EL COLEGIO DE MÉXICO
CENTRO DE ESTUDIOS ECONÓMICOS

LICENCIATURA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL TÍTULO DE LICENCIADO EN ECONOMÍA

ARBITRAGE COSTS AND MARKET INTEGRATION: EVIDENCE FROM THE MEXICAN RETAIL GASOLINE MARKET

CRISTIAN EDUARDO GUDIÑO GARCIA

PROMOCIÓN 2016 – 2020

ASESOR:

DRA. DIANA TERRAZAS SANTAMARÍA

ABRIL 2021

Agradecimientos

En primer lugar, quiero hacer un agradecimiento especial a quienes ya no están, pero que explican de dónde vengo. A mi abuela Bárbara, quien educó a mi padre con todo su amor. Y a mi abuelo Santos, quien educó a mi madre y me brindó un gran apoyo para hacer este sueño posible. Además, a mi abuela Isabel y abuelo Leobardo, quienes aún están conmigo.

Segundo, quiero agradecer a mis padres por ser mi fuente de inspiración y vida. A mi madre, Mónica, quien siempre cuidó y educó de mí con mucho amor, nunca me dejó solo y me apoyó en perseguir mis sueños. Así mismo, a mi padre, Leonel, quien me enseñó a luchar por lo que quiero y no rendirme nunca, no importa las adversidades, y ser leal a mis ideales.

Además, me gustaría agradecer a mis hermanas Diana y Fernanda con quienes puedo contar de manera incondicional, no importa la distancia que la vida nos llegue a separar. En particular, a mi hermana mayor, Diana, quien siempre me exigió ser mejor.

Sin duda, el amor – en todas su representaciones – es lo más importante de la vida, puesto que es fuente de inspiración. Es por esto que quiero agradecer a quien se convirtió en mi más grande compañía: Luisa Fernanda. Ella, quien me inspira y motiva cada día, mi suspiro en los momentos más difíciles y mi empuje a atreverme a lo desconocido.

A mis amigos Charleston, Daniel y Fernando por compartir los momentos más increíbles conmigo dentro y fuera del colegio. No siempre teníamos éxito, pero las risas nunca faltaron en *"el templo del saber"*. Además, quiero agradecer a todos mis compañeros de la generación 2016 – 2020, puesto que de ellos aprendí muchísimo, mi felicidad era infinita cada vez que compartíamos el salón de clase.

Quiero agradecer a la Dra. Diana Terrazas por haber sido la primera en creer y confiar en mí, y dirigir mi trabajo de investigación. Sus consejos, experiencia y pericia fueron indispensables para la realización de este trabajo. Además, quiero agradecer al Dr. Saúl Mendoza quien me apoyó en el momento más difícil y orientó para solucionar los problemas computacionales. Así mismo, quiero agradecer al Mtro. Jesús Carrillo, puesto que sin su ayuda, para la recolección de datos, este trabajo no hubiese sido posible.

Finalmente, pero no menos importante, quiero agradecer a El Colegio de México y El Centro de Estudios Económicos por haberme permitido cumplir uno de mis sueños, ser economista. También a la Fundación Colmex y a todos los que forman parte de este gran Colegio.

Abstract

In this work, I develop a computational algorithm to estimate a likelihood function, a model proposed by Spiller and Huang (1986), to analyze the market integration degree, the transaction costs, and the probability of observing binding arbitrage among nine regions. This methodology based on a switching regressions system, which is able to discriminate quite well between regimes of arbitrage and non-arbitrage, is applied to the retail gasoline market in Mexico.

Furthermore, to provide a theoretical framework, this work contains a brief literature review of the theory of competitive markets and arbitrage costs; and the applied literature in other markets. The model is estimated for the Mexican retail gasoline market from November 2017 to June 2019 – after the complete price liberalization – to analyze the degree of market integration subsequent to the Energy Reform of 2013. The evidence suggests that the transaction costs in the Mexican retail gasoline market are below 3.50% of price, increasing with the distance between regions. Furthermore, the probability of being in the same market decrease with distance. However, the Border region is not integrated with the nearest regions, such as North, Northeast, and Northwest.

Cristian E. Gudiño García

Contents

Arbitrage Costs and Market Integration: Evidence from the Mexican Retail Gasoline Market

Cristian E. Gudiño Garcia

El Colegio de México

Centro de Estudios Económicos

1 Introduction

Mexico started a sharp transformation in the energy sector in 2013 when President Enrique Peña Nieto sent an amendment to Congress. This constitutional reform established a new industry structure in the oil market where competition and the entry of new industry players were the key factors of this change. However, its implementation was gradually due to the monopolization of the markets.

Before the reform, the energy sector was controlled by two monopolies: Mexican Petroleum (PEMEX)[1] and the Federal Electricity Commission (CFE)[2]. The first entity controlled the production and distribution of crude oil as well as the refining and imports of hydrocarbons, while the last managed the production and distribution of electricity. Therefore, this structural change introduced competition in the electricity and gasoline markets with the entry of new industry players.

In Mexico, the gasoline price has three components: wholesale price, taxes – including value-added tax (IVA)[3], and a special tax on production and services (IEPS)[4] – and profit margin, where two of those elements can be changed by the Government, the IEPS and the IVA. Hence, prior to the reform coming into force in 2017, the final price of gasoline was

[1] Petróleos Mexicanos (PEMEX)

[2] Comisión Federal de Electricidad (CFE)

[3] Impuesto al Valor Agregado (IVA)

[4] Impuesto Especial sobre Producción y Servicios (IEPS)

fixed by the Federal Administration adjusting the special tax on production and services, avoiding price responses to shocks in supply and demand.

Since the Energy Reform introduced a new market structure in the oil market in Mexico, an important contribution of the reform is the transition to competitive markets in the energy sector. For the gasoline markets, the objective is to have prices which can respond to shocks in demand and supply, and reflect the real market conditions. Therefore, the implementation of the constitutional reform was necessary to schedule the liberalization of prices from a system of government intervention to a free price system.

According to the Federal Economic Competition Commission[5] (COFECE, 2019), the Energy Regulatory Commission (CRE)[6] set a calendar to liberalize prices gradually by regions. This program of gradual liberalization consisted of setting maximum prices with a specific validity. The beginning of this process of maximum prices started on January 1st, 2017, and finished in February 2017[7], where prices were valid for up to two weeks. After that, from March 2017 to October 2017, the entire price liberalization was completed when prices started to respond to shocks in demand and supply – changing daily. The main objective of price deregulation was to promote price competition.

Although this sharp transformation in the Mexican retail gasoline market and the importance of the energetic sector in the current Federal Administration of President Andrés Manuel López Obrador, there has been little research focused on examining the degree of market integration in the Mexican retail gasoline market.

Despite this, two recent works analyze the Mexican gasoline market in two different ways. On the one hand, Contreras et al. (2020) study the relationship between prices and the markups of the gasoline stations – using a spatial competition approach. Their main finding is that retail prices adjust slower to shocks in supply and demand than wholesale prices.

On the other hand, Liu et al. (2018) analyze the effectiveness of the Mexican institutional regulatory framework to enforce competition among service stations. They found that the probability of observing law violations in cities with a large number of gas stations is higher than in cities with a low number of service stations.

There are studies done with price information to test market integration in the United States. In the 1980s, the first studies with some more conclusive results (Slade, 1986;

[5]Comisión Federal de Competencia Económica (COFECE)
[6]Comisión Reguladora de Energía (CRE)
[7]See the Agreement 98/2016 of the Official Journal of the Federation (DOF 27/12/2016)

Spiller and Huang, 1986), found that nearby cities are more integrated than further-away cities. Likewise, Weiner (1993), analyzing patterns of price adjustment, examined the degree of integration and fragmentation of the world crude market – founding a high degree of regionalization in the world oil market.

Later on, other studies have been done to test market integration in natural gas markets (see De Vany and Walls, 1993). These authors concluded that there was an increase in market integration that led to a convergence of gas prices since the deregulation of the market – by using bivariate cointegration tests.

More recently, other works had analyzed the American gasoline markets. On the one hand, Rodney et al. (2001) analyzed the degree of market integration in the retail gasoline markets in the United States after it was completely deregulated – by using the Engel and Granger and Johansen cointegration tests. They found a high degree of market integration in the gasoline markets and a convergence in their prices; that is, their prices differ by up to transaction costs. On the other hand, Houde (2012) studied the relationship between transaction costs and spatial distribution of service stations; finding that arbitrage costs between gasoline stations have low spatial differentiation, and sales are poorly correlated with population density.

As I mentioned above, literature on the economic effects of deregulation and market integration is limited for the Mexican case. Hence, the purpose of this work is to apply a computational methodology to estimate the transaction costs required to arbitrage and the probability of observing binding arbitrage based on the switching regimes model proposed by Spiller and Huang (1986) to the case of the Mexican retail gasoline market.

Hence, the contribution of this work is twofold. On the one hand, I get back a methodology that has not been used before to analyze the degree of market integration in the gasoline markets in Mexico with an arbitrage cost approach. On the other hand, I propose a redefinition of the gasoline market in Mexico, including a ninth region, since the Border region has a different fiscal scheme and a close relationship with the American market.

Consequently, the importance of this work arises in two forms. Firstly, it resides in the methodology because it has not been used before in the Mexican gasoline market. Lastly, it remains in the way in which this methodology can help the regulator to define a market – based on the estimated arbitrage costs.

Since there are no studies that analyze the extent of the market in the Mexican retail gasoline market after the price liberalization reform, my contribution is, using this technique, to analyze the degree of market integration within the eight geographic regions, which are established by the CRE (see PwC (Strategy&), 2019). Moreover, I propose to add a ninth region into the Mexican retail gasoline market – Border – which is described below.

The organization of this work is as follows. The next section, competitive markets, contains some definitions of the economic market, antitrust market, and arbitrage costs. In section 3, the model is described. Section 4 contains an application of the model to the Mexican retail gasoline market. Finally, section 5, conclusions, contains some possible extensions and final comments.

2 Competitive markets

Based on a classical approach, a market is defined as "the entire territory of which the parts are so united by the relations of unrestricted commerce that prices there take the same level throughout, with ease and rapidity" (Cournot, 1960). This definition implies an idea of interdependent price determination between regions. According to Slade (1986), under an ideal scenario, where there are no restrictions to trade, and it is costless; if the product is homogeneous, such as regular gasoline in Mexico, the possibility of binding arbitrage implies uniform prices in the market.

Following the above idea, Spiller and Huang (1986) define an "economic market" as a set of agents whose prices are uniform. Furthermore, some other studies (Horowitz, 1981; Stigler and Sherwin, 1985) had been interested in price information to define economic markets. The main finding of that research is that a high degree of price correlation between two regions implies that any single region is not able to raise prices unless the price in the other region has been raised.

On the other hand, an "antitrust market" is that set of producers such that if they could fully coordinate their actions they would be able to raise price by a given minimum percentage (Spiller and Huang, 1986). Arbitrage costs are introduced in this work to relate "economic markets" with "antitrust markets." These arbitrage costs, also known as transaction costs, do not need to be only shipping costs. They can contain governmental policy costs, such as taxes and quotas.

Arbitrage among regions has an important role in defining "economic markets." Let B be a region that belongs to the economic market centered around region A if, when the price in A exceeds price in B, prices in both regions are linked by binding arbitrage conditions. Under this scenario of binding arbitrage, if producers in region A increase prices, then arbitrage from B takes place. That is to say, the price in the region A must exceed the price in the region B by exactly the transaction costs from B to A – having both regions in the same economic market centered around the region A.

Therefore, arbitrage costs reflect the maximum price differential that can be developed by the prices in the regions A and B. If the difference between the price in A and the price in B was less than the transaction costs, then prices in the two regions could vary independently of each other. However, if the price differential surpasses the arbitrage costs, arbitrage takes place.

According to Spiller and Huang (1986), in the market definition literature there is a wrong presumption that "economic markets" contain "antitrust markets."[8] Hence, whether a low price region is in the "antitrust" market centered in a high price region depends on two things. On the one hand, it depends on if it belongs to the respective "economic" market. On the other hand, if it does not, it depends on whether the arbitrage costs are less than the threshold price used to define an antitrust market plus the average price difference. Likewise, for a high price region to be part of the antitrust market centered in the low price region, arbitrage costs from the high to the low price region have to be less than the threshold minus the average price difference (see Spiller and Huang, 1986).

Although it is not possible to infer exactly whether two regions are in the same economic market, it is possible to estimate the probability that the two regions would be in the same economic market because of shocks in demand and supply shown by the behavior of prices. Hence, in the next section, I will describe the model used, in this work, to analyze the degree of market integration in the Mexican retail gasoline market. Furthermore, this model will allow estimating the transaction costs needed to observe binding arbitrage among regions.

[8] *Cf.* Spiller and Huang (1986). Consider two producer areas that, although they may not belong to the same "economic" market, they could if one area would increase its price by a relatively large amount. Then, the fact that the two regions were not in the same economic market does not allow me to infer both would be in the same market if one of the regions increases prices by a small percentage.

3 The model

Based on the model proposed by Spiller and Huang (1986), consider two regions i and j. Under a scenario of legal trading barriers, the autarky[9] equilibrium prices (P_t^{iA}, P_t^{jA}) at time t would be determined by shocks in supply and demand in each region. Without loss of generality, the reduced form equations of autarky prices are assumed to have constant means, $\pi^{i,j}$, and an error term, $\epsilon_t^{i,j}$, which represents shocks to the markets as follows:

$$P_t^{iA} = \pi^i + \epsilon_t^i \tag{1}$$

$$P_t^{jA} = \pi^j + \epsilon_t^j \tag{2}$$

As would be expected, under a scenario of legal trading barriers, there will not be binding arbitrage. However, in the absence of legal trading barriers, there will be two possible scenarios – no arbitrage opportunities and arbitrage. Let T_t be the transaction costs in each period t, and let (P_t^i, P_t^j) be the observed prices. Then, if the difference between autarky prices is lower than the transaction costs, no arbitrage opportunities emerge, and the observed prices[10] equal the autarky prices.

That is to say, assume that $P_t^{jA} > P_t^{iA}$. Then if,

$$0 < P_t^{jA} - P_t^{iA} < T_t \tag{3}$$

from equation (3) we have that $P_t^{iA} = P_t^i$ and $P_t^{jA} = P_t^j$. This implies that

$$0 < P_t^j - P_t^i < T_t \tag{4}$$

Likewise, if the difference between the autarky prices is higher than the transaction costs, arbitrage opportunities arise differing the observed prices by the transaction costs. In this case, contrary to the above scenario, both equilibrium prices are interrelated – a shock in one region is transmitted into the other. Thus, if:

$$0 < T_t < P_t^{jA} - P_t^{iA} \tag{5}$$

[9] Autarky prices are those prices that prevail if the trade is not feasible, either because of a binding legal rule or high transaction costs.

[10] (P_t^i, P_t^j) are independent; that is, an increase in the autarky price in one region should not affect the equilibrium price in the other.

then, we have that

$$0 < P_t^j - P_t^i = T_t \tag{6}$$

Now, let transaction costs T_t be a random variable with mean T, that is

$$T_t = Te^{V_t}$$

where V_t is normally distributed with zero mean and constant variance σ_v^2. Then, the probability of no arbitrage opportunities – the scenario where prices are described by the equation (4) – is a constant λ.

$$\begin{aligned}\mathbb{P}\{0 < P_t^j - P_t^i < T_t\} &= \mathbb{P}\{0 < P_t^{jA} - P_t^{iA} < Te^{V_t}\} \\ &= \mathbb{P}\{\ln[(\pi^j - \pi^i) + (\epsilon_t^j - \epsilon_t^i)] - V_t < \ln(T)\} \\ &= \lambda\end{aligned} \tag{7}$$

From the above equation, the constant λ is a function of $\pi^{i,j}$, T, the constant variance of V_t (σ_v^2), and the distribution parameters of $\epsilon_t^{i,j}$ in equations (1) and (2). On the other hand, the probability of arbitrage opportunities – the scenario where prices are modeled by the equation (6) – is $(1 - \lambda)$. Furthermore, this probability can be interpreted as a measure of how integrated the regions i and j are. That is to say, if $(1 - \lambda)$ is very close to one, then the two areas are almost always integrated. However, if $(1 - \lambda)$ is very close to zero, then both areas are almost never integrated.

Since there are two possible price regimes – arbitrage and no arbitrage – then the price equations (3) and (5) are a switching regressions system. The model is defined as follows:

$$\ln(P_t^j - P_t^i) = B + V_t - U_t \tag{8}$$

$$\ln(P_t^j - P_t^i) = B + V_t \tag{9}$$

where U_t is a positive random variable and $B = \ln(T)$. Furthermore, equation (8) has a probability λ, whereas equation (9) has a probability $(1 - \lambda)$. Hence, the first equation, (8), corresponds to the regime of no arbitrage opportunities or autarky state, and the second equation, (9), represents the arbitrage state.

Note that equation (8) is a composite error regression and corresponds to the standard stochastic frontier equation[11] described by Aigner *et al.* (1977). While λ represents the probability of being in the no arbitrage opportunities regime or autarky state, the positive random variable U_t corresponds to a measure of propensity to trade. Then, considering the equation (8), the higher is the value of U_t, the smaller is the propensity to trade.

To estimate the parameter vector denoted by $\theta = (B, \sigma_u^2, \sigma_v^2, \lambda)$, which represents the solution for the switching regressions system modeled by equations (8) and (9), it is necessary define a likelihood function. This function for n observations is:

$$L = \prod_{t=1}^{n} \left[\lambda f_t^i + (1-\lambda) f_t^j \right] \tag{10}$$

where f_t^i is the density function of equation (8), while f_t^j is the density function of (9). These density functions are:

$$f_t^i = \left(\frac{2}{\sqrt{(\sigma_u^2 + \sigma_v^2)}} \right) \phi \left(\frac{Y_t - B}{\sqrt{(\sigma_u^2 + \sigma_v^2)}} \right) \left[1 - \Phi \left(\frac{(Y_t - B)\frac{\sigma_u}{\sigma_v}}{\sqrt{(\sigma_u^2 + \sigma_v^2)}} \right) \right] \tag{11}$$

$$f_t^j = \frac{1}{\sigma_v} \phi \left(\frac{Y_t - B}{\sigma_v} \right) \tag{12}$$

where ϕ is the standard normal density function, Φ is the standard distribution function, and $Y_t = \ln(P_t^j - P_t^i)$. Although the maximum likelihood estimates of θ can be obtained by maximizing the likelihood function (10), it is possible obtain the solution $\hat{\theta}$ maximizing the logarithmic function of the likelihood function (see Kiefer, 1980).

It can be seen from the logarithmic function of equation (10) that the solution of the parameter λ is obtained by:

$$\frac{\partial \ln L}{\partial \lambda} = \sum_{t=1}^{n} \frac{f_t^i - f_t^j}{\lambda f_t^i + (1-\lambda) f_t^j} = 0$$

where $\frac{\partial \ln L}{\partial \lambda}$ is the only partial derivative that has a closed solution. Having this estimation problem to find the parameter vector $\hat{\theta}$, I developed the computational algorithm to estimate the parameter vector $\hat{\theta}$ – by maximizing the likelihood function (10).

[11] Aigner et al. (1977) define the standard stochastic frontier function as $\ln(Y) = \beta_0 + \beta_1 B + (v_t - u_t)$ where the last term corresponds to the composite error. Furthermore, U_t is assumed to be distributed independently of V_t with a one-sided half normal distribution, i.e $U_t \sim \mathbf{N}(0, \sigma_u^2)$.

Since a lack of statistical tools to estimate this model, I used MATLAB, a numeric computing environment, to develop the computational algorithm that allowed me to estimate the vector parameter $\hat{\theta}$ by maximizing the likelihood function (10). This computational algorithm was necessary to define the density functions, f_t^i and f_t^j, of equations (8) and (9) respectively, as well as the likelihood function (10).

Although this methodology has some disadvantages, such as the complexity of being estimated and the data used, which do not consider some externalities, for example the crude oil price, distance between stations, or indicators of quality, this methodology is able to discriminate quite well between regimes. That is, this model allows me to discriminate between arbitrage binding conditions regime and no-arbitrage conditions regime – by estimating the probability of being in one regime or another. Hence, this key advantage allows the model to be versatile and useful for other sectors and markets. Note that this model, being part of the dynamic models described by Hosken et al. (2008), is able to infer some conclusions from the price behavior.

4 An application to the Mexican retail gasoline market

In this section, I apply the methodology developed by Spiller and Huang (1986) and described in the previous section to analyze the degree of integration of the gasoline retail markets by regions in Mexico. The Energy Regulatory Commission (CRE) has divided the Mexican retail gasoline market into eight regions – Northwest, North, Northeast, West, Gulf, Center, Southeast, and South (see PwC (Strategy&), 2019). However, since in the northern border there is a different fiscal scheme for those municipalities that are near to the United States[12], I propose to create a ninth region – Border – for academic purposes of this work.

4.1 The data

From a transparency request, I obtained daily prices at service stations for all the municipalities of Mexico. However, since this work analyzes the degree of market integration in the Mexican retail gasoline market by region, I used weekly prices for nine regions - North, Northeast, Northwest, West, Gulf, Center, Southeast, South, and Border.

[12] See the Agreement 111/2018 of the Official Journal of the Federation (DOF 31/12/2018)

For each region, I chose the geographic division[13] used by the CRE to define the nine regions. Note that the Border[14] region is not part of the official Mexican market; it is a proposed region to analyze the market because it has a different fiscal scheme because it is near to the United States.

For each week, I used as the relevant price the lowest quotation for regular gasoline[15]. According to Spiller and Huang (1986), this price would be chosen for three reasons. Firstly, it seems to be that the lowest price is the most affected by the degree of competition in a region (Marvel (1978), cited by Spiller and Huang (1986)). Second, the lowest price is used to avoid the issue between the transaction price and retail price. Finally, the lowest price is set by those stations service who rarely discount[16].

I obtained weekly prices from January 1, 2017, to June 11, 2019. However, to estimate the model described in section 3, I used weekly prices from November 1, 2017, because the price liberalization was consolidated until October 31, 2017, to June 11, 2019. The red line in the four panels of figures 1 and 2 shows the beginning of a free price system in the Mexican retail gasoline market in the entire territory of Mexico. Although the data set includes 127 price quotations for each region, I used 87 price quotations for each region – from week 41, 2017, to week 24, 2019 – having, in total, 783 observations.

Figure 1 presents the gasoline prices in the Northern regions from January 2017 to June 2019; panel (a) shows prices for regular gasoline; panel (b) exposes prices for premium gasoline. In both panels, we can see how before price deregulation, the prices had a stable dynamic. That is, they do not vary significantly. However, after the price liberalization, the prices in the eight regions started to change more quickly – showing more sensitiveness to shocks in demand and supply as well as the market conditions.

[13] PwC (2019) describes how the CRE divides the Mexican retail gasoline market into eight regions and what entities are part of each of them. North contains Chihuahua and Durango. In the Northwest are Baja California, Baja California Sur, Sinaloa, Nayarit, and Sonora, while in the Northeast are Tamaulipas, Nuevo León, San Luis Potosí, and Coahuila. West contains Michoacán, Guanajuato, Colima, Jalisco, Aguascalientes, and Zacatecas. Center contains Querétaro, Hidalgo, Estado de México, Ciudad de México, Morelos, Tlaxcala, and Puebla. In the Gulf are Veracruz and Tabasco. Finally, Southeast contains Campeche, Quintana Roo, and Yucatán, whereas the South contains Guerrero, Oaxaca, and Chiapas.

[14] This region contains 386 service stations in the municipalities of Ensenada, Mexicali, Playas de Rosarito, Tecate, Tijuana, Agua Prieta, Altar, Caborca, Cananea, General Plutarco Elías Calles, Nogales, Puerto Peñasco, San Luis Río Colorado, Guadalupe y Calvo, Juárez, Ojinaga, Práxedis G. Guerrero, Acuña, Ocampo, Piedras Negras, Anáhuac, Matamoros, Nuevo Laredo, Reynosa, Río Bravo, and Valle Hermoso.

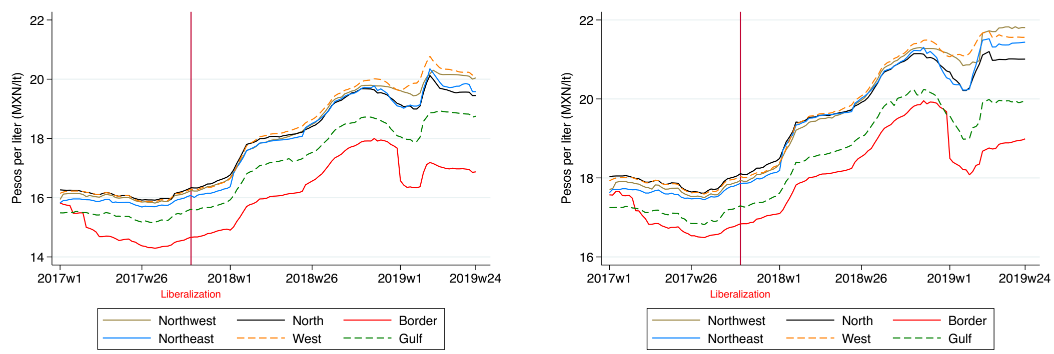
[15] For simplicity, I reported just estimations for regular gasoline. The estimations for premium gasoline do not differ significantly from regular gasoline estimations. Analysis and interpretation are the same. Therefore, to see premium gasoline estimations, see the appendix.

[16] Assuming gasoline differs in quality within a region, the selection of the lowest price implies that I chose those gas stations with the lowest quality – having an homogeneous good.

Figure 1: GASOLINE PRICES BY NORTHERN REGIONS

(a) Regular Gasoline Prices

(b) Premium Gasoline Prices

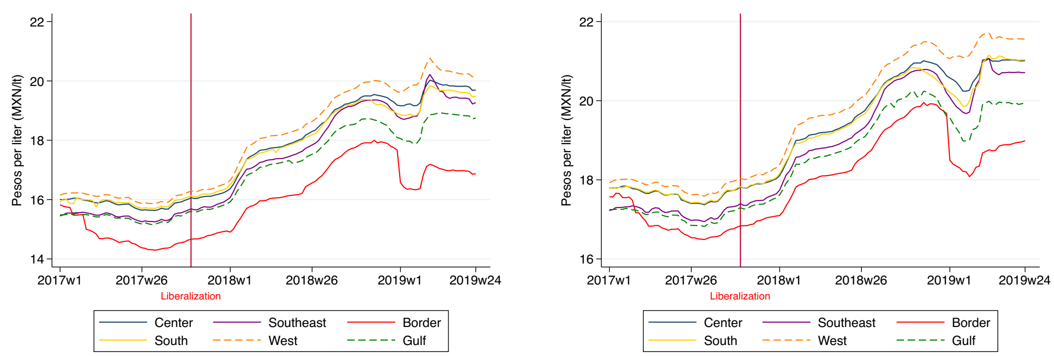


Notes: Both panels present gasoline prices for Northern regions from January 2017 to June 2019 in a weakly period. Panel (a) shows regular gasoline prices in the North of Mexico. Panel (b) shows premium gasoline prices in the North of Mexico.

Figure 2: GASOLINE PRICES BY SOUTHERN REGIONS

(a) Regular Gasoline Prices

(b) Premium Gasoline Prices



Notes: Both panels present gasoline prices for Southern regions from January 2017 to June 2019 in a weakly period. Panel (a) shows regular gasoline prices in the South of Mexico. Panel (b) shows premium gasoline prices in the South of Mexico.

Similarly, figure 2 shows the gasoline prices in the Southern regions through the same period; panel (a) exhibits prices for regular gasoline, while panel (b) presents prices for premium gasoline. As we can see in both figures, the Gulf region is the lower bound, while the West region is the upper bound – containing the rest of the regions. Furthermore, the price dynamics before and after price deregulation are the same as in figure 1. At the beginning of 2019, the gasoline prices went down abruptly, probably, because of a decrease in the crude oil prices and the policy implemented by President Andrés Manuel López Obrador against fuel theft in the Center and West regions.

Table 1 presents summary statistics of regular gasoline prices. Excluding the Border region, from January 2017 through June 2019, the Gulf region presents the lowest gasoline price, whereas the West region shows the highest retail price. In 2017, the mean price for the Gulf was 15.46 pesos per liter, and the mean for the West was 16.15 pesos per liter. In 2018, there was a significant increase in mean prices: the mean price in the Gulf equals 17.68 pesos per liter, while in the West, the mean price reaches 18.79 pesos per liter. Thus there was an increase of 14.35% and 16.34%, respectively.

This dynamic price was still increased until 2019 when the mean of price gasoline in the Gulf was 18.56 pesos per liter, while the mean in the West was 20.18 pesos per liter – having the largest increase from the beginning of price liberalization. Since the Gulf region has the lowest price quotation during the whole period of analysis, it should be expected that its price is limiting the price of the other regions – by bounding low. That is, under trading between the Gulf and Center, the price in Center cannot exceed the Gulf price by more than the arbitrage costs. Furthermore, the Center price will be bound by the prices in surrounding cities. As a consequence, arbitrage can occur between any pair of regions. According to Spiller and Huang (1986), it is expected to have a negative relationship between distance and arbitrage cost; that is, the nearest two regions are, the higher the arbitrage cost they have.

From figures 1 and 2, although the price dynamic in the Border region is completely similar to the other regions, it can be seen that there is a significant price gap between the Gulf and the Border. This price differential between both regions can be explained by the proximity between the American market and the Mexican Border market; that is, the Border region could be integrated into the South American gasoline market. However, a full study of market integration in the borderline between Mexico and the United States is beyond the scope of this work.

Table 1: SUMMARY STATISTICS FOR REGULAR GASOLINE

	Mean $MXN	**Median** $MXN	**SD**[1] $MXN	**Min** $MXN	**Max** $MXN
2017					
Border[2]	14.77	14.67	0.42	14.29	15.82
Northwest	16.09	16.07	0.19	15.82	16.58
North	16.19	16.17	0.20	15.92	16.69
Northeast	15.92	15.91	0.16	15.69	16.31
West	16.15	16.15	0.18	15.86	16.56
Gulf	15.46	15.47	0.18	15.16	15.85
Center	15.92	15.93	0.17	15.63	16.29
Southeast	15.52	15.52	0.18	15.25	15.99
South	15.97	15.97	0.18	15.70	16.39
2018					
Border	16.74	16.59	0.93	14.90	18.00
Northwest	18.64	18.51	0.93	16.67	19.80
North	18.61	18.43	0.81	16.77	19.68
Northeast	18.56	18.53	0.90	16.37	19.78
West	18.79	18.67	0.94	16.65	20.01
Gulf	17.68	17.55	0.78	15.93	18.72
Center	18.39	18.32	0.90	16.37	19.54
Southeast	18.09	17.87	0.98	16.07	19.36
South	18.30	18.21	0.83	16.52	19.38
2019					
Border	16.81	16.96	0.30	16.34	17.19
Northwest	19.94	20.10	0.29	19.44	20.30
North	19.47	19.56	0.32	18.99	20.12
Northeast	19.60	19.73	0.39	19.02	20.35
West	20.18	20.24	0.30	19.61	20.77
Gulf	18.56	18.82	0.41	17.88	18.93
Center	19.64	19.82	0.31	19.16	20.03
Southeast	19.33	19.42	0.44	18.70	20.22
South	19.41	19.60	0.37	18.82	19.84

Source: Author's own elaboration. This table shows the summary statistics for regular gasoline prices per year, from 2017 to 2019, by region – the first eight regions are fixed by the CRE, while the last region is the proposed region by the author. *1.* Deviation Standard. *2.* Proposed region by the author, which contains the municipalities of Ensenada, Mexicali, Playas de Rosarito, Tecate, Tijuana, Agua Prieta, Altar, Caborca, Cananea, General Plutarco Elías Calles, Nogales, Puerto Peñasco, San Luis Río Colorado, Guadalupe y Calvo, Juárez, Ojinaga, Práxedis G. Guerrero, Acuña, Ocampo, Piedras Negras, Anáhuac, Matamoros, Nuevo Laredo, Reynosa, Río Bravo, and Valle Hermoso.

Table 2 shows the correlation matrix for prices of the nine regions, both in levels and first differences. In their levels, all prices are highly correlated; certainly, they have the same dynamic price from 2017 to 2019. However, in the first differences, price correlations decreased moderately. The lowest correlations are between the Border and the rest of the regions. Nevertheless, price correlations are not a proper indicator to infer whether two regions are in the same market (see Spiller and Huang, 1986). Hence, in the next section, I will use the methodology described in section 3, the model, to estimate the probability of seeing binding arbitrage among regions.

Table 2: CORRELATION MATRIX - LEVEL/FIRST DIFFERENCE

	Northwest	***North***	***Northeast***	***West***	***Gulf***	***Center***	***Southeast***	***South***
Northwest	-							
North	0.996/0.860	-						
Northeast	0.997/0.783	0.999/0.876	-					
West	0.999/0.833	0.996/0.925	0.997/0.863	-				
Gulf	0.996/0.866	0.997/0.891	0.998/0.823	0.995/0.865	-			
Center	0.999/0.891	0.996/0.929	0.997/0.876	0.999/0.947	0.997/0.933	-		
Southeast	0.997/0.815	0.997/0.916	0.998/0.852	0.998/0.909	0.996/0.879	0.998/0.901	-	
South	0.998/0.807	0.997/0.816	0.998/0.739	0.998/0.800	0.998/0.859	0.999/0.828	0.998/0.796	-
Border	0.933/0.584	0.947/0.573	0.941/0.511	0.930/0.541	0.944/0.545	0.936/0.560	0.938/0.569	0.934/0.479

Source: Own construction with data from CRE. This table shows the price correlation coefficient in levels and first differences by regions for regular gasoline. The region Border contains municipalities of the states of Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas, which differs from the other regions by its different fiscal scheme.

4.2 THE ESTIMATION

I estimated for each region pair the model described in section three by equations (10) – (12). Note that T reflects the real transaction costs. Then, to estimate the model, I did not deflate the prices because during the analysis period price were relatively stable and they do not change abruptly. Furthermore, to have a defined estimation it is necessary to satisfy the condition $P_t^j - P_t^i > 0$ – contrary Y_t is not defined. Hence, the estimation should be made in two directions. That is, the sample is divided in two for each region pair. On the one hand, the set of all observations for which the price in one of the regions is the highest. On the other hand, the set that comprises those observations for which the price of the other region is the highest[17].

This division is made to represent all the possibilities where the binding arbitrage would take place from the region with the lower price to the region with the higher price because according to Spiller and Huang (1986), transaction costs may differ from one direction to other. Estimating the model it is possible to estimate the probability of binding arbitrage between two regions, and their transaction costs. The estimation is performed by maximizing the likelihood function (9)[18]. The results of this maximization are shown in Table 3.

For each region pair, Table 3 presents the parameter vector $\hat{\theta}$. That is the maximum likelihood estimates which contain the logarithm of the transaction costs (B), the variances of U and V, σ_u^2 and σ_v^2 respectively, the probability of not being bound by binding arbitrage conditions (λ), the number of observations across region-pairs in which the price in one region exceeds the price in the other (N), and the transaction costs (TC) involved in each region-pairs.

Table 3 provides the most relevant results of this work.[19] I made this table with the clearest region-pairs examples. With these examples, I will explain why is the Border region a feasible option to redefine the Mexican retail gasoline market. Furthermore, the results are in line as we would expect. Lastly, the information in Table 3 allows me to explain how the interconnection and distribution of gasoline among regions could explain the extent of integration.

[17] *Cf.* Spiller and Huang (1986). The model is estimated only for those cases in which the price in one region exceeds the price in the other region in at least 50 weeks to avoid computational mistakes. That is, the estimation is not always made in two directions.

[18] It is possible to made the estimation by maximizing the log likelihood function, and the solution will be the same as I mentioned in section three, the model.

[19] Without loss of generality this table contains only estimates for nine pair-regions. These nine pair-regions estimates are the clearest results to explain the main findings. The rest estimations are in the Appendix.

Table 3: PARAMETER ESTIMATES FOR REGULAR GASOLINE PRICES

	Northeast Border	***North Border***	***Northwest Border***	***West South***	***Northwest South***	***West Center***	***North Center***	***Northeast Gulf***	***Southeast Gulf***
B	0.67	0.99	1.15	-0.52	-0.67	-0.81	-0.87	-0.05	-0.35
σ_u^2	0.001	0.36	0.52	0.89	0.59	0.28	1.64	0.46	1.30
σ_v^2	0.24	0.04	0.01	0.28	0.31	0.21	0.03	0.16	0.26
λ	0.72	0.85	0.80	0.33	0.91	0.41	0.87	0.37	0.65
N	87	87	87	87	87	87	67	87	87
TC	1.95	2.69	3.15	0.59	0.51	0.45	0.42	0.95	0.71

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

According to Spiller and Huang (1986), region-pairs with high probabilities of being in the same market (low λ) seem to be those that are nearby. Such is the case of West-Center ($1-\lambda = 0.53$), Northeast-Gulf ($1-\lambda = 0.63$), and West-South($1-\lambda = 0.67$). However, the last region-pair is special – this case will be explained below. On the other hand, distant region-pairs seem to be those that have low probabilities of being in the same market (high λ): Northwest-South ($1-\lambda = 0.09$), and North-Center ($1-\lambda = 0.13$).

Considering this relationship between the distance among regions and the probability of being in the same market, it is expected to have a market integration between the Border region and the north of Mexico because they are nearby. However, the average probability of being in the same market for those regions in the Northern is low and equal to 0.21. Note that North-Border has a λ equal to 0.85 ($1-\lambda = 0.15$), while the probability of being in the same market of Northwest-Border is equal to $1-\lambda = 0.20$, and the Northeast-Border has the highest probability of being integrated ($1-\lambda = 0.28$). Hence, since the Border region is not integrated into the Mexican retail gasoline market or any of the surrounding regions – North, Northeast, and Northwest – I can infer that the Border region could be integrated into the American retail gasoline market. However, the test of that hypothesis is beyond the scope of this work. Furthermore, it makes sense to consider nine regions in the Mexican retail gasoline market instead of eight regions.

Although the regions of West and South are distant, the high probability of being in the same market is explained by the gasoline distribution method. The gasoline in the region pair of West-South is transported by cargoes – from the port of Salina Cruz to the ports of Acapulco, Lázaro Cárdenas, and Manzanillo. Since this is the most important way of transportation between both regions, these markets are interrelated (PwC (Strategy&), 2019).[20]

Finally, the evidence indicates, as Spiller and Huang (1986), that region-pairs with low transaction costs have relatively high probabilities of being in the same market. That is, those region-pairs with low λ values, such as West-Center (TC = 0.45), and West-South (TC = 0.59). However, the average transaction costs in the region-pairs of North-Border, Northeast-Border, and Northwest-Border is equal to 2.60% of price. Furthermore, binding arbitrage conditions usually cause the price of the Gulf region to bound the price of the other regions. Hence, the Gulf region is a relevant player in the retail gasoline market.

[20] *Cf.* Spiller and Huang (1986). Binding arbitrage among regions is more expensive when it is indirect rather than direct.

5 Conclusions

In this work, I used the methodology proposed by Spiller and Huang (1986) to estimate the transaction costs among nine regions and the probability of being in the same market. This methodology consists of a switching regressions system, which models two different regimes. In one regime the prices between two regions differ by the transaction (arbitrage) costs. In the other regime, these prices differ by less than the arbitrage costs.

I applied this methodology to the Mexican retail gasoline market in nine regions. Examining weekly pricing for 87 weeks, from November 2017 to June 2019, my main findings are as follows. First, the retail prices of gasoline in the Mexican market raised an average of 10% from the beginning of price liberalization per year. It is explained by a higher sensitivity of retail gasoline prices to shocks in supply and demand – showing the real market conditions. Second, the Mexican gasoline market is well-defined since in almost all regions the probability of being in the same market is low. However, the Border region could be considered a ninth region in the Mexican market because it is not integrated into the nearby regions – North, Northeast, and Northwest. Third, arbitrage costs are relatively low – these costs are below 3.5% of the price. Hence, those regions with smaller transaction costs are more integrated than those regions with higher arbitrage costs; nearby regions are more integrated than distant regions.

As I mentioned, this methodology quite well discriminates between regimes – arbitrage and non-arbitrage. Thus, it is useful for other markets and sectors. This versatility allows the regulator to define or redefine any market, as well as to know the relevant economic agent in that market based on the estimation of transaction costs. As a consequence, it is possible to infer when an economic market is contained in an antitrust market.

Finally, given the importance of the energy sector for the current federal government and the research focused on the Mexican retail gasoline market, I consider the retail gasoline market as a potential area for future research. This study could be complemented by analyzing the degree of market integration within a region based on the gasoline brands. This will allow us to know the price behavior of the gasoline retailers and their price strategies to compete or coordinate (Hosken et al., 2008). Some possible extensions of this work are: to analyze the market integration degree between the Border region and the American market; in the same way, to estimate the effect of Andrés Manuel López Obrador administration on the market integration degree in the Mexican retail gasoline market.

6 Appendix

6.1 Estimations for Regular Gasoline

Table 4: SUMMARY STATISTICS FOR REGULAR GASOLINE

	Mean $MXN	**Median** $MXN	**SD**[1] $MXN	**Min** $MXN	**Max** $MXN
2017					
Border[2]	14.77	14.67	0.42	14.29	15.82
Northwest	16.09	16.07	0.19	15.82	16.58
North	16.19	16.17	0.20	15.92	16.69
Northeast	15.92	15.91	0.16	15.69	16.31
West	16.15	16.15	0.18	15.86	16.56
Gulf	15.46	15.47	0.18	15.16	15.85
Center	15.92	15.93	0.17	15.63	16.29
Southeast	15.52	15.52	0.18	15.25	15.99
South	15.97	15.97	0.18	15.70	16.39
2018					
Border	16.74	16.59	0.93	14.90	18.00
Northwest	18.64	18.51	0.93	16.67	19.80
North	18.61	18.43	0.81	16.77	19.68
Northeast	18.56	18.53	0.90	16.37	19.78
West	18.79	18.67	0.94	16.65	20.01
Gulf	17.68	17.55	0.78	15.93	18.72
Center	18.39	18.32	0.90	16.37	19.54
Southeast	18.09	17.87	0.98	16.07	19.36
South	18.30	18.21	0.83	16.52	19.38
2019					
Border	16.81	16.96	0.30	16.34	17.19
Northwest	19.94	20.10	0.29	19.44	20.30
North	19.47	19.56	0.32	18.99	20.12
Northeast	19.60	19.73	0.39	19.02	20.35
West	20.18	20.24	0.30	19.61	20.77
Gulf	18.56	18.82	0.41	17.88	18.93
Center	19.64	19.82	0.31	19.16	20.03
Southeast	19.33	19.42	0.44	18.70	20.22
South	19.41	19.60	0.37	18.82	19.84

Source: Author's own elaboration. This table shows the summary statistics for regular gasoline prices per year, from 2017 to 2019, by region – the first eight regions are fixed by the CRE, while the last region is the proposed region by the author. *1.* Deviation Standard. *2.* Proposed region by the author, which contains the municipalities of Ensenada, Mexicali, Playas de Rosarito, Tecate, Tijuana, Agua Prieta, Altar, Caborca, Cananea, General Plutarco Elías Calles, Nogales, Puerto Peñasco, San Luis Río Colorado, Guadalupe y Calvo, Juárez, Ojinaga, Práxedis G. Guerrero, Acuña, Ocampo, Piedras Negras, Anáhuac, Matamoros, Nuevo Laredo, Reynosa, Río Bravo, and Valle Hermoso.

Table 5: CORRELATION MATRIX - LEVEL/FIRST DIFFERENCE

	Northwest	***North***	***Northeast***	***West***	***Gulf***	***Center***	***Southeast***	***South***
Northwest	-							
North	0.996/0.860	-						
Northeast	0.997/0.783	0.999/0.876	-					
West	0.999/0.833	0.996/0.925	0.997/0.863	-				
Gulf	0.996/0.866	0.997/0.891	0.998/0.823	0.995/0.865	-			
Center	0.999/0.891	0.996/0.929	0.997/0.876	0.999/0.947	0.997/0.933	-		
Southeast	0.997/0.815	0.997/0.916	0.998/0.852	0.998/0.909	0.996/0.879	0.998/0.901	-	
South	0.998/0.807	0.997/0.816	0.998/0.739	0.998/0.800	0.998/0.859	0.999/0.828	0.998/0.796	-
Border	0.933/0.584	0.947/0.573	0.941/0.511	0.930/0.541	0.944/0.545	0.936/0.560	0.938/0.569	0.934/0.479

Source: Own construction with data from CRE. This table shows the price correlation coefficient in levels and first differences by regions for regular gasoline. The region Border contains municipalities of the states of Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas, which differs from the other regions by its different fiscal scheme.

Table 6: PARAMETER ESTIMATES FOR REGULAR GASOLINE PRICES

	Northwest Gulf	***North Gulf***	***Northeast Gulf***	***West Gulf***	***Center Gulf***	***Southeast Gulf***	***South Gulf***	***West North***	***Northwest North***
B	-0.004	-0.12	-0.05	0.34	0.004	-0.35	-0.44	-0.33	-0.65
σ_u^2	0.003	0.002	0.46	0.32	0.45	1.30	0.002	1.56	1.80
σ_v^2	0.27	0.14	0.16	0.22	0.18	0.26	0.21	0.10	0.11
λ	0.71	0.71	0.37	0.83	0.87	0.65	0.72	0.76	0.72
N	87	87	87	87	87	87	87	69	55
TC	0.99	0.87	0.95	1.40	1.00	0.71	0.64	0.72	0.52

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Table 7: PARAMETER ESTIMATES FOR REGULAR GASOLINE PRICES

	Gulf Border	***South Border***	***Southeast Border***	***Center Border***	***West Border***	***Northeast Border***	***North Border***	***Northwest Border***
$\boldsymbol{B}$	0.63	0.56	0.42	0.61	0.82	0.67	0.99	1.15
σ_u^2	0.68	0.002	0.001	0.001	0.001	0.001	0.36	0.52
σ_v^2	0.01	0.25	0.32	0.27	0.26	0.24	0.04	0.01
λ	0.88	0.71	0.71	0.70	0.71	0.72	0.85	0.80
$\boldsymbol{N}$	87	87	87	87	87	87	87	87
$\boldsymbol{TC}$	1.87	1.75	1.52	1.84	2.27	1.95	2.69	3.15

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by $\boldsymbol{TC}$. $\boldsymbol{N}$ represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Table 8: PARAMETER ESTIMATES FOR REGULAR GASOLINE PRICES

	Northwest Southeast	***North Southeast***	***Northeast Southeast***	***West Southeast***	***South Southeast***	***Center Southeast***	***West Northeast***	***Northwest Northeast***
B	-0.24	-0.35	-0.50	-0.34	-0.98	-0.90	-0.69	-0.81
σ_u^2	0.52	1.11	0.60	0.01	1.50	0.73	0.78	1.88
σ_v^2	0.10	0.06	0.70	0.20	0.13	0.08	0.41	0.14
λ	1	0.67	0.93	0.70	0.53	0.53	1	0.88
N	86	83	86	87	75	84	87	78
TC	0.77	0.70	0.60	.071	0.38	0.41	0.50	0.45

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

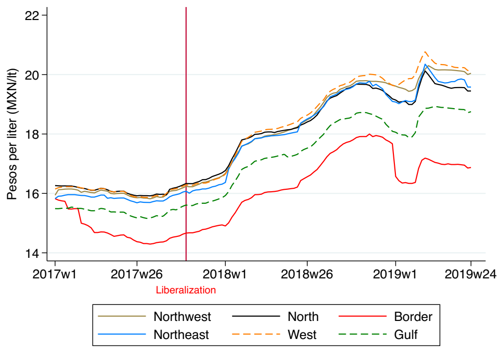
Table 9: PARAMETER ESTIMATES FOR REGULAR GASOLINE PRICES

	Center South	***West South***	***North South***	***Northeast South***	***Northwest South***	***West Center***	***North Center***	***Northeast Center***	***Northwest Center***
B	-1.11	-0.52	-0.81	-0.94	-0.67	-0.81	-0.87	-1.56	-1.14
σ_u^2	1.61	0.89	0.72	0.77	0.59	0.28	1.64	1.98	0.35
σ_v^2	0.05	0.28	0.07	0.21	0.31	0.21	0.03	0.16	0.14
λ	0.87	0.33	1	1	0.91	0.41	0.87	0.46	0.72
N	70	87	76	74	87	87	67	68	87
TC	0.33	0.59	0.45	0.40	0.51	0.45	0.42	0.21	0.32

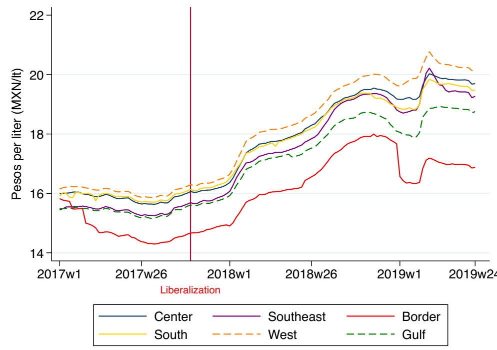
Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Figure 3: REGULAR GASOLINE PRICES BY REGIONS

(a) Northern Regions



(b) Southern Regions

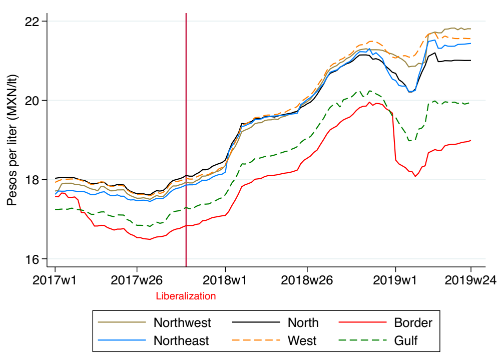


Notes: Both figures present regular gasoline prices by regions from January 2017 to June 2019 in a weakly period. Panel (a) shows regular gasoline prices in the North of Mexico. Panel (b) shows regular gasoline prices in the South of Mexico.

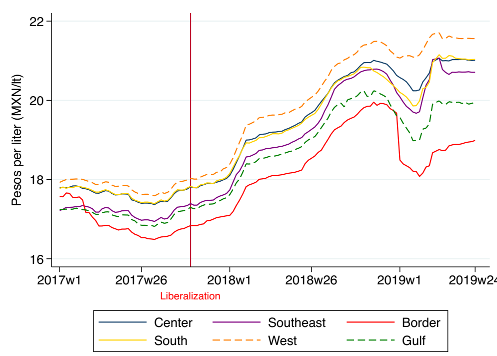
6.2 Estimations for Premium Gasoline

Figure 4: PREMIUM GASOLINE PRICES BY REGIONS

(a) Northern Regions



(b) Southern Regions



Notes: Both figures present premium gasoline prices by regions from January 2017 to June 2019 in a weakly period. Panel (a) shows premium gasoline prices in the North of Mexico. Panel (b) shows premium gasoline prices in the South of Mexico.

Table 10: SUMMARY STATISTICS FOR PREMIUM GASOLINE

	Mean $MXN	**Median** $MXN	**SD**[1] $MXN	**Min** $MXN	**Max** $MXN
2017					
Border[2]	16.91	16.83	0.34	16.49	17.66
Northwest	17.81	17.81	0.19	17.50	18.27
North	17.96	17.95	0.21	17.61	18.43
Northeast	17.70	17.68	0.18	17.45	18.14
West	17.91	17.94	0.18	17.59	18.30
Gulf	17.16	17.19	0.19	16.82	17.54
Center	17.69	17.72	0.18	17.37	18.03
Southeast	17.25	17.27	0.18	16.94	17.65
South	17.71	17.73	0.17	17.39	18.06
2018					
Border	18.73	18.57	0.86	17.10	19.95
Northwest	20.17	20.00	0.89	18.36	21.30
North	20.12	19.93	0.74	18.49	21.15
Northeast	20.13	20.05	0.79	18.19	21.31
West	20.27	20.10	0.88	18.39	21.49
Gulf	19.19	19.07	0.77	17.62	20.24
Center	19.86	19.72	0.85	18.11	21.01
Southeast	19.54	19.30	0.93	17.74	20.79
South	19.77	19.65	0.79	18.16	20.84
2019					
Border	18.61	18.73	0.31	18.08	18.99
Northwest	21.42	21.70	0.42	20.84	21.83
North	20.79	20.98	0.31	20.21	21.20
Northeast	21.01	21.32	0.50	20.22	21.52
West	21.42	21.56	0.23	21.07	21.71
Gulf	19.63	19.90	0.38	18.98	19.99
Center	20.77	21.00	0.31	20.24	21.06
Southeast	20.42	20.70	0.45	19.67	21.07
South	20.65	21.02	0.49	19.86	21.15

Source: Author's own construction. This table shows the summary statistics for premium gasoline prices per year, from 2017 to 2019, by region – the first eight regions are fixed by the CRE, while the last region is the proposed region by the author. *1.* Deviation Standard. *2.* Proposed region by the author, which contains the municipalities of Ensenada, Mexicali, Playas de Rosarito, Tecate, Tijuana, Agua Prieta, Altar, Caborca, Cananea, General Plutarco Elías Calles, Nogales, Puerto Peñasco, San Luis Río Colorado, Guadalupe y Calvo, Juárez, Ojinaga, Práxedis G. Guerrero, Acuña, Ocampo, Piedras Negras, Anáhuac, Matamoros, Nuevo Laredo, Reynosa, Río Bravo, and Valle Hermoso.

Table 11: CORRELATION MATRIX – LEVEL/FIRST DIFFERENCE

	Northwest	***North***	***Northeast***	***West***	***Gulf***	***Center***	***Southeast***	***South***
Northwest	-							
North	0.992/0.717	-						
Northeast	0.993/0.631	0.997/0.892	-					
West	0.997/0.622	0.994/0.871	0.993/0.796	-				
Gulf	0.985/0.818	0.996/0.824	0.992/0.763	0.985/0.722	-			
Center	0.997/0.803	0.997/0.916	0.995/0.819	0.998/0.878	0.993/0.894	-		
Southeast	0.994/0.757	0.997/0.905	0.996/0.831	0.994/0.794	0.994/0.844	0.998/0.892	-	
South	0.996/0.855	0.995/0.859	0.996/0.772	0.993/0.761	0.992/0.884	0.997/0.906	0.997/0.881	-
Border	0.918/0.486	0.945/0.466	0.930/0.420	0.923/0.458	0.962/0.455	0.940/0.522	0.942/0.508	0.933/0.493

Source: Own construction with data from CRE. This table shows the price correlation coefficient in levels and first differences by regions for premium gasoline. The region Border contains municipalities of the states of Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas, which differs from the other regions by its different fiscal scheme.

Table 12: PARAMETER ESTIMATES FOR PREMIUM GASOLINE PRICES

	Northwest Gulf	***North Gulf***	***Northeast Gulf***	***West Gulf***	***Center Gulf***	***Southeast Gulf***	***South Gulf***	***West North***	***Northwest North***
B	0.63	-0.03	0.8	0.15	-0.31	-0.33	-0.41	-0.45	-0.22
σ_u^2	0.68	0.001	0.24	0.002	0.001	1.20	0.001	1.75	1.86
σ_v^2	0.006	0.13	0.23	0.30	0.30	0.23	0.28	0.20	0.02
λ	0.88	0.71	0.56	0.70	0.71	0.73	0.71	0.76	0.84
N	87	87	87	87	87	87	87	87	87
TC	1.88	0.97	1.08	1.16	0.73	1.39	0.65	0.63	0.80

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Table 13: PARAMETER ESTIMATES FOR PREMIUM GASOLINE PRICES

	Gulf Border	***South Border***	***Southeast Border***	***Center Border***	***West Border***	***Northeast Border***	***North Border***	***Northwest Border***
B	0.10	0.19	-0.06	0.78	1.05	0.43	0.44	1.08
σ_u^2	1	0.003	0.003	0.66	0.62	0.001	0.001	0.69
σ_v^2	0.007	0.34	0.44	0.04	0.07	0.30	0.23	0.007
λ	0.90	0.73	0.71	0.82	0.88	0.71	0.72	0.92
N	86	87	87	87	87	87	87	87
TC	1.11	1.21	0.99	2.18	2.86	1.54	1.56	2.94

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Table 14: PARAMETER ESTIMATES FOR PREMIUM GASOLINE PRICES

	Northwest Southeast	***North Southeast***	***Northeast Southeast***	***West Southeast***	***South Southeast***	***Center Southeast***	***West Northeast***	***Northwest Northeast***
B	-0.37	-0.24	-0.27	-0.26	-1	-0.90	-1.85	-0.44
σ_u^2	0.001	0.76	0.46	0.001	1.11	0.71	0.01	-0.44
σ_v^2	0.28	0.03	0.02	0.22	0.12	0.16	0.84	0.05
λ	0.71	0.73	0.80	0.70	0.37	0.42	0.70	0.90
N	87	87	87	87	77	85	84	62
TC	0.69	0.79	0.76	0.77	0.37	0.41	0.16	0.64

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by ***TC***. ***N*** represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

Table 15: PARAMETER ESTIMATES FOR PREMIUM GASOLINE PRICES

	Center South	***West South***	***North South***	***Northeast South***	***Northwest South***	***West Center***	***North Center***	***Northeast Center***	***Northwest Center***
$\boldsymbol{B}$	-2.20	-0.54	-0.73	-0.60	-0.13	-0.74	-0.92	-1.1	-1.14
σ_u^2	0.48	0.34	0.51	0.98	0.98	0.29	1.51	1.05	0.01
σ_v^2	1.05	0.41	0.01	0.14	0.11	0.29	0.07	0.20	0.72
λ	0.86	0.70	0.93	1	0.94	0.48	0.71	0.35	0.72
$\boldsymbol{N}$	60	87	76	87	87	87	74	77	87
$\boldsymbol{TC}$	0.11	0.58	0.48	0.55	0.89	0.47	0.40	0.33	0.32

Source: Author's own construction. This table shows the estimated parameters of the equation (10) by pair regions using the maximum likelihood estimation method. Transactions costs are represented by $\boldsymbol{TC}$. $\boldsymbol{N}$ represents the number of observations that were used to estimate the vector parameters $\hat{\theta}$. First region has the highest price.

References

Aigner, D., Lovell, C. K. & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21–37.

COFECE. (2019). THE TRANSITION TO COMPETITIVE RETAIL GASOLINE AND DIESEL, Reporte, Ciudad de México, México.

Contreras, B., Leal, R., Salcedo, A. & Mosqueda, J. (2020). Competition and coordination in the mexican retail market for gasoline, Working Paper, Banco de México, México.

Cournot, A. (1960). Mathematical principles of the theory of wealth, Translated by Nathaniel T. Bacon(Augustus M. Kelley, New York).

De Vany, A. & Walls, W. D. (1993). Access and Market Integration in the Natural Gas Industry: Evidence from Cointegration Tests. *The Energy Journal*, *14*(4), 1–19.

Horowitz, I. (1981). Market Definition in Antitrust Analysis: A Regression-Based Approach. *Southern Economic Journal*, *48*(1), 1–16.

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

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, *102*(5), 2147–2182.

Kiefer, N. M. (1980). A Note on Switching Regressions and Logistic Discrimination. *Econometrica*, *48*(4), 1065–1069.

Liu, X., Kirwan, B. & Martens, A. (2018). Regulatory compliance, information disclosure and peer effects: Evidence from the mexican gasoline market. *Journal of Regulatory Economics*, (54), 53–80.

Marvel, H. (1978). The economics of information and retail gasoline prices. *Review of Economics and Statistics*, 252–258.

PwC (Strategy&). (2019). Almacenamiento de petrolíferos en México : retos y oportunidades, Reporte, Ciudad de México, México.

Rodney, J. P., Dragan, M. & Viju, I. (2001). Market integration in US gasoline markets. *Applied Economics*, *33*(10), 1335–1340.

Slade, M. E. (1986). Exogeneity Tests of Market Boundaries Applied to Petroleum Products. *The Journal of Industrial Economics*, *34*(3), 291–303.

Spiller, P. T. & Huang, C. J. (1986). On the Extent of the Market : Wholesale Gasoline in the Northeastern United States. *The Journal of Industrial Economics*, *35*(2), 131–145.

Stigler, G. J. & Sherwin, R. A. (1985). The Extent of the Market. *The Journal of Law and Economics*, *28*(3), 555–585.

Weiner, R. J. (1993). Is the world oil market "one great pool"? *The Energy Journal*, *12*(3), 121–130.

List of Figures

List of Tables