UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

PRICE CHANGES AND COMPETITION BETWEEN FRENCH RETAILERS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS APLICADAS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

PABLO NICOLÁS URIBE PIZARRO

PROFESOR GUÍA:
ALEJANDRO JOFRÉ CÁCERES
PROFESOR CO-GUÍA:
ALEJANDRO BERNALES SILVA

COMISIÓN:
MARCELA VALENZUELA BRAVO

Este trabajo ha sido parcialmente financiado por:
CMM ANID PIA AFB170001, CMM ANID BASAL ACE210010
y CMM ANID BASAL FB210005,
Institute for Research in Market Imperfections and Public Policy
(ICM IS130002, Ministerio de Economía de Chile),
y Fondecyt project 1190162.

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR AL TÍTULO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA
POR: **PABLO NICOLÁS URIBE PIZARRO**
FECHA: 2022
PROF. GUÍA: Alejandro Jofré Cáceres

PRICE CHANGES AND COMPETITION BETWEEN FRENCH RETAILERS

El análisis de los cambios de precios es fundamental para la evaluación de diferentes aspectos de la macroeconomía y, en particular, para el diseño de políticas monetarias óptimas [16]. En el contexto del mercado minorista, muchos estudios han analizado empíricamente la frecuencia y la magnitud de los cambios de precios [1, 21, 6, 10, 16], mientras que diferentes modelos han abordado teóricamente la dispersión y los ajustes de precios en ambientes competitivos [18, 19, 22, 11]. No obstante, hay pocos estudios que aborden la competencia entre cadenas de retail de manera empírica [3]. Por otro lado, la competencia y los cambios de precio también se han estudiado en el mercado de opciones mediante modelos que estiman empíricamente las componentes asociadas a dichos cambios [7, 13, 9, 14]. Inspirados en estas ideas, el objetivo principal de esta tésis es estudiar la competencia entre cadenas de retail a través de un modelo para analizar cómo las tiendas responden a eventos de cambio de precio en función de la distancia geográfica entre ellas, mediante la descomposición del impacto de un cambio en tres componentes: *Init*, *Close* y *Far*. La primera es percibida por las tiendas que inician una competencia cambiando su precio de venta, mientras que las dos últimas son percibidas por las tiendas cercanas y lejanas respectivamente. Para estimar estas componentes, analizamos empíricamente datos diarios de 1000 productos y más de 1000 tiendas pertenecientes a 11 cadenas de retail diferentes en Francia. Nuestros resultados evidencian que las tiendas de diferentes cadenas de retail tienden a competir más con tiendas ubicadas en su cercanía que con tiendas ubicadas lejos de ellas. Además, analizamos a través de la misma metodología otros atributos relacionados con cada tienda, como el número de diferentes cadenas de retail presentes en la zona, la concentración local del mercado y la cuota local del mercado, así como también, el efecto del precio de los productos en la competencia. Descubrimos que estos atributos producen efectos competitivos análogos a aquellos encontrados al comparar la competencia entre tiendas cercanas y lejanas.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN CIENCIAS
DE LA INGENIERÍA
POR: **PABLO NICOLÁS URIBE PIZARRO**
FECHA: 2022
PROF. GUÍA: Alejandro Jofré Cáceres

PRICE CHANGES AND COMPETITION BETWEEN FRENCH RETAILERS

The analysis of price changes is crucial for the assessment of different issues in macroeconomics, and in particular for the design of optimal monetary policies [16]. In the context of the retail market, many studies have analysed empirical data to study price changes' frequency and size [1, 21, 6, 10, 16], while different models have theoretically addressed price dispersion and adjustments in competitive settings [18, 19, 22, 11]. Nonetheless, there is little work addressing competition between retailers through empirical data [3]. On the other hand, price changes and competition have also been addressed in the equity options market by models that empirically estimate components of price changes [7, 13, 9, 14]. Inspired by these ideas, the main purpose of this thesis is to study competition between retailers through a novel model for analysing how stores respond to price change events based on the geographical distance between them, by decomposing the price impact of a price change into three components: *Init*, *Close* and *Far*. The first is perceived by initial stores that start a competition by changing their selling price for a given product, while the last two are faced by close and far stores respectively. To estimate these components, we analyse daily data from 1000 products and more than 1000 stores belonging to 11 different retail chains across France. Our results provide evidence that stores from different retail chains tend to compete more with stores situated in their proximity than with stores located far away. Furthermore, we analyse through the same methodology other features related to each store, such as the local number of different retails, the local market concentration and the local market share, and also the effect of products' prices. We find that these features produce analogous competitive effects during price change events than those found between *Close* and *Far* stores.

Ils disent que l'amour rend aveugle
Mais il t'a redonné la vue
Il t'as fait muer quand ta rage était sourde
Il a fait fredonner la rue

— Nekfeu

Agradecimientos

Gracias:

A mi familia por apoyarme siempre y preocuparse que nunca me faltara nada.

A los cabros del DIM: Mati, Nano, Flora, Freddy, Feña, PL, Danner, Kike, Juanito. A Alonso, por ser mi vecino, mi amigo y mi hermano. Nunca más nos quedaremos dormidos para una prueba.

A Agustin, Seba y Roco, mis primeros amigos de la universidad que alegraron mis días de mechón.

À Clément et Daniel, mes potes à Centrale avec lesquels j'ai vécu l'une des plus belles expériences de ma vie et un tiers de mon parcours universitaire.

A los cabros de escalada: Benja, Srdjan, Anto, Provi, Gabo. Si sumara el esfuerzo de todas las rutas y boulders ya tendríamos varias tésis. Gracias por apañar al desestrés.

A Barbi, por ser un apoyo incondicional y darme siempre tu amor. Gracias por motivarme a hacer todos los esfuerzos para lograr esta tésis.

A todos quienes me guiaron en mi experiencia laboral siendo tutores y referentes para mí. Matthieu Favreau, Corrado Carbone, Pedro Ruiz, Abelino Jiménez.

A todos quienes me acompañaron en este proceso, infinitas gracias.

Table of Contents

Introduction **1**

1. Background **4**
1.1. Stochastic Processes . . . 4
1.2. Basics of Graph Theory . . . 4

2. Price Changes and Competition **6**
2.1. The Idea . . . 6
2.1.1. Relationship with other works . . . 8
2.2. Mathematical Formulation . . . 9
2.2.1. Price Process and Public Information . . . 10
2.2.2. Distance between stores . . . 10
2.2.3. Initial Price Changes . . . 12
2.2.4. Competitive, Item, Time and Retail Effects . . . 15
2.2.5. Weak and Strong Response Conditions . . . 15
2.3. Other Proxys of Competition . . . 16
2.3.1. N Retail . . . 17
2.3.2. HHI . . . 18
2.3.3. Market Share . . . 19
2.3.4. Luxury Products . . . 19

3. Data and Implementation **21**
3.1. The Dataset . . . 21
3.1.1. Price Change Frequency . . . 22
3.1.2. Regular Prices . . . 23
3.1.3. Initial Price Changes . . . 24
3.1.4. Response Probability . . . 25
3.2. Implementation . . . 25
3.2.1. Chunking . . . 25
3.2.2. Pipeline . . . 27

4. Results **29**
4.1. Competitive Environment . . . 29
4.1.1. Close and Far Effects . . . 29
4.1.2. N Retail . . . 32
4.1.3. HHI . . . 35
4.2. Market Power . . . 37

4.3. Luxury Products 39

Conclusion **42**

Bibliography **43**

List of Tables

3.1. Summary Statistics of Price Changes for each category. Price changes were obtained using parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$ and filtered using the *week* conditional formulation from subsection 2.2.5. Observations below the 2nd and above 98th percentiles are removed, as well as products with less than 10 price changes. 21

3.2. Missing days and mean price change frequencies. Results were obtained by computing the change frequency for each product and retail chain on daily data, and then averaging by retail chain. 22

3.3. Response probabilities for conditional categories. In this experiment $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. Note that values in the n_{weak} column do not coincide with those in Table 3.1 since extreme values and items are not filtered for this calculation. 25

4.1. α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*. 30

4.2. α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*. 30

4.3. α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 7$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*. 31

4.4. α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 40$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*. 31

4.5. α coefficients of equations (2.44) & (2.45) and summary statistics for N Retail competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{retail} and L_{retail} store groups obtained using the median of N_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{retail}) are different from the respective coefficients in the right panel (L_{retail}) 33
4.6. α coefficients of equations (2.44) & (2.45) and summary statistics for N Retail competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{retail} and L_{retail} store groups obtained using the 1st and 3rd quartiles of N_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{retail}) are different from the respective coefficients in the right panel (L_{retail}) 34
4.7. α coefficients of equations (2.44) & (2.45) and summary statistics for *HHI* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{HHI} and L_{HHI} store groups obtained using the median of HHI_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{HHI}) are different from the respective coefficients in the right panel (L_{HHI}) 36
4.8. α coefficients of equations (2.44) & (2.45) and summary statistics for *HHI* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{HHI} and L_{HHI} store groups obtained using the 1st and 3rd quartiles of HHI_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{HHI}) are different from the respective coefficients in the right panel (L_{HHI}) 37
4.9. α coefficients of equations (2.54) & (2.55) and summary statistics for *MS* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{MS} and L_{MS} store groups obtained using the median of MS_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{MS}) are different from the respective coefficients in the right panel (L_{MS}) 38
4.10. α coefficients of equations (2.54) & (2.55) and summary statistics for *MS* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{MS} and L_{MS} store groups obtained using the 1st and 3rd quartiles of MS_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{MS}) are different from the respective coefficients in the right panel (L_{MS}) 39

4.11. α coefficients and summary statistics for P competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_P and L_P item groups obtained using the median of P_i in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_P) are different from the respective coefficients in the right panel (L_P) 40
4.12. α coefficients and summary statistics for P competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_P and L_P item groups obtained using the 1st and 3rd quartiles of P_i in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_P) are different from the respective coefficients in the right panel (L_P) 41

List of Figures

2.1. Price Changes of *Init*, *Close* and *Far* stores. 7
2.2. Geographical configuration of stores *Init*, *Close* and *Far*. 7
2.3. Example of a store s and its competitive perimeter. Colours (shapes) represent stores from different retails. The closest stores from each retail within distance δ_c are marked with a dotted circumference, representing the close stores of s. On the other hand, the furthest stores from each retail between δ_c and δ_f are marked with a circumference, representing the far stores of s. 11
2.4. Example of Initial Price Changes. We assume that the competitive perimeter of store A is given by stores B and C. The classification is done from store A's reference. Store A (blue, thick) starts with an initial price change, and stores B (red, dashed) and C (violet, dotted) respond within period dt, changing their price too. Afterwards, when period dt is over, store B makes another price change, but it happens within the cool down time span t_c of the last price change, not fulfilling requirement 2. The same happens for the next price change done by store A. Finally, store C makes a price change after the cool down period t_c of the last price change, so it is considered initial. Note that from stores A and B references, the classification of the same price change events can be different depending on the stores belonging to their competitive perimeter. 13
3.1. Example of missing values for one product and four different stores in our dataset. Solid lines represent known prices, while dotted lines represent missing values. 22
3.2. Example of latent regular prices derived from price trajectories. 24
3.3. Example of initial price changes 24
3.4. Connected Components of the Distance Graph G using $\delta_f = 20$ and approximate borders of France. 26
3.5. Connected Components of the Distance Graph G using $\delta_f = 60$ and approximate borders of France. 27
3.6. Pipeline describing each step for obtaining the final model coefficients. 28

Introduction

The analysis of price changes plays an important role for the assessment of different issues in macroeconomics, such as the welfare consequences of business cycles, the behaviour of real exchange rates, and optimal monetary policy [16]. In particular, sticky prices are an important element of Keynesian economics [1], as they affect the short-term impact of nominal interest rates on real activity and the response to inflation to monetary policy. Thus, many economical models include nominal rigidity as a feature, assuming that companies are unable to freely adjust their prices [5].

In the case of retailers, many studies have analysed empirical data to better understand the nature of price changes' frequency and size. Blinder et al. (1998) [1] found that the median price in the United States changes once a year, being consistent with the work from Taylor (1999) [21], and Druant et al. (2005) [6], who found the same result for the euro area. However, evidence has suggested that considering sales and temporary promotions is relevant when measuring the frequency of price changes. For example, Kehoe and Midrigan (2007) [10] found that, when excluding sales, the duration of prices varies from four to five months, compared to three weeks if sales are included. On the other hand, Nakamura and Steinsson (2008) [16] found that the median duration of prices varies from eight to eleven months when excluding sales. Other studies have focused on understanding the different factors that affect price changes. For instance, evidence of the complexity of changing prices has been provided by Levy et al. (1997) [12], showing that a supermarket with higher menu costs changes prices two and a half less frequently than other chains. Moreover, evidence has shown how retailers tend to change their prices as consequence of economical factors such as taxes [17], inflation variability and regulation [2]. Nonetheless, internal features of the market can also make prices change, as recent research has shown that prices tend to fluctuate more as products become more homogeneous and more players enter into the industry. Therefore, understanding the relationship between price rigidity, product homogeneity and competition between retailers crucial for designing antitrust and monetary policies [20].

Particularly, competition has been theoretically addressed by several models of price dispersion and adjustment. For example, Salop and Stiglitz (1976) [18] formulated a spatial price dispersion model, this is, a model in which several stores offer a same product simultaneously at a different price. In their model, heterogeneity of consumers rationality and costly-information gathering makes firms that equilibrium of the market does not happen at the perfectly competitive price, but for some configuration of parameters equilibrium is found where some fraction of stores sell at the competitive price, while others sell at a higher price. Moreover, Shilony et al. (1977)[19] proposed a game-theoretic approach in which consumers can buy without extra cost from neighbourhood stores, but incur into a transportation cost if they decide to visit distant stores in search for lower prices. On the other hand, stores use sales as a marketing device for stimulating consumers venture to their location. Shilony et al. found no Nash equilibrium for the game in pure pricing strategies, however they demonstrated the existence of an equilibrium mixed strategy, in which firms randomise their prices causing temporal price dispersion. Following these ideas, Varian adressed in 1980 the question of sales equilibria through a model that may be regarded as a combination of both models described previously [22]. Another game-theoretical study was done by Lal and Matutes (1989) [11]. In their setting, they analysed the effects of a multiproduct competition on prices and profits in a static duopoly with complete information in equilibrium. The resulting price dispersion for this equilibrium shows that stores collectively discriminate between two types of consumers: the rich (who has high willingness to pay but with high opportunity cost of time) and the poor (who has low willingness to pay but zero opportunity cost of time), providing a possible explanation to the phenomenon of loss leaders, a special promotional activity widely used by retailers.

Price changes and competition has also been addressed in the equity options market. For example Madhavan and Smidt (1991) [13] proposed a model of intraday price formation where market makers' beliefs evolve following Bayes' rule. In their work, they estimate the model's components empirically using a dataset from a New York Stock Exchange (NSYE) specialist. Moreover, Huang and Stoll (1997) [9] constructed and estimated a basic trade indicator model for identifying three components of the spread: order processing, adverse information and inventory holding cost, finding that these components are a function of the trade size. They used a dataset of all trades and quotes for 20 large NYSE stocks in 1992 for estimating these components. Later, Madhavan et al. (1997) [14] proposed an intraday price formation model that incorporates both public information shocks and microstructure effects in the security market. Their work helps to understand the sources of intraday price volatility and the effect on information flows on stock prices over the day. They estimate the components of their model using transaction level data for stocks listed on the NYSE. Overall, these works present different methodologies for measuring empirically the different components of price changes.

However, most of previous research focuses on studying the frequency of price changes of retail products from monthly and weekly data, or addressing competition through game-theoretic approaches that explain static and temporal price dispersion in equilibrium. In this thesis, we study the competition between retailers through a novel model for analysing how stores respond to competitive price changes based on competition effects depending on the geographical distance between them. To do so, we introduce the concept of *initial price change events*, which refers to the specific events in which the change of a product's price generates competitive responses among other stores. Explicitly, we address the following questions:

- *How does the distance between stores affects their responses to competitive price change events?.*
- *Which other factors affect retailers' responses to competitive price change events?.*

Our model decomposes the price impact of a price change into three components: *Init*, *Close* and *Far*. The first is perceived by initial stores that start a competition by changing their selling price for a given product, while the last two are faced by close and far stores respectively, which receive instant information about the competitive price change event and adjust their selling prices accordingly. In this work, our goal is to estimate these components empirically for the retail market. By analysing daily data from 1000 products and more than 1000 stores belonging to 11 different retail chains across France, our results provide evidence that stores from different retail chains tend to compete more with stores situated in their proximity than with stores located far away. Moreover, we examine how other features related to the competitive environment of each store, such as the local number of different retails and the market concentration affect the respective responses to initial price changes. Finally, we analyse the local market share and the product price as features for the same purpose.

This thesis is organised as following. In Chapter 1 we introduce the mathematical foundations of our model. Subsequently, in Chapter 2 we present the main idea of our model and we formalise initial price change events mathematically. Later, in Chapter 3 we explain how we implemented our model for our particular dataset. Afterwards, in Chapter 4 we present the results obtained by our model, and finally we conclude and discuss about our findings.

Chapter 1

Background

The idea of this chapter is to introduce the mathematical concepts used in the development of this theses. We start by reviewing the basic concepts of stochastic processes, describing in particular the notion of *Stopping Time*. Next, we review the basic ideas of graph theory, and particularly the notion of *Connected Component* of a graph.

1.1. Stochastic Processes

Definition 1.1 ***(Stochastic Process)*** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and T a set. A stochastic process indexed by T with values in a measurable space (S, Σ) is a collection of random variables $(X_t)_{t \in T}$ defined in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in (S, Σ).*

Definition 1.2 ***(Filtration)*** *A filtration of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing sequence $(\mathcal{F}_t)_{t \in T}$ of sub sigma-algebras of $\mathcal{F}$ indexed by a totally ordered set T. In this case $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathbb{P})$ is called a filtered probability space.*

Definition 1.3 ***(Adapted Process)*** *A process $(X_t)_{t \in T}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the measurable space (S, Σ) is adapted to the filtration $(\mathcal{F}_t)_{t \in T}$ if X_t is a $(\mathcal{F}_t, \Sigma)$-measurable function for each $t \in T$. In that case, we say that $(X_t)_{t \in T}$ is a $(\mathcal{F}_t)_{t \in T}$-adapted stochastic process.*

Definition 1.4 ***(Stopping Time)*** *A random variable τ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathbb{P})$ with values in $\mathbb{N} \cup \{+\infty\}$ is called a stopping time (with respect to the filtration $\mathbb{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$, if $\{\tau = n\} \in \mathcal{F}_n$ for all n.*

1.2. Basics of Graph Theory

In this section we present we basic notions of graph theory that were used for this work. In particular, the following definitions were closely followed from the book *Graph Theory* by Reinhard Diestel [4].

Definition 1.5 ***(Graph)*** *A graph is a pair $G = (V, E)$ of sets such that $E \subseteq V \times V$. The elements of V are called the vertices of the graph G, while the elements of E are its edges. Moreover, a graph with vertex set V is said to be a graph on V. $V(G)$ denotes the vertex set* *of a graph G, and $E(G)$ its edge set.*

Definition 1.6 ***(Subgraph)*** *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs. If $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$, then G_1 is a subgraph of G_2 (and G_2 a supergraph of G_1), written as $G_1 \subseteq G_2$.*

Definition 1.7 ***(Induced Subgraph)*** *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs. If $G_1 \subseteq G_2$ and G_1 contains all the edges $xy \in E_2$ such that $x, y \in V_1$, then G_1 is an induced subgraph of G. We say that V_1 induces or spans G_1 in G_2, and write $G_1 =: G_2[V_1]$.*

Definition 1.8 ***(Path)*** *A path is a non-empty graph $P = (V, E)$ of the form:*

$$V = \{x_0, x_1, ..., x_k\}$$

$$E = \{x_0x_1, x_1x_2, ..., x_{k-1}x_k\}$$

Given sets A, B of vertices, we call $P = x_0...x_k$ an A-B path if $V(P) \cap A = \{x_0\}$ and $V(P) \cap B = \{x_k\}$.

Definition 1.9 ***(Connectivity)*** *A non-empty graph G is called connected if any two of its vertices are linked by a path in G. If $U \subseteq V(G)$ and $G[U]$ is connected, we also call U itself connected (in G).*

Definition 1.10 ***(Connected Component)*** *Let $G = (V, E)$ be a graph. A maximal connected subgraph of G is called a connected component of G.*

Chapter 2

Price Changes and Competition

In this Chapter, we introduce our model of *Price Changes and Competition between retailers* based on the mathematical notions described previously. First, we introduce the main idea of our model through a case example. Afterwards, we explain the relationship of our model with other works, and finally we formalise our model using the notions introduced in Chapter 1.

2.1. The Idea

This subsection explains the intuition behind the proposed model, while the next section introduces it formally. Consider the following simplified framework. There are three competitive stores belonging to (for simplicity) different retail chains that sell item i at a given (for simplicity, equal) price at day $t-1$. Imagine that one store, which we call store *Init*, at day t increases the price of item i by Δp_t^{Init}:

$$\Delta p_t^{Init} = p_t^{Init} - p_{t-1}^{Init} \tag{2.1}$$

Let's assume that store *Init* does not change the price of item i again between days t and $t+dt$, and therefore the price change in period dt, given by $\underset{dt}{\Delta} p_t^{Init}$, is equal to Δp_t^{Init}. This is:

$$\Delta p_t^{Init} = \underset{dt}{\Delta} p_t^{Init} := p_{t+dt}^{Init} - p_{t-1}^{Init} \tag{2.2}$$

Moreover, let's call the other two stores respectively *Close* and *Far*, based on their graphical distance with respect to store *Init*: store *Close* (*Far*) is closer to (far away from) store *Init*. The other stores learn from the price change in store *Init* at time t. Thus, stores *Close* and *Far* adjust their prices to reflect the new information in period dt by $\underset{dt}{\Delta} p_{Close}$ and $\underset{dt}{\Delta} p_{Far}$ respectively:

$$\underset{dt}{\Delta} p_t^{Close} := p_{t+dt}^{Close} - p_{t-1}^{Close} \tag{2.3}$$

$$\underset{dt}{\Delta} p_t^{Far} := p_{t+dt}^{Far} - p_{t-1}^{Far} \tag{2.4}$$

Figure 2.2 describes this setup graphically, while Figure 2.1 reflects price changes of the stores over time.

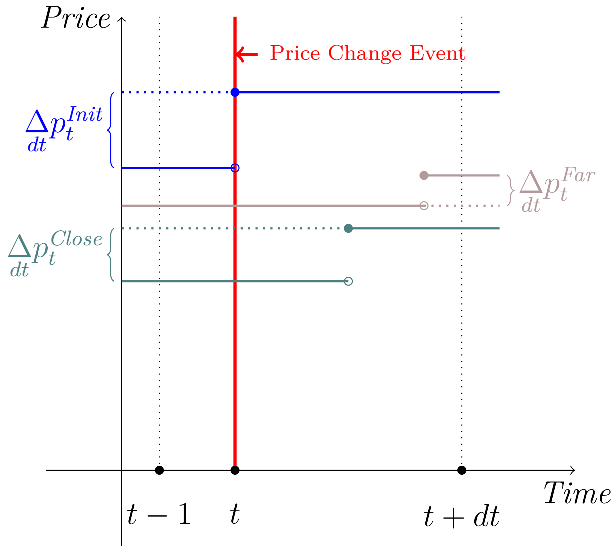

Figure 2.1: Price Changes of *Init*, *Close* and *Far* stores.

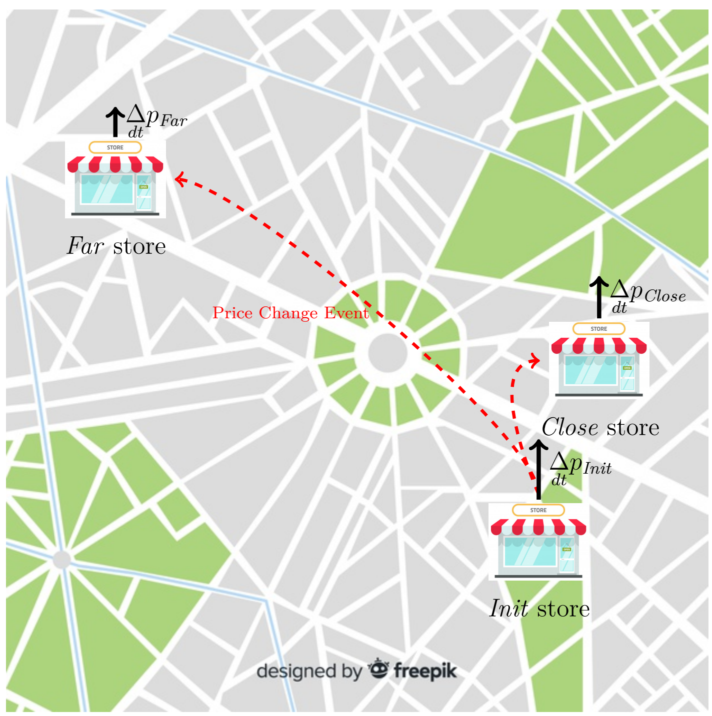

Figure 2.2: Geographical configuration of stores *Init*, *Close* and *Far*.

Given this framework, our model allows the event of a price change to have a competitive effect associated to the distance to the initial price change (α), after controlling for time, retail and item fixed effects (FE_t, FE_r and FE_i respectively), separating between positive and negative price changes. This is, in the case of a positive initial price change:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{+Init} = \alpha^{+Init} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.5}$$

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{+Close} = \alpha^{+Close} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.6}$$

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{+Far} = \alpha^{+Far} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.7}$$

Where $+Close$ and $+Far$ are the close and far responses to a positive initial price change (but not necessarily positives themselves). Symmetrically, in the case of a negative change on prices, the model is:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{-Init} = \alpha^{-Init} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.8}$$

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{-Close} = \alpha^{-Close} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.9}$$

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{t,i}^{-Far} = \alpha^{-Far} + \text{FE}_i + \text{FE}_t + \text{FE}_r + \epsilon \tag{2.10}$$

Where $-Close$ and $-Far$ are the close and far responses to a negative initial price change (but not necessarily negatives themselves). This example considers a single price change, but, because individual price responses are very noisy, taking an average over a large number of price changes is required to estimate the components. The same general framework can be adapted to investigate competition between and even within retail chains, as well as heterogeneous effects resulting from producers' market power or consumer biases (e.g. rational inattention toward less often consumed items).

Note that the division between positive (price increases) and negative (price cuts) changes is crucial for the analysis. Indeed, we expect that stores in the group *Close* should be more competitive than stores in the group *Far*, and therefore, that $|\alpha^{-Close}| > |\alpha^{-Far}|$, meaning that *Close* stores respond more aggressively to price cuts than stores in the group *Far* to offer competitive prices. In addition, we expect that $\alpha^{+Close} < \alpha^{-Far}$, meaning that stores in the group *Close* respond less aggressively to price increases than stores in the group *Far* to offer competitive prices. This asymmetry in price change responses allows to measure competitiveness and can be applied to other proxys of competition (see Section 2.3).

2.1.1. Relationship with other works

The work of Shilony et al. (1977) [19] was probably one of the first models in which the physical distance between stores plays a key role in competition and price setting strategies. As mentioned previously, in their model consumers incur into transportation costs when they decide to buy from distant stores, while stores use sales to attract consumers. However, in their setting, this transportation cost is fixed and unique, which may differ from reality since transportation efforts are bigger when distances are longer. In our model, we account for this limitation explicitly by finding empirical evidence that *Close* and *Far* stores have different pricing strategies in response to competitive price changes.

Concerning the price change dynamics, in many models for the equity options market (see for example [15]), the event of a trade by a market-maker produces an impact on ask prices of market makers. Thus, when a trade occurs at time t, both the market-maker responsible of the trade and other market-makers are able to respond between time t and $t + dt$. In contrast, in our model stores face an impact on their selling price due to the event of a price change between time $t-1$ and t by a store *Init* of the competition. Therefore, since store *Init* has already changed its selling price by time t, to set equal response periods for *Init*, *Close* and *Far* stores, price responses are measured between time $t - 1$ and $t + dt$. In this way, we capture both the price change of store *Init* between $t - 1$ and t, and the price changes of stores *Close* and *Far* between t and $t+dt$. Moreover, if we admit some regularity of prices, we make sure that store *Init* does not change its price again between t and $t+dt$ (see subsection 3.1.2).

2.2. Mathematical Formulation

In this section we formalise the idea presented in section 2.1 mathematically. In our model, stores' selling prices are stochastic processes that change randomly over time. Thus, competitive price change events will be captured by a stopping time process that will indicate moments in which stores face competitive impacts on their prices. Moreover, distance between stores determine which of them compete with each other. Finally, we use a fixed effects model to estimate the different competitive components introduced previously. We start this section with some preliminary definitions. Let:

- $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.
- $\mathcal{S} = \{s_1, ..., s_N\}$ be a finite set of stores.
- $\mathcal{I} = \{i_1, ..., i_M\}$ be a finite set of items sold in each store.
- $\mathcal{T} = \{0, 1, 2, ..., T\}$ be an index set corresponding to the discrete time interval from 0 to T.
- $\mathcal{R}$ be a partition of $\mathcal{S}$, called the set of retail chains. For each store s, $[s]$ denotes the equivalence class of s under $\mathcal{R}$ (i.e. the retail chain of s).
- $d : \mathcal{S} \times \mathcal{S} \to \mathbb{R}^+$ be a distance function between stores such that $\forall s_1, s_2, s_3 \in \mathcal{S} : s_2 \neq s_3 \Rightarrow d(s_1, s_2) \neq d(s_1, s_3)$. This means that for a given store s_1, there are no two distinct stores s_2, s_3 with equal distance to s_1.
- $Categories = \{+Init, -Init, +Close, -Close, +Far, -Far\}$ be the set of competitive categories.

2.2.1. Price Process and Public Information

Definition 2.1 ***(Price Process and Price Changes)*** *For each item $i \in \mathcal{I}$ and store $s \in \mathcal{S}$, let $\{p_{t,i,s}\}_{t\in\mathcal{T}}$ be a stochastic process on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{R}^+$ and indexed by $\mathcal{T}$, called the price process of item i at store s. For each $1 < t \leq T$, we define the (normalised) price change at time t by:*

$$\left(\frac{\Delta p}{p}\right)_{t,i,s} = \frac{p_{t,i,s} - p_{t-1,i,s}}{p_{t-1,i,s}} \tag{2.11}$$

With the initial condition:

$$\left(\frac{\Delta p}{p}\right)_{0,i,s} = 0 \tag{2.12}$$

In fact, to exclude price changes due to sales and inventory issues from the analysis, the price change process $\{p_{t,i,s}\}_{t\in\mathcal{T}}$ can be replaced in this definition and every following definition by the derived latent regular price process $\{r_{t,i,s}\}_{t\in\mathcal{T}}$.

Definition 2.2 ***(Public Information)*** *The public information at time t is defined as the sigma algebra generated by the price processes of all items and stores until time t:*

$$F_t = \sigma(p_{k,i,s} : k \leq t, i \in \mathcal{I}, s \in \mathcal{S}) \tag{2.13}$$

Moreover, $\mathbb{F} := (F_t)_{t\in\mathcal{T}}$ is a filtration, which we call the filtration of public information.

Property 2.2.1. $\left\{\left(\frac{\Delta p}{p}\right)_{t,i,s}\right\}_{t\in\mathcal{T}}$ *is an $\mathbb{F}$-adapted stochastic process.*

2.2.2. Distance between stores

In this work, we argue that distance between stores plays an important roll in competition. Thus, for each store $s \in \mathcal{S}$, the idea is to separate its environment in two groups: *Close* and *Far* stores. However, stores *too* far from s should not be considered (for instance, a store in Paris does not compete with stores in Marseille). In this subsection we formalise these ideas through the following definitions:

Definition 2.3 ***(Close and Far Stores)*** *For each $s \in \mathcal{S}$ we define the close stores of s as:*

$$c_s = \left\{\hat{s} \in \mathcal{S} : d(s, \hat{s}) \leq \delta_c\right\} \tag{2.14}$$

Where $\delta_c \in \mathbb{R}^+$ is a fixed parameter. We define the far stores of s as:

$$f_s = \left\{\hat{s} \in \mathcal{S} : \delta_c < d(s, \hat{s}) \leq \delta_f\right\} \tag{2.15}$$

Where $\delta_f \in \mathbb{R}^+$ is a fixed parameter such that $\delta_f > \delta_c$.

An important fact is that not every store in these sets should be considered in the competition of s. Indeed, to avoid biases from global company policies of price changes across the country, we assume that competitive stores in *Close* and *Far* groups belong to different retail chains with respect to s. Additionally, to avoid over-weighting bigger retail chains, we consider only one store per retail chain. In case there are several stores belonging to the same retail chain in the *Close* group, we keep only the closest from store s, and in case there are several stores belonging to the same retail chain in the *Far* group, we keep only the furthest one from store s. We reformulate the previous definitions by incorporating these considerations, introducing *Competitive Close* and *Far* stores:

Definition 2.4 ***(Competitive Close and Far Stores)*** *For each $s \in \mathcal{S}$ we define the competitive close stores of s as:*

$$\mathcal{C}_s = \left\{ \hat{s} \in c_s : \hat{s} \notin [s] \wedge \hat{s} = \arg \min_{s' \in c_s \cap [\hat{s}]} d(s, s') \right\} \tag{2.16}$$

Where $\delta_c \in \mathbb{R}^+$ is a fixed parameter. We define the competitive far stores of s as:

$$\mathcal{F}_s = \left\{ \hat{s} \in f_s : \hat{s} \notin [s] \wedge \hat{s} = \arg \max_{s' \in f_s \cap [\hat{s}]} d(s, s') \right\} \tag{2.17}$$

Furthermore, we define the competitive perimeter of s as:

$$\mathcal{B}_s = \mathcal{C}_s \cup \mathcal{F}_s \cup \{s\} \tag{2.18}$$

Note that in the previous definition minimums and maximum are unique (if they exist) since for each store s, there are no two distinct stores s_2, s_3 with equal distance to s. On the other hand, $\mathcal{B}_s$ represents the set of stores that will produce competitive price changes from s perspective. Figure 2.3 shows a visual example of a store s and its competitive perimeter.

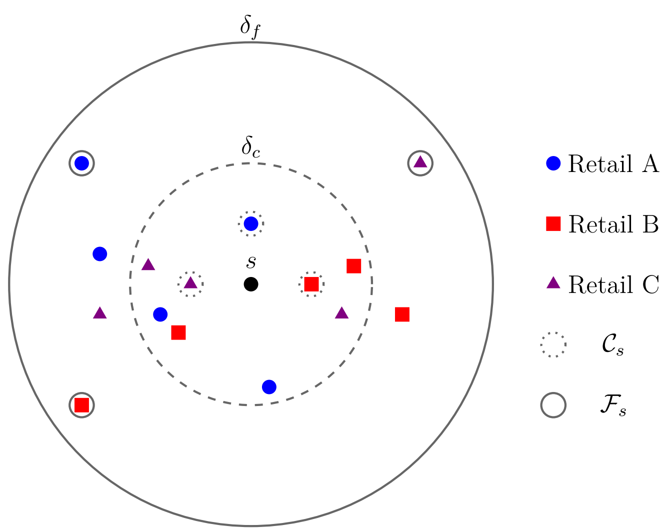

Figure 2.3: Example of a store s and its competitive perimeter. Colours (shapes) represent stores from different retails. The closest stores from each retail within distance δ_c are marked with a dotted circumference, representing the close stores of s. On the other hand, the furthest stores from each retail between δ_c and δ_f are marked with a circumference, representing the far stores of s.

2.2.3. Initial Price Changes

In this section, our goal is to define the times $\tau \in \mathcal{T}$ corresponding to price changes that lead to competition effects, which we call *initial* price changes. One important fact is that since stores have bounded competition perimeters (by parameter δ_f), the distinction of initial price change times is relative to each store. For this purpose we first define the sets of stores in the competition perimeter of a store s that change their price at a given time t:

Definition 2.5 ***(Price Change stores)*** *For each store $s \in \mathcal{S}$, we define $\mathcal{P}^{+}_{t,i,s}$ the positive price change stores at time t of item i with respect to s as:*

$$\mathcal{P}^{+}_{t,i,s} = \left\{ \hat{s} \in \mathcal{B}_s : \left(\frac{\Delta p}{p}\right)_{t,i,\hat{s}} > 0 \right\} \tag{2.19}$$

And the negative price change stores at time t of item i with respect to s as:

$$\mathcal{P}^{-}_{t,i,s} = \left\{ \hat{s} \in \mathcal{B}_s : \left(\frac{\Delta p}{p}\right)_{t,i,\hat{s}} < 0 \right\} \tag{2.20}$$

Finally, the set of price change stores at time t of item i with respect to s is defined as:

$$\mathcal{P}_{t,i,s} = \mathcal{P}^{+}_{t,i,s} \cup \mathcal{P}^{-}_{t,i,s} \tag{2.21}$$

Note that the event $\{\mathcal{P}_{t,i,s} \neq \varnothing\}$ indicates that store s has perceived a price change for item i within its competition perimeter at time t. In the following lines we explain our procedure for determining which of these times are actually related to initial price changes. Indeed, we assume that an initial price change should fulfil two requirements:

- **Req. 1:** An initial price change can only occur after (mininum) a response period dt from the respective last initial price change.
- **Req. 2:** An initial price change can only occur after (mininum) a cool down period t_c from the last price change (not necessarily initial).

Intuitively, the first requirement implies that price changes that occur right after an initial price change are responses to this change, which should be later classified in the *Close* or *Far* categories. The second requirement indicates that initial price changes should not be close in time to previous price changes. This allows to dismiss price changes that might be responses to previous changes, rather than lead to new competitions. Figure 2.4 shows an example of price changes for three different stores.

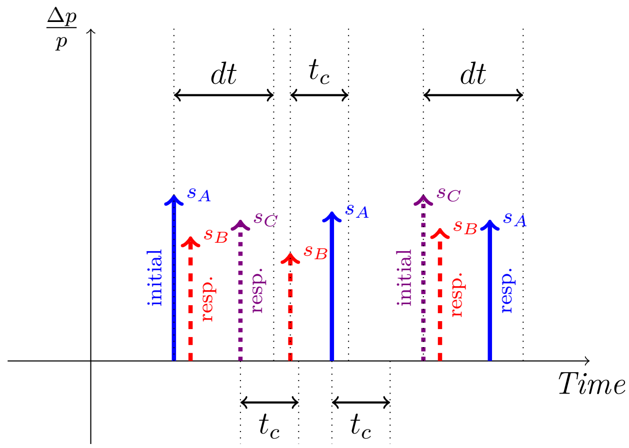

Figure 2.4: Example of Initial Price Changes. We assume that the competitive perimeter of store A is given by stores B and C. The classification is done from store A's reference. Store A (blue, thick) starts with an initial price change, and stores B (red, dashed) and C (violet, dotted) respond within period dt, changing their price too. Afterwards, when period dt is over, store B makes another price change, but it happens within the cool down time span t_c of the last price change, not fulfilling requirement 2. The same happens for the next price change done by store A. Finally, store C makes a price change after the cool down period t_c of the last price change, so it is considered initial. Note that from stores A and B references, the classification of the same price change events can be different depending on the stores belonging to their competitive perimeter.

The previous ideas motivate the following definitions:

Definition 2.6 ***(Last Price Change Time)*** *For each $t \in \mathcal{T}$, we define $\mathcal{L}_{t,i,s}$, the last price change time of item $i \in \mathcal{I}$ with respect to store $s \in \mathcal{S}$ at time t as:*

$$\mathcal{L}_{t,i,s} = \max\left\{\hat{t} < t : \mathcal{P}_{\hat{t},i,s} \neq \varnothing\right\} \tag{2.22}$$

Where $\max \emptyset = -\infty$.

Property 2.2.2. *For all $t \in \mathcal{T}$, $\mathcal{L}_{t,i,s}$ is F_t-measurable.*

Using this previous definition, we can now define initial price change times:

Definition 2.7 ***(Initial Price Change Times)*** *For each $1 \leq n \leq T$, we define $\tau_{n,i,s}$, the n^{th}-initial price change time of item $i \in \mathcal{I}$ with respect to store $s \in \mathcal{S}$ as:*

$$\tau_{n,i,s} = \min\left\{t > \tau_{n-1,i,s} + dt : t > \mathcal{L}_{t,i,s} + t_c \wedge \mathcal{P}_{t,i,s} \neq \varnothing\right\} \tag{2.23}$$

Where $\min \emptyset = \infty$ and the initial condition is set to $\tau_{0,i,s} = -\infty$.

Note that the event $\{t > \tau_{n-1,i,s} + dt\}$ represents the first requirement (**Req. 1**), while the event $\{t > \mathcal{L}_{t,i,s} + t_c\}$ the second one (**Req. 2**). The following property assures that stores are able to identify initial price changes within their competitive perimeter in real time:

Property 2.2.3. *For each $1 \leq n \leq T$, $\tau_{n,i,s}$ is a stopping time with respect to $\mathbb{F}$.*

Now that we have defined the initial price changes that a store s perceives, our goal is to determine for each initial price change time τ whether s was the initiator (*Init*) of the price change, or whether it was a *Close* or *Far* store responding to another store.

Definition 2.8 *We define the set of finite initial price change times of item $i \in \mathcal{I}$ with respect to store $s \in \mathcal{S}$ as:*

$$\mathcal{T}_{i,s} = \Big\{\tau_{k,i,s} : k \leq T, |\tau_{k,i,s}| < \infty\Big\} \tag{2.24}$$

Definition 2.9 *For each category cat $\in$ Categories, we define the set of (cat)-price change times of item $i \in \mathcal{I}$ with respect to store $s \in \mathcal{S}$ respectively as:*

$$\mathcal{T}_{i,s}^{+Init} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}^{+}_{\tau,i,s}} d(\hat{s}, s) = s\Big\} \tag{2.25}$$

$$\mathcal{T}_{i,s}^{+Close} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}_{\tau,i,s}} d(\hat{s}, s) \in \mathcal{C}_s \cap \mathcal{P}^{+}_{\tau,i,s}\Big\} \tag{2.26}$$

$$\mathcal{T}_{i,s}^{+Far} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}_{\tau,i,s}} d(\hat{s}, s) \in \mathcal{F}_s \cap \mathcal{P}^{+}_{\tau,i,s}\Big\} \tag{2.27}$$

And in the case of negative price change times:

$$\mathcal{T}_{i,s}^{-Init} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}^{-}_{\tau,i,s}} d(\hat{s}, s) = s\Big\} \tag{2.28}$$

$$\mathcal{T}_{i,s}^{-Close} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}_{\tau,i,s}} d(\hat{s}, s) \in \mathcal{C}_s \cap \mathcal{P}^{-}_{\tau,i,s}\Big\} \tag{2.29}$$

$$\mathcal{T}_{i,s}^{-Far} = \Big\{\tau \in \mathcal{T}_{i,s} : \arg\min_{\hat{s} \in \mathcal{P}_{\tau,i,s}} d(\hat{s}, s) \in \mathcal{F}_s \cap \mathcal{P}^{-}_{\tau,i,s}\Big\} \tag{2.30}$$

Note that:

- τ is an *Init* positive (negative) price change time for store s and item i, if τ is an initial price change time and s increased (decreased) the price of i at time τ.
- τ is a *Close* positive (negative) price change time for store s and item i, if τ is an initial price change time, and the closest store within the competitive perimeter of s that changed the price of i at time τ belongs to the *Competitive Close* stores of s, and it increased (decreased) the price of i.
- τ is a *Far* positive (negative) price change time for store s and item i, if τ is an initial price change time, and the closest store within the competitive perimeter of s that changed the price of i at time τ belongs to the *Competitive Far* stores of s, and it increased (decreased) the price of i.

The intuition is that if two stores within the competitive perimeter of s change their price at the same time, we consider that s responds to the closest one. Furthermore, note that being positive or negative price change times depends on the direction of the 'initiator' of the initial price change, and not on the direction of the response. Finally, the classification and identification of initial price change events depends completely on the reference of each store s.

Property 2.2.4. *For all $s \in \mathcal{S}$ and $i \in \mathcal{I}$, $\left\{\mathcal{T}_{i,s}^{cat}\right\}_{cat \in Cat}$ is a partition of $\mathcal{T}_{i,s}$.*

The previous property assures that the classification of initial price change times is well done, in the sense that every initial price change time belongs to one and only one of the competitive categories.

Definition 2.10 *For each category $cat \in$ Categories, store $s \in \mathcal{S}$ and item $i \in \mathcal{I}$, we define the cat-price change process, of item i with respect to store s as:*

$$\left\{ \left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} \right\}_{\tau \in \mathcal{T}_{i,s}^{cat}} \tag{2.31}$$

Where for all τ such that $\tau + dt \leq T$:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \frac{p_{\tau+dt,i,s} - p_{\tau-1,i,s}}{p_{\tau-1,i,s}} \tag{2.32}$$

And for τ such that $\tau + dt > T$ we carry forward the last price observation:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \frac{p_{T,i,s} - p_{\tau-1,i,s}}{p_{\tau-1,i,s}} \tag{2.33}$$

2.2.4. Competitive, Item, Time and Retail Effects

As explained previously in section 2.1, our model allows the event of a price change to have a competitive effect, after controlling for time, item and retail fixed effects. For instance, for each store $s \in \mathcal{S}$, item $i \in \mathcal{I}$ and each competitive category $cat \in$ *Categories*, we formulate the following equation relating price changes and competitive effects:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \alpha^{cat} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in \mathcal{T}_{i,s}^{cat} \tag{2.34}$$

Where the term α^{cat} correspond to the *competitive effect* of category cat, FE_i is an *item fixed effect* associated with item i, and FE_τ is a *time fixed effect* of time τ and FE_r is a fixed effect associated with $r = [s]$, the retail chain of s. $\epsilon_{\tau,i,s}$ is an error term reflecting the fact that price responses may be affected by other factors.

2.2.5. Weak and Strong Response Conditions

While equation 2.34 holds for every competitive category, empirical evidence from our data shows that stores in categories *Close* and *Far* are very unlikely to respond to initial price changes (see Table 3.3), which is consistent with the low degree of price change synchronisation of retail products in the euro area [2]. This means that competitive, item, time and retail effects can be underestimated by $\mathbb{P}((\underset{dt}{\Delta}p/p)_{t,i,s} \neq 0)$. Thus, we reformulate equation (2.34) to account for this probability. We propose a weak conditional version, in which we exclude cases in which responses are zero, and a strong conditional version in which we also exclude cases in which responses do not follow the same direction as the initial price change. The weak conditional formulation is then:

$$weak(\mathcal{T}_{i,s}^{cat}) = \left\{\tau \in \mathcal{T}_{i,s}^{cat} : \left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} \neq 0\right\} \tag{2.35}$$

While the strong conditional formulation is:

$$strong(\mathcal{T}_{i,s}^{+Close}) = \left\{\tau \in \mathcal{T}_{i,s}^{+Close} : \left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} > 0\right\} \tag{2.36}$$

$$strong(\mathcal{T}_{i,s}^{-Close}) = \left\{\tau \in \mathcal{T}_{i,s}^{-Close} : \left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} < 0\right\} \tag{2.37}$$

$$strong(\mathcal{T}_{i,s}^{+Far}) = \left\{\tau \in \mathcal{T}_{i,s}^{+Far} : \left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} > 0\right\} \tag{2.38}$$

$$strong(\mathcal{T}_{i,s}^{-Far}) = \left\{\tau \in \mathcal{T}_{i,s}^{-Far} : \left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} < 0\right\} \tag{2.39}$$

Model (2.34) is then reformulated as:

$$\left(\frac{\underset{dt}{\Delta}p}{p}\right)_{\tau,i,s} = \alpha^{cat} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall\tau \in weak(\mathcal{T}_{i,s}^{cat}) \tag{2.40}$$

2.3. Other Proxys of Competition

In this section we propose other proxys of competition between stores from different retails. In particular, we define two additional features related to the store's competitive environment: local number of different retails (*N Retail*) and the local market concentration (measured through the *HHI*), one feature related to the store's market power: the market share (*MS*) measured in each store's proximity, and one feature related to the product characteristics: the product price (P). All of these features are considered *proxys* of competition, in the sense that they allow to differentiate different pricing strategies based on their values (high or low), similarly to the differences found between *Close* and *Far* stores given by distance. More details are presented in the following subsections:

2.3.1. N Retail

We propose the local number of different retails present in the proximity (*N Retail*) of each store as a measure of competition. The idea behind is that the higher is the local diversity of retail chains in the market, the higher is the competition between retailers in that particular zone. Therefore, stores with high *N Retail* should respond more aggressively to price cuts (and less aggressively to price increases) than stores with low *N Retail*. Note that this measure is related to the competitive environment of each store and the characteristics of the store's location.

Definition 2.11 ***(Number of retail chains)*** *For each store $s \in \mathcal{S}$, we define the number of retail chains of s as:*

$$N_s = \Big|\mathcal{R}_s\Big| \tag{2.41}$$

Where $\mathcal{R}_s := \Big\{[\hat{s}] : \hat{s} \in c_s\Big\}$ is the set of different retails present in the close stores of s.

Definition 2.12 ***(High and Low Retail stores)*** *We define the set of High Retail stores as:*

$$H_{retail} = \Big\{s \in \mathcal{S} : N_s \geq q_{High}\Big\} \tag{2.42}$$

And the set of Low Retail stores as:

$$L_{retail} = \Big\{s \in \mathcal{S} : N_s \leq q_{Low}\Big\} \tag{2.43}$$

Where $q_{High}, q_{Low} \in \mathbb{R}$ such that $q_{High} > q_{Low}$.

We formulate the following model for analysing differences in *Low* and *High Retail* stores' competitive pricing behaviour. For all $s \in H_{retail}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \alpha^{cat}_{H_{retail}} + \text{FE}_i + \text{FE}_\tau + \text{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.44}$$

And for all $s \in L_{Retail}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \alpha^{cat}_{L_{retail}} + \text{FE}_i + \text{FE}_\tau + \text{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.45}$$

Where $\alpha^{cat}_{H_{retail}}$ and $\alpha^{cat}_{L_{retail}}$ are the competitive parameters associated with category $cat \in \{+Init, -Init, +Close, -Close\}$ for *High* and *Low Retail* stores respectively, and fixed effects FE_i, FE_τ and FE_r are shared for equations 2.44 and 2.45.

2.3.2. HHI

We use the *HHI* for measuring the local market concentration in the proximity of each store, and thus, the level of competition of the store's environment. Indeed, since there exist an inverse relationship between market concentration and competition, stores in zones with lower *HHI* should be more competitive than stores in zones with higher *HHI*, and thus, react more aggressively to price cuts (and less aggressively to price increases). Note that similarly to *N Retail*, this measure is related to the competitive environment of each store and the characteristics of the store's location.

Definition 2.13 ***(Herfindahl–Hirschman Index)*** *For each store $s \in \mathcal{S}$, we define the number of Herfindahl–Hirschman Index (HHI) of s as:*

$$HHI_s = \sum_{r \in \mathcal{R}_s} \frac{|c_s \cap r|^2}{|c_s|^2} \tag{2.46}$$

Note that $c_s \cap r$ is the set of stores from retail r within the close stores of s, so that $\frac{|c_s \cap r|}{|c_s|}$ is the market share of retail r in the close stores of s.

Definition 2.14 ***(High and Low HHI stores)*** *We define the set of High HHI stores as:*

$$H_{HHI} = \left\{ s \in \mathcal{S} : HHI_s \geq q_{High} \right\} \tag{2.47}$$

And the set of Low HHI stores as:

$$L_{HHI} = \left\{ s \in \mathcal{S} : HHI_s \leq q_{Low} \right\} \tag{2.48}$$

Where $q_{High}, q_{Low} \in \mathbb{R}$ such that $q_{High} > q_{Low}$.

We formulate the following model for analysing differences in *Low* and *High HHI* stores' competitive pricing behaviour. For all $s \in H_{HHI}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p} \right)_{\tau,i,s} = \alpha^{cat}_{H_{HHI}} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.49}$$

And for all $s \in L_{HHI}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p} \right)_{\tau,i,s} = \alpha^{cat}_{L_{HHI}} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.50}$$

Where $\alpha^{cat}_{H_{HHI}}$ and $\alpha^{cat}_{L_{HHI}}$ are the competitive parameters associated with category $cat \in \{+Init, -Init, +Close, -Close\}$ for *High* and *Low HHI* stores respectively, and fixed effects FE_i, FE_τ and FE_r are shared for equations 2.49 and 2.50.

2.3.3. Market Share

We propose the local market share of each store (MS) as a measure of competition related to the market power of the respective retail chain. Intuitively, retail chains that compete to locally possess the market in certain geographical sectors, also compete to offer attractive prices in those particular zones. Thus, stores from retail chains with high market share should respond more aggressively to price cuts (and less aggressively to price increases) than stores from retails with low market share.

Definition 2.15 ***(Market Share)*** *For each store $s \in \mathcal{S}$, we define market share of s as:*

$$MS_s = \frac{|c_s \cap [s]|}{|c_s|} \tag{2.51}$$

Note that $c_s \cap [s]$ is the set of stores from the same retail of s within the close stores of s.

Definition 2.16 ***(High and Low MS stores)*** *We define the set of High MS stores as:*

$$H_{MS} = \left\{ s \in \mathcal{S} : MS_s \geq q_{High} \right\} \tag{2.52}$$

And the set of Low MS stores as:

$$L_{MS} = \left\{ s \in \mathcal{S} : MS_s \leq q_{Low} \right\} \tag{2.53}$$

Where $q_{High}, q_{Low} \in \mathbb{R}$ such that $q_{High} > q_{Low}$.

We formulate the following model for analysing differences in *Low* and *High MS* stores' competitive pricing behaviour.. For all $s \in H_{MS}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p} \right)_{\tau,i,s} = \alpha^{cat}_{H_{MS}} + \text{FE}_i + \text{FE}_\tau + \text{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.54}$$

And for all $s \in L_{MS}$:

$$\left(\frac{\frac{\Delta p}{dt}}{p} \right)_{\tau,i,s} = \alpha^{cat}_{L_{MS}} + \text{FE}_i + \text{FE}_\tau + \text{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}^{cat}_{i,s}) \tag{2.55}$$

Where $\alpha^{cat}_{H_{MS}}$ and $\alpha^{cat}_{H_{MS}}$ are the competitive parameters associated with category $cat \in \{+Init, -Init, +Close, -Close\}$, and *High* and *Low MS* stores respectively, and fixed effects FE_i, FE_τ and FE_r are shared for equations 2.54 and 2.55.

2.3.4. Luxury Products

We propose the price of the products as a proxy for determining competition. Indeed, we can think that stores compete more for regular daily products than for items with inelastic demands, such as luxury products. Therefore, we expect that stores respond more aggressively to price cuts (and less aggressively to price increases) of low priced products than to price cuts of expensive items.

Definition 2.17 ***(Mean Product Price)*** *For each item $i \in \mathcal{I}$, we define the mean product price of i as:*

$$P_i = \frac{1}{|\mathcal{T}|}\frac{1}{|\mathcal{S}|}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}} p_{t,i,s} \tag{2.56}$$

Definition 2.18 ***(High and Low Price Items)*** *We define the set of High Price items as:*

$$H_P = \Big\{ i \in \mathcal{I} : P_i \geq q_{High} \Big\} \tag{2.57}$$

And the set of Low Price items as:

$$L_P = \Big\{ i \in \mathcal{I} : P_i \leq q_{Low} \Big\} \tag{2.58}$$

Where $q_{High}, q_{Low} \in \mathbb{R}$ such that $q_{High} > q_{Low}$.

We formulate the following model for analysing differences in competitive pricing strategies for *Low* and *High Price* items. For all $i \in H_P$:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \alpha_{H_P}^{cat} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}_{i,s}^{cat}) \tag{2.59}$$

And for all $i \in L_P$:

$$\left(\frac{\frac{\Delta p}{dt}}{p}\right)_{\tau,i,s} = \alpha_{L_P}^{cat} + \mathrm{FE}_i + \mathrm{FE}_\tau + \mathrm{FE}_r + \epsilon_{\tau,i,s}, \quad \forall \tau \in weak(\mathcal{T}_{i,s}^{cat}) \tag{2.60}$$

Where $\alpha_{H_P}^{cat}$ and $\alpha_{H_P}^{cat}$ are the competitive parameters associated with category $cat \in \{+Init, -Init, +Close, -Close\}$, and *High* and *Low Price* items respectively, and fixed effects FE_i, FE_τ and FE_r are shared for equations 2.59 and 2.60.

Chapter 3

Data and Implementation

In this Chapter we explain the main features of our dataset, the price change frequencies, and details of the implementation of our model presented in Chapter 2.

3.1. The Dataset

Our dataset consists of the daily prices of 1000 products of 1546 stores belonging to 11 different retail chains of the french territory between 2011 and 2012. This dataset was provided by the *Bank of France*, and built by recurrently web-scrapping the prices of the different products for all the shops considered for this study. Moreover, using the information about the route-distances between shops we were able to reconstruct the distance function d from Section 2.2. An important fact is that, since our data was obtained by recurrently web-scrapping the daily prices of different products, price trajectories can posses missing values for days in which prices were not web-scrapped (see Figure 3.1). Moreover, our dataset does not contain information on weather a product was on sale. Thus, we used a *transient price filter* (see subsection 3.1.2) to derive latent regular prices from price processes. Initial price changes and responses were derived (see Figure 3.3) and filtered using the *week* conditional formulation from subsection 2.2.5. Furthermore, extreme values were removed by discarding observations below the 2nd and above 98th percentiles. Finally, items with less than 10 price changes were removed. Table 3.1 presents the summary statistics for the price changes obtained in each category.

Table 3.1: Summary Statistics of Price Changes for each category. Price changes were obtained using parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$ and filtered using the *week* conditional formulation from subsection 2.2.5. Observations below the 2nd and above 98th percentiles are removed, as well as products with less than 10 price changes.

Price Change Category	N	Mean	Std.	Skewness	Kurtosis
$+Init$	586183	$3.239 \cdot 10^{-2}$	$3.075 \cdot 10^{-2}$	1.237	4.407
$-Init$	615445	$-3.130 \cdot 10^{-2}$	$2.954 \cdot 10^{-2}$	1.234	4.375
$+Close$	23370	$2.111 \cdot 10^{-3}$	$4.445 \cdot 10^{-2}$	$1.955 \cdot 10^{-2}$	3.824
$-Close$	21533	$-1.149 \cdot 10^{-3}$	$4.364 \cdot 10^{-2}$	$1.438 \cdot 10^{-1}$	3.919
$+Far$	27456	$4.609 \cdot 10^{-3}$	$4.449 \cdot 10^{-2}$	$1.767 \cdot 10^{-2}$	3.585
$-Far$	25686	$-5.321 \cdot 10^{-4}$	$4.550 \cdot 10^{-2}$	$1.171 \cdot 10^{-1}$	3.833

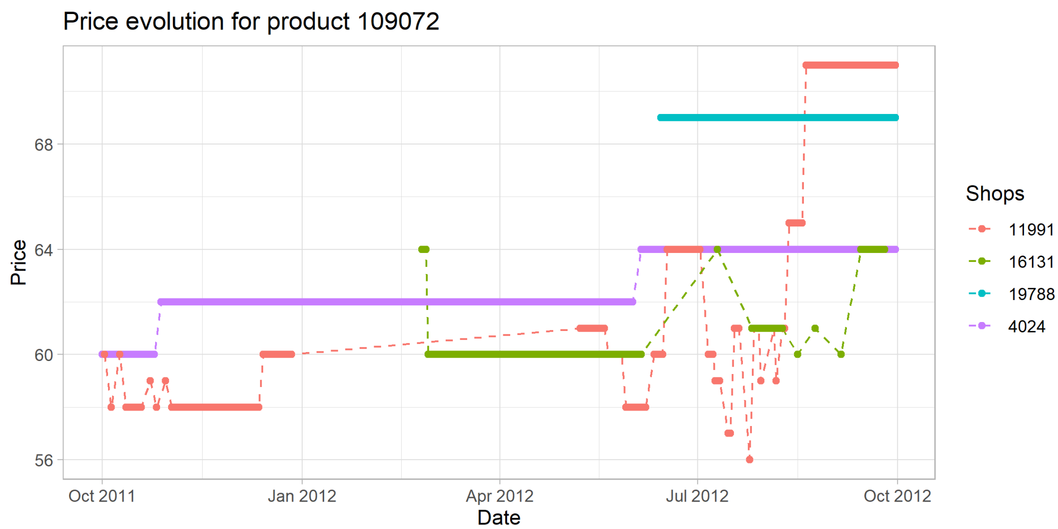

Figure 3.1: Example of missing values for one product and four different stores in our dataset. Solid lines represent known prices, while dotted lines represent missing values.

3.1.1. Price Change Frequency

To better comprehend the behaviour of price changes, we analysed their frequency when including and excluding sales. Table 3.2 resumes the mean price change frequency for each retail chain. Price change frequency ranges from 1.4% to 3.3%, so that price changes occur each 30 to 71 days. Moreover, when using the transient price filter for excluding sales, price change frequencies vary from 0.3% to 1.3%, implying a mean price duration that varies from 74 to 285 days.

Table 3.2: Missing days and mean price change frequencies. Results were obtained by computing the change frequency for each product and retail chain on daily data, and then averaging by retail chain.

Retail Chain	Mean Price Change Frequency	Mean Price Change Frequency (w/o sales)
1 H	3.2%	1.09%
1 H/S	1.5%	0.6%
2 H	2.8%	0.6%
2 H/S	1.4%	0.3%
3 H	2.1%	0.4%
3 H/S	2.8%	0.5%
4 H	1.6 %	0.8%
4 H/S	1.5%	0.8%
4 S	1.4%	0.6%
5 H/S	3.3%	1.3%
6 H/S	2.1%	0.8%

3.1.2. Regular Prices

The model presented in Chapter 2 allows us to identity competitive price changes by analysing the price process of different products and shops. However, according to Nakamura and Steinsson (2008), different types of price adjustments have substantially different macroeconomic implications [16]. In their work, authors distinguish three types of price changes: (1) regular price changes, (2) temporary sales, and (3) price changes due to product substitution. In this thesis, we assume that sales are observable and measurable in real time due to sales flags that explicitly indicate them, so that all shops know at time t whether $p_{t,i,s}$ is or not a regular price. However, since our dataset does not contain explicit information about sales a *transient price filter* is proposed to avoid temporary price changes due to inventory issues or fast sales, and correctly identify the latent regular price during these events:

Definition 3.1 *Let* $\{p_{t,i,s}\}_{t\in\mathcal{T}}$ *the price change process of item* i *at store* s*. We define the transient filtered regular price process* $\left\{r^{Transient}_{t,i,s}\right\}_{t\in\mathcal{T}}$ *as:*

$$r^{Transient}_{0,i,s} = p_{0,i,s} \tag{3.1}$$

And for all $0 < t \leq T$*:*

$$r^{Transient}_{t,i,s} = \begin{cases} p_{t,i,s} & , \text{ if } \exists k \leq dt : \forall j \leq dt : p_{t-k,i,s} = p_{t-k+j,i,s} \\ r^{Transient}_{t-1,i,s} & \sim \end{cases} \tag{3.2}$$

Intuitively, the *transient price filter* removes price changes that do not hold for a minimum span time of dt. Thus, by using this filter we can ensure that $\underset{dt}{\Delta} p^{Init}_t$ is equal to Δp^{Init}_t. Our model can be directly reformulated in terms of $\left\{r^{\text{Transient}}_{t,i,s}\right\}_{t\in\mathcal{T}}$ instead of $\{p_{t,i,s}\}_{t\in\mathcal{T}}$. See Figure 3.2 for an example of the resulting latent regular prices derived from the price processes. Figure 3.2 illustrates an example of how regular prices inferred from price trajectories behave in comparison to the original prices. Solid lines represent the derived regular price processes of four different stores for a particular product, while dotted lines represent the raw price processes. Note that the red dotted line differs from the solid one, as the transient filter is able to exclude both symmetric *V-shaped* patterns (i.e., prices that fully revert after a short period of time) and asymmetric patterns (i.e., prices that partially revert after a short period of time). See Figure I in Nakamura and Steinsson (2008) for reference [16].

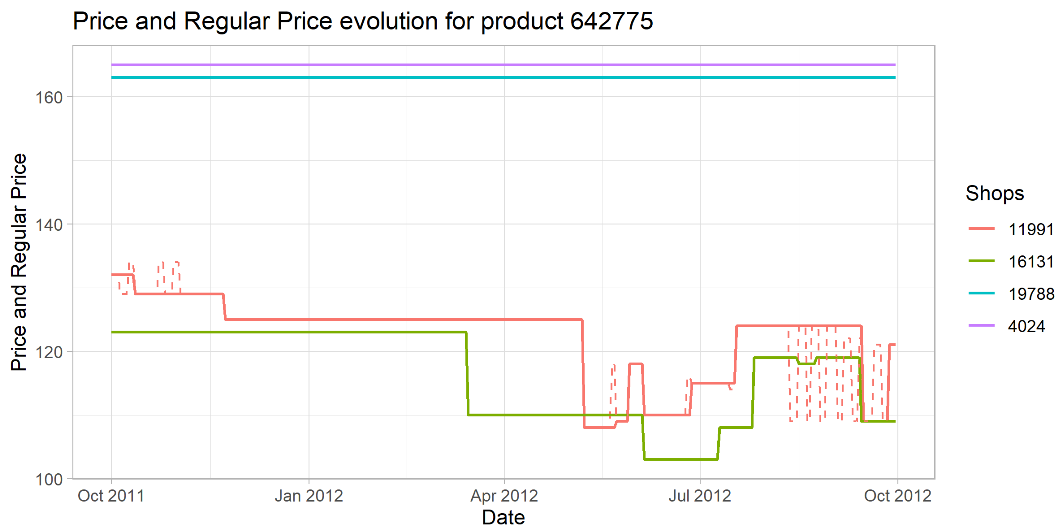

Figure 3.2: Example of latent regular prices derived from price trajectories.

3.1.3. Initial Price Changes

The model presented in section 2.2 allows us to recognise initial price changes from price trajectories and moreover to determine whether each shop behaved as an *Init* changer, or a response belonging to the *Close* or *Far* groups. Figure 3.3 shows an example of how price changes are derived from price trajectories and classified into the different categories.

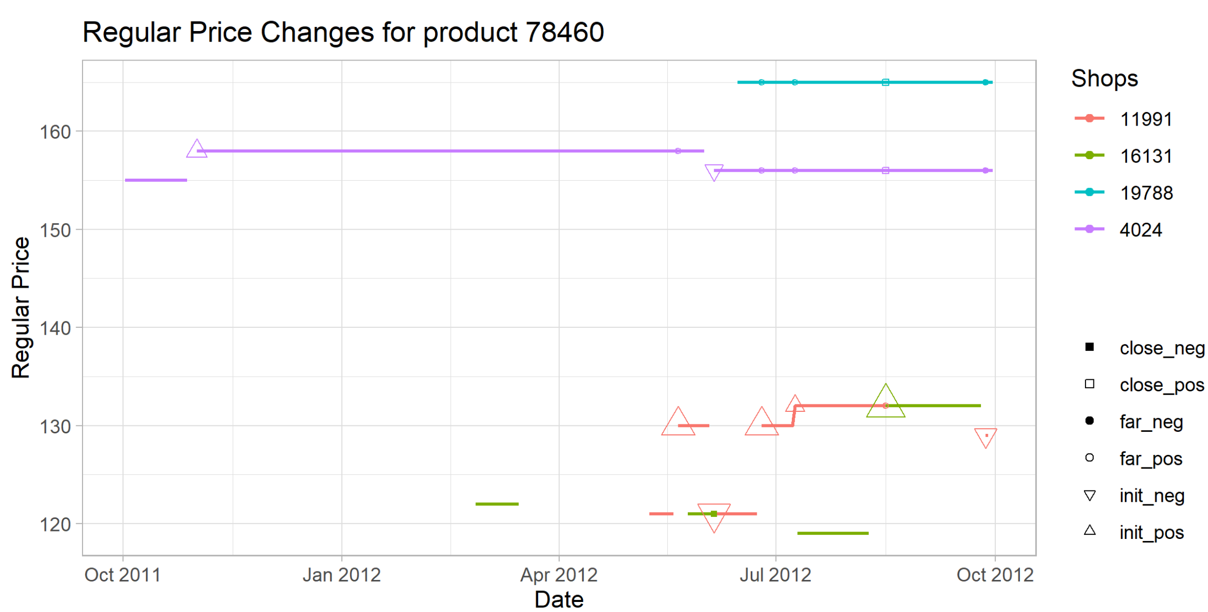

Figure 3.3: Example of initial price changes

3.1.4. Response Probability

We analysed the proportion of price changes that belonged to the *weak* and *strong* conditional categories proposed in subsection 2.2.5. Thus, we calculated (before applying any other filter) for each $cat \in Categories$ the quantities:

$$p_{weak} = \frac{|weak(\mathcal{T}_{i,s}^{cat})|}{|\mathcal{T}_{i,s}^{cat}|}$$

And

$$p_{strong} = \frac{|strong(\mathcal{T}_{i,s}^{cat})|}{|\mathcal{T}_{i,s}^{cat}|}$$

Table 3.3 resumes the response probabilities obtained for each category for a fixed set of hyper-parameters. Note that p_{weak} is quite low (less than 3%) for *Close* and *Far* categories, which shows that stores are not likely to respond to price changes when they perceive them. However, it is important to mention that *Close* stores are slightly more likely to respond than *Far* stores. Finally, we can appreciate that in all cases $p_{strong} \approx p_{weak}/2$, meaning that stores respond to a competitive price change by increasing or decreasing its selling price with (almost) equal probability.

Table 3.3: Response probabilities for conditional categories. In this experiment $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. Note that values in the n_{weak} column do not coincide with those in Table 3.1 since extreme values and items are not filtered for this calculation.

Category	n	n_{weak}	n_{strong}	p_{weak}	p_{strong}
$+$*Init*	618357	618357	618357	1	1
$-$*Init*	646302	646302	646302	1	1
$+$*Close*	828716	24472	13003	0.0295	0.0156
$-$*Close*	839121	22431	11779	0.0267	0.0140
$+$*Far*	1085140	28845	15850	0.0265	0.0146
$-$*Far*	1132195	26932	14149	0.0237	0.0124

3.2. Implementation

In this section we explain our procedure for chunking our dataset, deriving latent regular prices and running our model across the whole dataset.

3.2.1. Chunking

To implement our model, we designed an algorithm able to analyse the dataset through independent chunks of data. To do so, we run our model independently across two dimensions: products, as we assumed that price changes in one product do not affect the competition in other products, and distance, as we assumed that stores excesively far from each other do not compete. Intuitively, this last point states that we can, for example, run our model on data corresponding to the Parisian region's stores first, and afterwards on data from Marseille's stores. In order to identify groups of stores that correspond to independent geographical clusters, we construct a distance graph, whose nodes correspond to the stores, and edges correspond to pairs of stores within the distance threshold δ_f. Connected components of this graph constitute the different and independent store clusters (see Figures 3.4 and 3.4 for examples).

Definition 3.2 ***(Distance Graph)*** *We define the distance graph as the undirected graph* $G = (V, E)$*, where* $V = \mathcal{S}$ *and* $E = \{(s_1, s_2) \in \mathcal{S} \times \mathcal{S} : d(s_1, s_2) \leq \delta_f\}$.

The following property assures that we can partition our dataset into indepent chunks for computing price change times:

Property 3.2.1. *Let* C *be the set of connected components of the distance graph* G*. Then,* $\forall c \in C,\ s \in c \Rightarrow \mathcal{B}_s \subseteq c$.

Indeed, let $D = \{p_{i,t,s} : i \in \mathcal{I}, t \in \mathcal{T}, s \in \mathcal{B}_s\}$ be our dataset, and $D_{(c,i)} = \{p_{i,t,s} : t \in \mathcal{T}, s \in c\}$ the sub-part (chunk) of the dataset containing the price processes of item i and stores of the connected component c. Since for each store $s \in \mathcal{S}$ and item $i \in \mathcal{I}$ the competitive price change times $\mathcal{T}_{i,s}$ depend only on $\{p_{i,t,s} : t \in \mathcal{T}, s \in \mathcal{B}_s\}$, it suffices to use $D_{(c,i)}$ for calculating them.

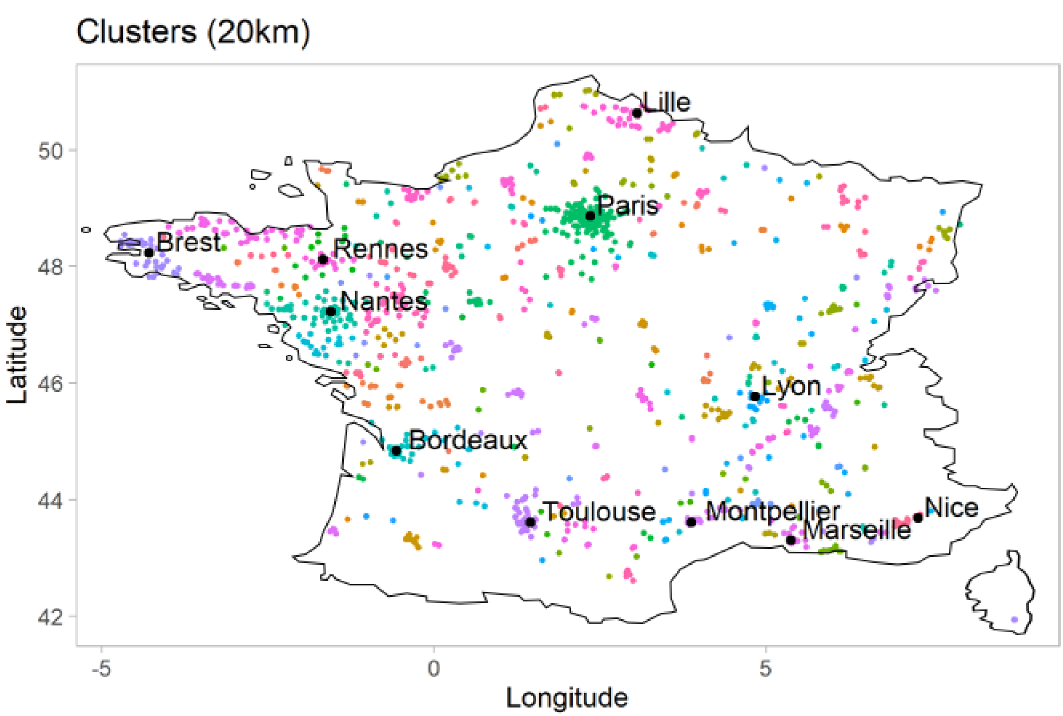

Figure 3.4: Connected Components of the Distance Graph G using $\delta_f = 20$ and approximate borders of France.

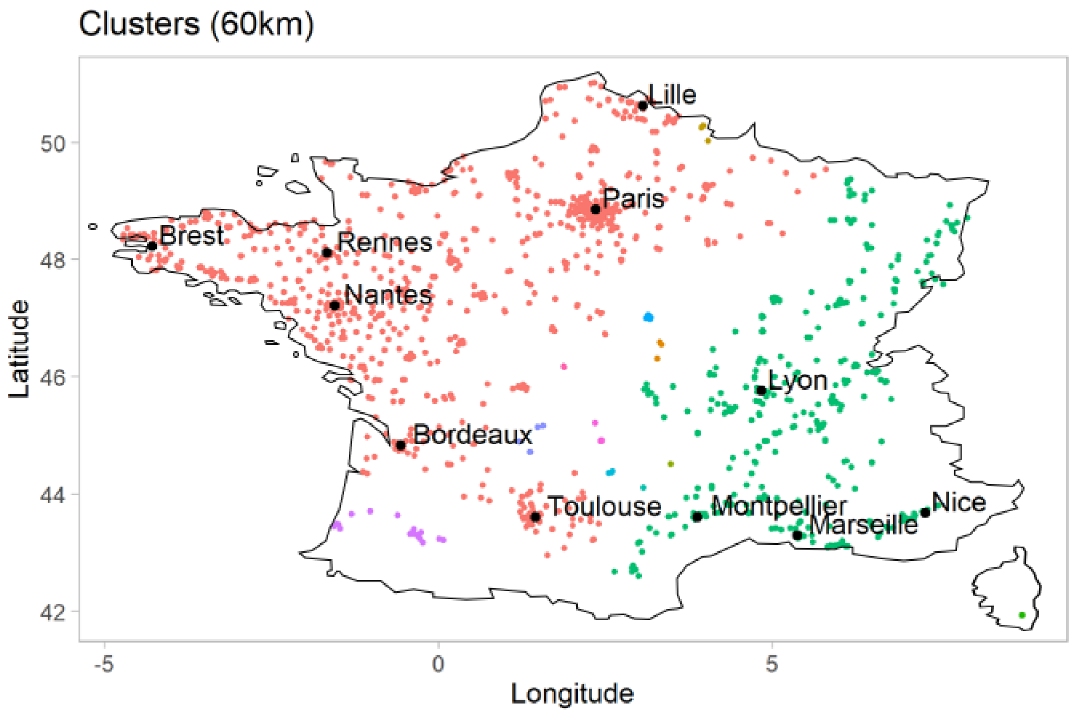

Figure 3.5: Connected Components of the Distance Graph G using $\delta_f = 60$ and approximate borders of France.

3.2.2. Pipeline

Our methodology for estimating the competition components consists of a six steps pipeline illustrated in Figure 3.6. Each step is detailed below:

1. **Obtain Distance Graph:** Once the set hyper parameters is defined, we obtain the distance graph G introduced in definition 3.2, as well as its connected components C. Chunks with only one store are excluded since they induce no competition among other stores.

2. **Chunk the Dataset** Using the set of connected components C, we partition our dataset D into independent chunks $D_{(c,i)}$ corresponding to each component $c \in C$ and item $i \in \mathcal{I}$.

3. **Preprocessing:** Each chunk $D_{(c,i)}$ is preprocessed separately before obtaining price changes. This consists on deriving latent regular prices from price trajectories using the transient price change filter described in subsection 3.1.2, and filtering price changes with not enough previous observations. Explicitly, we remove price changes for which the previous seven days are missing data.

4. **Obtain Competitive Price Changes:** Once every chunk $D_{(c,i)}$ is preprocessed, we obtain the set of competitive price changes of each chunk $P_{(c,i)}$. Afterwards, we build the dataset of competitive price changes (P) by collecting the competitive price changes derived from each chunk.

5. **Filter Price Changes:** For estimating the coefficients of our model, we only keep price changes satisfying the week conditional version presented in 2.2.5. Next, we exclude the extreme values of $\underset{dt}{\Delta} p/p$ in P. This is done by discarding observations below the 2nd and above the 98th percentiles. Moreover, we exclude all products for which we find less than 10 observations to avoid the overfitting of the item fixed effect parameter FE_i.

6. **Obtain Regression Coefficients:** Finally, we obtain the coefficients of the model described in equation 2.40 by fitting it to the set of all price changes remaining.

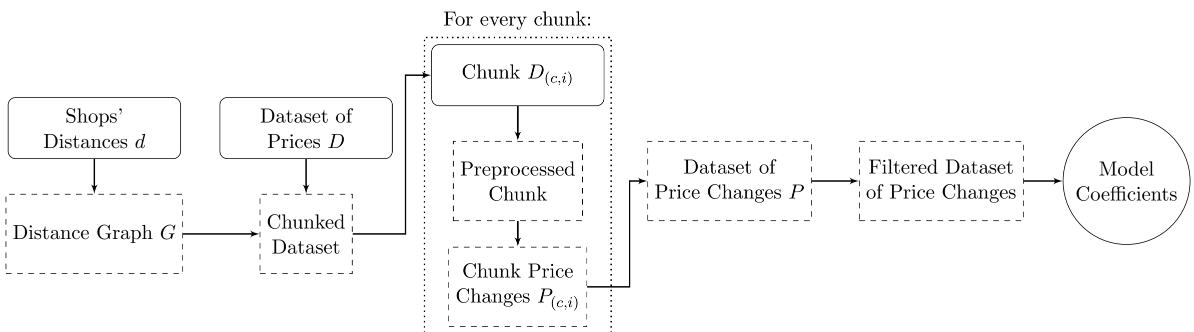

Figure 3.6: Pipeline describing each step for obtaining the final model coefficients.

Chapter 4

Results

4.1. Competitive Environment

In this section we present the results of models using features related to the competitive environment of stores, such as the distance, the local number of different retails, and the local market concentration.

4.1.1. Close and Far Effects

We run model (2.40) for analysing the competitive effects associated with close and far stores responding to initial price changes for four different sets of parameters. Tables 4.1, 4.2, 4.3, and 4.4 show the values obtained by our model for the competitive effects, as well as some summary statistics of the samples in each category. Note that in all tables:

- $\alpha^{+Close}, \alpha^{+Far} > 0$
- $\alpha^{-Close}, \alpha^{-Far} < 0$
- $\alpha^{+Close} < \alpha^{+Far}$ (and significantly differents)
- $|\alpha^{-Close}| > |\alpha^{-Far}|$ (and significantly differents)
- All α coefficients are significantly different from 0.

These results show that, even though stores in groups *Close* and *Far* do not always respond in the same direction of the initiator (see p_{strong} in Table 3.3), competitive effects coefficients do follow the direction of the initiator. Furthermore, stores in group *Close* increase less their prices in response to price increases than stores in group *Far*. This can be interpreted as that, on average, stores in group *Close* try to maintain low prices when competitive stores increase the cost of a product. Moreover, when negative price changes occur, stores in group *Close* react by decreasing their prices more than stores in group *Far*. Again, this can be interpreted as that, on average, stores in group *Close* are more sensible to price cuts than stores in group *Close*, evidencing a stronger competitive behaviour. The asymmetry in positive and negative price responses reflects the competitiveness of *Close* stores with respect to *Far* stores, showing that distance is an important feature for explaining price changes in competitive settings.

Table 4.1: α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*.

	Init	*Close*	*Far*
α^+ estimate	$3.18 \cdot 10^{-2}$ ***	$1.223 \cdot 10^{-3}$ ***,aaa	$3.432 \cdot 10^{-3}$ ***,aaa
(Std. Error)	$(1.503 \cdot 10^{-4})$	$(2.506 \cdot 10^{-4})$	$(2.381 \cdot 10^{-4})$
Mean obs.	$3.239 \cdot 10^{-2}$	$2.111 \cdot 10^{-3}$	$4.609 \cdot 10^{-3}$
Std. obs.	$3.075 \cdot 10^{-2}$	$4.445 \cdot 10^{-2}$	$4.449 \cdot 10^{-2}$
N obs.	586183	23370	27456
α^- estimate	$-3.276 \cdot 10^{-2}$ ***	$-2.268 \cdot 10^{-3}$ ***,a	$-1.848 \cdot 10^{-3}$ ***,a
(Std. Error)	$(1.506 \cdot 10^{-4})$	$(2.572 \cdot 10^{-4})$	$(2.428 \cdot 10^{-4})$
Mean obs.	$-3.130 \cdot 10^{-2}$	$-1.149 \cdot 10^{-3}$	$-5.321 \cdot 10^{-4}$
Std. obs.	$2.954 \cdot 10^{-2}$	$4.364 \cdot 10^{-2}$	$4.550 \cdot 10^{-2}$
N obs.	615445	21533	25686

Table 4.2: α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*.

	Init	*Close*	*Far*
α^+ estimate	$3.180 \cdot 10^{-2}$ ***	$1.267 \cdot 10^{-3}$ ***,aaa	$3.435 \cdot 10^{-3}$ ***,aaa
(Std. Error)	$(1.504 \cdot 10^{-4})$	$(2.509 \cdot 10^{-4})$	$(2.382 \cdot 10^{-4})$
Mean obs.	$3.239 \cdot 10^{-2}$	$2.167 \cdot 10^{-3}$	$4.625 \cdot 10^{-3}$
Std. obs.	$3.075 \cdot 10^{-2}$	$4.442 \cdot 10^{-2}$	$4.549 \cdot 10^{-2}$
N obs.	584559	23337	27415
α^- estimate	$-3.270 \cdot 10^{-2}$ ***	$-2.197 \cdot 10^{-3}$ ***,a	$-1.740 \cdot 10^{-3}$ ***,a
(Std. Error)	$(1.508 \cdot 10^{-4})$	$(2.574 \cdot 10^{-4})$	$(2.430 \cdot 10^{-4})$
Mean obs.	$-3.128 \cdot 10^{-2}$	$-1.152 \cdot 10^{-3}$	$-4.955 \cdot 10^{-4}$
Std. obs.	$2.954 \cdot 10^{-2}$	$4.365 \cdot 10^{-2}$	$4.451 \cdot 10^{-2}$
N obs.	613156	21507	25642

Table 4.3: α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 7$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*.

	Init	*Close*	*Far*
α^+ estimate	$3.178 \cdot 10^{-2}$ ***	$1.310 \cdot 10^{-3}$ ***,aaa	$3.353 \cdot 10^{-3}$ ***,aaa
(Std. Error)	$(1.859 \cdot 10^{-4})$	$(2.736 \cdot 10^{-4})$	$(2.620 \cdot 10^{-4})$
Mean obs.	$3.240 \cdot 10^{-2}$	$2.224 \cdot 10^{-3}$	$4.570 \cdot 10^{-3}$
Std. obs.	$3.076 \cdot 10^{-2}$	$4.442 \cdot 10^{-2}$	$4.546 \cdot 10^{-2}$
N obs.	581827	23282	27377
α^- estimate	$-3.276 \cdot 10^{-2}$ ***	$-2.323 \cdot 10^{-3}$ ***,a	$-1.845 \cdot 10^{-3}$ ***,a
(Std. Error)	$(1.861 \cdot 10^{-4})$	$(2.796 \cdot 10^{-4})$	$(2.666 \cdot 10^{-4})$
Mean obs.	$-3.126 \cdot 10^{-2}$	$-1.155 \cdot 10^{-3}$	$-4.864 \cdot 10^{-4}$
Std. obs.	$2.952 \cdot 10^{-2}$	$4.368 \cdot 10^{-2}$	$4.454 \cdot 10^{-2}$
N obs.	608492	21401	25574

Table 4.4: α coefficients of equation (2.40) and summary statistics for distance competitive effects with parameters: $\delta_f = 40$, $\delta_c = 20$, $dt = 7$, $t_c = 3$. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the panel *Close* are different from the respective coefficients in the panel *Far*.

	Init	*Close*	*Far*
α^+ estimate	$3.192 \cdot 10^{-2}$ ***	$1.585 \cdot 10^{-3}$ ***,aaa	$2.602 \cdot 10^{-3}$ ***,aaa
(Std. Error)	$(1.497 \cdot 10^{-4})$	$(2.480 \cdot 10^{-4})$	$(2.576 \cdot 10^{-4})$
Mean obs.	$3.243 \cdot 10^{-2}$	$2.259 \cdot 10^{-3}$	$3.381 \cdot 10^{-3}$
Std. obs.	$3.076 \cdot 10^{-2}$	$4.456 \cdot 10^{-2}$	$4.489 \cdot 10^{-2}$
N obs.	585856	23774	21298
α^- estimate	$-3.276 \cdot 10^{-2}$ ***	$-2.003 \cdot 10^{-3}$ ***,aaa	$-1.031 \cdot 10^{-4}$ ***,aaa
(Std. Error)	$(1.501 \cdot 10^{-4})$	$(2.547 \cdot 10^{-4})$	$(2.633 \cdot 10^{-4})$
Mean obs.	$-3.129 \cdot 10^{-2}$	$-1.035 \cdot 10^{-3}$	$1.121 \cdot 10^{-4}$
Std. obs.	$2.953 \cdot 10^{-2}$	$4.381 \cdot 10^{-2}$	$4.478 \cdot 10^{-2}$
N obs.	615515	21888	19954

4.1.2. N Retail

Results in the previous subsection indicate that stores in group *Close* are more competitive than stores in group *Far*. By focusing in the first group, in this subsection we present the results of model (2.44) & (2.45). Using the price changes in categories $\pm Init$ and $\pm Close$ derived with parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$, we constructed H_{retail} and L_{retail} store groups by taking values above/equal to and below the median of the number of retail chains N_s in our dataset of price changes respectively (see Table 4.5). We also did this taking values above/equal to the 3rd quartile and below/equal to the 1st quartile (see Table 4.6). In order to balance the observations, the median, 1st, and 3rd quartiles were calculated independently for $\pm Init$ and $\pm Close$ categories, and groups were divided accordingly. Note that in Tables 4.5 and 4.6:

- $\alpha^{+Close}_{L_{retail}} > 0$ (and significantly different from 0)
- $\alpha^{-Close}_{H_{retail}} < 0$ (and significantly different from 0)
- $\alpha^{+Close}_{H_{retail}} < \alpha^{+Close}_{L_{retail}}$ (and significantly different)
- $|\alpha^{-Close}_{H_{retail}}| > |\alpha^{-Close}_{L_{retail}}|$ (and significantly different)

Results show that most of the positive effect of α^{+Close} is represented by L_{retail} stores, while most of the negative part of α^{-Close} is contributed by H_{retail} stores. Similarly to *Close* and *Far* distance coefficients, we note that stores in group H_{retail} are more competitive when responding to competitive price changes than stores in group L_{retail}, showing the same asymmetry between positive and negative price change responses. Results illustrate that when prices increase, stores with low *N Retail* respond less aggressively than stores with high *N Retail*. In contrast, stores with high *Retail* respond more aggressively than stores with high *Retail* when prices decrease. This reflects the fact that, the higher is the local diversity of retail chains in the market, the higher is the competition between retailers in that particular zone.

Table 4.5: α coefficients of equations (2.44) & (2.45) and summary statistics for N Retail competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{retail} and L_{retail} store groups obtained using the median of N_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{retail}) are different from the respective coefficients in the right panel (L_{retail})

	Observations above the 50th percentile # retail chains	Observations below the 50th percentile # retail chains
α^{+Close} estimate	$-7.906 \cdot 10^{-4}$ ^{*,aaa}	$3.915 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(3.079 \cdot 10^{-4})$	$(3.268 \cdot 10^{-4})$
Mean obs.	$-1.213 \cdot 10^{-4}$	$4.722 \cdot 10^{-3}$
Std. obs.	$4.361 \cdot 10^{-2}$	$4.528 \cdot 10^{-2}$
N obs.	12597	10773
α^{-Close} estimate	$-2.753 \cdot 10^{-3}$ ^{***,aa}	$-1.655 \cdot 10^{-3}$ ^{***,aa}
(Std. Error)	$(3.169 \cdot 10^{-4})$	$(3.382 \cdot 10^{-4})$
Mean obs.	$-1.735 \cdot 10^{-3}$	$-4.571 \cdot 10^{-4}$
Std. obs.	$4.330 \cdot 10^{-2}$	$4.404 \cdot 10^{-2}$
N obs.	11671	9862
α^{+Init} estimate	$3.224 \cdot 10^{-2}$ ^{***,aaa}	$3.146 \cdot 10^{-2}$ ^{***,aaa}
(Std. Error)	$(1.510 \cdot 10^{-4})$	$(1.576 \cdot 10^{-4})$
Mean obs.	$3.266 \cdot 10^{-2}$	$3.196 \cdot 10^{-2}$
Std. obs.	$3.078 \cdot 10^{-2}$	$3.069 \cdot 10^{-2}$
N obs.	362334	223849
α^{-Init} estimate	$-3.284 \cdot 10^{-2}$ ***,	$-3.271 \cdot 10^{-2}$ ***,
(Std. Error)	$(1.512 \cdot 10^{-4})$	$(1.574 \cdot 10^{-4})$
Mean obs.	$-3.147 \cdot 10^{-2}$	$-3.104 \cdot 10^{-2}$
Std. obs.	$2.964 \cdot 10^{-2}$	$2.938 \cdot 10^{-2}$
N obs.	382860	232585

Table 4.6: α coefficients of equations (2.44) & (2.45) and summary statistics for N Retail competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{retail} and L_{retail} store groups obtained using the 1st and 3rd quartiles of N_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{retail}) are different from the respective coefficients in the right panel (L_{retail})

	Observations above the 75th percentile # retail chains	Observations below the 25th percentile # retail chains
α^{+Close} estimate	$-4.111 \cdot 10^{-4}$ ^{,aaa}	$7.000 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(4.383 \cdot 10^{-4})$	$(4.185 \cdot 10^{-4})$
Mean obs.	$-1.726 \cdot 10^{-4}$	$7.338 \cdot 10^{-3}$
Std. obs	$4.116 \cdot 10^{-2}$	$4.545 \cdot 10^{-2}$
N obs.	6041	6538
α^{-Close} estimate	$-1.982 \cdot 10^{-3}$ ^{***,a}	$-1.139 \cdot 10^{-3}$ ^{***,a}
(Std. Error)	$(4.570 \cdot 10^{-4})$	$(4.261 \cdot 10^{-4})$
Mean obs.	$-1.714 \cdot 10^{-3}$	$-2.790 \cdot 10^{-4}$
Std. obs	$4.153 \cdot 10^{-2}$	$4.466 \cdot 10^{-2}$
N obs.	5413	6225
α^{+Init} estimate	$3.274 \cdot 10^{-2}$ ^{***,aaa}	$3.192 \cdot 10^{-2}$ ^{***,aaa}
(Std. Error)	$(1.991 \cdot 10^{-4})$	$(1.990 \cdot 10^{-4})$
Mean obs.	$3.280 \cdot 10^{-2}$	$3.196 \cdot 10^{-2}$
Std. obs	$3.053 \cdot 10^{-2}$	$3.069 \cdot 10^{-2}$
N obs.	185911	223849
α^{-Init} estimate	$-3.288 \cdot 10^{-2}$ ^{***,aa}	$-3.226 \cdot 10^{-2}$ ^{***,aa}
(Std. Error)	$(1.995 \cdot 10^{-4})$	$(1.990 \cdot 10^{-4})$
Mean obs.	$-3.194 \cdot 10^{-2}$	$-3.104 \cdot 10^{-2}$
Std. obs	$2.977 \cdot 10^{-2}$	$2.938 \cdot 10^{-2}$
N obs.	193665	232585

4.1.3. HHI

In this subsection we present the results of model (2.44) & (2.45). Using the price changes derived with parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$, we constructed H_{HHI} and L_{HHI} store groups analogously to what is explained in subsection 4.1.2 . Note that in Tables 4.7 and 4.8:

- $\alpha^{+Close}_{H_{HHI}} > 0$ (and significantly different from 0)
- $\alpha^{-Close}_{L_{HHI}} < 0$ (and significantly different from 0)
- $\alpha^{+Close}_{L_{HHI}} < \alpha^{+Close}_{H_{HHI}}$ (and significantly different)
- $|\alpha^{-Close}_{L_{HHI}}| > |\alpha^{-Close}_{H_{HHI}}|$

Particularly, coefficients in the last point are significantly different in Table 4.8. Results show that most of the positive effect of α^{+Close} is represented by H_{HHI} stores, while most of the negative part of α^{-Close} is contributed by L_{HHI} stores. Similarly to *Close* and *Far* distance coefficients, we note that stores in group L_{HHI} are more competitive when responding to competitive price changes than stores in group H_{HHI}, showing the same asymmetry between positive and negative price change responses. In particular, when prices increase, stores with low *HHI* respond less aggressively than stores with high *HHI*. On the other hand, when price cuts happen, stores with high *HHI* respond more aggressively than stores with high *HHI*. These results are intuitive and consistent since low *HHI* values correspond to lower market concentration, and therefore to a more competitive market.

Table 4.7: α coefficients of equations (2.44) & (2.45) and summary statistics for HHI competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{HHI} and L_{HHI} store groups obtained using the median of HHI_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{HHI}) are different from the respective coefficients in the right panel (L_{HHI})

	Observations above the 50th percentile HHI	Observations below the 50th percentile HHI
α^{+Close} estimate	$3.549 \cdot 10^{-3}$ ***,*aaa*	$-7.954 \cdot 10^{-4}$ *,*aaa*
(Std. Error)	$(3.159 \cdot 10^{-4})$	$(3.178 \cdot 10^{-4})$
Mean obs.	$4.394 \cdot 10^{-3}$	$-1.841 \cdot 10^{-4}$
Std. obs	$4.499 \cdot 10^{-2}$	$4.379 \cdot 10^{-2}$
N obs.	11716	11654
α^{-Close} estimate	$-2.182 \cdot 10^{-3}$ ***,	$-2.311 \cdot 10^{-3}$ ***,
(Std. Error)	$(3.266 \cdot 10^{-4})$	$(3.271 \cdot 10^{-4})$
Mean obs.	$-9.749 \cdot 10^{-4}$	$-1.323 \cdot 10^{-3}$
Std. obs	$4.385 \cdot 10^{-2}$	$4.344 \cdot 10^{-2}$
N obs.	10737	10796
α^{+Init} estimate	$3.196 \cdot 10^{-2}$ ***,	$3.195 \cdot 10^{-2}$ ***,
(Std. Error)	$(1.543 \cdot 10^{-4})$	$(1.528 \cdot 10^{-4})$
Mean obs.	$3.236 \cdot 10^{-2}$	$3.243 \cdot 10^{-2}$
Std. obs	$3.090 \cdot 10^{-2}$	$3.059 \cdot 10^{-2}$
N obs.	298214	287969
α^{-Init} estimate	$-3.277 \cdot 10^{-2}$ ***,	$-3.279 \cdot 10^{-2}$ ***,
(Std. Error)	$(1.540 \cdot 10^{-4})$	$(1.532 \cdot 10^{-4})$
Mean obs.	$-3.122 \cdot 10^{-2}$	$-3.139 \cdot 10^{-2}$
Std. obs	$2.956 \cdot 10^{-2}$	$2.953 \cdot 10^{-2}$
N obs.	310052	305393

Table 4.8: α coefficients of equations (2.44) & (2.45) and summary statistics for *HHI* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{HHI} and L_{HHI} store groups obtained using the 1st and 3rd quartiles of HHI_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{HHI}) are different from the respective coefficients in the right panel (L_{HHI})

	Observations above the 75th percentile HHI	Observations below the 25th percentile HHI
α^{+Close} estimate	$6.411 \cdot 10^{-3}$ ^{***,aaa}	$-1.017 \cdot 10^{-3}$ ^{*,aaa}
(Std. Error)	$(4.576 \cdot 10^{-4})$	$(4.613 \cdot 10^{-4})$
Mean obs.	$6.887 \cdot 10^{-3}$	$-8.776 \cdot 10^{-4}$
Std. obs	$4.505 \cdot 10^{-2}$	$4.456 \cdot 10^{-2}$
N obs.	5752	5662
α^{-Close} estimate	$-1.501 \cdot 10^{-3}$ ^{**,aa}	$-2.887 \cdot 10^{-3}$ ^{***,aa}
(Std. Error)	$(4.648 \cdot 10^{-4})$	$(4.636 \cdot 10^{-4})$
Mean obs.	$-4.856 \cdot 10^{-4}$	$-2.357 \cdot 10^{-3}$
Std. obs	$4.391 \cdot 10^{-2}$	$4.415 \cdot 10^{-2}$
N obs.	5497	5573
α^{+Init} estimate	$3.206 \cdot 10^{-2}$ ^{***,aa}	$3.287 \cdot 10^{-2}$ ^{***,aa}
(Std. Error)	$(2.316 \cdot 10^{-4})$	$(2.296 \cdot 10^{-4})$
Mean obs.	$3.201 \cdot 10^{-2}$	$3.284 \cdot 10^{-2}$
Std. obs	$3.103 \cdot 10^{-2}$	$3.056 \cdot 10^{-2}$
N obs.	169992	147139
α^{-Init} estimate	$-3.230 \cdot 10^{-2}$ ***,	$-3.269 \cdot 10^{-2}$ ***,
(Std. Error)	$(2.313 \cdot 10^{-4})$	$(2.305 \cdot 10^{-4})$
Mean obs.	$-3.112 \cdot 10^{-2}$	$-3.188 \cdot 10^{-2}$
Std. obs	$2.958 \cdot 10^{-2}$	$2.972 \cdot 10^{-2}$
N obs.	177746	153794

4.2. Market Power

In this section we present the results of model (2.54) & (2.55) related to the market power of each store. Using the price changes derived with parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$, we constructed H_{MS} and L_{MS} store groups analogously to what is explained in subsection 4.1.2. Note that in Tables 4.9 and 4.10:

- $\alpha^{+Close}_{L_{MS}} > 0$ (and significantly different from 0)
- $\alpha^{-Close}_{H_{MS}} < 0$ (and significantly different from 0)
- $\alpha^{+Close}_{H_{MS}} < \alpha^{+Close}_{L_{MS}}$ (and significantly different)
- $|\alpha^{-Close}_{H_{MS}}| > |\alpha^{-Close}_{L_{MS}}|$ (and significantly different)

Results show that most of the positive effect of α^{+Close} is represented by L_{MS} stores, while most of the negative part of α^{-Close} is contributed by H_{MS} stores. Similarly to *Close* and *Far* distance coefficients, we note that stores in group H_{MS} are more competitive when responding to competitive price changes than stores in group L_{MS}, showing the same asymmetry between negative and positive price change responses. Particularly, note that when prices increase, stores with low *MS* respond less aggressively than stores with high *MS*. In contrast, stores with high *MS* respond more aggressively than stores with high *MS* when prices decrease. This can be explained by thinking that retail chains that compete to locally possess the market, also make efforts to offer competitive prices in those geographical zones.

Table 4.9: α coefficients of equations (2.54) & (2.55) and summary statistics for *MS* competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{MS} and L_{MS} store groups obtained using the median of MS_s in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. *aaa*, *aa*, and *a* denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{MS}) are different from the respective coefficients in the right panel (L_{MS})

	Observations above the 50th percentile Market Share	Observations below the 50th percentile Market Share
α^{+Close} estimate	$-6.208 \cdot 10^{-5}$ ^{,aaa}	$3.148 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(3.056 \cdot 10^{-4})$	$(3.277 \cdot 10^{-4})$
Mean obs.	$6.478 \cdot 10^{-4}$	$3.912 \cdot 10^{-3}$
Std. obs	$4.363 \cdot 10^{-2}$	$4.539 \cdot 10^{-2}$
N obs.	12893	10477
α^{-Close} estimate	$-2.636 \cdot 10^{-3}$ ^{***,a}	$-1.809 \cdot 10^{-3}$ ^{***,a}
(Std. Error)	$(3.182 \cdot 10^{-4})$	$(3.364 \cdot 10^{-4})$
Mean obs.	$-1.631 \cdot 10^{-3}$	$-5.884 \cdot 10^{-4}$
Std. obs	$4.235 \cdot 10^{-2}$	$4.511 \cdot 10^{-2}$
N obs.	11594	9939
α^{+Init} estimate	$3.192 \cdot 10^{-2}$ ***,	$3.196 \cdot 10^{-2}$ ***,
(Std. Error)	$(1.544 \cdot 10^{-4})$	$(1.529 \cdot 10^{-4})$
Mean obs.	$3.249 \cdot 10^{-2}$	$3.230 \cdot 10^{-2}$
Std. obs	$3.107 \cdot 10^{-2}$	$3.044 \cdot 10^{-2}$
N obs.	288406	297777
α^{-Init} estimate	$-3.277 \cdot 10^{-2}$ ***,	$-3.282 \cdot 10^{-2}$ ***,
(Std. Error)	$(1.542 \cdot 10^{-4})$	$(1.531 \cdot 10^{-4})$
Mean obs.	$-3.131 \cdot 10^{-2}$	$-3.130 \cdot 10^{-2}$
Std. obs	$2.960 \cdot 10^{-2}$	$2.949 \cdot 10^{-2}$
N obs.	312488	302957

Table 4.10: α coefficients of equations (2.54) & (2.55) and summary statistics for MS competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_{MS} and L_{MS} store groups obtained using the 1st and 3rd quartiles of MS_s in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_{MS}) are different from the respective coefficients in the right panel (L_{MS})

	Observations above the 75th percentile Market Share	Observations below the 25th percentile Market Share
α^{+Close} estimate	$-6.551 \cdot 10^{-4}$ ^{,aaa}	$5.504 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(4.269 \cdot 10^{-4})$	$(4.235 \cdot 10^{-4})$
Mean obs.	$3.042 \cdot 10^{-4}$	$5.945 \cdot 10^{-3}$
Std. obs	$4.399 \cdot 10^{-2}$	$4.516 \cdot 10^{-2}$
N obs.	5810	5794
α^{-Close} estimate	$-4.307 \cdot 10^{-3}$ ^{***,aaa}	$-3.394 \cdot 10^{-4}$ ^{,aaa}
(Std. Error)	$(4.394 \cdot 10^{-4})$	$(4.275 \cdot 10^{-4})$
Mean obs.	$-2.848 \cdot 10^{-3}$	$7.139 \cdot 10^{-4}$
Std. obs	$4.302 \cdot 10^{-2}$	$4.604 \cdot 10^{-2}$
N obs.	5419	5664
α^{+Init} estimate	$3.198 \cdot 10^{-2}$ ^{***,a}	$3.154 \cdot 10^{-2}$ ^{***,a}
(Std. Error)	$(1.667 \cdot 10^{-4})$	$(1.607 \cdot 10^{-4})$
Mean obs.	$3.265 \cdot 10^{-2}$	$3.189 \cdot 10^{-2}$
Std. obs	$3.146 \cdot 10^{-2}$	$3.021 \cdot 10^{-2}$
N obs.	174982	243798
α^{-Init} estimate	$-3.331 \cdot 10^{-2}$ ^{***,aa}	$-3.272 \cdot 10^{-2}$ ^{***,aa}
(Std. Error)	$(1.661 \cdot 10^{-4})$	$(1.608 \cdot 10^{-4})$
Mean obs.	$-3.156 \cdot 10^{-2}$	$-3.102 \cdot 10^{-2}$
Std. obs	$3.001 \cdot 10^{-2}$	$2.920 \cdot 10^{-2}$
N obs.	188723	250122

4.3. Luxury Products

In this subsection we present the results of model (2.54) & (2.55). Using the price changes derived with parameters $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$, we constructed H_P and L_P product groups analogously to what is explained in subsection 4.1.2. Note that in Table 4.11:

- $\alpha^{+Close}_{H_P} > 0$ (and significantly different from 0)
- $\alpha^{-Close}_{L_P} < 0$ (and significantly different from 0)
- $\alpha^{+Close}_{L_P} < \alpha^{+Close}_{H_P}$ (and significantly different)
- $|\alpha^{-Close}_{L_P}| > |\alpha^{-Close}_{H_P}|$ (and significantly different)

Results show that most of the positive effect of α^{+Close} is represented by H_P stores, while most of the negative part of α^{-Close} is contributed by L_P stores. Similarly to *Close* and *Far* distance coefficients, we note that stores compete more for items in group L_P than for items in group H_P, showing the same asymmetry between positive and negative price change responses. Particularly, note that when prices increase, stores with respond less aggressively to expensive products than to cheaper products. In contrast, when prices decrease, stores respond more aggressively to inexpensive items than to expensive ones. This can be explained thinking that stores compete more for regular daily products than for items with inelastic demands, such as luxury products. However, results from Table 4.12 are not consistent, therefore, further analysis is needed to confirm this hypothesis, for example, by incorporating product categories.

Table 4.11: α coefficients and summary statistics for P competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_P and L_P item groups obtained using the median of P_i in our dataset of price changes. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_P) are different from the respective coefficients in the right panel (L_P)

	Observations above the 50th percentile Product Price	Observations below the 50th percentile Product Price
α^{+Close} estimate	$4.000 \cdot 10^{-3}$ ^{***,aaa}	$-1.368 \cdot 10^{-3}$ ^{,aaa}
(Std. Error)	$(9.777 \cdot 10^{-4})$	$(9.851 \cdot 10^{-4})$
Mean obs.	$2.707 \cdot 10^{-3}$	$1.424 \cdot 10^{-3}$
Std. obs	$4.365 \cdot 10^{-2}$	$4.536 \cdot 10^{-2}$
N obs.	12522	10848
α^{-Close} estimate	$8.626 \cdot 10^{-4}$ ^{,aaa}	$-5.453 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(9.835 \cdot 10^{-4})$	$(9.842 \cdot 10^{-4})$
Mean obs.	$-1.654 \cdot 10^{-4}$	$-2.141 \cdot 10^{-3}$
Std. obs	$4.298 \cdot 10^{-2}$	$4.428 \cdot 10^{-2}$
N obs.	10805	10728
α^{+Init} estimate	$3.245 \cdot 10^{-2}$ ***,	$3.147 \cdot 10^{-2}$ ***,
(Std. Error)	$(9.406 \cdot 10^{-4})$	$(9.427 \cdot 10^{-4})$
Mean obs.	$3.074 \cdot 10^{-2}$	$3.411 \cdot 10^{-2}$
Std. obs	$3.078 \cdot 10^{-2}$	$3.063 \cdot 10^{-2}$
N obs.	298798	287385
α^{-Init} estimate	$-2.908 \cdot 10^{-2}$ ^{***,aaa}	$-3.659 \cdot 10^{-2}$ ^{***,aaa}
(Std. Error)	$(9.407 \cdot 10^{-4})$	$(9.417 \cdot 10^{-4})$
Mean obs.	$-2.969 \cdot 10^{-2}$	$-3.294 \cdot 10^{-2}$
Std. obs	$2.973 \cdot 10^{-2}$	$2.926 \cdot 10^{-2}$
N obs.	309405	306040

Table 4.12: α coefficients and summary statistics for P competitive effects with parameters: $\delta_f = 60$, $\delta_c = 20$, $dt = 7$, $t_c = 0$. H_P and L_P item groups obtained using the 1st and 3rd quartiles of P_i in our dataset of price changes respectively. ***, **, and * denote significance at 1%, 5%, and 10%, respectively. aaa, aa, and a denote significance at 1%, 5%, and 20%, respectively, for whether the coefficients from the left panel (H_P) are different from the respective coefficients in the right panel (L_P)

	Observations above the 75th percentile Product Price	Observations below the 25th percentile Product Price
α^{+Close} estimate	$-5.917 \cdot 10^{-4}$ ^{,aaa}	$4.378 \cdot 10^{-3}$ ^{***,aaa}
(Std. Error)	$(1.036 \cdot 10^{-3})$	$(1.046 \cdot 10^{3})$
Mean obs.	$3.264 \cdot 10^{-3}$	$1.222 \cdot 10^{-4}$
Std. obs	$4.334 \cdot 10^{-2}$	$4.779 \cdot 10^{-2}$
N obs.	6549	5636
α^{-Close} estimate	$-3.788 \cdot 10^{-3}$ ^{***,aa}	$-9.567 \cdot 10^{-6}$ ^{,aa}
(Std. Error)	$(1.050 \cdot 10^{-3})$	$(1.044 \cdot 10^{-3})$
Mean obs.	$3.551 \cdot 10^{-4}$	$-2.618 \cdot 10^{-3}$
Std. obs	$4.108 \cdot 10^{-2}$	$4.626 \cdot 10^{-2}$
N obs.	5380	5648
α^{+Init} estimate	$2.551 \cdot 10^{-2}$ ^{***,aaa}	$3.878 \cdot 10^{-2}$ ^{***,aaa}
(Std. Error)	$(9.674 \cdot 10^{-4})$	$(9.674 \cdot 10^{-4})$
Mean obs.	$2.905 \cdot 10^{-2}$	$3.563 \cdot 10^{-2}$
Std. obs	$3.077 \cdot 10^{-2}$	$3.110 \cdot 10^{-2}$
N obs.	152147	143143
α^{-Init} estimate	$-3.202 \cdot 10^{-2}$ ***,	$-3.215 \cdot 10^{-2}$ ***,
(Std. Error)	$(9.675 \cdot 10^{-4})$	$(9.675 \cdot 10^{-4})$
Mean obs.	$-2.754 \cdot 10^{-2}$	$-3.449 \cdot 10^{-2}$
Std. obs	$2.924 \cdot 10^{-2}$	$2.920 \cdot 10^{-2}$
N obs.	155726	156521

Conclusion

In this thesis, we have presented an empirical model for determining how the distance between stores affect competition in french retailers, formalising initial price change events, and introducing a framework that allows us to analyse large amounts of data. By controlling by item, time and retail fixed effects, we estimated competitive effects of close and far stores, providing evidence that close stores adopt stronger competitive pricing strategies in response to price changes than stores located far away from the competition. We further explored other features related to the competitive environment of the stores, such as the number of retailers present in their proximity, and the local market concentration of the retail, finding similar competitive behaviours than those found for *Close* and *Far* stores. Indeed, stores compete more in settings in which various retailers participate than in zones where the number of different retail chains present is rather low. Similarly, we found that stores are more competitive when the market concentration of their environment is low rather than when it is high. We also showed that other features, such as the local market share of each store and the price of each product, produce analogous responses to initial price changes. Our results demonstrate that stores from retail chains with high market share react more aggressively to competitive price changes than stores from low market share retail chains, and that competition is stronger for low priced products than for expensive ones.

Nonetheless, this work is not exempt of limitations. In fact, since the competitive price change event classification relies on the price processes from surrounding stores to correctly identify *Init*, *Close* and *Far* categories, our model requires consistent data from all the stores present in the particular geographical zone under analysis. However, despite the fact that missing information in our dataset may have introduced some noise in the model, the results presented in this thesis were consistent and coherent as we analysed huge amounts of data, capturing on average the competitive effects between retailers. Other limitations are related to the bias of the *weak* conditional formulation of our model, in which we selected non-zero price change responses in order to estimate the coefficients of our model. Future work considers extending this thesis by proposing other ways for measuring the impact of a competitive price change event, including the fraction of stores that respond after each competitive price change, and the speed of these responses. Furthermore, differentiating stores that consistently respond from those that rarely respond to competitive price changes would serve to correct bias from the *weak* conditional formulation, for instance, by using Heckman correction [8].

Finally, we hope that this thesis contributes to the understanding of price changes and competition between retail chains, and to promote the interest in models of competition that incorporate features of the stores' environment to analyse competitive price change events.

Bibliography

[1] Alan Blinder, Elie RD Canetti, David E Lebow, and Jeremy B Rudd. *Asking about prices: a new approach to understanding price stickiness.* Russell Sage Foundation, 1998.

[2] Emmanuel Dhyne, Luis J Alvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lunnemann, Fabio Rumler, and Jouko Vilmunen. Price changes in the euro area and the united states: Some facts from individual consumer price data. *Journal of Economic Perspectives*, 20(2):171–192, 2006.

[3] Peter R Dickson and Joel E Urbany. Retailer reactions to competitive price changes. *Journal of Retailing*, 70(1):1–21, 1994.

[4] Reinhard Diestel, Alexander Schrijver, and Paul D Seymour. Graph theory. In *MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH REPORT NO. 16/2007*. Citeseer, 2007.

[5] Colin Ellis et al. Do supermarket prices change from week to week. *Bank of England*, 378:3–27, 2009.

[6] Silvia Fabiani, Martine Druant, Ignacio Hernando, Claudia Kwapil, Bettina Landau, Claire Loupias, Fernando Martins, Thomas Mathä, Roberto Sabbatini, Harald Stahl, et al. The pricing behaviour of firms in the euro area: New survey evidence. 2005.

[7] Lawrence R Glosten and Lawrence E Harris. Estimating the components of the bid/ask spread. *Journal of financial Economics*, 21(1):123–142, 1988.

[8] James Heckman. Shadow prices, market wages, and labor supply. *Econometrica: journal of the econometric society*, pages 679–694, 1974.

[9] Roger D Huang and Hans R Stoll. The components of the bid-ask spread: A general approach. *The Review of Financial Studies*, 10(4):995–1034, 1997.

[10] Patrick J Kehoe, Virgiliu Midrigan, et al. *Sales and the real effects of monetary policy.* Federal Reserve Bank of Minneapolis, Research Department, 2007.

[11] Rajiv Lal and Carmen Matutes. Price competition in multimarket duopolies. *The RAND Journal of Economics*, pages 516–537, 1989.

[12] Daniel Levy, Mark Bergen, Shantanu Dutta, and Robert Venable. The magnitude of menu costs: direct evidence from large us supermarket chains. *The Quarterly Journal of Economics*, 112(3):791–824, 1997.

[13] Ananth Madhavan and Seymour Smidt. A bayesian model of intraday specialist pricing. *Journal of Financial Economics*, 30(1):99–134, 1991.

[14] Ananth Madhavan, Matthew Richardson, and Mark Roomans. Why do security prices change? a transaction-level analysis of nyse stocks. *The Review of Financial Studies*, 10 (4):1035–1064, 1997.

[15] Dmitriy Muravyev. Order flow and expected option returns. *The Journal of Finance*, 71(2):673–708, 2016.

[16] Emi Nakamura and Jón Steinsson. Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4):1415–1464, 2008.

[17] James M Poterba. Retail price reactions to changes in state and local sales taxes. *National Tax Journal*, 49(2):165–176, 1996.

[18] Steven Salop and Joseph Stiglitz. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *The Review of Economic Studies*, 44(3):493–510, 1977.

[19] Yuval Shilony. Mixed pricing in oligopoly. *Journal of Economic Theory*, 14(2):373–388, 1977.

[20] Shailender Singh and Chen Guan Ru. Price rigidity, market competition, and product differentiation. *Economic research-Ekonomska istraživanja*, 32(1):2935–2952, 2019.

[21] John B Taylor. A historical analysis of monetary policy rules. In *Monetary policy rules*, pages 319–348. University of Chicago Press, 1999.

[22] Hal R Varian. A model of sales. *The American economic review*, 70(4):651–659, 1980.