Price Dispersion and Wholesale Costs Shocks in the Colombian Retail Gasoline Markets

By: Alex Perez
Juan Sebastián Vélez-Velásquez

Borradores de ECONOMÍA

No. 1220
2022

BANCO DE LA REPÚBLICA COLOMBIA

tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col

Price Dispersion and Wholesale Costs Shocks in the Colombian Retail Gasoline Markets[†]

Alex Perez[‡] and Juan Sebastián Vélez-Velásquez[§]

Abstract

Price dispersion is a prevalent feature even of markets where, arguably, homogeneous good are traded. At the heart of the causes of price dispersion lie the firms' strategic interactions with their customers and rivals. Consumers' eagerness and ability to search for lower prices tends to reduce dispersion because it enhances competition. Firms inability to sustain tacit collusion, on the other hand, increases price dispersion. Wholesale costs shock can impact both. We use data on retail gasoline markets from Colombia to assess whether changes in price dispersion following wholesale cost shocks are explained by consumer searching or by firms breaking away from collusive equilibria. We also explore the role played by market structure on the extent of price dispersion. Our findings suggest that changes in price dispersion following wholesale cost shocks are driven by consumers searching more intensely for lower prices. We also find a positive correlation between the number of service stations in a market and how disperse prices are. Our results are robust to alternative ways of measuring price dispersion and alternative ways of defining relevant markets.

Keywords: *Price dispersion, gasoline markets, wholesale costs, retail prices.*

JEL Classification: *D82, D83, Q49, L11, L94.*

[†]We thank Margarita Gáfaro, Juan Esteban Carranza, Jorge Flórez and other participants at the CEPCO Brown Bag sessions. All remaining errors are our own.

[‡]Professional, Banco de la República de Colombia. Mailto: aperezli@banrep.gov.co

[§]Researcher, Banco de la República de Colombia. Mailto: jvelezve@banrep.gov.co

Dispersión de precios y choques de costos mayoristas en el mercado minorista de gasolina en Colombia[†]

Alex Perez[‡] y Juan Sebastián Vélez-Velásquez[§]

Resumen

La dispersión de precios es una característica prevalente de mercados en los que se transan bienes que son presuntamente homogéneos. Entre las causas de dicha dispersión de precios se destacan las interacciones estratégicas entre las firmas y entre éstas y sus clientes. La propensión de los consumidores a buscar un precio más bajo reduce la dispersión porque incrementa la competencia. La inhabilidad de las firmas para sostener colusión tácita, por otra parte, incrementa la dispersión. Los costos mayoristas pueden afectar a ambas. En este documento usamos datos de los mercados de gasolina minoristas colombianos para averiguar si los patrones de dispersión que se observan después de un choque de costos son explicados porque los consumidores buscan más intensivamente o porque las firmas se desvían de un equilibrio colusivo. Además, exploramos el papel que juega la estructura de mercado en el nivel de dispersión de precios. Nuestros hallazgos sugieren que los cambios en el nivel de dispersión de precios después de un choque de costos son determinados por consumidores que buscan más intensamente por menores precios. También encontramos una relación positiva entre el número de estaciones de servicio en el mercado y que tan dispersos son los precios. Nuestros resultados son robustos a distintas formas de definir el precio y los mercados relevantes.

Palabras clave: *Dispersión de precios, mercados de gasolina, costos mayoristas, precios minoristas.*

Clasificación JEL: *D82, D83, Q49, L11, L94.*

[†] Agradecemos a Margarita Gáfaro, Juan Esteban Carranza, Jorge Flórez y demás participantes en las sesiones Brown Bag del CEPCO. Todos los errores restantes son nuestros.

[‡] Profesional, Banco de la República de Colombia. Correo: aperezli@banrep.gov.co

[§] Investigador, Banco de la República de Colombia. Correo: jvelezve@banrep.gov.co

1 Introduction

In simple models of price competition, the law of one price suggests that competition between firms selling a homogeneous good leads to a single equilibrium price. Despite this theoretical prediction, price dispersion is a recurrent characteristic of retail markets with many competing firms. This stylized fact motivated the development of models that rationalize the existence of price dispersion in homogeneous goods markets like the one developed by Varian (1980). Retail gasoline markets have captured the attention of many researchers because they are a good example of markets where price dispersion is persistent despite the product being a homogeneous good. In this article we study how price dispersion is affected by unanticipated wholesale costs shocks. Besides, we link our empirical findings to two branches of theory literature that offer opposite predictions about the impact that wholesale shocks have on price dispersion in a bid to find which branch better explains the dispersion of prices in our markets.

There are different potential causes for the observed price dispersion in markets with homogeneous goods. For instance, some inquiries on the determinants of price dispersion in gasoline markets have looked into the role played by market structure, firm characteristics, branding, geographic differentiation, market concentration, and search costs (Eckert, 2003; Barron et al., 2004; Eckert and West, 2004; Hosken et al., 2008; Chandra and Tappata, 2011; Lewis, 2011) in determining the level of dispersion. Other studies have considered pricing strategies based on Edgeworth cycles (Maskin and Tirole, 1988), whereby managers make significant price increases followed by a sequence of slight price reductions to attract the attention of search-intensive consumers. This pricing strategy has been extensively studied in the literature (Castanias and Johnson, 1993; Zimmerman et al., 2013; Lewis and Noel, 2011; Byrne and De Roos, 2019; Gonzalez and Hurtado, 2021).

The theoretical predictions about how cost shocks affect price dispersion are not straightforward. Consumer search models, for example, predict an inverse relationship between marginal cost and price dispersion. In such models, firms anticipate that an increase in prices would make consumers search for the lowest price more intensely which reduces the margins and price dispersion as a result. On the other hand, tacit collusion models predict a direct relationship between cost shocks and dispersion. An increase in costs reduces the margins, which reduces the expected punishment for deviating from the collusive equilibrium, which in turn reduces the incentives of the firms to remain in the collusive equilibrium. By deviating from the collusive equilibrium the firms increase price dispersion. So theory alone cannot guide our understanding of the relationship between cost shocks and price dispersion because there are at least two different families of models with opposite predictions.

In this paper, we follow the estimation strategy proposed by Lewis (2011) to test the theoretical predictions derived from his consumer search costs model. One of the predictions of his model, of interest to us, is the relationship between price dispersion and profit margins. An increase in marginal costs induces firms to rise their prices, but their ability to increase prices is constrained by the degree of search intensity of the consumers in the market. This model thus predicts that both effects, change in marginal costs and change in search intensity, determine the observed relationship between price dispersion and profit margins. Additionally, Lewis (2011) proposes an empirical strategy to test these predictions in the particular context of a retail gasoline market.

To the best of our knowledge, this is the first attempt at carefully characterizing the causes of prevalent price dispersion in the Colombian retail gasoline markets. Articles related to the Colombian gasoline markets have focused on studying how changes in international prices are transmitted to domestic prices (Hofstetter and Tovar, 2010). Our work is, as far as we know, the first to explore the relationship between price dispersion and wholesale costs in the country. We use detailed price data from the Colombian retail gasoline market to test whether changes in price dispersion following cost shocks are better explained by consumer search or by tacit collusion. We compute from our data several measures of prices dispersion for each market and regress these measures on covariates that can be linked to either consumer search or tacit collusion models. Statistic tests then are used to to assess whether the data favors one type of explanation for price dispersion or the other.

Our article improves on earlier work such as Chandra and Tappata (2011) by having a more precise measure of wholesale costs that includes transportation costs, and by being able to control for national and local taxes that affect the retailer's unit cost. Lewis (2011) uses a richer dataset for his empirical strategy but focuses on a single urban agglomeration. Our dataset includes information on gasoline retailers in the four biggest metropolitan areas representing the main regional markets of Colombia. Our data contain information on the retail prices of service stations operating in those four cities, along with measures of the wholesale costs faced by these stations as well as some characteristics such as the brand they operate under. The data is collected monthly by the Ministerio de Minas y Energía (MME) and our sample includes all service stations selling fuel between 2014 and 2019. Our findings suggest that changes in price dispersion following wholesale cost shocks are driven by consumers searching more intensely for lower prices. We also find a positive relationship between the number of service stations in a market and how disperse prices are. Although defining relevant markets is always a contentious issue, our results are robust to alternative ways of measuring price dispersion and alternative ways of defining relevant markets.

In the second section of this paper, we explain in more detail the theoretical aspects that govern the relationship between price dispersion and production costs. In the third section we present a brief description of the Colombian market, in the fourth section we present our empirical approach. In the fifth section we present the results, and in the sixth section we conclude.

2 Theoretical framework of price dispersion in retail markets

The existence of price dispersion in homogeneous goods markets has motivated a variety of studies. In his seminal paper Stigler (1961) rationalizes the presence of price dispersion, even in homogeneous goods markets, using consumers that differ in their ability or willingness to search. In Stigler's framework dispersion arises when some firms can extract rents from uninformed consumers keeping high prices while other firms can set prices to compete for consumers that are eager searchers. Analogously, Varian (1980), writes a model whereby firms reach a mixed-strategies equilibrium by randomizing their prices. These random prices allow the firms to price discriminate between price sensitive consumers, who purchase the product only when the prices is low, and price inelastic consumers who buy the good at any moment. At any given time, the observed price dispersion is the result of these mixed-strategies.

According to Hong and Lee (2018), there are two theoretical branches making predictions about the relationship between dispersion of retail prices of homogeneous good and changes in wholesale costs. One branch is comprised of consumer search models and the other is comprised of focal point tactical collusion models[1]. These two theoretical approaches offer contrasting predictions for the effect of wholesale shocks on price dispersion. On the one hand consumer search models predict that the relationship between price dispersion and wholesale costs is negative, because firms anticipate that an increase in wholesale costs will lead to higher prices and consumers will search more, which enhances competition. On the other hand, tacit collusion models predict a positive relationship because an increase in wholesale costs reduces the margins that help sustain the collusive equilibrium thus reducing the punishment from deviating from the collusion.

Chandra and Tappata (2011) develop a consumer search model to study how consumer information affects price dispersion. They characterize an equilibrium relationship between dispersion of prices and relevant variables such as production cost, the number of firms in the market, and consumer search cost. Their predictions indicate that price dispersion increases with the number of firms in the market, decreases with the

[1] Some articles that link their work to search models are Borenstein et al. (1997); Yang and Ye (2008); Tappata (2009); Lewis (2011). Examples of articles framed around focal point tacit collusion areGreen and Porter (1984); Rotemberg and Saloner (1986); Haltiwanger and Harrington (1991); Borenstein and Shepard (1996).

production cost, and increases with search costs. Their model assumes that consumers know the firms' production costs, thus, the consumers' response to changes in production costs is fully anticipated. Similarly, Pennerstorfer et al. (2020) provide evidence of an inverse-U shaped relationship between the share of informed consumers, which is akin to Chandra's search intensity, and price dispersion. Their work suggests that price dispersion is significantly smaller in markets where firms face a large share of either informed or uninformed consumers.

Lewis (2011) proposes a model with heterogeneous consumers that differ in their search costs and relaxes the extent of knowledge they have about the firms' production costs. Although consumers do not know the production costs, they form expectations about them based on the history of prices. In this scenario, the model's predictions evoke results found in contexts with complete information on production costs. When consumers form expectations about prices and a positive cost shock occurs, firms know that consumers will anticipate the increase in prices and more consumers will search for the lowest price. More consumers searching for lower prices enhances competition in a way that prevents sellers from profiting from prices that are too different than their rivals' prices, hence reducing dispersion. When a negative shock occurs the sellers tend to reduce their prices. However, not all sellers will lower the prices by the same amount as it depends on the proportion of non-searching consumers that they serve, thus dispersion increases. Sellers that serve non-search consumers will adjust their prices little, while those that serve search consumers will reduce their prices largely. Lewis' model, then, predicts predicts a negative (positive) relationship between wholesale costs (margins) and dispersion. Lewis uses data from American fuel stations to test the predictions of his model but is unable to control for search intensity. He shows that prices respond faster to cost changes during periods when margins are high

Among the first to use models of focal point tacit collusion is Green and Porter (1984) who write a model in which collusion can be sustained if, and only if, the market price is above a threshold level or trigger price. Inspired by Green and Porter, Rotemberg and Saloner (1986) write a tacit collusion model that leads to price wars. In their model a positive demand shock breaks collusion because with high demand a single firm's incentives to deviate captures a large part of the market until the others are able to change their prices. On the other hand punishment form deviating is less affected by the state of demand if punishments are distributed in the future. Borenstein et al. (1997) hypothesize that past prices could serve as focal points that allow retailers to coordinate. Borenstein and Shepard (1996) reinterpret demand shocks as dynamic cost changes and find that the profit margins associated with collusion increase with negative cost shocks and decrease with positive shocks. In their model, the ability of firms to sustain collusion is directly related to margin changes resulting from wholesale cost fluctuations. When a positive wholesale cost shock occurs, retailers would no longer collude and set independent prices because, otherwise, margin profits would become negative. In contrast, collusion is easier to sustain with cost declines because maintaining past prices benefits the margin profits for all firms. Another relevant variable to explain collusion, and price dispersion, is the number of firms in the market. Borenstein and Shepard (1996) claim that tacit collusion models predict greater sustainability of coordination when fewer firms exist in the market. Therefore, when there are few firms in a market it is reasonable to expect low dispersion of prices as coordination is easier with fewer firms.

As mentioned above, focal point tacit collusion models predict a positive relationship between the dispersion of retail prices and wholesales costs. When the wholesale cost increases price dispersion is higher because firms have incentives to break the tacit collusion and set prices independently because as margins fall so do expected punishments from deviating. On the other hand, when wholesale costs decrease (margins increase) it is profitable (more costly) for firms to set coordinated prices at a focal point (to deviate), and the result is a less (more) disperse distribution of prices. This implies a prediction that contradicts those made by search models, that is, a negative relationship between price dispersion and profit margins.

These two sets of theories predict testable relationships between price dispersion and profit margins that go in different directions. The relationship is positive for the consumer search cost models, and negative for the focal point tacit collusion models. In relation to the number of firms in the market, both theories predict a direct relationship between price dispersion and the number of firms.

3 Colombian gasoline market

Colombia's retail gasoline market comprises stations that sell fuel to boats, airplanes, automobiles and motorcycles. In this article we focus exclusively on service stations that fuel automobiles and motorcycles. Most aspects of this market are regulated by the Ministry of Mines and Energy (MME). Most importantly, prices have been subject to regulation for many years. For instance, before 1998 the MME would set the fuel prices charged in every station. In 1998 the market was liberalized so that gasoline prices would reflect the price movements of related commodities (Caicedo-García and Tique-Calderón, 2012). Based on the MME's assessed level of competition in Colombian cities, the current regulation effectively divides the country in two subsets of cities. One subset corresponds to cities that were deemed competitive enough to allow stations to set prices freely. The other subset corresponds to cities where, according to the MME's assessment, competition alone was insufficient to achieve efficient pricing. Thus, stations operating in the latter group of cities face a price ceiling.

In 2011, the price regulation experienced a substantial change. The new regulation had to balance two objectives that often go in oppoiste directions. First, retail prices should better reflect the movement of prices of related commodities to ensure producers were properly compensated and would continue having enough incentives to provide fuel in the domestic market. Second, consumers should be shielded from the largest variations in international oil prices. In subsequent years the regulation underwent several changes spurred by substantial changes in international gasoline prices as the regulator tried to protect consumers. Under the current scheme the price consumers pay at the pump consists of four main components: 1) producer's revenue, 2) taxes, 3) transportation costs, and 4) margins. The producer's revenue refers to the opportunity cost faced by the producers of gasoline–Ecopetrol being the lone producer of gasoline– and ethanol when selling their products domestically. Effectively the producers' revenues are monotonic functions of international gasoline prices–specifically the price of Gulf Coast UNL 87 Gasoline M2 (Platts)–and international sugar prices–as the input for producing ethanol is sugar cane–. Several taxes affect the price paid by consumers at the pump like a national tax on gasoline and several regional and municipal surcharges. Transportation costs include the cost of moving gasoline from the refinery to the fuel rack through pipelines and from racks to stations in trucks. The other component of the price determination is the margins. For wholesalers these margins are defined as a fixed amount of pesos for every gallon of gasoline sold to the retailers. For retailers the margin is the amount of pesos over the cost of having a gallon of fuel delivered on site. Depending on where the station operates the retailer's margin might be subject to regulation. As we mentioned above, the regulator segregates cities into two groups. In some cities the regulator defines a maximum retailer's margins, so in practice, stations that operate in those cities face a price ceiling. In cities where stations are free to set prices, each station sets retail prices by adding an arbitrary–albeit constrained by its competitors– amount of pesos to the cost of a gallon of gasoline. An additional institutional feature is that the MME's publicates monthly a "reference price" for the country's main 23 free-market cities (Hofstetter and Tovar, 2010). This price is published the first week of each month, it is city-specific and not mandatory for retailers, which means that stations can set prices above or below the reference price. The MME states that this price represents all costs in the production and distribution chain that allow retailers to earn a reasonable profit.

4 Methodology

4.1 Data

We use data on the price per gallon of gasoline charged by all service stations operating in Colombia. Stations must report these retail prices to the MME in the first five days of the month–regardless of whether they keep charging the same price of the previous month–and every time they change it during a given month. Our data cover the period January 2014 through December 2019. The dataset includes the name and address of the service station, the department and municipality where it is located, and its brand (Texaco, Mobil, Terpel, etc.). These stations sell four types of liquid fuel (diesel, regular gasoline, and premium gasoline) and some sell autogas (liquified petroleum gas). Since stations report several prices during the month, we use the monthly average retail price of gasoline–regular and premium–. We also focus on cities belonging to the four largest urban areas of Colombia: Bogotá, Medellín, Cali and Barranquilla[2]. We use data from these urban agglomerations, first, because stations in them are not subject to price ceiling. We drop cities where stations are constrained by price ceilings because markets in such cities are bound to exhibit less price dispersion in a way that is induced by the regulation instead of being caused by the consumers eagerness to search or by tacit collusion. Another reason why we use only stations operating in the largest metropolitan areas is the reliability of our geolocation algorithm, which is more accurate for addresses originated from large cities. Third, we focus on this metro areas because their high population density alone eases any concerns about how the market definition could affect our econometric results. Finally, markets in these large cities have enough cross-section variation in the number of competing stations to allows us to identify the role that market structure plays on price discrimination.

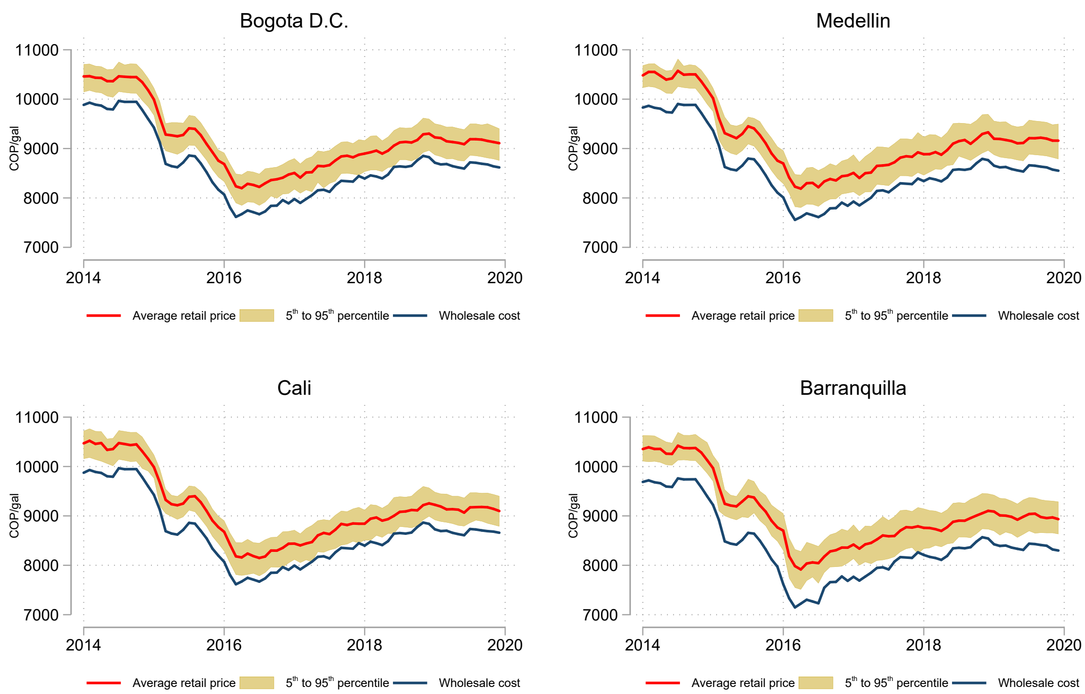
We also obtained cost data from the upstream part of the business. These include producer's revenue, pipeline transportation costs, and truck transportation costs from the wholesaler to the retailer. Since transportation costs, both via pipeline and via truck, are determined by the distance, they vary between cities. We add these costs to obtain a measure of the wholesale cost of gasoline for the service stations, which in turn we use to obtain a measure of the profit margins of the stations.

In Figure 1 we show our measure of wholesale cost, the distribution of retail prices and the reference price of gasoline by cities deflated by CPI.There are two noteworthy periods. In the first one, going from 2014 to the first half of 2016, the retail prices and costs decrease due to a fall in international fuel prices. During the second noteworthy period, going from the second half of 2016 to 2019, the wholesale cost shows a growing trend that coincides with a remarkable growth in the global demand for fuel. We exploit these shocks to understand how the dispersion of retail prices responds to wholesale cost variation. The retail and reference prices exhibit a temporary behavior similar to that of the wholesale cost. A notable feature is that price dispersion, measured as the band surrounding the average retail price, increases after 2016 when the

[2]From the Bogotá urban area we include Soacha. From the Medellín urban area we include Medellín, Copacabana, Bello, Envigado, Itagüí, Sabaneta, La Estrella and Caldas. From the Cali urban area, Cali, Yumbo and Jamundi are included. Finally, we include Barranquilla and Soledad form the urban area of Barranquilla.

wholesale cost starts trending upwards. This seems to indicate that there is a positive relationship between dispersion and costs. This is the kind of feature we want to formally test with the econometric approach presented below that allows us to link this initial empirical conjecture to a sound theoretical framework.

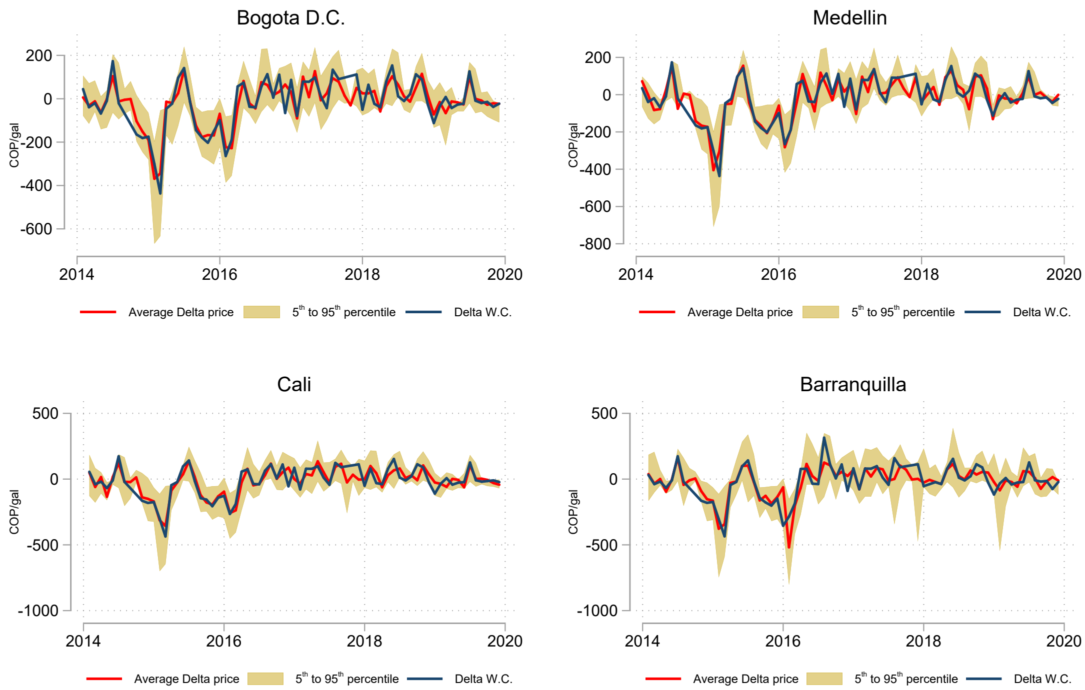
Figure 1: Retail prices and wholesale cost



In Figure 2 we show, by city, the distribution of the monthly changes in retail prices and wholesale costs for all stations in the data. The series of average change in retail prices and average change in wholesale costs follow a similar dynamic. Both series show strong falls during the earlier periods and a more stable behavior after 2016.

We use the Google Maps API to geolocate the stations in our data. Stations report their location through a municipal address system and as well as their commercial names. With these two characteristics we can uniquely identify a station within a city. Our algorithm fails to geolocate 37 out of 830 stations. Those stations are geolocated manually. Next we want to define relevant markets within a city to determine the number of competitors that each station faces. Following Chandra and Tappata (2011) and Lewis (2008), we define the the relevant market for each station as the number of competitors within a one or two kilometer radius. This definition implies that a station competes in several markets at the same time, reflecting the competitive environment faced by stations. Since we are interested in calculating the price dispersion of

Figure 2: Change in retail prices and wholesale cost



gasoline prices, we drop markets with fewer than three competitors. We end up with 382 stations operating in the metropolitan area of Bogotá, 168 in Medellín, 173 in Cali and 107 in Barrranquilla. Our definition of overlapping markets implies the existence of high correlation between the unobserved characteristics of the markets. We alternatively design a strategy to define non-overlapping markets by creating clusters of stations based on Euclidean distance within the metropolitan areas. We select the number of clusters considered within agglomerations using hierarchical agglomerative cluster analysis and Duda-Hart rule. We obtained 17 non-overlapping markets for Bogotá, 13 for Medellín, 8 for Cali 8 and 5 for Barranquilla.

We use three different measures of price dispersion: the standard deviation, the max-min range, and the interquartile range. First we calculate dispersion for the posted prices. Second we construct adjusted prices as in Lewis (2011) by discounting the average premium from the observed price charged by the station[3]. This way we control for unobserved time-invariant differences between stations, like amenities, that allow stations to consistently charge more or less than others on average. The implicit assumption about the interaction between product differentiation and imperfect information is that they are additive separable (Chandra and Tappata, 2011).

[3]We calculate the adjusted price as $p_{it}^{adj} = p_{it} - \frac{1}{T}\sum_{t=1}^{T}(p_{it} - \frac{1}{N}\sum_{i=1}^{N} p_{it})$, that is, the price discounting the average over time of the difference between the price of the station and the average price of the city in each month.

In Table 1, for two alternative market radius, we present a descriptive summary of price dispersion as measured by the average standard deviation of observed prices, the average number of competitors faced by each station, the average number of different brands in the market, and average margins. The average dispersion of prices is greatest in Medellín and lowest in Bogotá but due to the a high variability of this dispersion, the difference is not statistically significant. The standard deviation of prices in markets defined at a 2 km radius tends to be 250 COP/gal, while it is 210 COP/gal at a 1 km radius. The number of competitors, which is around 16 for a radius of 2 km and 5 for a radius of 2 km, does not vary between cities. The number of different brands operating within a 2 km radius ranges between 6–for Barranquilla– and 4–for Medellin. In a radius of 1 km, the differences in the number of competitors between cities is less salient with stations in Barranquilla facing on average 4.4 competitors within 1 km and stations in Cali competing with 2.9 stations. Regarding profit margins, stations in Barranquilla earn on average 150 COP per gallon more than stations in Cali.

Table 1: Descriptive statistics for overlapping markets

	Number of markets	Price dispersion - S.D. (COP/gal)		Number of Competitors		Number of Brands		Profit margins (COP/gal)	
Market radius		2 km	1 km	2 km	1 km	2 km	1 km	2 km	1 km
Barranquilla	107	244,30	203,63	16,68	5,78	6,19	4,42	648,12	652,36
		(178,93)	(206,07)	(8,95)	(3,98)	(1,17)	(2,05)	(149,76)	(206,85)
Bogotá D.C.	382	242,64	214,98	15,05	4,33	5,14	3,36	521,03	524,06
		(178,41)	(211,42)	(7,52)	(2,94)	(1,2)	(1,31)	(116,44)	(169,62)
Medellín	168	276,58	236,61	16,50	4,59	4,02	2,92	593,16	589,90
		(234,49)	(282,04)	(6,88)	(2,46)	(0,78)	(0,93)	(126,73)	(200,99)
Cali	173	249,77	205,69	16,04	4,86	4,06	2,89	489,74	494,23
		(196,99)	(225,08)	(9,08)	(2,95)	(0,85)	(0,97)	(142,82)	(204,61)

Note: Price dispersion is measured as the standard deviation of the distribution of prices within a market. Each column shows the mean values of the variable and standard deviation in parentheses. Because we define overlapping markets, there are as many markets as are stations.

Table 2 show similar descriptive statistic as table 1 but for non-overlapping markets. The features of average price dispersion and margins resemble those obtained for overlapping markets. However the non-overlapping markets are larger. For instance, there are on average 33 stations operating in non-overlapping markets in Barranquilla, 25 in Bogotá, 24 in Cali and 12 in Medellín. Similarly, the average number of different brands within a market is seven for Barranquilla, six for Bogotá, and three in Cali and Medellín.

The scatter plot in Figure 3 shows the changes in dispersion of retail gasoline prices against the changes in the wholesale cost for the four metropolitan areas. As discussed above, consumer search models imply

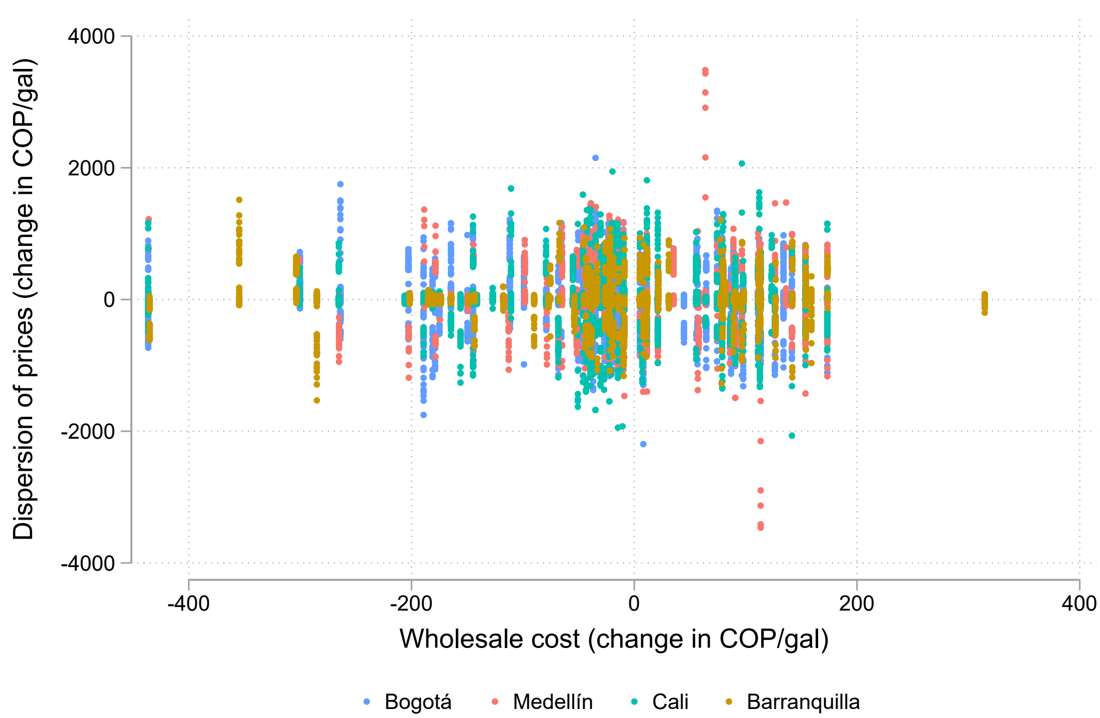
Table 2: Descriptive statistics for non-overlapping markets

	Number of markets	Price dispersion - S.D. (COP/gal)	Number of stations	Number of Brands	Profit margins (COP/gal)
Barranquilla	3	237.49	33.18	6.91	661.74
		(191.27)	(27.44)	(1.58)	(150.85)
Bogotá D.C.	14	229.47	25.04	5.64	528.41
		(153.55)	(19.25)	(1.16)	(99.03)
Medellín	14	253.72	12.05	3.24	567.91
		(336.93)	(12.17)	(1.15)	(202.55)
Cali	7	219.22	23.94	3.24	473.59
		(264.21)	(29.72)	(1.59)	(214.86)

Note: Price dispersion measured as the standard deviation of the distribution of prices within a market. Each column shows the mean value of the variable and standard deviation in parentheses.

a negative relationship between these two variables. In contrast, tacit collusion models predict a positive relationship between cost shocks and price dispersion. Nevertheless, the evidence provided by the scatter plot regarding the relationship between these two variables is inconclusive at best.

Figure 3: Dispersion of price and Wholesale costs



4.2 Empirical strategy

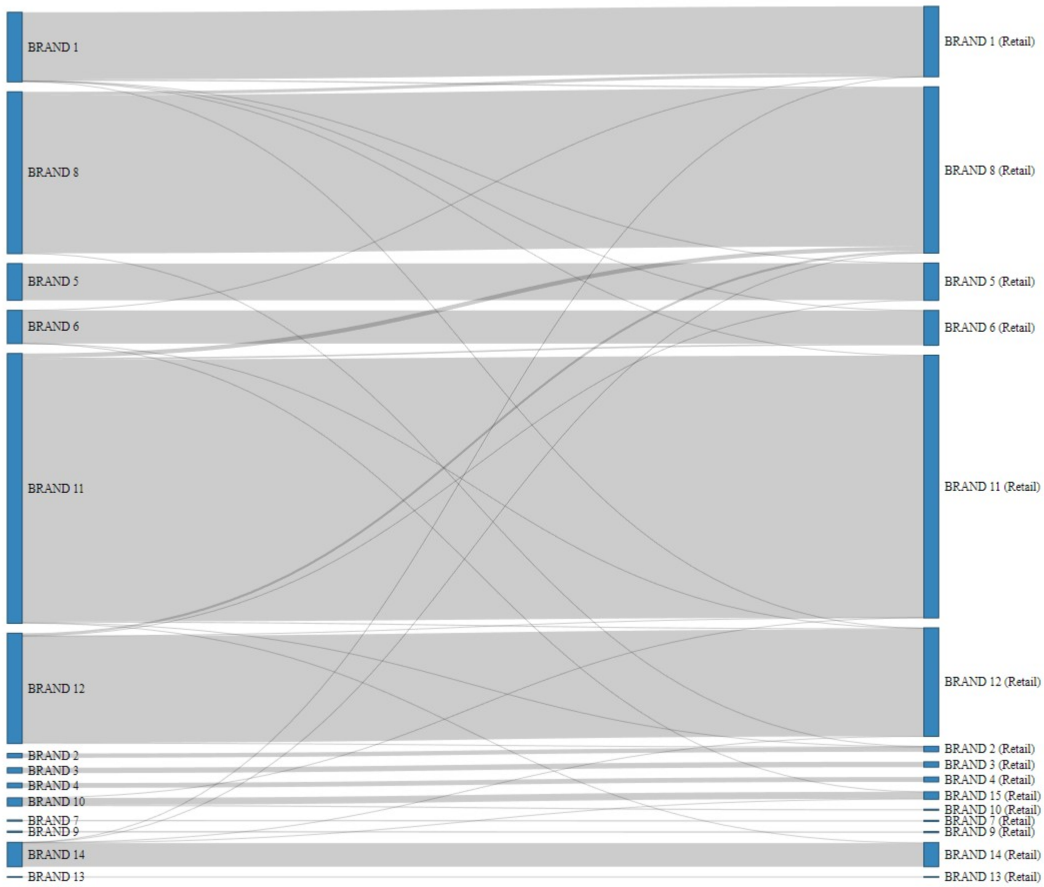
To test the sign of the the relationship between the dispersion of retail prices and wholesale costs, Lewis (2011) proposes the following econometric specification

$$\sigma_{mt} = \beta_0 + \beta_1 M_{m,t-1} + \beta_2 \Delta(P_{mt}) + \theta_m + \epsilon_{mt}, \tag{1}$$

where σ is the price dispersion measure of market m in time t, M is the average profit margins of market, $\Delta(P)$ is the change in the average retail price, θ is a vector of market fixed effects, and ϵ_{mt} are other unobserved things affecting the price dispersion that we do not account for. M enters equation 1 lagged, because the contemporaneous markups are likely to be endogenous. Consumer search models predict $\beta_1 > 0$, that is, increases in margins are associated with high price dispersion whereas decreases in margins lead to lower price dispersion. Tacit collusion models predict $\beta_1 < 0$ because an increase in wholesale costs would reduce margins and should make tacit collusion less profitable, therefore, there is a tendency for stations to break away from collusion. The result is a more disperse price distribution. The average change in retail price acts as a control for temporary price differences that usually occur during market-wide price movements simply because some stations adjust their prices faster than others. These larger overall price movements create temporary dispersion of prices that would not explained by margins adjustment and lack the strategic component described above. The fixed effects, θ, control for time-invariant unobserved features of the market such as the station being close to a highway, a high traffic area, or another feature that drives up demand such as malls and stadiums. We estimate equation 1 via OLS.

The industry exhibits a high degree of vertical integration and there is very little upstream competition. This vertical integration can take two forms. On the one hand, wholesalers own and operate some of the stations. For instance, of the 2,653 station operating in 2021, 157 (6%) are operated by a single wholesaler (Brand 11). On the other hand, independently owned and operated stations sign exclusive distribution contracts with a wholesaler. In practical terms this means that an individual station operating under Brand B can only buy fuel from the wholesaler B unless wholesaler B is not able to fulfill the order. Thus, most of the fuel bought upstream comes from wholesalers that operate under the same brand as the station. To illustrate the lack of interaction between wholesalers and retailers of different brands, we show in Figure 4 the flows of fuel between wholesalers and retailers for March 2021. The thickness of each line corresponds to the amount of gallons traded during the month. Moreover, individual stations differ on how much agency their managers have to make strategic decisions, such as pricing. The difference in a manager's ability do take strategic decisions arises from the different types of contracts that govern the relationship wholesaler-station. Some stations are owner-operated (franchised) while others are directly operated by the wholesaler. These

Figure 4: Vertical integration between Retailers and Wholesalers



Note: Amount of fuel in gallons flowing between wholesalers and retailers during March 2021.

characteristics, type of ownership and vertical integration, should have an impact on price dispersion. For instance a market where all stations are of the same brand, and hence share common costs, should be less disperse than a market with many brands that don't have a common upstream monopolist. Similarly one should observe less dispersion in markets with a high incidence of stations that make pricing decisions in a centralized manner. For these reasons we modify the baseline specification to study the effect that the number of brands has on price dispersion.To do this, in each market defined around a station, we identify the number of brands (B) and include it as an explanatory variable, as well as its interaction with profit margins. Our prior is that tacit collusion is more difficult to sustain in markets with more brands. Thus, we expect $\gamma_2 > 0$, and a negative coefficient to the interaction between brands and margins (γ_3).

$$\sigma_{mt} = \gamma_0 + \gamma_1 M_{m,t-1} + \gamma_2 B_{mt} + \gamma_3 M_{m,t-1} \times B_{mt} + \gamma_4 \Delta(P_{mt}) + \eta_m + u_{mt}. \quad (2)$$

5 Results

In this section, we present the results of estimating the equations presented above with data from the Colombian gasoline market to study the relationship between price dispersion and wholesale costs. A first approach is simply regressing measures of price dispersion on wholesale costs and other relevant controls. In Table 3 we present the results of this regression for different specifications. In columns 1 through 4 the dependent variable is the standard deviation of prices. In columns 5 through 7 the dependent variable is the standard deviation of adjusted prices defined above. Columns 1 through 5 use markets defined as competitors within a 2 km radius of a given station, while columns 6 and 7 use the alternative 1 km radius. In column 1 we only include wholesale cost as covariate, column 2 adds the number of competitors, column 3 adds fixed effects, and column 4 adds the number of different brands in the market as control. Standard errors are clustered at the metropolitan area-time level. In general, we find evidence of a negative relationship between dispersion, measured as the standard deviation of prices within the market, and wholesale costs. The only exception is the estimates of wholesale costs in columns 5 and 7 where the standard deviation is calculated using adjusted prices and the radius defining the market is 2 km.We interpret this negative relationship as evidence in favor of consumer search models. Finally, we find no evidence of a relationship between price dispersion and the number of competitors or the number of different brands in a market. However this is expected because we are including market fixed-effects and the market structure, be it number of brands or number of competitors, is relatively constant over time.

Table 3: Price dispersion and wholesale costs

	(1) Std. Dev. Obs. Prices	(2) Std. Dev. Obs. Prices	(3) Std. Dev. Obs. Prices	(4) Std. Dev. Obs. Prices	(5) Std. Dev. Adj. Prices	(6) Std. Dev. Obs. Prices	(7) Std. Dev. Adj. Prices
Wholesale cost	-0.0173***	-0.0167***	-0.0168***	-0.0168***	0.00955*	-0.0161***	0.0102**
	(0.00546)	(0.00542)	(0.00547)	(0.00546)	(0.00552)	(0.00422)	(0.00404)
Number of competitors		2.640***	1.643	-0.126	-4.451	2.261	1.741
		(0.352)	(3.163)	(3.390)	(3.630)	(6.281)	(6.623)
Brands				13.16	-1.657	4.024	-13.69
				(8.055)	(8.711)	(8.705)	(8.788)
Constant	399.7***	351.4***	368.5***	334.6***	181.0**	328.4***	99.88**
	(46.61)	(47.14)	(68.96)	(71.71)	(76.41)	(44.24)	(44.58)
FE by market	No	No	Yes	Yes	Yes	Yes	Yes
Market radius	2 km	2 km	2 km	2 km	2 km	1 km	1 km
Observations	52,699	52,699	52,695	52,695	52,695	49,844	49,844
R-squared	0.003	0.016	0.120	0.121	0.092	0.153	0.107

Note: Standard errors are corrected by city-time level. The dependent variable is the standard deviation of retail prices.
*** p<0.01, ** p<0.05, * p<0.1

An important determinant of the price dispersion of a market is how willing to search consumers are. For instance, a market where all consumers are willing to search for lower price intensely should exhibit lower dispersion than a market where search intensity is low. In that sense the previous specification might suffer from omitted variable bias, since we are not accounting for the search intensity in each market. To address this potential bias, we link our specification to the theoretical relationship implied by consumer search or tacit collusion models by including the margins instead.

We present the results of estimating equation 1 using OLS with margins instead of wholesale costs in Table 1 for different specifications. As before, standard errors are clustered to allow f or correlation between stations operating in given city-month. Columns 1 through 5 define markets as competitors within a radius of 2 km from a given station, and columns 6 through 8 define them within of 1 km. In column 1 only the margins are used to explain dispersion. The estimate on column 1 suggests that there is a negative relationship between dispersion and margins, as expected from Figure 1. However, when we control for temporary price differences and number of brands (columns 2), the relationship between dispersion and margins becomes positive. This implies that the heterogeneous adjustment of stations is a relevant factor determining price dispersion. In some specifications, namely columns 2 and 5, the number of brands is statistically significant corroborating our prior about dispersion increasing with the number of different brands operating in a market. When we include market fixed effects, the magnitude of the coefficient that relates dispersion and margins increases (columns 3 and 6) when markets are defined with a radius of 2 km. This means that the effect that margins have on dispersion is dampened when we pool all the markets. In other words, markets with systematically low dispersion correspond with markets where margins tend to be high which hints at collusion. In contrast, by adding market fixed-effects we are effectively comparing changes in margins and changes in dispersion within the same market. Columns 4 and 7 show that the interaction of margins and the number of brands is not statistically significant for either definition of relevant market. Finally, columns 5 and 8, where we use adjusted prices to compute the dependent variable, confirm the positive impact that margins have on dispersion.

Since there are several ways of assessing the dispersion of the distribution of prices in a market we present in Table 5 the results of estimating equation 1 using alternative measures of dispersion. Columns 1 through 4 use markets defined within a radius of 2 km from the station, and columns 5-8 define markets within a radius of 1 km. In odd numbered columns the dependent variable is the max-min range, while in even numbered columns it is the interquartile range. Columns 3, 4, 7 and 8 use adjusted prices. All columns include market fixed effects, and the standard errors are clustered at market-month level. Our results provide strong evidence that supports the theoretical predictions of consumer search models when markets are defined within 1 km

Table 4: Price dispersion and profit margins

	(1) Std. Dev. Obs. Prices	(2) Std. Dev. Obs. Prices	(3) Std. Dev. Obs. Prices	(4) Std. Dev. Obs. Prices	(5) Std. Dev. Adj. Prices	(6) Std. Dev. Obs. Prices	(7) Std. Dev. Obs. Prices	(8) Std. Dev. Adj. Prices
Profit margins	-0.0802***	0.188***	0.510***	0.492***	0.671***	0.614***	0.643***	0.702***
	(0.0241)	(0.0356)	(0.0709)	(0.121)	(0.0757)	(0.0478)	(0.0696)	(0.0492)
Δ(p)		0.675***	0.798***	0.798***	0.816***	0.837***	0.838***	0.847***
		(0.0616)	(0.0790)	(0.0790)	(0.0904)	(0.0554)	(0.0555)	(0.0606)
Brands		5.379**	12.82	10.78	-6.902	16.22**	20.60*	-6.026
		(2.166)	(8.520)	(14.34)	(9.200)	(7.138)	(10.89)	(7.249)
Profit margins×Brands				0.00401			-0.00919	
				(0.0211)			(0.0149)	
Constant	296.4***	137.0***	-72.32	-62.80	-138.4**	-157.7***	-171.4***	-199.5***
	(13.86)	(23.65)	(61.27)	(82.28)	(67.97)	(41.82)	(50.61)	(45.06)
FE by market (station)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Market radius	2 km	2 km	2 km	2 km	2 km	1 km	1 km	1 km
Observations	51,414	51,414	51,413	51,413	51,413	48,649	48,649	48,649
R-squared	0.003	0.204	0.352	0.352	0.326	0.464	0.464	0.450

Note: Standard errors are corrected by city-time level. Profit margins are lagged one month.
*** p<0.01, ** p<0.05, * p<0.1

as we obtain positive estimates for both measures of dispersion–range and interquartile range– and for both types of prices–observed and adjusted.

For markets defined at a 2 km radius the relationship between price dispersion and profit margins is positive, except when dispersion is measured as the interquantile range of observed prices which indicates that most of the dispersion is driven by outlier prices prices. This is confirmed by the result in column 1 which uses the range of observed prices as the measure of dispersion. In relation to the variable controlling for temporary price changes, we find a positive relationship across all specifications. This implies that the dispersion of the prices can be explained, in part, by the heterogeneity in the time individual stations take to adjust their prices. When there is a positive wholesale cost shock, some stations may adjust their prices faster than others stations, thus creating a temporary price dispersion that is not explained by the profit margin adjustment. Finally, the evidence provided in the table suggests that the number of brands has a positive relationship with price dispersion of observed prices that is robust to market definition.

As stated before the definition of overlapping markets raises concerns about the regressors in equation 1 being highly correlated, which implies that unobservable characteristics would also be highly correlated. For non overlapping markets, whose results are shown in table 6, the data suggest a positive relationship between price dispersion and margins and such result is robust to the way dispersion is measured. The estimate on the variable controlling for the heterogeneity in adjustment times to aggregate shocks is robust to the definition of non-overlapping markets. As expected, under this specification we find no relationship between the

Table 5: Measures of price dispersion

	(1) Range Obs. Prices	(2) Range Q Obs. Prices	(3) Range Adj. Prices	(4) Range Q Adj. Prices	(5) Range Obs. Prices	(6) Range Q Obs. Prices	(7) Range Adj. Prices	(8) Range Q Adj. Prices
Profit margins	2.134***	-0.139***	2.488***	0.0744***	1.240***	0.425***	1.393***	0.497***
	(0.283)	(0.0287)	(0.292)	(0.0282)	(0.105)	(0.0532)	(0.106)	(0.0540)
Δ(p)	3.122***	0.0770**	3.055***	0.00921	1.671***	0.616***	1.664***	0.580***
	(0.308)	(0.0304)	(0.331)	(0.0407)	(0.117)	(0.0629)	(0.125)	(0.0677)
Brands	44.87	18.85***	-28.87	-5.274*	93.97***	37.78***	17.56	-1.438
	(43.54)	(2.912)	(44.40)	(2.868)	(19.83)	(5.987)	(19.68)	(5.675)
Constant	-388.5	266.4***	-430.0	121.7***	-438.5***	-87.88**	-417.9***	-108.8***
	(273.2)	(22.89)	(286.7)	(24.25)	(94.36)	(37.54)	(98.61)	(39.71)
Market radius	2 km	2 km	2 km	2 km	1 km	1 km	1 km	1 km
Observations	51,413	51,413	51,413	51,413	51,413	51,413	51,413	51,413
R-squared	0.327	0.441	0.305	0.180	0.425	0.388	0.383	0.316

Note: Standard errors are corrected by city-time level. All regressions include fixed effects by market. Profit margins are lagged one month.
*** p<0.01, ** p<0.05, * p<0.1

number of brands in the market and price dispersion, because after including market-fixed effects there is virtually no variation in the number of brands of the market.

Table 6: Price dispersion for non-overlapping markets

	(1) Std. Dev. Obs. Prices	(2) Range Obs. Prices	(3) Range Q Obs. Prices	(4) Std. Dev. Adj. Prices	(5) Range Adj. Prices	(6) Range Q Adj. Prices
Profit margins	0.559***	1.569***	0.597***	0.733***	2.011***	0.658***
	(0.0987)	(0.204)	(0.196)	(0.0924)	(0.181)	(0.207)
Δ(p)	0.876***	2.369***	0.911***	0.892***	2.396***	0.808***
	(0.113)	(0.185)	(0.249)	(0.114)	(0.192)	(0.275)
Brands	15.70	65.52	26.61***	2.326	10.91	-4.419
	(11.05)	(54.46)	(10.09)	(12.15)	(57.31)	(8.914)
Constant	-120.4	-228.9	-170.5	-214.9***	-382.0	-173.2
	(80.43)	(278.4)	(119.9)	(81.21)	(286.6)	(121.2)
Observations	2,595	2,595	2,595	2,595	2,595	2,595
R-squared	0.481	0.431	0.433	0.494	0.407	0.450

Note: Standard errors are corrected by city-time level. All regressions include fixed effects by non-overlapping market. Profit margins are lagged one month.
*** p<0.01, ** p<0.05, * p<0.1

The evidence presented so far indicates that our data favors the predictions associated with consumer search models. As those models predict, we find evidence of a positive relationship between the dispersion of gasoline retail prices and margins. Markets with average high (low) profit margins exhibit high (low) price dispersion, that is, stations tend to set more (less) similar prices as competition between them relaxes (increase), in response to higher (lower) profit margins. We also find that this positive relationship is robust to how markets are defined–overlapping or not, 2 km radius or 1 km radius– and how dispersion is measured–standard deviation, range or interquantile range. We have also explored the relationship between the number of different brands operating in a market. We find that price dispersion increases with the number of brands.

6 Conclusions

Gasoline is arguably a homogeneous good and yet price dispersion is a constant feature of retail gasoline markets around the world. In a bid to understand more about the causes of price dispersion we use data about the pricing decisions of individual service stations in Colombia to gauge the effect that wholesale costs shocks have on the price dispersion of gasoline markets. We find that, at least for Colombian gasoline markets, the response of price dispersion to wholesale shocks evokes the predictions of consumer search models. We also find that price dispersion increases with the number of stations competing in the market. Our results stand numerous robustness checks.

There are other causes of price dispersion that we can explore in the future to better our understanding of this phenomenon. For instance we can potentially look at how the type of ownership affects price dispersion. In our data we observe whether stations are operated independently or are operated by wholesaler agent. We could exploit variation across markets in the prevalence of either type of ownership to identify its impact on the distribution of prices. Finally, in our effort to link our empirical results to a solid theoretical background we have considered that either consumer search or tacit collusion could explain the relationship between changes in costs and dispersion. However, this warrants a more careful consideration because empirically one does not preclude the other.

References

Barron, J. M., Taylor, B. A., and Umbeck, J. R. (2004). Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization*, 22(8-9):1041–1066.

Borenstein, S., Cameron, C., and Gilbert, R. (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *Quarterly Journal of Economics*, 112:305–339.

Borenstein, S. and Shepard, A. (1996). Dynamic pricing in retail gasoline markets. *RAND Journal of Economics*, 27:429–451.

Byrne, D. P. and De Roos, N. (2019). Learning to coordinate: A study in retail gasoline. *American Economic Review*, 109(2):591–619.

Caicedo-García, E. and Tique-Calderón, E. Y. (2012). La nueva fórmula de la gasolina y su potencial impacto inflacionario en Colombia. *Borradores de Economía*, (698). Banco de la República de Colombia.

Castanias, R. and Johnson, H. (1993). Gas wars: Retail gasoline price fluctuations. *Review of Economics and Statistics*, pages 171–174.

Chandra, A. and Tappata, M. (2011). Consumer search and dynamic price dispersion: an application to gasoline markets. *RAND Journal of Economics*, 42(4):681–704.

Eckert, A. (2003). Retail price cycles and the presence of small firms. *International Journal of Industrial Organization*, 21(2):151–170.

Eckert, A. and West, D. S. (2004). Retail gasoline price cycles across spatially dispersed gasoline stations. *Journal of Law and Economics*, 47(1):245–273.

Gonzalez, J. and Hurtado, C. (2021). Price Differences Within Retail Gasoline Market. Available at SSRN.

Green, E. J. and Porter, R. H. (1984). Noncooperative collusion under imperfect price information. *Econometrica*, 52:87–100.

Haltiwanger, J. and Harrington, E. J. (1991). The impact of cyclical demand movements on collusive behavior. *RAND Journal of Economics*, 22:89–106.

Hofstetter, M. and Tovar, J. (2010). Common Knowledge Reference Price and Asymmetric Price Adjustments. *Review of Industrial Organization*, 37(2):141–159.

Hong, W. H. and Lee, D. (2018). Asymmetric pricing dynamics with market power: investigating island data of the retail gasoline market. *Empirical Economics*, 27:1–41.

Hosken, D. S., McMillan, R. S., and Taylor, C. T. (2008). Retail gasoline pricing: What do we know? *International Journal of Industrial Organization*, 26(6):1425–1436.

Lewis, M. (2008). Price dispersion and competition with differentiated sellers. *Journal of Industrial Economics*, 56(3):654–678.

Lewis, M. (2011). Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics and Management Strategy*, 20:409–449.

Lewis, M. and Noel, M. (2011). The speed of gasoline price response in markets with and without Edgeworth cycles. *Review of Economics and Statistics*, 93(2):672–682.

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles. *Econometrica*, pages 571–599.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., and Yontcheva, B. (2020). Information and price dispersion: Theory and evidence. *International Economic Review*, 61(2):871–899.

Rotemberg, J. J. and Saloner, G. (1986). A supergame-theoretic model of price wars during booms. *American Economic Review*, 76:390–407.

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.

Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *RAND Journal of Economics*, 40:673–687.

Varian, H. R. (1980). A model of sales. *American economic review*, 70(4):651–659.

Yang, H. and Ye, L. (2008). Search with learning: Understanding asymmetric price adjustments. *RAND Journal of Economics*, 39:547–564.

Zimmerman, P. R., Yun, J. M., and Taylor, C. T. (2013). Edgeworth price cycles in gasoline: Evidence from the United States. *Review of Industrial Organization*, 42(3):297–320.


BANCO DE LA REPÚBLICA
COLOMBIA