.755***	14.834***	17.891***	21.944***	25.812***	38.182***	67.798***
	-0.036	-0.039	-0.059	-0.064	-0.073	-0.097	-0.128	-0.161	-0.192	-0.273	(0.358)
N	183341	183341	183341	183341	183341	183341	183341	183341	183341	183341	184277
R2	0.752	0.657	0.749	0.761	0.764	0.733	0.736	0.749	0.735	0.687	0.645

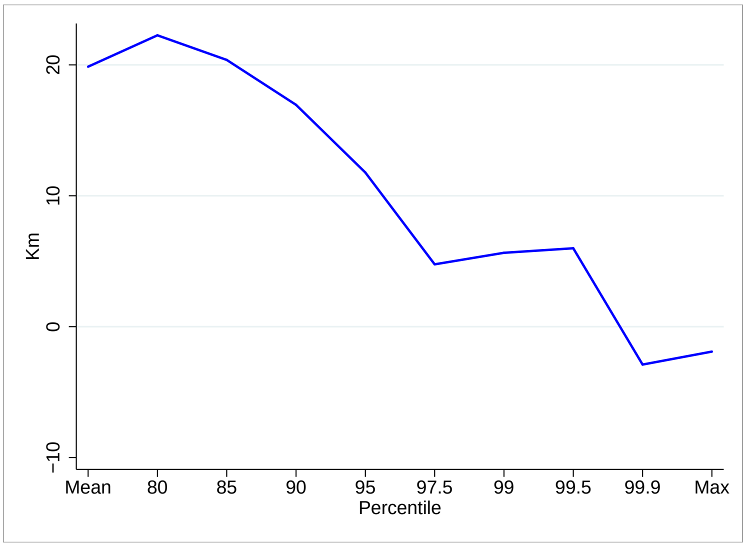
Non Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	-1.220**	-3.840***	2.562***	3.890***	3.888***	5.123***	23.404***	44.674***	67.140***	190.276***	496.586***
	-0.506	-0.553	-0.835	-0.909	-1.027	-1.376	-1.812	-2.276	-2.716	-3.838	(4.885)
City	0.544***	0.439***	0.973***	1.014***	1.045***	0.886***	1.347***	2.148***	2.999***	5.676***	10.710***
	-0.024	-0.026	-0.039	-0.043	-0.049	-0.065	-0.086	-0.108	-0.128	-0.181	(0.231)
City*Dist	2.632***	5.349***	-0.989	-2.257**	-2.080**	-2.727**	-20.608***	-41.664***	-64.293***	-187.771***	-481.214***
	-0.507	-0.554	-0.836	-0.91	-1.029	-1.378	-1.815	-2.28	-2.72	-3.845	(4.893)
Chain	-6.099***	-5.312***	-9.823***	-10.949***	-12.355***	-14.911***	-18.000***	-21.820***	-24.680***	-33.997***	-47.680***
	-0.019	-0.021	-0.032	-0.035	-0.039	-0.053	-0.069	-0.087	-0.104	-0.147	(0.180)
$Dist^2$	33.048***	45.971***	16.168***	10.248**	15.630***	17.760**	-63.107***	-178.667***	-309.537***	-1027.722***	-2772.443***
	-2.908	-3.175	-4.797	-5.221	-5.9	-7.905	-10.406	-13.076	-15.599	-22.048	(28.102)
$Dist^3$	0.067***	0.068***	0.072***	0.075***	0.090***	0.139***	0.176***	0.174***	0.147***	0.134***	0.619***
	-0.002	-0.002	-0.004	-0.004	-0.005	-0.006	-0.008	-0.01	-0.012	-0.017	(0.022)
City*$Dist^2$	-33.610***	-46.552***	-16.786***	-10.896**	-16.387***	-18.860**	61.740***	177.259***	308.252***	1026.488***	2766.629***
	-2.908	-3.175	-4.797	-5.221	-5.9	-7.905	-10.406	-13.076	-15.599	-22.048	(28.102)
Const	5.202***	3.992***	8.610***	10.003***	11.825***	14.912***	17.574***	21.062***	24.291***	33.159***	54.267***
	-0.038	-0.042	-0.063	-0.069	-0.078	-0.104	-0.138	-0.173	-0.206	-0.291	(0.372)
N	183341	183341	183341	183341	183341	183341	183341	183341	183341	183341	184277
R2	0.755	0.661	0.751	0.762	0.766	0.734	0.737	0.75	0.736	0.691	0.669

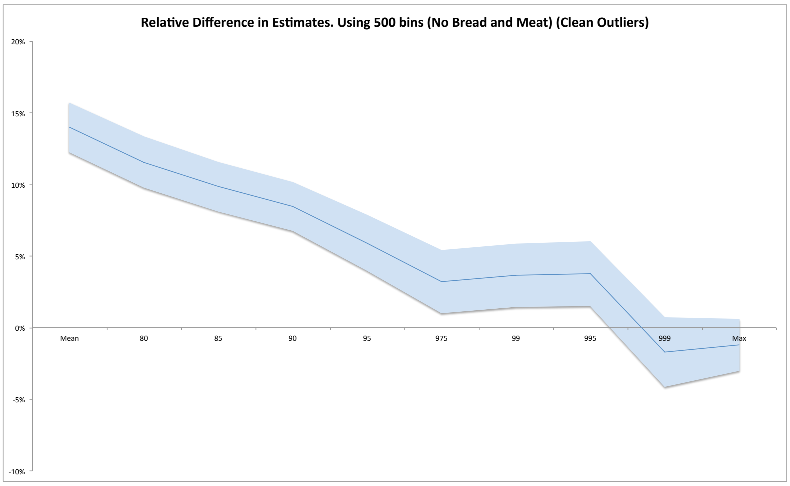
Note: * significant at 10%; ** significant at 5%; *** significant at 1%.

Figure 2.12: Estimation of the city border effect excluding meat, bread, and outliers

(a) Implied Kilometers
Additional Km implied by City Border Effect for Stores 10 Km Apart



(b) Relative Increase in Price Dispersion
of City Borders for Stores 10 Km Apart



Note: Panel (a) shows the additional km implied by the city border effect for the linear specification, excluding meat, bread, as well as outliers, using 500 bins. Panel (b) shows the relative increase in the degree of segmentation, with its 95th percent confidence band, for the same specification.

Table 2.6: Price Differential (excluding Meat, Bread and Outliers, 500 bins, Weighted Least Square)

Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	3.892***	3.291***	4.753***	5.128***	6.249***	7.930***	13.006***	15.211***	17.272***	24.002***	45.887***
	-0.169	-0.185	-0.279	-0.305	-0.348	-0.478	-0.637	-0.789	-0.939	-1.374	(1.820)
City	1.149***	1.143***	1.520***	1.545***	1.674***	1.721***	1.923***	2.382***	2.772***	1.754***	3.603***
	-0.015	-0.016	-0.025	-0.027	-0.031	-0.043	-0.057	-0.07	-0.084	-0.123	(0.162)
City*Dist	-3.760***	-3.128***	-4.621***	-5.000***	-6.149***	-7.872***	-13.035***	-15.237***	-17.375***	-24.488***	-44.766***
	-0.169	-0.185	-0.279	-0.305	-0.348	-0.478	-0.638	-0.79	-0.939	-1.374	(1.821)
Chain	-5.954***	-5.151***	-9.579***	-10.700***	-12.104***	-14.697***	-17.884***	-21.690***	-24.328***	-33.595***	-48.239***
	-0.019	-0.021	-0.031	-0.034	-0.039	-0.054	-0.072	-0.089	-0.106	-0.155	(0.198)
Const	5.140***	3.873***	8.674***	10.103***	11.893***	14.988***	18.073***	22.063***	25.784***	38.307***	67.890***
	-0.034	-0.037	-0.056	-0.061	-0.07	-0.097	-0.129	-0.159	-0.19	-0.277	(0.369)
N	163729	163729	163729	163729	163729	163729	163729	163729	163729	163729	165257
R2	0.718	0.611	0.715	0.728	0.731	0.698	0.718	0.74	0.714	0.65	0.62

Non Linear Specification

	Average	50	80	85	90	95	97.5	99	99.5	99.9	Max
Distance	-1.484***	-4.044***	1.702**	2.972***	2.602**	3.160**	21.621***	41.226***	61.861***	189.136***	503.975***
	-0.496	-0.541	-0.818	-0.895	-1.021	-1.406	-1.874	-2.32	-2.759	-4.02	(5.173)
City	0.458***	0.352***	0.824***	0.852***	0.872***	0.677***	1.103***	1.907***	2.762***	5.301***	10.530***
	-0.023	-0.026	-0.039	-0.042	-0.048	-0.066	-0.089	-0.11	-0.13	-0.19	(0.244)
City*Dist	2.850***	5.523***	-0.18	-1.398	-0.876	-0.885	-18.941***	-38.225***	-58.906***	-186.650***	-488.696***
	-0.496	-0.542	-0.819	-0.896	-1.023	-1.408	-1.877	-2.324	-2.763	-4.026	(5.181)
Chain	-5.940***	-5.137***	-9.560***	-10.679***	-12.083***	-14.670***	-17.840***	-21.623***	-24.239***	-33.397***	-47.521***
	-0.019	-0.021	-0.031	-0.034	-0.039	-0.054	-0.072	-0.089	-0.106	-0.154	(0.191)
$Dist^2$	32.907***	44.890***	18.684***	13.212**	22.336***	29.204***	-52.666***	-159.090***	-272.702***	-1010.075***	-2805.846***
	-2.851	-3.113	-4.705	-5.148	-5.876	-8.088	-10.782	-13.35	-15.875	-23.13	(29.808)
$Dist^3$	0.068***	0.071***	0.074***	0.078***	0.092***	0.139***	0.178***	0.184***	0.163***	0.137***	0.600***
	-0.002	-0.002	-0.004	-0.004	-0.005	-0.006	-0.008	-0.01	-0.012	-0.018	(0.023)
City*$Dist^2$	-33.471***	-45.482***	-19.309***	-13.863***	-23.085***	-30.284***	51.313***	157.639***	271.333***	1008.830***	2800.138***
	-2.852	-3.113	-4.705	-5.148	-5.876	-8.088	-10.782	-13.35	-15.875	-23.13	(29.808)
Const	5.296***	4.086***	8.760***	10.162***	11.996***	15.122***	17.809***	21.281***	24.448***	33.390***	54.252***
	-0.037	-0.04	-0.061	-0.066	-0.076	-0.104	-0.139	-0.172	-0.204	-0.298	(0.384)
N	163729	163729	163729	163729	163729	163729	163729	163729	163729	163729	165257
R2	0.721	0.616	0.716	0.73	0.733	0.699	0.719	0.741	0.715	0.655	0.646

Note: * significant at 10%; ** significant at 5%; *** significant at 1%.

2.9 Data Details

Product	**Brand**	**Specification**	**Share in CPI (percent)**	**Category**
Beer	Pilsen	0.96 L	0.3	Alcohol
Wine	Roses	1 L	0.34	Alcohol
Wine	Santa Teresa Clasico	1 L	0.34	Alcohol
Wine	Tango	1 L	0.34	Alcohol
Beef (peceto)	No brand	1 Kg	0.9	Food
Beef (nalga)	With bone, no brand	1 Kg	0.43	Food
Beef (nalga)	Boneless, no brand	1 Kg	0.43	Food
Beef (aguja)	With bone, no brand	1 Kg	0.86	Food
Beef (aguja)	With bone, no brand	1 Kg	0.86	Food
Beef (paleta)	With bone, no brand	1 Kg	n/i	Food
Beef (rueda)	With bone, no brand	1 Kg	n/i	Food
Ground beef	Up to 20 percent fat	1 Kg	0.29	Food
Ground beef	Up to 5% fat	1 Kg	0.29	Food
Bread	No brand	1 unit ($\approx$ 0.215 Kg)	1.21	Food
Brown eggs	El Ecologito	1/2 dozen	0.34	Food
Brown eggs	El Jefe	1/2 dozen	0.34	Food
Brown eggs	Prodhin	1/2 dozen	0.34	Food
Butter	Calcar	0.2 Kg	0.15	Food
Butter	Conaprole sin sal	0.2 Kg	0.15	Food
Butter	Lacteria	0.2 Kg	0.15	Food
Cacao	Copacabana	0.5 Kg	0.04	Food
Cacao	Vascolet	0.5 Kg	0.04	Food
Cheese	Cerros del Este	1 Kg	0.23	Food
Cheese	Dispnat	1 Kg	0.23	Food
Chicken	Avicola del Oeste	1 Kg	0.64	Food
Chicken	Tenent	1 Kg	0.64	Food
Coffee	Aguila	0.25 Kg	0.1	Food
Coffee	Chana	0.25 Kg	0.1	Food
Dulce de leche	Conaprole	1 Kg	0.14	Food
Dulce de leche	Los Nietitos	1 Kg	0.14	Food
Dulce de leche	Manjar	1 Kg	0.14	Food
Flour	Canuelas	1 Kg	0.16	Food
Flour	Cololo	1 Kg	0.16	Food
Flour	Puritas	1 Kg	0.16	Food
Frankfurters	Cattivelli	8 units ($\approx$ 0.340 Kg)	0.26	Food
Frankfurters	Ottonello	8 units ($\approx$ 0.340 Kg)	0.26	Food
Frankfurters	Schneck	8 units ($\approx$ 0.340 Kg)	0.26	Food
Grated cheese	Conaprole	0.08 Kg	0.15	Food
Grated cheese	El Trebol	0.08 Kg	0.15	Food
Grated cheese	Milky	0.08 Kg	0.15	Food
Semolina noodles	Adria	0.5 Kg	n/i	Food
Semolina noodles	Las Acacias	0.5 Kg	n/i	Food
Ham	Centenario	1 Kg	0.21	Food
Ham	La Constancia	1 Kg	0.21	Food
Ham	Schneck	1 Kg	0.21	Food
Margarine	Danica dorada	0.2 Kg	0.02	Food
Margarine	Doriana nueva	0.25 Kg	0.02	Food
Margarine	Primor	0.25 Kg	0.02	Food
Mayonnaise	Fanacoa	0.5 Kg	0.09	Food
Mayonnaise	Hellmans	0.5 Kg	0.09	Food

Product	Brand	Specification	Share in CPI (percent)	Category
Beer	Patricia	0.96 L	0.3	Alcohol
Mayonnaise	Uruguay	0.5 Kg	0.09	Food
Noodles	Cololo	0.5 Kg	0.3	Food
Peach jam	Dulciora	0.5 Kg	0.17	Food
Peach jam	Limay	0.5 Kg	0.17	Food
Peach jam	Los Nietitos	0.5 Kg	0.17	Food
Peas	Arcor	0.35 Kg	0.05	Food
Peas	El Hogar	0.35 Kg	0.05	Food
Peas	Trofeo	0.35 Kg	0.05	Food
Quince jam	Los Nietitos	0.4 Kg	n/i	Food
Rice	Aruba tipo Patna	1 Kg	0.2	Food
Rice	Blue Patna	1 Kg	0.2	Food
Rice	Green Chef	1 Kg	0.2	Food
Rice	Pony	1 Kg	0.2	Food
Rice	Vidarroz	1 Kg	0.2	Food
Crackers	El Trigal	0.15 Kg	0.17	Food
Crackers	Famosa	0.14 Kg	0.17	Food
Crackers	Maestro Cubano	0.12 Kg	0.17	Food
Salt	Sek	0.5 Kg	0.05	Food
Salt	Torrevieja	0.5 Kg	0.05	Food
Salt	Urusal	0.5 Kg	0.05	Food
Semolina pasta	Adria	0.5 Kg	n/i	Food
Semolina pasta	Las Acacias	0.5 Kg	n/i	Food
Soybean oil	Condesa	0.9 L	n/i	Food
Sugar	Azucarlito	1 Kg	0.25	Food
Sugar	Bella Union	1 Kg	0.25	Food
Sunflower oil	Optimo	0.9 L	0.25	Food
Sunflower oil	Uruguay	0.9 L	0.25	Food
Tea	Hornimans	Box (10 units)	0.09	Food
Tea	La Virginia	Box (10 units)	0.09	Food
Tea	Lipton	Box (10 units)	0.09	Food
Tomato paste	Conaprole	1 L	0.08	Food
Tomato paste	De Ley	1 L	0.08	Food
Tomato paste	Qualitas	1 L	0.08	Food
Yerba	Canarias	1 Kg	0.34	Food
Yerba	Del Cebador	1 Kg	0.34	Food
Yerba	Sara	1 Kg	0.34	Food
Yogurt	Conaprole	0.5 Kg	0.06	Food
Yogurt	Parmalat (Skim)	0.5 Kg	0.06	Food
Bleach	Agua Jane	1 L	0.08	Personal
Bleach	Sello Rojo	1 L	0.08	Personal
Bleach	Solucion Cristal	1 L	0.08	Personal
Dishwashing detergent	Deterjane	1.25 L	0.2	Personal
Dishwashing detergent	Hurra Nevex Limon	1.25 L	0.2	Personal
Laundry soap	Drive	0.8 Kg	n/i	Personal
Laundry soap	Nevex	0.8 Kg	n/i	Personal
Laundry soap	Skip, Paquete azul	0.8 Kg	n/i	Personal
Laundry soap, in bar	Bull Dog	0.3 Kg (1 unit)	0.45	Personal
Laundry soap, in bar	Nevex	0.2 Kg (1 unit)	0.45	Personal
Shampoo	Fructis	0.35 L	n/i	Personal
Shampoo	Sedal	0.35 L	n/i	Personal
Shampoo	Suave	0.93 L	n/i	Personal
Soap	Astral	0.125 Kg	0.16	Personal

Product	Brand	Specification	Share in CPI (percent)	Category
Beer	Patricia	0.96 L	0.3	Alcohol
Soap	Palmolive	0.125 Kg	0.16	Personal
Soap	Suave	0.125 Kg	0.16	Personal
Toilet paper	Higienol Export	4 units (25 M each)	0.24	Personal
Toilet paper	Personal	4 units (25 M each)	0.24	Personal
Toilet paper	Sin Fin	4 units (25 M each)	0.24	Personal
Toothpaste	Closeup Triple	0.09 Kg	0.49	Personal
Toothpaste	Colgate Total	0.09 Kg	0.49	Personal
Toothpaste	Kolynos	0.09 Kg	0.49	Personal
Cola	Coca Cola	1.5 L	1.94	Soft drinks
Cola	Nix	1.5 L	1.94	Soft drinks
Cola	Pepsi	1.5 L	1.94	Soft drinks
Sparkling water	Matutina	2 L	0.7	Soft drinks
Sparkling water	Nativa	2 L	0.7	Soft drinks
Sparkling water	Salus	2.25 L	0.7	Soft drinks

Chapter 3

Retail Price Setting in Uruguay

Abstract

We analyze the behavior of prices in Uruguay using a unique database of 30 millions daily prices. We find that prices change about 5 times a year with no seasonal patter. Prices changes are highly synchronized and concentrated on the first day of the month. Our paper is the first to present evidence of high synchronization of prices, which in turn could be explained mainly by the data periodicity. Overall the analysis seems to be consistent with state-dependent pricing models, although we found some interesting features of prices that could not be explained by these models. [1]

3.1 Introduction

In recent years there has been a large increase in the empirical literature of price behavior. As new and detailed datasets become available we observe an important number of studies on the microeconomic fundamentals of price setting of firms - mainly retailers - and their impact on inflation. This analysis allows a better understanding of the behavior, dispersion and volatility of prices.

In this paper, we use a rich and unique dataset of 30 million daily prices in grocery stores and supermarkets across the country to analyze stylized facts about consumer price behavior. Our findings are as follows: i) The median duration of prices is two and one-half months. Therefore, retail prices in Uruguay are less sticky than in the U.S. and Brazil, but stickier than in Chile and the U.K. ii) We do not find evidence of a seasonal pattern in the likelihood of price adjustments. iii) The frequency of price adjustment is only correlated with expected inflation for the personal care product category. However, for the food category we find that firms change the percentage points of the adjustment and not their frequency. iv) The probability of price change on the first day of the month is nine times higher than on any another day. v) The probability of a price change is not constant over time. vi) There exists a high synchronization of price changes in our database, either at the city level or chain level. Overall, our analysis seems to be consistent with time dependent models, although the high synchronization of price changes on the first day of the month awaits a better theoretical formalization.

[1]This paper was written with Fernando Borraz and published as “Retail Price Setting in Uruguay", Economia 12, 2 (2012), pp. 77-109

3.1.1 A brief review of the empirical literature

Although there are different theoretical models that explain these issues in the macroeconomic literature - such as menu cost models, sticky price, sticky information models, and time or state-dependent pricing strategies -, the stylized facts pointed out in the literature avoid a unique formalization. Klenow and Malin (2010) provide an up-to- date and concise overview of the empirical evidence, and confront the data with different theoretical models. They stress ten facts of the microeconomic behavior of prices. The primary facts are that prices do change at least once a year; that the main instrument for downward price adjustment is sales; that most markets have a stickier reference price; that goods prices differ in their frequency of adjustment and their changes are asynchronous between them; that there exist microeconomic forces which explain the behavior of prices that differ from aggregate inflation and, finally, that prices adjust mainly when wages change.

Gopinath and Rigobon (2008) study the stickiness of traded goods using micro data on U.S. import and export prices at-the-dock for the period 1994-2005. They find long price duration of traded goods - 10.6 months for imports, and 12.8 months for exports -; great heterogeneity in price stickiness across goods at the disaggregated level; a declining probability of price adjustment over time for imports; and a rather low exchange rate pass-through into U.S. import prices.

Nakamura and Steinsson (2008) use the Consumer Price Index (CPI) and the Producer Price Index (PPI) from the Bureau of Labor Statistics (BLS) in the U.S. for the period 1988-2005 to study price stickiness. Their results show that there is a duration of regular prices of between 8 and 11 months, after excluding price sales; that temporary sales are an important source of price flexibility - mainly downward price flexibility -; that, excluding sales, roughly one-third of price changes are price decreases; that price increases function strongly as covariates with inflation, but price decreases do not; and that price changes are highly seasonal - mainly in the first quarter. Finally, they find that the hazard function of price changes, which estimates the probability of a price change after t periods without changing, is slightly downward sloping, which implies that the probability of a price change occurring decreases the longer the time span since the last change.

Some of these conclusions are relativized by Klenow and Kryvtsov (2008). Using monthly price information from the BLS for the period 1988- 2004, they find that prices change quite frequently, every 3.7 months if sales are included and up to 7.2 months if excluded. They compare their results with those of other papers for the U.S. and conclude that different methodologies on how to include or not include sales and how to take into account prices of substituted goods, change the estimated rigidity of prices. Price changes are quite large, up to an average of 10% a year in their sample. Also, they find a large number of small price changes: nearly 44% of price changes are smaller than 5% in absolute value, with 12% being smaller than 1%. The distribution of the size of price changes is similar between price increases and decreases. Hazard rate estimates for a given item are quite flat, after taking into account the mix of heterogeneous hazard rates for different goods, that is, survival bias.

Ellis (2009) studies the behavior of prices using weekly data for the U.K. He finds low price rigidities in the U.K. retailing industry. Prices change frequently (the mean duration is about two weeks) even after discarding promotions and sales. When analyzing the sign of the price change in price reversals - that is, price changes that later reverted to the original price -, he finds that there is a prevalence of price decreases, which is consistent with sales. Also the range of price changes is very wide: there are some products that display large changes in prices, and a large number that show small changes. Lastly, he finds that all products have declining hazard functions, as do Nakamura and Steinsson (2008).

Studies for Latin America are scarce due to the lack of available scan data, and they have concentrated on micro CPI data. Barros, Bonomo, and Santos (2009) and Medina, Rappoport, and Soto (2007) analyze price formation in Brazil and Chile, respectively. They show that the frequency of adjustment is different from the one obtained using macro data. They estimate median duration of 4 and 3 months for Brazil and Chile, respectively. Because their data is monthly, they cannot capture price changes within a month. Also, the CPI data must deal with a higher measurement error than does scan data. Chaumont, Fuentes, Labbé, and Naudon (2011) study price setting behavior in Chile using weekly scan data. They find significant heterogeneity in price behavior by supermarkets. One salient finding is the relative price flexibility of Chilean supermarkets in their database; price duration is about 1.3 weeks, even lower than in the U.K., see Ellis (2009). In contrast to Nakamura (2008), they find that nearly 35% of price changes are idiosyncratic to product or chain shocks, and 65% of prices changes are common shocks that affect all products in a category and all stores in the country at the same time. The only paper that compares price rigidities across Latin American countries is that of Cavallo (2010). He uses scraped online data from Argentina, Brazil, Chile, Colombia, and Uruguay. He finds price stickiness in Chile and relative price flexibility in Brazil.

To the best of our knowledge, our paper is the first to analyze price behavior of retailers in a small open economy using daily price data from across all country regions. The objective of this study is to describe stylized facts of price formation in Uruguay and to compare them with those of the existing literature. The paper is organized as follows: The next section provides a detailed description of the database. After that, we present the main findings of the analysis, and offer a brief comparison with the available evidence. Then, we discuss the implication of our findings for the existing theoretical literature. Finally, the last section shows the study's main conclusions.

3.2 Data

We analyze a micro dataset with a daily frequency compiled by The General Directorate of Commerce (DGC, by its Spanish acronym) which includes more than 300 grocery stores all over the country and 155 products (see Annex 3.6 for a map with the cities covered in the dataset). The product brands were chosen to be the most representative of the product being described, and they were selected as the best selling brand in each category. The products in the sample represent at least 12.6% of the goods and services in the CPI basket (see Annex II).

The DGC is the authority responsible for the enforcement of the Consumer Protection Law at the Ministry of Economy and Finance. In 2006 a new tax law was passed by the legislature which changed the tax base and rates of the value added tax (VAT). The basic rate was reduced from 23% to 22% and its minimum rate (staple foods, hotel rooms (high season), certain health related services and electricity for public consumption) from 14% to 10%. In addition, exemptions were eliminated (e.g. health sector, passengers transport, sales of new homes). A tax on intermediate consumption of goods at a 3% rate (COFIS) was eliminated. The tax reform also reduced the asymmetries between sectors of activity regarding the employer contribution to social security and introduced a personal income tax.

As the Ministry of Economy and Finance is concerned about incomplete pass- through from tax reduction to consumer prices, it publishes an open public dataset of prices in different grocery stores and supermarkets in order to inform consumers. In this regard, the DGC issued Resolution Number 061/006 which mandates that grocery stores and supermarkets must report the daily prices for a list of products if they fulfill the following two conditions: i) they sell more than 70% of the products listed in Annex II of said Resolution, and ii) they have more than four grocery stores under the same name, or have more than three cashiers in a store. The information sent by each supermarket is a swann statement, which means that they are subject to penalties in case of misreport.

The DGC makes the information public through a web page that publishes the average monthly prices of each product for each store in the defined basket (see http://www.dgc-mef.gub.uy/publico/). This information is available within the first ten days of the next month. It should be noted that there is no further use for the information; e.g. no price control, nor are any further policies implemented to control supermarkets or producers. The idea is to give consumers adequate information about prices so they can do their shopping at the cheapest store.

The products that are to be reported to the DGC were initially established per the results of a survey distributed to the main supermarket chains inquiring about their annual sales for each item and brand. After discarding supermarkets' own brands, the three highest-selling brands were chosen to be reported for each item. Most items had to be homogenized in order to be comparable, and each supermarket must always report the same item. For example, bottled sparkling water of the SALUS brand is reported in its 2.25 liter variety by all stores. If this specific variety is not available at a store, then no price is reported.

Each item is defined by its universal product code (UPC) with the exception of meat, eggs, ham, some types of cheese, and bread. In some instances, as in the case of meat and various types of cheese, general definitions were set, but because of the nature of the products,the items could not be homogenized. In the case of bread, most grocery stores buy frozen bread and bake it, rather than produce it at the store. Grocery stores differ in the kinds of bread they sell, so in some cases the reported bread does not coincide with the definition, and grocery stores prorate the price submitted to the DGC; i.e. if the store sells bread that is 450 grams per unit, and the requested bread is 225grams, it submits half the price of its own bread.

Each month, the DGC issues a brief report with general details of the price evolution. This report counts the number of products that increase or decrease their prices. The prices used for these calculations are the simple average market prices for each product. The database records begin in March 2007, and the new tax base was put into place in July 2007. A few months

Table 3.1: Number of Daily Price Observations, by Product Category (April 2007 - December 2010)

Category	**Number**	**Percentage**
Food	20,380,541	66
Soft Drinks	1,814,628	6
Alcohol	1,486,176	5
Personal products	7,038,089	23
Total	30,719,434	100

Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

later, new products were added to the database, after a push of inflation in basic consumer products in 2008. The government made "voluntary sectoral price agreements" with producers in the salad oil, rice and meat markets. Additionally, in the second semester of 2010, newer goods were added to the dataset in order to expand its representation.

Within two days of the end of the month, each supermarket uploads its price information to the DGC. After that, it begins a process of 'price consistency checking.' This process starts by calculating the average price for each item in the basket. Each price 40% greater or less than the average price is selected. Then, the supermarket is contacted in order to check whether the submitted price is right. If there is no answer from the supermarket, or if the supermarket confirms the price submitted, the price is posted online as reported. If the supermarket corrects the price, which is an exception, the price is corrected in the database and posted online.

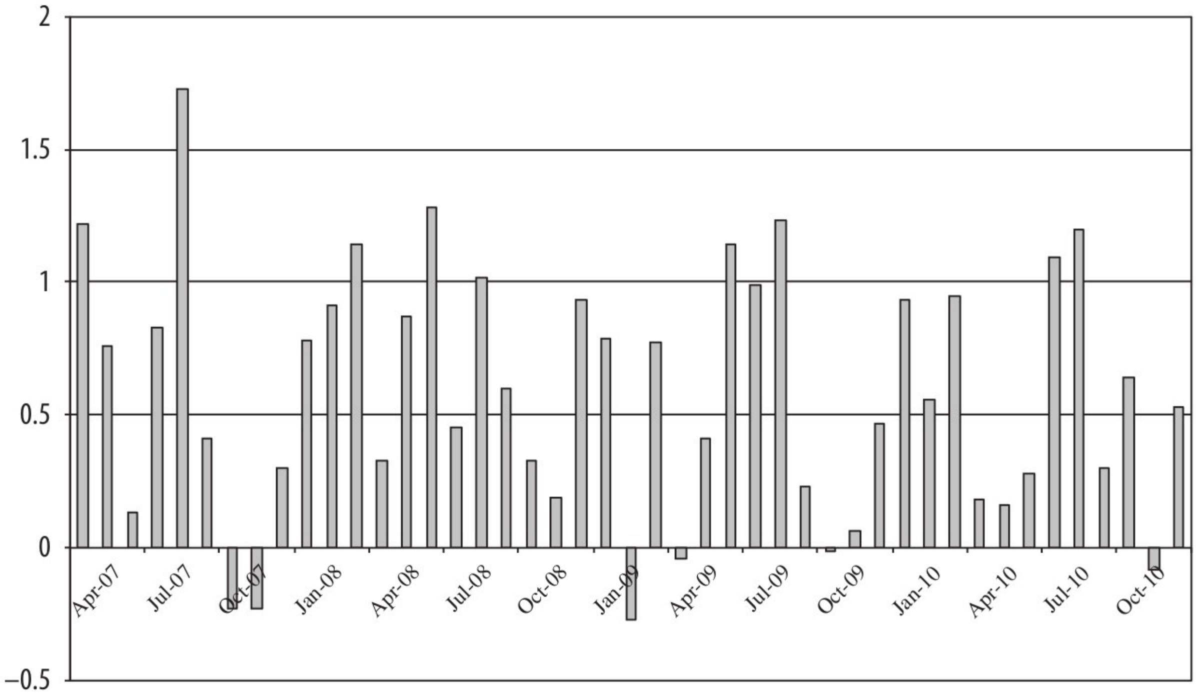
Our database contains daily prices from April 2007 to December 2010 on 155 items. From the database, we eliminated: i) those items that were not correctly categorized (marked as 'XXX' and '0'); ii) ham, as different products mistakenly share the same UPC; and iii) one brand of cheap ham "Leonesa" and meat that also share the same UPC . The complete list of products can be found in Annex II. We also eliminated March 2007 observations, because they were preliminary and had not been posted online. Finally, we eliminated those products - and supermarkets - for which there are no observations for more than half of the period.

We end up with data for 117 products in 303 grocery stores from 45 cities in the 19 Uruguayan departments (see Appendix 3.6). These cities represent 80% of the total population of Uruguay. The capital city, Montevideo, with 45% of the population contains 60% of the supermarkets in the sample.

Table 3.1 summarizes the total number of price observations (30 million) according to four product categories: food, soft drinks, alcohol, and personal care and cleaning items (named personal). Food is the main category, followed by products of personal cleaning, and lastly beverages.

Finally, as our results could be driven by differences in the overall inflation in the sample, we plot the monthly variation of prices (see Figure 3.1). This period is characterized by inflation pushes (the median monthly inflation rate is 0.56 percent), as the government was worried that inflation would reach a high level in the medium term.

Figure 3.1: Monthly Inflation Rate (Percent)



Source: National Institute of Statistics

3.3 Results

We review the frequency of price adjustments by supermarkets and examine seasonality in price adjustments and the nexus between individual price changes and expected overall inflation. We also analyze price changes by day of the month, which is new in the literature. We then compute the joint hazard rate of price changes and examine the synchronization of prices at the chain and city level.

3.3.1 Frequency of Price Adjustments

As is standard in the literature, we first study the rigidity of prices by computing the median probability of daily price changes and the median duration of prices in months and by contrasting the results of price increases and decreases. It should be noted that we study the whole sample and do not differentiate between sales and the absence of sales. From a theoretical point of view, a price decrease because of a sale shows evidence of price flexibility, and we do not want to eliminate such an observation (see Klenow and Kryvtsov, 2008).

The median daily price change for the whole sample is a nontrivial 1.3 percent. That implies a medium price change every 75 days, or every 2.5 months, on average, which is considerably lower than the estimates in Nakamura and Steinsson (2008) and Nakamura (2008) but higher than the results in Chaumont, Fuentes, Labbé, and Naudon (2011) for Chile and those in Ellis (2009). This result is slightly less than the median duration of three and four months found in Barros, Bonomo, and Santos (2009) and Medina, Rappoport, and Soto (2007) for Brazil and

Table 3.2: Price Variation and Duration, by Product Category

Category	**Median probability of daily variation**	**Percentage decrease**	**Monthly duration**
Food	0.013	40.6	2.5
Soft Drinks	0.010	33.3	3.2
Alcohol	0.009	30.0	3.5
Personal products	0.017	42.0	1.9
Total	0.013	40.4	2.5

Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

Chile, respectively.

We offer two explanations for our result. First, this is a period of relatively high inflation, so one could expect prices to change more quickly: the median monthly inflation during the period in Uruguay was 0.56 percent. Second, because our database has daily prices, we can calculate price changes more accurately than in previous studies that use weekly or monthly data. In this case, we can detect earlier price changes and our measure of price rigidity would be more sensitive to them. That would result in less price stickiness for our database.

In line with Nakamura and Steinsson (2008), 40 percent of the price changes are price decreases. Table 3.2 presents the median probability of price changes, the percentage of price decreases, and the median monthly duration by product category. Our results show that prices change most frequently in the personal products category and least frequently in the alcohol category. There is significant variation in price stickiness across product categories, ranging from 1.9 months for personal products to 3.5 months for alcohol.

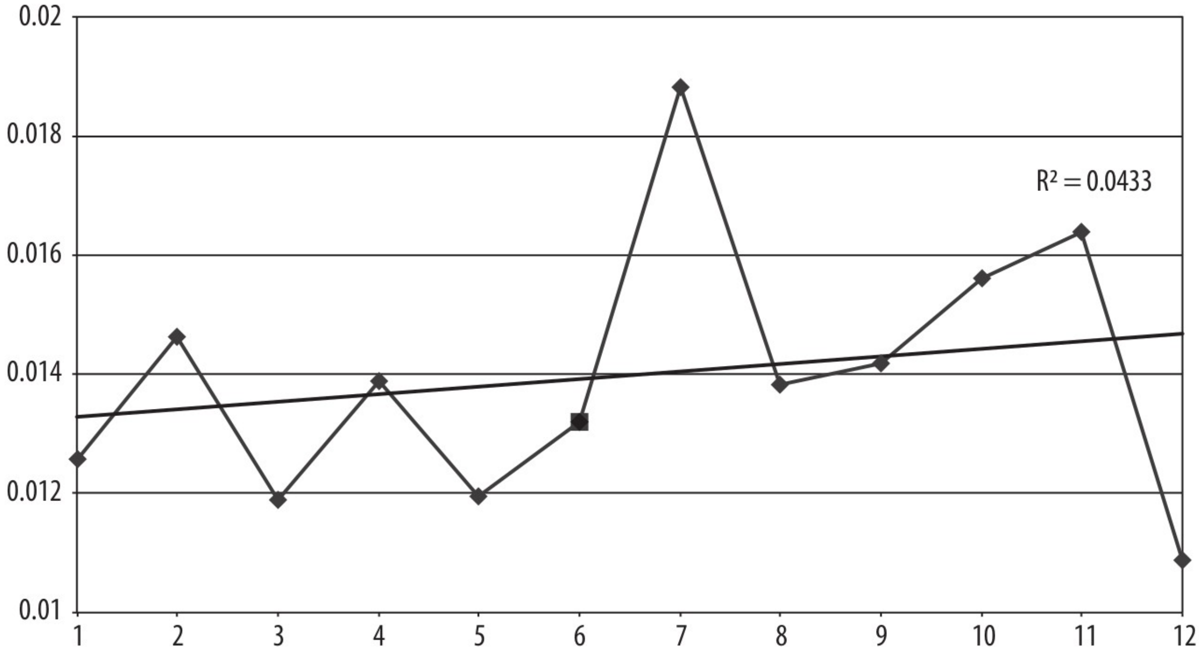
Appendix 3.8 presents a detailed analysis of the results for each product in the sample. There is a high variability of results across products. For example, we find products that change prices quite frequently, such as cheese of the "Disnapt" and "Cerros del Este" brands, for which prices change five and two times a month, respectively. Prices of other products change more slowly, like "El Ecologito" brand brown eggs and "Torrevieja" brand salt, whose prices can remain the same up to five months.

3.3.2 Seasonality of Price Changes

Second, we study seasonal adjustment patterns of prices. Nakamura and Steinsson (2008) finds that price changes in the United States are highly seasonal; they are concentrated in the first quarter and then decrease. This finding is consistent with the authors' price rigidity calculation of about eight months. In contrast, Ellis (2009) finds no monthly seasonality, a result in line with the author's finding of just two weeks of price rigidity. As we find a price duration of 2.5 months, we should expect to find no seasonality in the data.

Figure 3.2 shows that there is not a clear pattern of seasonality in the price adjustments. In addition, we do not find a seasonal pattern in price changes when we look at quarterly data.

Figure 3.2: Probability of Price Change, by Month



Source: Authors' calculations based on data from the Ministry of Economy and Finance.

Table 3.3: Seasonal Probability of Price Change, by Product Category

Quarter	Food	Soft Drinks	Alcohol	Personal products
1	0.013	0.008	0.006	0.013
2	0.012	0.009	0.008	0.017
3	0.016	0.012	0.010	0.018
4	0.015	0.010	0.009	0.019

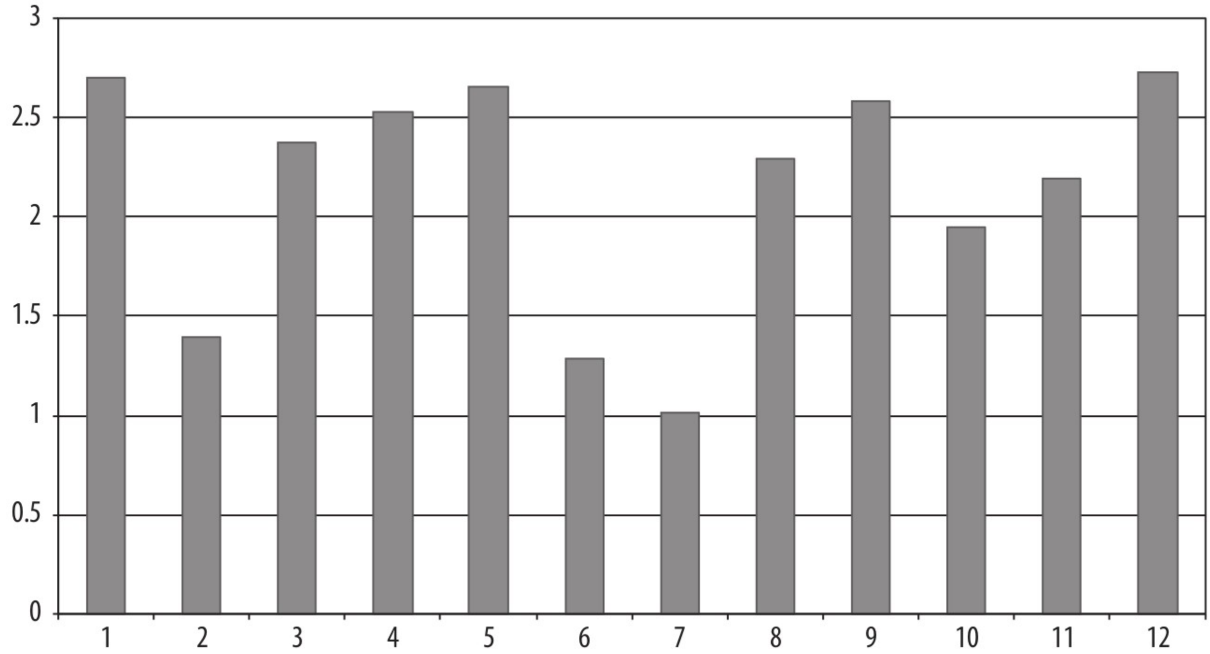
Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

The percentage of daily price changes in the first quarter is 1.28, 1.29 in the second, 1.58 in the third, and 1.49 in the fourth. The greatest price change seems to be concentrated in the third quarter. Table 3.3 shows that all categories but personal products have the greatest number of price changes in the third quarter, although there is no clear tendency in the data. Therefore, we cannot conclude that seasonality exists in the frequency of price adjustments.

Nor do we observe a clear pattern of seasonality in the level of price adjustments. Figure 3.3 shows the rate of price growth conditional on price change by month. It should be stated that in Uruguay workers receive an extra half- month's wages in June and December. Also, during December's New Year festivities, supermarket sales generally receive a boost.[2] In summary, we

[2]In Uruguay, sales usually soar the day before supermarkets close for a holiday. January 1 and 6, May 1, and December 25 are usually the days when supermarkets do not open.

Figure 3.3: Price Growth Rate Giving Price Change, by Month (Percent)



Source: Authors' calculations based on data from the Ministry of Economy and Finance.

do not find demand-driven seasonal price changes in the data.

3.3.3 Individual Price Changes and Inflation Expectations

One interesting issue is whether price changes and inflation expectations move together. Ellis suggests a positive relationship between the frequency of price changes in his sample and respondents' expectations of inflation in a survey conducted by the Bank of England Ellis (2009). Table 3.4 shows the result of an ordinary least squares (OLS) regression estimation in which the dependent variable is the median probability of price change and the exploratory variables are expected inflation and indicator variables for the July 2007 tax reform. The expected inflation variable is the median forecast from a survey of experts conducted by the Central Bank of Uruguay. We include an indicator variable before and after the tax reform to capture anticipated effects of the reform.

The regression shows no correlation between changes in prices and inflation perceptions. If prices tended to be stickier, then inflation should not be expected to accelerate. It is interesting to point out that we observe a correlation between inflation and the percent variation in individual prices only when considering price decreases. The tax reform indicator variables suggest that supermarkets anticipated the reform and changed prices before the implementation of the reform in July 2007.

Table 3.4: Individual Price Changes and Inflation Perceptions: OLS Regression (April 2007 - December 2010[a])

	Dependent variable			
		Price change (percent)		
Variable	*Probability of price change*	*All*	*Increases*	*Decreases*
Expected yearly inflation	0.001	−0.024	0.449	−0.640***
	(0.001)	(0.412)	(0.369)	(0.194)
Tax reform indicator variable				
May 2007	0.008*	3.052*	3.659**	−1.043
	(0.004)	(1.792)	(1.604)	(0.844)
June 2007	0.012**	−4.102**	2.500	−0.288
	(0.004)	(1.790)	(1.602)	(0.843)
July 2007	0.011**	−1.371	−4.849***	2.740***
	(0.004)	(1.789)	(1.602)	(0.843)
August 2007	−0.018***	3.396*	−0.550	−1.401
	(0.004)	(1.793)	(1.605)	(0.845)
September 2007	−0.009***	−0.390	0.183	0.479
	(0.003)	(1.293)	(1.158)	(0.609)
Constant	−0.001	1.520	5.090**	−4.304***
	(0.007)	(2.780)	(2.488)	(1.309)
Observations	45	45	45	45
R^2	0.733	0.229	0.405	0.399

Source: Authors' calculations based on data from the Ministry of Economy and Finance.
a. Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3.5: Individual Price Changes and Inflation Expectations: OLS Regression by Product Category (April 2007 - December 2010[a])

	Dependent variable			
		Price change (percent)		
Category	*Probability of price change*	*All*	*Increases*	*Decreases*
Food	0.001	–0.168	0.700	–0.771***
	(0.001)	(0.522)	(0.456)	(0.221)
Soft drinks	–0.001	–1.644*	–1.678	0.393
	(0.001)	(0.924)	(1.997)	(0.513)
Alcohol	0.003	0.298	0.256	–0.064
	(0.002)	(0.790)	(0.781)	(0.552)
Personal products	0.003**	0.839	0.195	–0.602
	(0.001)	(0.527)	(0.477)	(0.361)
Observations	45	45	45	45

Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance and the Central Bank of Uruguay.
a. Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

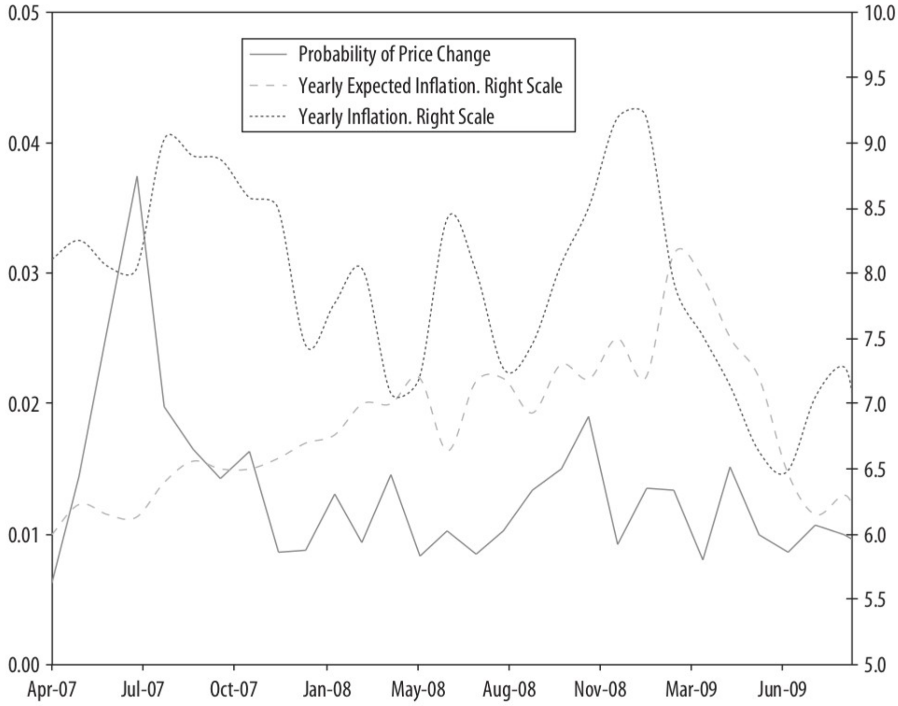
Source: Authors' calculations based on data from the Ministry of Economy and Finance.
a. Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

For a better understanding of the relationship between individual daily prices and inflation, we estimate the previous equation by product category. Table 3.5 shows the results of the coefficient on expected inflation. Interestingly, results indicate that there is a positive association between probability of price changes and expected inflation only for the personal product category. For the other product categories, the correlation is zero. That means that expectations about future inflation do not influence the price strategies of supermarkets in those markets. We do find an association between changes in prices and the average rate of price decreases in the food product category. To provide more evidence for this topic, figure 3.4 plots the probability of price adjustment (left scale) and the inflation and expected inflation rates (right scale). We observe no association between price changes and inflation perceptions.

3.3.4 Price Changes by Day of the Month

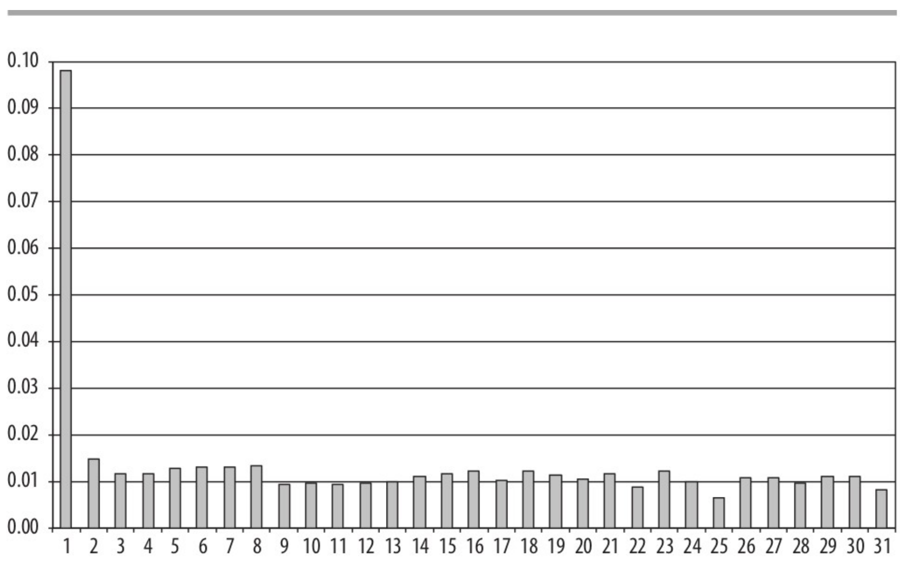
Given that we have daily data, we can analyze the pricing decisions of super- markets by day of the month. Figure 3.5 shows the probability of a price change by day of the month. Interestingly, the probability of price change is nine times higher on the first day of the month than on any other day. Figure 3.6 plots the daily probability of a price change from the second day to the last day of the month. In this case, we do not observe a clear pattern in the data.

Figure 3.4: Probability of Price Change, Inflation, and Expected Inflation



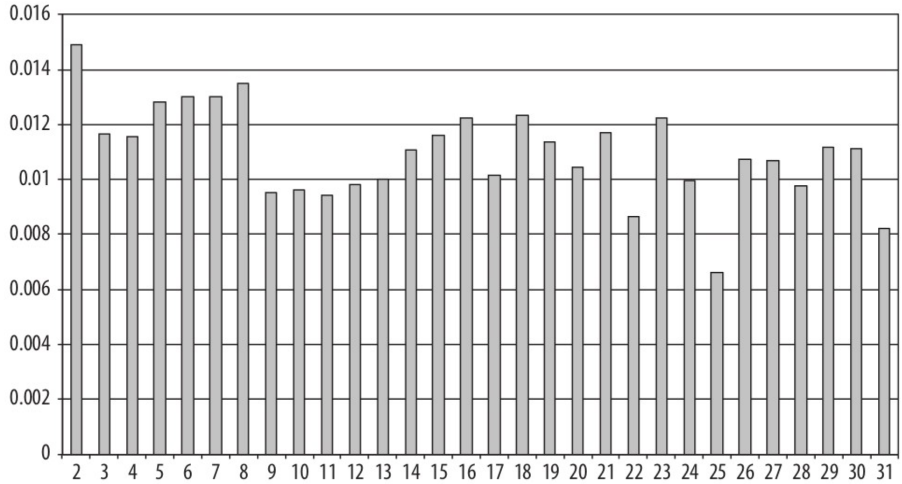
Source: Authors' calculations base on data from the Ministry of Economy and Finance and the Central Bank of Uruguay.

Figure 3.5: Probability of Price Change, by Day of Month



Source: Authors' calculations base on data from the Ministry of Economy and Finance.

Figure 3.6: Probability of Price Change, by Day 2 to Day 31



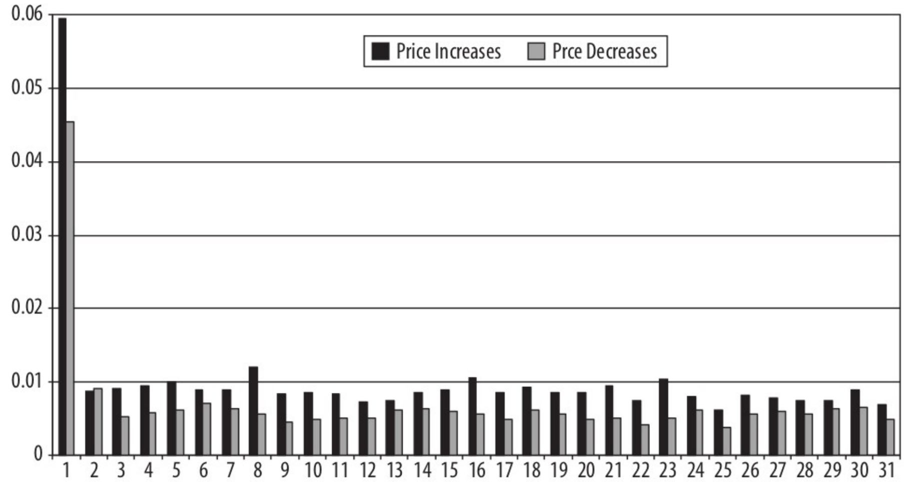
Source: Authors' calculations base on data from the Ministry of Economy and Finance.

Figure 3.7 shows that price increases and decreases also are concentrated on the first day of the month. In addition, figure 3.8 shows that the finding that price changes are concentrated on the first day of the month is a general result, valid for all product categories. This is one of the most remarkable findings of our study, since to the best of our knowledge no other study analyzes the distribution of price changes by day of the month. One supermarket manager told us that this pricing behavior is related to producers, who tend to adjust their prices the first day of the month. In this case, the observed behavior could be a response to cost increases by supermarkets. The pattern is the same for price increases and price decreases. As price decreases are associated with sales, this implies that supermarkets tend to follow a pattern of price changes that concentrates most of them in one day, which may indicate the existence of menu costs associated with price change for supermarkets or some other rigidity that prevents the supermarkets from changing prices.

3.3.5 Hazard Rate Estimates

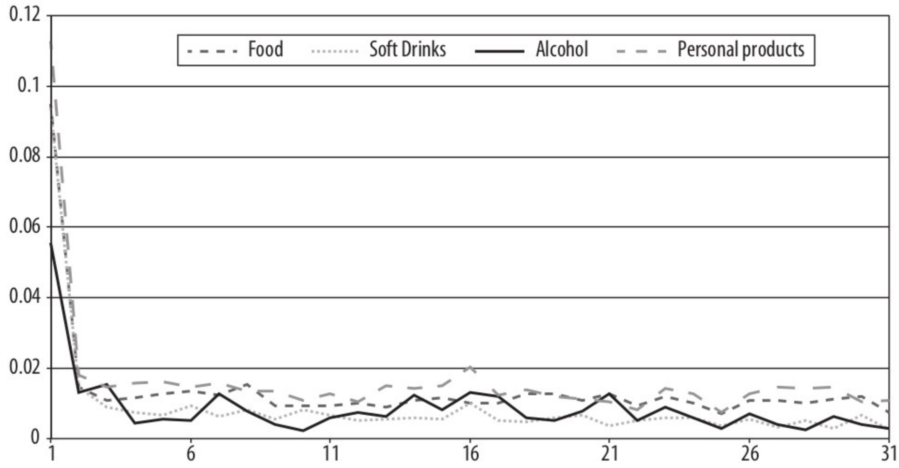
In order to study whether price changes are time dependent, we estimate the hazard rate. The hazard rate at moment t is calculated as the quotient of the number of prices that change at t, given that they do not change until that moment, over the number of prices that have not changed until moment t. As the greatest price duration is half a year (see appendix 3.8) we calculate the hazard function up to 200 days. Figure 3.9 shows the smoothed hazard rates. We observe a hazard rate that is not constant over time. This result is consistent with results in Nakamura (2008) and Ellis (2009), although the authors find hazard rates to be decreasing and we find them to be increasing. The upward-sloping hazard rate is consistent with state-dependent pricing. This fact invalidates the modeling of a constant probability of price change

Figure 3.7: Probability of Price Increases and Decreases, by Day of Month



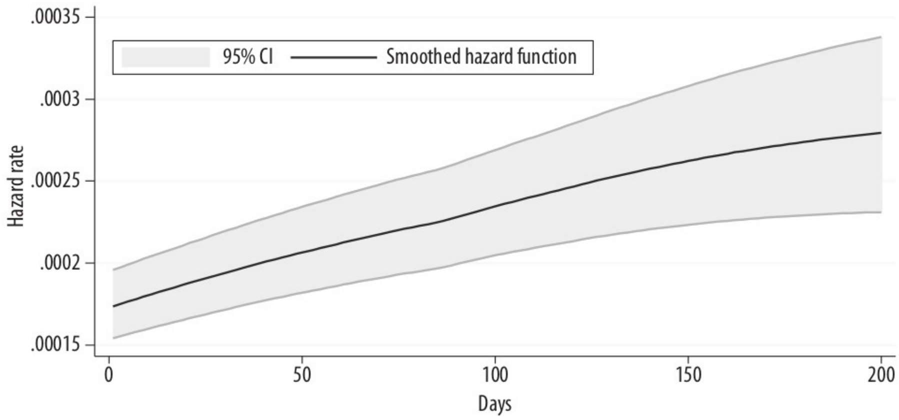
Source: Authors' calculations base on data from the Ministry of Economy and Finance.

Figure 3.8: Daily Probability of Price Change, by Product Category



Source: Authors' calculations base on data from the Ministry of Economy and Finance.

Figure 3.9: Smoothed Hazard Estimate



Source: Authors' calculations base on data from the Ministry of Economy and Finance.

and implies that supermarkets do not follow a time-dependent strategy for price setting. In turn, this result is in line with our finding of no seasonality in price changes.

3.3.6 Price Synchronization

We estimate price synchronization in two ways: across stores that belong to the same chain and across stores in each city. To estimate price synchronization we calculate the Fisher and Konieczny (2000) estimator (FK). Table 3.6 indicates that price changes across supermarkets of the same chain are highly synchronized.[3] For this result, two remarks are in order. First, our database consists of daily observations, and we find that prices change on average after about 2.5 months. Second, we also find that price changes are concentrated on the first day of the month. Therefore, our database has a great deal of synchronized "no price changes" and consequently a high FK. To control for this effect, we also estimate the FK synchronization indicator, conditional on price change (see table 3.7).

In this case, the synchronization estimates are lower than before, but the main result of high synchronization of price adjustments in supermarkets that belong to the same chain remains. This result is in contrast to that in Chaumont, Fuentes, Labbé, and Naudon (2011), which finds much lower price synchronization for Chile. In addition, we estimate the FK synchronization indicator across the cities in our sample. Figure 3.10 shows the FK estimator for each city. As can be seen, synchronization is by itself large, with a minimum of 0.63 for Montevideo—which has the greatest number of supermarkets—and 1 for a large number of cities that have few supermarkets.

[3]We estimate the FK indicator just for the major chains: those that have more than five stores and more than three cashiers per store on average.

Table 3.6: Price Synchronization across Stores That Belong to the Same Chain

Chain	**Fisher and Konieczny indicator**
Devoto	0.94
Tienda Inglesa	0.92
Macromercado	0.96
El Dorado	0.92
Multiahorro	0.91
Disco	0.96
Ta Ta	0.84

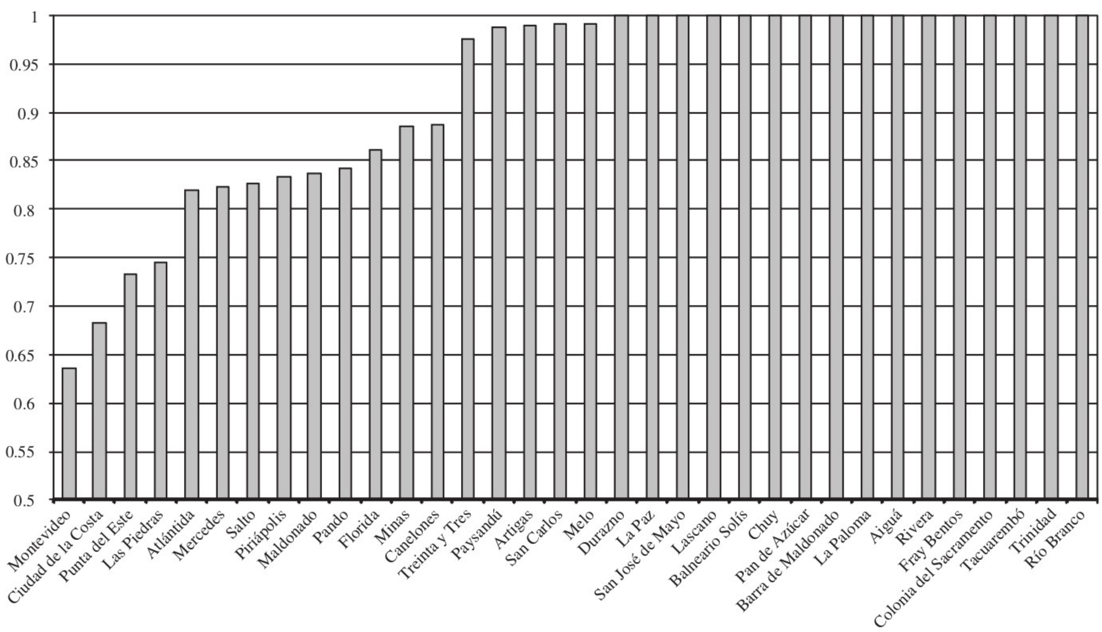
Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

Table 3.7: Adjusted Price Synchronization across Stores That Belong to the Same Chain, Conditional on Price Change

Chain	**Synchronization indicator**
Devoto	0.54
Tienda Inglesa	0.56
Macromercado	0.75
El Dorado	0.51
Multiahorro	0.56
Disco	0.61
Ta Ta	0.36

Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

Figure 3.10: Fisher and Konieczny Synchronization Indicator, by City



Source: Authors' calculations base on data from the Ministry of Economy and Finance.

Table 3.8: Stylized Facts and Model Features

Fact	Consistent Features
Price changes are somewhat flexible	Small menu costs
No seasonality of price changes	State-dependent models
Price changes are mainly on the first day of the month	Time-dependent models
Upward-sloping hazard rates	State-dependent models
Price changes are highly synchronized	State-dependent models/common shocks/strategic complementarities

Source: Authors' calculations based on data from the Uruguayan Ministry of Economy and Finance.

3.4 Comparing Results with Theory

Here we compare the results of the analysis with the main theoretical predictions of menu costs and time-dependent and state-dependent theories, discussing each stylized fact in the analysis and how it fits the theoretical explanations. Table 3.8 presents a brief summary of the analysis, in a vein similar to that of table 14 inKlenow and Malin (2010). As can be seen in the table, the empirical evidence seems to point to state-dependent models as the main explanation for the inflation phenomena in Uruguay. The flexibility of prices remains a disputed issue in the empirical literature; as we have considered sales in our database, the relative flexibility could be less if we take them out.

Our results, unlike those in the empirical literature, found high synchronization of prices even at the chain and city level. That result could be driven by the particularity of our database, which consists of daily observations. In the same vein, we discovered that prices tend to change on the first day of the month. This result suggests that common shocks may be an important part of price adjustment policies of supermarkets.

We think that this result could not be explained in full using macro models. As all the items in our database are the highest-selling brands and most markets are oligopolies—even in the supermarket industry—price-setting behavior needs to be analyzed using micro modeling. As for the matter of prices changing mostly on the first day of the month, we think that this could serve as a reference point for price setting by supermarkets. Setting prices on this particular day, in turn, could reduce menu costs in the event of price changes.

3.5 Conclusions

This paper presents evidence on price formation at the retail level in Uruguay, drawn from a rich and unique data set of 30 million daily prices in grocery stores and supermarkets across the country, to analyze the behavior of consumer prices. We find that retail prices in Uruguay change frequently. Prices are less sticky than in the United States and Brazil but stickier than in the United Kingdom and Chile. The median duration of prices in Uruguay is 2.5 months.

We do not find evidence of a seasonal pattern in the adjustment of prices. The probability of price changes varies positively with expected inflation only for the personal products category. However, for the food category we find an association between price changes and the percentage rate of price decreases. In addition, we find that the probability of price changes on the first day of the month is nine times higher than on any other day of the month and the probability of price adjustments is not constant over time. Finally, we find very high synchronization of price changes.

This evidence seems to point to a state-dependent model of price changes. Nonetheless, the high synchronization of price changes is a newer element in the empirical literature, which could be the result of analyzing daily data. Last, the high concentration of price changes on the first day of the month needs further theoretical analysis, as one possible interpretation could be that this day serves as a reference point for price adjustment.

Bibliography

Barros, R., M. Bonomo, and S. Santos (2009): "Price Setting in a Variable Macroeconomic Environment: Evidence from Brazilian CPI," Discussion paper, Getulio Vargas Foundation and Federal Reserve Bank of New York.

Cavallo, A. (2010): "Scraped Data and Sticky Prices," *MIT Sloan Working Paper 4976-12*.

Chaumont, G., M. Fuentes, F. Labbé, and A. Naudon (2011): "A Reassessment of Flexible Price Evidence using Scanner Data: Evidence from an Emerging Economy," Documento de Trabajo 641, Banco Central de Chile.

Ellis, C. (2009): "Do supermarket prices change from week to week?," Bank of England working papers 378, Bank of England.

Fisher, T. C., and J. D. Konieczny (2000): "Synchronization of price changes by multiproduct firms: evidence from Canadian newspaper prices," *Economics Letters*, 68(3), 271 – 277.

Gopinath, G., and R. Rigobon (2008): "Sticky Borders," *The Quarterly Journal of Economics*, 123(2), 531–575.

Klenow, P. J., and O. Kryvtsov (2008): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?," *The Quarterly Journal of Economics*, 123(3), 863–904.

Klenow, P. J., and B. A. Malin (2010): "Microeconomic Evidence on Price-Setting," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3 of *Handbook of Monetary Economics*, chap. 6, pp. 231–284. Elsevier.

Medina, J. P., D. Rappoport, and C. Soto (2007): "Dynamics of Price Adjustment: Evidence from Micro Level Data for Chile," Discussion Paper Working Paper 432, Central Bank of Chile.

Nakamura, E. (2008): "Pass-Through in Retail and Wholesale," *American Economic Review*, 98(2), 430–37.

Nakamura, E., and J. Steinsson (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.

Appendix

3.6 Cities included in the Study

3.7 List of Products

Product	Brand	Specification	Share in CPI (percent)	Category
Beer	Pilsen	0.96 L	0.3	Alcohol
Wine	Roses	1 L	0.34	Alcohol
Wine	Santa Teresa Clasico	1 L	0.34	Alcohol
Wine	Tango	1 L	0.34	Alcohol
Beef (peceto)	No brand	1 Kg	0.9	Food
Beef (nalga)	With bone, no brand	1 Kg	0.43	Food
Beef (nalga)	Boneless, no brand	1 Kg	0.43	Food
Beef (aguja)	With bone, no brand	1 Kg	0.86	Food
Beef (aguja)	With bone, no brand	1 Kg	0.86	Food
Beef (paleta)	With bone, no brand	1 Kg	n/i	Food
Beef (rueda)	With bone, no brand	1 Kg	n/i	Food
Ground beef	Up to 20 percent fat	1 Kg	0.29	Food
Ground beef	Up to 5% fat	1 Kg	0.29	Food
Bread	No brand	1 unit (≈ 0.215 Kg)	1.21	Food
Brown eggs	El Ecologito	1/2 dozen	0.34	Food
Brown eggs	El Jefe	1/2 dozen	0.34	Food
Brown eggs	Prodhin	1/2 dozen	0.34	Food
Butter	Calcar	0.2 Kg	0.15	Food
Butter	Conaprole sin sal	0.2 Kg	0.15	Food
Butter	Lacteria	0.2 Kg	0.15	Food
Cacao	Copacabana	0.5 Kg	0.04	Food
Cacao	Vascolet	0.5 Kg	0.04	Food
Cheese	Cerros del Este	1 Kg	0.23	Food
Cheese	Dispnat	1 Kg	0.23	Food
Chicken	Avicola del Oeste	1 Kg	0.64	Food
Chicken	Tenent	1 Kg	0.64	Food
Coffee	Aguila	0.25 Kg	0.1	Food
Coffee	Chana	0.25 Kg	0.1	Food
Dulce de leche	Conaprole	1 Kg	0.14	Food
Dulce de leche	Los Nietitos	1 Kg	0.14	Food
Dulce de leche	Manjar	1 Kg	0.14	Food
Flour	Canuelas	1 Kg	0.16	Food
Flour	Cololo	1 Kg	0.16	Food
Flour	Puritas	1 Kg	0.16	Food
Frankfurters	Cattivelli	8 units (≈ 0.340 Kg)	0.26	Food
Frankfurters	Ottonello	8 units (≈ 0.340 Kg)	0.26	Food
Frankfurters	Schneck	8 units (≈ 0.340 Kg)	0.26	Food

Product	Brand	Specification	Share in CPI (percent)	Category
Beer	Patricia	0.96 L	0.3	Alcohol
Grated cheese	Conaprole	0.08 Kg	0.15	Food
Grated cheese	El Trebol	0.08 Kg	0.15	Food
Grated cheese	Milky	0.08 Kg	0.15	Food
Semolina noodles	Adria	0.5 Kg	n/i	Food
Semolina noodles	Las Acacias	0.5 Kg	n/i	Food
Ham	Centenario	1 Kg	0.21	Food
Ham	La Constancia	1 Kg	0.21	Food
Ham	Schneck	1 Kg	0.21	Food
Margarine	Danica dorada	0.2 Kg	0.02	Food
Margarine	Doriana nueva	0.25 Kg	0.02	Food
Margarine	Primor	0.25 Kg	0.02	Food
Mayonnaise	Fanacoa	0.5 Kg	0.09	Food
Mayonnaise	Hellmans	0.5 Kg	0.09	Food
Mayonnaise	Uruguay	0.5 Kg	0.09	Food
Noodles	Cololo	0.5 Kg	0.3	Food
Peach jam	Dulciora	0.5 Kg	0.17	Food
Peach jam	Limay	0.5 Kg	0.17	Food
Peach jam	Los Nietitos	0.5 Kg	0.17	Food
Peas	Arcor	0.35 Kg	0.05	Food
Peas	El Hogar	0.35 Kg	0.05	Food
Peas	Trofeo	0.35 Kg	0.05	Food
Quince jam	Los Nietitos	0.4 Kg	n/i	Food
Rice	Aruba tipo Patna	1 Kg	0.2	Food
Rice	Blue Patna	1 Kg	0.2	Food
Rice	Green Chef	1 Kg	0.2	Food
Rice	Pony	1 Kg	0.2	Food
Rice	Vidarroz	1 Kg	0.2	Food
Crackers	El Trigal	0.15 Kg	0.17	Food
Crackers	Famosa	0.14 Kg	0.17	Food
Crackers	Maestro Cubano	0.12 Kg	0.17	Food
Salt	Sek	0.5 Kg	0.05	Food
Salt	Torrevieja	0.5 Kg	0.05	Food
Salt	Urusal	0.5 Kg	0.05	Food
Semolina pasta	Adria	0.5 Kg	n/i	Food
Semolina pasta	Las Acacias	0.5 Kg	n/i	Food
Soybean oil	Condesa	0.9 L	n/i	Food
Sugar	Azucarlito	1 Kg	0.25	Food
Sugar	Bella Union	1 Kg	0.25	Food
Sunflower oil	Optimo	0.9 L	0.25	Food
Sunflower oil	Uruguay	0.9 L	0.25	Food
Tea	Hornimans	Box (10 units)	0.09	Food
Tea	La Virginia	Box (10 units)	0.09	Food
Tea	Lipton	Box (10 units)	0.09	Food

Product	Brand	Specification	Share in CPI (percent)	Category
Beer	Patricia	0.96 L	0.3	Alcohol
Tomato paste	Conaprole	1 L	0.08	Food
Tomato paste	De Ley	1 L	0.08	Food
Tomato paste	Qualitas	1 L	0.08	Food
Yerba	Canarias	1 Kg	0.34	Food
Yerba	Del Cebador	1 Kg	0.34	Food
Yerba	Sara	1 Kg	0.34	Food
Yogurt	Conaprole	0.5 Kg	0.06	Food
Yogurt	Parmalat (Skim)	0.5 Kg	0.06	Food
Bleach	Agua Jane	1 L	0.08	Personal
Bleach	Sello Rojo	1 L	0.08	Personal
Bleach	Solucion Cristal	1 L	0.08	Personal
Dishwashing detergent	Deterjane	1.25 L	0.2	Personal
Dishwashing detergent	Hurra Nevex Limon	1.25 L	0.2	Personal
Laundry soap	Drive	0.8 Kg	n/i	Personal
Laundry soap	Nevex	0.8 Kg	n/i	Personal
Laundry soap	Skip, Paquete azul	0.8 Kg	n/i	Personal
Laundry soap, in bar	Bull Dog	0.3 Kg (1 unit)	0.45	Personal
Laundry soap, in bar	Nevex	0.2 Kg (1 unit)	0.45	Personal
Shampoo	Fructis	0.35 L	n/i	Personal
Shampoo	Sedal	0.35 L	n/i	Personal
Shampoo	Suave	0.93 L	n/i	Personal
Soap	Astral	0.125 Kg	0.16	Personal
Soap	Palmolive	0.125 Kg	0.16	Personal
Soap	Suave	0.125 Kg	0.16	Personal
Toilet paper	Higienol Export	4 units (25 M each)	0.24	Personal
Toilet paper	Personal	4 units (25 M each)	0.24	Personal
Toilet paper	Sin Fin	4 units (25 M each)	0.24	Personal
Toothpaste	Closeup Triple	0.09 Kg	0.49	Personal
Toothpaste	Colgate Total	0.09 Kg	0.49	Personal
Toothpaste	Kolynos	0.09 Kg	0.49	Personal
Cola	Coca Cola	1.5 L	1.94	Soft drinks
Cola	Nix	1.5 L	1.94	Soft drinks
Cola	Pepsi	1.5 L	1.94	Soft drinks
Sparkling water	Matutina	2 L	0.7	Soft drinks
Sparkling water	Nativa	2 L	0.7	Soft drinks
Sparkling water	Salus	2.25 L	0.7	Soft drinks

3.8 Detailed Price Changes and Duration, by Product

Product	Brand	Probability of daily variation	Monthly price duration	Percentage decrease
Beer	Pilsen	0.009	3.5	23.2
Wine	Roses	0.008	4.0	22.1
Wine	Santa Teresa Clasico	0.012	2.7	38.3
Wine	Tango	0.011	2.9	39.4
Beef (peceto)	No brand	0.026	1.2	40.3
Beef (nalga)	With bone, no brand	0.027	1.2	43.1
Beef (nalga)	Boneless, no brand	0.015	2.2	34.2
Beef (aguja)	With bone, no brand	0.018	1.8	34.7
Beef (aguja)	With bone, no brand	0.027	1.2	40.1
Beef (paleta)	With bone, no brand	0.028	1.2	39.9
Beef (rueda)	With bone, no brand	0.013	2.5	34.2
Ground beef	Up to 20 percent fat	0.022	1.5	37.5
Ground beef	Up to 5% fat	0.019	1.7	36.6
Bread	No brand	0.011	2.9	28.6
Brown eggs	El Ecologito	0.007	5.0	24.7
Brown eggs	El Jefe	0.008	4.2	29.5
Brown eggs	Prodhin	0.012	2.8	33.8
Butter	Calcar	0.018	1.8	41.8
Butter	Conaprole sin sal	0.016	2.0	32.3
Butter	Lacteria	0.007	4.7	43.2
Cacao	Copacabana	0.011	2.9	34.4
Cacao	Vascolet	0.019	1.7	40.7
Cheese	Cerros del Este	0.068	0.5	45.0
Cheese	Dispnat	0.145	0.2	48.4
Chicken	Avicola del Oeste	0.041	0.8	42.8
Chicken	Tenent	0.039	0.8	44.6
Coffee	Aguila	0.009	3.7	34.0
Coffee	Chana	0.007	4.6	42.6
Dulce de leche	Conaprole	0.013	2.5	33.3
Dulce de leche	Los Nietitos	0.013	2.6	40.0
Dulce de leche	Manjar	0.013	2.6	31.4
Flour	Canuelas	0.027	1.2	43.7
Flour	Cololo	0.024	1.4	39.6
Flour	Puritas	0.015	2.2	36.3
Frankfurters	Cattivelli	0.010	3.2	45.7
Frankfurters	Ottonello	0.012	2.7	42.4
Frankfurters	Schneck	0.015	2.1	36.1
Grated cheese	Conaprole	0.009	3.8	25.1
Grated cheese	El Trebol	0.009	3.5	36.9
Grated cheese	Milky	0.015	4.4	30.0
Semolina noodles	Adria	0.015	2.2	36.6

Product	Brand	Probability of daily variation	Monthly price duration	Percentage decrease
Beer	Patricia	0.008	3.9	20.4
Semolina noodles	Las Acacias	0.019	1.7	40.2
Ham	Centenario	0.008	4.2	29.0
Ham	La Constancia	0.034	1.0	46.1
Ham	Schneck	0.015	2.2	35.8
Margarine	Danica dorada	0.012	2.7	39.0
Margarine	Doriana nueva	0.013	2.6	42.6
Margarine	Primor	0.016	2.1	41.2
Mayonnaise	Fanacoa	0.011	3.0	39.5
Mayonnaise	Hellmans	0.021	1.5	41.9
Mayonnaise	Uruguay	0.024	1.3	42.3
Noodles	Cololo	0.017	1.9	38.8
Peach jam	Dulciora	0.012	2.6	35.9
Peach jam	Limay	0.008	4.1	30.4
Peach jam	Los Nietitos	0.011	3.0	37.9
Peas	Arcor	0.010	3.3	42.9
Peas	El Hogar	0.009	3.5	25.3
Peas	Trofeo	0.017	1.9	44.4
Quince jam	Los Nietitos	0.011	2.9	38.6
Rice	Aruba tipo Patna	0.018	1.8	43.4
Rice	Blue Patna	0.025	1.4	41.4
Rice	Green Chef	0.027	1.2	42.6
Rice	Pony	0.009	3.5	41.1
Rice	Vidarroz	0.012	2.7	49.3
Crackers	El Trigal	0.009	3.6	32.4
Crackers	Famosa	0.010	3.2	29.5
Crackers	Maestro Cubano	0.012	2.6	41.1
Salt	Sek	0.011	3.1	41.9
Salt	Torrevieja	0.007	4.7	30.4
Salt	Urusal	0.012	2.7	41.7
Semolina pasta	Adria	0.015	2.2	35.6
Semolina pasta	Las Acacias	0.018	1.9	41.1
Soybean oil	Condesa	0.029	1.1	56.2
Sugar	Azucarlito	0.017	1.9	35.3
Sugar	Bella Union	0.017	2.0	34.7
Sunflower oil	Optimo	0.033	1.0	42.1
Sunflower oil	Uruguay	0.032	1.0	40.9
Tea	Hornimans	0.009	3.5	46.5
Tea	La Virginia	0.010	3.2	46.8
Tea	Lipton	0.009	3.8	40.6
Tomato paste	Conaprole	0.017	1.9	36.3
Tomato paste	De Ley	0.012	2.7	34.4
Tomato paste	Qualitas	0.012	2.8	45.8
Yerba	Canarias	0.013	2.5	38.1

Product	**Brand**	**Probability of daily variation**	**Monthly price duration**	**Percentage decrease**
Beer	Patricia	0.008	3.9	20.4
Yerba	Del Cebador	0.013	2.5	36.4
Yerba	Sara	0.015	2.2	40.4
Yogurt	Conaprole	0.013	2.6	29.5
Yogurt	Parmalat (Skim)	0.012	2.8	34.1
Bleach	Agua Jane	0.018	1.7	37.7
Bleach	Sello Rojo	0.015	2.2	33.6
Bleach	Solucion Cristal	0.018	1.8	43.3
Dishwashing detergent	Deterjane	0.025	1.3	44.1
Dishwashing detergent	Hurra Nevex Limon	0.025	1.4	43.3
Laundry soap	Drive	0.015	2.2	43.1
Laundry soap	Nevex	0.023	1.4	44.8
Laundry soap	Skip, Paquete azul	0.018	1.8	45.3
Laundry soap, in bar	Bull Dog	0.016	2.0	39.6
Laundry soap, in bar	Nevex	0.015	2.2	39.8
Shampoo	Fructis	0.022	1.5	44.5
Shampoo	Sedal	0.016	2.1	47.3
Shampoo	Suave	0.011	3.0	45.0
Soap	Astral	0.018	1.8	46.3
Soap	Palmolive	0.023	1.4	50.0
Soap	Suave	0.013	2.5	46.6
Toilet paper	Higienol Export	0.016	2.1	32.7
Toilet paper	Personal	0.013	2.5	31.8
Toilet paper	Sin Fin	0.021	1.6	41.8
Toothpaste	Closeup Triple	0.009	3.7	38.1
Toothpaste	Colgate Total	0.023	1.4	39.1
Toothpaste	Kolynos	0.013	2.5	34.6
Cola	Coca Cola	0.010	3.3	25.5
Cola	Nix	0.008	4.0	34.6
Cola	Pepsi	0.010	3.2	31.7
Sparkling water	Matutina	0.011	3.0	43.0
Sparkling water	Nativa	0.007	4.6	27.0
Sparkling water	Salus	0.013	2.6	35.0

Source: Authors' elaboration based on data from the Ministry of Economy and Finance.