Title page

Do classes of gas stations contribute differently to fuel prices? Evidence to foster effective competition in Spain

Jacint Balaguer[a], Jordi Ripollés[a]

[a]*Department of Economics. Universitat Jaume I*
Av. de Vicent Sos Baynat, s/n 12071 Castelló, Spain.

Abstract

Despite the relatively large number of gas stations reached in Spain after decades of sectorial reforms, pre-tax fuel prices in the country remain systematically among the highest in the EU. The literature provides evidence suggesting that a low intensity of competition in the retail distribution could contribute to these casual observations. With the purpose of shedding light on ways to design effective competition measures, we conduct an empirical analysis of more than ten million observations containing information about prices, brands, and locations at the station level. This allows us to know whether the exit (entry) of some classes of stations have the ability to reduce the prices of nearby competitors. Our results suggest that the presence in a local market of a station belonging to the network of the dominant market companies will tend to generate prices above the average. This is not only because these stations set higher prices but also because their presence will give rise to overpricing by local competitors. The opposite occurs with the self-advertised as "low-cost" stations. Policy measures promoting the gradual exit of stations associated with the dominant companies seem quite reasonable in view of the commitment to the transition toward transport decarbonization.

Keywords: Transportation fuel, Spain, retail prices, competition policy.

JEL codes: L13; Q41; Q48.

Email addresses: jacint.balaguer@uji.es (Jacint Balaguer), jripolle@uji.es (Jordi Ripollés)

1. Introduction

The aim of this paper is to contribute to the better design of competition policies in the retail fuel industry by exploring to what extent certain classes of gas stations operating in Spain could have different effects on their competitors' prices. Our study focuses on this country because its transport fuel prices (net of taxes) are systematically among the highest in Europe, although policy authorities have been developing sectorial rules aimed at improving the degree of competition.

More specifically, the importance of this problem can be illustrated by comparing the Spanish prices with those of the main fuel-consuming countries in the European Union (EU-28). For instance, according to a recent report on the sector (December, 2018),[1] it can be observed that pre-tax prices for diesel (gasoline) are 11.71% (17.81%) higher than in Germany, 9.67% (9.19%) higher than in France, and 16.02% (20%) higher with respect to the United Kingdom. This is surprising considering that over the last two decades the industry has experienced substantial legislative changes aimed at limiting the expansion of networks belonging to the dominant operators, as well as fostering the entry of fuel sellers. These legislative changes have been accompanied by a 33% increase in the number of new stations since the mid-2000s,[2] which has led to a relatively high number of them. This can easily be highlighted by a comparison with the three main fuel-consuming EU-28 countries indicated above. Thus, while Germany, France, and the United Kingdom have about 227, 285, and 223 stations per million vehicles, respectively, Spain has 349 stations per million.[3]

Reducing pre-tax prices on the basis of the number of sellers is not a straightforward policy task. In fact, it does not depend only on the price set by the entrant (outgoing) seller, but also on their impact on competitors' prices. The latter obviously cannot be directly observed from raw data on prices. If we turn to economic

[1]See the national competition authority report CNMC (2018).

[2]According to the information provided by the Spanish Association of Petroleum Products Operators (AOP), corresponding to the period from 2006 to 2017.

[3]Data refer to December 2017. It can also be observed that the proportion of stations per inhabitant in Spain is noticeably greater than in the other countries mentioned. For more information on data sources and the evolution of these ratios, see Appendix A.

theory, the predictions are not unambiguous. For instance, the conventional models in the spirit of Chamberlin's monopolistic competition indicate that a higher density of sellers implies lower prices. However, the opposite could also occur. According to Rosenthal (1980), the existence of captive consumers by some class of sellers associated to well-positioned brands can lead toward a price-increasing competition result. Moreover, following Chen and Riordan (2008), it is also conceivable to think that this last outcome becomes more likely as consumers' preferences for certain varieties increase.

The empirical literature on competition and retail fuel prices regularly found that a greater number of stations operating within predefined geographical areas implies lower prices (e.g., Barron et al., 2004; Hosken et al., 2008; Clemenz and Gugler, 2006; Pennerstorfer, 2009). However, the estimated importance of this causal relationship is quite heterogeneous. Both the incorporation of methodological improvements and the context studied seem critical in the empirical results. In fact, on the one hand, it should be kept in mind that outstanding advances have emerged and have occasionally been considered since the early empirical works. So, for example, Tappata and Yan (2017) address the endogeneity problem in the classical regression models used, while Perdiguero and Borrell (2018) delineated the relevant geographical markets for each seller based on a travel-time isochrone around each of them, which seems more realistic than using Euclidean distances. On the other hand, the Spanish context that we are concerned with here seems quite particular. It has been exposed to intensive liberalization and competition measures, but they do not seem to have led to a very satisfactory performance (e.g., García, 2010; Bello et al., 2018; Bernardo, 2018). Literature provides, however, some optimistic outcomes regarding the entry of a minority class of stations. This refers to findings in Bernardo et al. (2014) from data collected from a metropolitan area which suggest that the entry of the so-called "low-cost" stations and those linked to supermarkets significantly decreases prices at their nearby stations. It therefore remains to be confirmed whether this beneficial impact for prices occurs for the whole of Spain. Moreover, in view of the expected evolution of fuel demand due to the commitments to decarbonization, there is still a need to further explore whether it is possible that the exit of some classes of stations could help to reduce prices. In line with the theoretical papers mentioned above (i.e., Rosenthal, 1980; Chen and Riordan, 2008), it seems reasonable to further focus our attention on those stations associated with the best-positioned brands.

The rest of this paper is organized as follows. In the next section, we provide an overview of the sectorial policies carried out to date with a description of the context that was analyzed in each case. Additionally, we present relevant aspects of empirical literature on the topic. In Section 3, we will specify an empirical model in accordance with our aim and make the appropriate econometric considerations for its estimation. Further, we will describe the database that we will use as well as the construction of the variables. In Section 4 we will discuss our empirical outcomes, specifically, those related to the diagnostic tests, baseline results, and some tests for robustness. Finally, in Section 5 we will present the concluding remarks and policy implications under the commitment to the transition toward decarbonization in road transport.

2. Background

2.1. Characteristics of the context

Since the abolition of the *CAMPSA* monopoly in 1992, its commercial network has been distributed among other operating firms and the Spanish fuel oil sector began a new stage in the long process toward liberalization. Since then, the policy authorities have been generating new sectorial rules aimed at increasing the level of competition in the distribution of liquid fuels, especially since the beginning of the past decade. For example, they have been encouraging the entry of new stations in the retail market. More particularly, the restriction of maintaining a separation of at least 20 kilometers between stations included in the service areas of state roads was eliminated in 2001 (Royal Decree 114/2001), and the administrative procedures for obtaining new licenses were divided into two stages: the first for large shopping centers, in 2000, (Royal Decree-Law 6/2000, article 3) and the second for both shopping malls and areas of industrial activity, since 2013 (Royal Decree-Law 4/2013).

Moreover, in recent years restrictions have also been established on opening new retail locations and perpetuating existing ones for the major oil companies in the country. Since the year 2000 (Royal Decree-Law 6/2000), a temporary containment measure has been implemented to promote competition. Those wholesale operators for which the number of stations in their distribution network exceeded 30% of the total in the national territory could not increase the number of stations for a period of five years. If the percentage was between 15% and 30%, they could not expand the number of stations for three years. These two constraints affected stations linked with the oil companies Repsol and Cepsa, respectively, whether under ownership or by an exclusive contractual arrangement.[4] Because the supply contracts were more recently considered important barriers to expansion and the entry of alternative operators, they were limited to a maximum duration of one year (extendable twice) as of 2014 (Royal Decree-Law 4/2013). The expansion constraints were again tightened from 2016 onward (Law 8/2015), which ultimately affected both major operators. Particularly, new rules indicated that those firms whose distribution network sells over 30% in a specific province cannot open or acquire stations in that territory.

The set of measures outlined above have been accompanied by an increase in the number of stations, which has been quite remarkable in recent years, in spite of the important impact that the international financial crisis has had on demand. In particular, as we can see in Table 1, the total number of stations grew by more than 9% between 2010 and 2016, which contrasts with what happened in other major European countries. For example, the number of stations fell by about -1.59% in Germany, -7.06% in France, and -3.54% in United Kingdom during the above referenced period.[5] Most of these recent entries in the Spanish retail market belong to brands other than those that have traditionally been operating in the country. Thus, the overall market structure has been gradually changing with regard to the composition of the types of sellers. On the one hand, the pronounced rise in the number of independent retailers, self-advertised as "low cost", has been remarkable. On the other hand, there has been only a slight reduction in the number of stations associated with the major operators (Repsol and Cepsa). Specifically, the figure fell from 50% in 2010 to 45% in 2016. This is not surprising taking into account the initial advantage derived from the network that these oil companies received in 1992 from the state monopoly. Interestingly, a non-negligible proportion of about a fifth of these stations operate under exclusive

[4]Dealer-owned stations can be associated with the wholesaler through an exclusive supply and image contract. In these cases stations can be operated directly by dealers or by a wholesaler under a rental agreement.

[5]This evolution can be obtained from the Mineralölwirtschaftsverband, Union Française des Industries Pétrolières, UK Petroleum Industry Association, and Fuels Europe.

supply contracts,[6], which provide oil operators with enough flexibility to sequentially partner with different retailers in the territories in which the law allows them to do so as their contractual agreements expire or are terminated.

Table 1: Number of sellers by brand in Spain

	2010	2011	2012	2013	2014	2015	2016
Repsol	3,600	3,620	3,615	3,615	3,585	3,544	3,501
Cepsa	1,483	1,487	1,516	1,470	1,477	1,512	1,518
"Low-cost" brands:	285	294	373	403	473	534	584
Ballenoil	1	5	14	20	53	72	75
Petromiralles	16	13	63	63	61	62	63
Petroprix	0	0	0	3	9	20	33
Supermarkets	268	275	295	308	323	341	358
Other "low cost"	0	1	1	9	27	39	55
Others	4,870	4,908	4,920	5,129	5,177	5,357	5,585
Total	10,238	10,309	10,424	10,617	10,712	10,947	11,188

The total number of sellers and those associated with Repsol, Cepsa and supermarkets have been collected from the annual reports of the Spanish Association of Petroleum Products Operators (AOP). Data for Ballenoil, Petromiralles, Petroprix, and other "low-cost" brands have been collected from the Hydrocarbons Geoportal of the Spanish Ministry for the Ecological Transition. Supermarkets include Alcampo, Bonarea, Carrefour, E.Leclerc, Eroski, Esclatoil, Gmoil, Makro, and Simply. Other "low-cost" sellers include those whose commercial label contains the words "low cost". Information refer to December 31 of each respective year.

Finally, it is interesting to note than pre-tax price differences between the stations associated with the major operators and the "low-cost" brands are notable. For example, according to the Spanish Ministry for the Ecological Transition (December 2018), diesel prices at Repsol and Cepsa stations are on average 10% and 12% higher than

[6] According to rough estimates based on the information contained in CNMC (2009), ruling 06060/2009 of the Spanish Court for the Defense of Competition, and the Spanish Association of Petroleum Products Operators (AOP).

those fixed by "low-cost" stations. Initially, this cannot be attributed to the heterogeneity of fuels, since their basic chemical composition is stipulated by current law (Royal Decree 61/2006).[7]

2.2. Empirical literature

The retail fuel sector is one of the industries that has traditionally received most attention from economists. This is not surprising given its high economic weight in most countries. As can be seen in a survey by Eckert (2013), there are interesting issues largely debated in the industrial economics field that have been analyzed from the reality of this sector. Thus, for example, some of these empirically examined topics have been cost pass-through asymmetries (e.g., Bacon, 1991; Bachmeier and Griffin, 2003; Deltas, 2008), Edgeworth cycles (e.g., Noel, 2007; Noel, 2009; Lewis and Noel, 2011), merger effects (e.g., Coloma, 2002; Simpson and Taylor, 2008; Houde, 2012), regulatory impacts (e.g., Vita, 2000; Taylor and Fischer, 2003; Bernardo, 2018), competition effect on price dispersion (e.g., Lewis, 2008; Chandra and Tappata, 2011; Balaguer and Ripollés, 2018a), as well as the effect of competition on price levels. Since the early works, the literature on this last topic has incorporated substantial methodological advances. Next, we highlight improvements in controlling for station-level characteristics, determining local competitors, and treating endogeneity bias.

Collecting precise data on station-specific characteristics, in order to control for price differences that go beyond the impact derived from local competition, has attracted much attention from researchers. Thus, besides detailed information on prices set by stations, empirical works have also commonly collected identifications of brands (e.g., Barron et al., 2004; Sen, 2005; Tappata and Yan, 2017). In several of them, an additional effort was made to use data on other potentially relevant characteristics. Distance to refinery (Pennerstorfer, 2009), local per-capita income (Zimmerman, 2012), local population (Tappata and Yan, 2017), or indicators of the type of seller such as

[7]Dominant brands frequently attribute these differences to the introduction of voluntary additives in their products. Even assuming a certain degree of product differentiation, the present work is not so much interested in the observable price differences but in knowing how different classes of sellers affect their competitors' prices. This last unobservable aspect is what will shed light for designing appropriate rules oriented toward reducing the level of prices set by each type of station.

convenience store or repair station (Barron et al., 2004) have also been considered. Literature in this research area has also benefited from a modest but steadily growing number of studies that have exploited panel data (e.g., Hastings, 2004; Hosken et al., 2008; Lach and Moraga-González, 2017; Bernardo, 2018). While more degrees of freedom and sample variability is expected to improve the efficiency of the estimates, the most important advantage of this latest generation of studies is the control of a broad set of the unobserved characteristics.

Useful improvements have also been made in determining the relevant competitors for each operating station. The procedure followed by Shepard (1991) in a paper mainly devoted to price discrimination in retail fuel has been later applied by much of the empirical research on the topic that we are addressing here. Specifically, those relevant competitors are determined by generating a circle around each station. Sizes of the circle are generally defined arbitrarily by authors, but they are kept reasonably small to ensure that the sellers included within it can be viewed as substitutes for consumers. For example, the influential work by Barron et al. (2004) computed the number of neighboring gas sellers around each station within a 1.5-mile radius. While other authors have alternatively considered grosser measures to define the relevant market for each station, such as municipalities, administrative districts or commuter routes (e.g., Van Meerbeeck, 2003; Clemenz and Gugler, 2006; Cooper and Jones, 2007), the empirical strategy adopted in Barron et al. (2004) has been incorporated in many papers (e.g., Hosken et al., 2008; Pennerstorfer, 2009; Albalate and Perdiguero, 2015). Recently, however, a more sophisticated approach based on geo-information technologies has been exploited in a limited number of works. Specifically, both driving distances (e.g., Tappata and Yan, 2017; Kvasnička et al., 2018) and driving time (e.g., Perdiguero and Borrell, 2018; Perdiguero and Borrell, 2019) have allowed authors to define local markets in a more realistic way than by using simple Euclidean distances. The paper by Perdiguero and Borrell (2019) can be considered of special importance for our purpose since it seeks to delimit the relevant market for each gas station in Spain. Authors specifically found that it is delineated by a 5- to 6-min travel-time isochrone around each seller.

Avoiding potential endogeneity bias has been another interesting challenge in this literature. Although the problem derived of possible reverse causality from prices to the number of sellers has often been explicitly recognized (e.g., Barron et al., 2004, Hosken et al., 2008), it has rarely been addressed. Authors are likely to expect the bias to be small or, as suggested in Clemenz and Gugler (2006), insignificant. But this does not always seems to be true. In fact, the paper by Tappata and Yan (2017) indicates that ignoring endogeneity leads to serious underestimation of the effect of sellers on local average prices. Specifically, depending on the model specification used, it is shown that bias is around 55% to 70%.

Regarding the results on the effect of competitive pressure on price level, empirical literature offers consistent findings, at least with respect to the sign of this relationship. These results regularly indicate that more competition derived from the number of stations in a delimited geographical area is negatively associated with retail fuel prices or gross margins on wholesale prices (e.g., Barron et al., 2004; Clemenz and Gugler, 2006; Cooper and Jones, 2007; Sen and Townley, 2010; Nowakowski and Karasiewicz, 2016; Tappata and Yan, 2017; Bernardo, 2018). For example, estimates in Barron et al., 2004 indicate that a 50% increase in the number of competitors within a 1.5-mile radius around a station implies a decrease in the price of that station by about 0.5%. However, the magnitude of this impact is quite heterogeneous in the literature, suggesting that their importance is largely dependent on particularities of the context under analysis.

The Spanish context is particularly characterized by a relatively high number of stations which mostly belong to two operating brands (as can be seen in sub-section 2.1). This situation probably induces less effective competition than one would expect from the observed number of stations.[8] At least this is what emerges from previous research. The dynamics of pricing in retailing constitute one of the essential supports for this idea. Pricing behavior is found to be basically dependent on the strategy of the two dominant companies, which together respond faster to changes in wholesale fuel prices than other competitors (Balaguer and Ripollés, 2018a). Their price leadership is quite consistent with early research highlighting the capacity to generate collusive agreements in this market (García, 2010). In addition, it may have facilitated two

[8]Regardless of any further absence of effective competition in upstream markets (those including the activities of refining, transportation, storage, and distribution to the pump location of fuel products).

particular pricing anomalies that took place, at least for relatively recent periods, at the beginning of 2010. First, there was a behavior consisting in cutting prices on Mondays, the day on which the European Commission collected data to monitor the sector, and then sharply increasing them again on Tuesdays (Jiménez and Perdiguero, 2014). Second, there was also an asymmetric price response to wholesale price changes which took the form of the well-known "rockets and feathers" phenomenon during the first week of adjustments (Balaguer and Ripollés, 2016).

Evidence on the performance of the retail market and its evolution also support a low degree of effective competition. On the one hand, literature offers us two studies indicating that neither the effects of the economic crisis as of 2008 nor the relaxation of entry restrictions a few years later have been able to cut the substantial gross retail margins to any significant degree.[9] Specifically, the paper by Bello et al. (2018) reveals that those stations belonging to companies with broad market power have tended to increase their gross margins during the beginning of the recessive period in order to offset the drop in consumption, which resulted in the generation of greater price differences between the various sorts of stations. The paper by Bernardo (2018) further suggests that the reforms undertaken in 2013 (Royal Decree-Law 4/2013) had a limited impact on the gross margins of stations. Indeed, the entry of more than one station in local areas has hardly any marginal effects, and any that had an initial impact ended up being diluted to a great extent. Specifically, it was revealed that the average reduction in gross retail margins has been only 0.75% one year after the first entry of a new seller in local markets.

Lastly, we also benefit from findings regarding specific stations linked to supermarkets and those designated as "low cost". These findings refer to those obtained in Bernardo et al. (2014), from data on the metropolitan area of Barcelona (Spain). It is suggested that this class of sellers impose a significant discipline on the prices of the stations near them. Although as the authors themselves point out this outcome should be taken with caution due to possible endogeneity bias, it seems quite reasonable and interesting for our purpose. In fact, it is broadly consistent with outcomes previously

[9] The gross retail margin can be defined as the percentage difference between retail (pre-tax) prices and wholesale spot prices quoted in reference markets.

obtained from similar classes of stations operating in countries such as Austria (Pennerstorfer, 2009) and the USA (Zimmerman, 2012).

3. Methodology and data

Inspired by the empirical studies that use a circular approach around a seller to define their relevant competitors in the local market (e.g., Shepard, 1991; Barron et al., 2004; Lewis, 2008; Balaguer and Pernías, 2013; Lach and Moraga-González, 2017), we ask ourselves about the effect of a particular type of seller on the average prices of the surrounding competitors. Therefore, we formulate a baseline specification where the logarithm of the average price fixed by the competitors of a seller i at time t can be expressed as:

$$ln(\overline{P}_{it}) = \alpha + \sum_{m=1}^{M} \beta_m CLASS_{m,it} + \sum_{m=1}^{M} \gamma_m ln(N_{m,it}) + \theta_t + \lambda_z + u_{it} \qquad (1)$$

where α is a constant term, and $CLASS_m$ represents a set of dummy variables $m = 1, 2..., M$ that take a value of 1 if seller i belongs to class m, and 0 otherwise. Their associated coefficients (i.e., β_m) represent the impact of each seller's class. Moreover, the price in the local environment surrounding i is also expected to depend on the market size and composition. With the aim of controlling for both these factors, we have introduced information on the number of competitors from each of the different classes. Therefore, $ln(N_m)$ refers to the logarithm of 1 plus the number of sellers belonging to class m that operate in the local environment surrounding seller i. Time fixed effects, θ_t, can be useful to capture the impact from the regular wholesale price changes, while spatially clustered (ZIP-code) fixed effects, λ_Z, control for the characteristics of the neighborhood where the seller is located (e.g., transportation costs from the supply center or local demand idiosyncrasy). Finally, u_{it} is the random disturbance term, which captures the influence of other unobserved variables.

The estimation of the above specification requires some considerations. First, to avoid perfect collinearity, we have chosen to refer β_m with respect the mean effect caused by all classes of sellers. That is, $\sum_{m=1}^{M} \beta_m = 0$. This strategy will facilitate the interpretation of coefficients. So, by introducing this consideration in Eq. (1), the specification can be finally redefined as:

$$ln(\overline{P}_{it}) = \alpha + \sum_{m=1}^{M-1} \beta_m class_{m,it} + \sum_{m=1}^{M} \gamma_m ln(N_{m,it}) + \theta_t + \lambda_z + u_{it} \quad (2)$$

where $class_{m,it}$ (i.e. $CLASS_{m,it} - CLASS_{M,it}$) represents a set of dummy variables $m = 1, 2..., M-1$, that takes a value of 1 if seller i belongs to class m, a value of -1 if seller i belongs to class M, and 0 otherwise. Therefore, the coefficient associated to the sellers' class M is now dropped.

Second, although for the sake of simplicity Eq.(2) will be initially estimated by ordinary least squares (OLS) as in much of the literature in this research area, we acknowledge that the validity of the exogeneity of right-hand variables with respect to prices could not be fulfilled. Namely, both price and sellers in a market could be simultaneously determined since high (low) price levels could appeal (drive away) prospective entrants, and conversely.[10] To overcome this concern, we will also estimate Eq.(2) by using instrumental variables (IV). Finally, as can occur in this sort of model, we will further take into account that u_{it} could be heteroskedastic, as well as spatially and temporally correlated.

Regarding data, they are mainly collected from the Hydrocarbons Geoportal of the Spanish Ministry for the Ecological Transition (`www.geoportalgasolineras.es`). This website contains information provided by each gas station operating on the Spanish peninsula, all of which are required to submit their retail prices every Monday and when price changes take place (Ministerial Order ITC/2308/2007). Although the website only provides real-time information and historical series are not publicly available for confidentiality reasons, we have collected complete information for every gas station daily between December 2010 and July 2016. The resulting dataset is an unbalanced panel comprising information on diesel prices, coordinates, and brand identity for a maximum of 10,876 gas stations over the sampled time period.[11] Because we are only

[10]Despite this simultaneity problem also being acknowledged in the literature, it is not always addressed empirically (e.g., Barron et al., 2004, Kwoka and Shumilkina, 2010, Zimmerman, 2012). In favor of these works we can say that, according to Clemenz and Gugler (2006), this issue may not be particularly relevant in this sort of model.

[11]Some observations have obviously been lost due to temporary closures (e.g., holidays and/or repair work) or departures of stations.

interested in the pricing behavior of stations, all taxes have been excluded in accordance with the information available from the Spanish Tax Agency.[12]

In order to build the variables of Eq. (2) from our data, it is necessary to group the stations by brand type and define the relevant local market. With regard to the first purpose, we believe that two aspects should be conciliated, namely, obtaining a final model that is sufficiently parsimonious to avoid further complication in the estimation process, and introducing groups that are sufficiently homogeneous, at least in the brands of interest, to provide a useful answer to our objective. So, we consider it essential to obtain information about the effect of three classes of brands on local markets. On the one hand, the first group of brands includes Repsol and Cepsa. We consider that they are the most well-positioned brands. Both brands belong to traditional oil firms, which are vertically integrated, and have the largest network of stations on the Spanish peninsula. On the other hand, the brands that are advertised as “low cost” are presumably another relatively homogeneous group in terms of pricing behavior, whose possible effect on local competitors could be of special interest. Finally, the remaining stations will be considered as another group.

The other challenge we face is to define as accurately as possible the relevant local market around each sampled gas station. With this purpose in mind, we take as a reference the recent empirical paper by Perdiguero and Borrell (2018), whose findings suggest that the relevant market in the Spanish fuel sector is delineated by a 5-min driving time isochrone surrounding each gas station.[13] Then, information on the coordinates from our dataset is exploited to define 5-min isochrones around each sampled station by using the Open Source Routing Machine software (`www.project-osrm.org`), which is based on the shortest car route in accordance with the road networks and speed

[12]The Spanish taxes for diesel fuel include indirect and special excise duties. The former is given by the value added tax (VAT), ranging from 16% to 18% from July 1st 2010, and from 18% to 21% since September 1st 2012. The latter is composed of a general rate (0.307 euros/liter), a State rate (0.024 euros/liter), and a regional rate that depends on each autonomous government. A summary of the regional tax rates prevailing during the sample period is given in Table B1 from Appendix B.

[13]This implies an approximate travel distance of 4 and 10 kilometers in urban and interurban roads with speed limits of 50 and 120 km/h, respectively. It is quite consistent with the result obtained by the circular approach followed in Balaguer and Ripollés (2018b).

limits available at OpenStreetMap.

Once the sellers have been grouped and the relevant local market defined, we identify those neighboring stations with which each seller competes on a daily basis within their corresponding isochrone. With this information, on the one hand, we obtain the average prices of the competitors ($ln(\overline{P}_{it})$) and, on the other hand, we quantify the number of surrounding stations belonging to each class considered (i.e., Repsol, Cepsa, "low cost", and others). For descriptive purposes, in Table 2 we show the sample annual averages for prices and number of local competitors in the defined markets. As can be seen, stations associated to Repsol and Cepsa are surrounded by a relatively smaller number of competitors, which are also more expensive than the overall average. The opposite is observed in the case of "low-cost" stations, which are surrounded by relatively more competitors with lower prices. Even so, one can appreciate how the presence of competitors around all classes of sellers has progressively increased, without exception, over the period of time considered. Hence, for example, the average number of local competitors has increased by about 22% for Repsol and Cepsa, 52% for "low-cost" brands, and 29% for the remaining class of sellers. Finally, we also examine the stationary properties of these variables. To do so, the Im et al. (2003) unit root test has been applied on the demeaned series, which accounts for certain forms of cross-sectional dependence. The corresponding test results suggest that our series are stationary (as can be seen in Table C1 in Appendix C).

Table 2: Average prices (and number of competitors) 5-min around each class of seller

	Repsol	Cepsa	"Low cost"	Others
December 2010	0.950 (1.669)	0.949 (1.698)	0.947 (2.430)	0.948 (1.972)
2011	1.038 (1.702)	1.038 (1.748)	1.035 (2.545)	1.037 (2.018)
2012	1.096 (1.751)	1.096 (1.788)	1.093 (2.667)	1.095 (2.082)
2013	1.068 (1.795)	1.067 (1.802)	1.060 (2.806)	1.065 (2.164)
2014	1.018 (1.861)	1.018 (1.877)	1.008 (3.114)	1.013 (2.276)
2015	0.865 (1.951)	0.865 (1.988)	0.856 (3.518)	0.861 (2.433)
January - July 2016	0.762 (2.034)	0.761 (2.094)	0.755 (3.697)	0.757 (2.538)

Authors' elaboration based on data from the Spanish Ministry for the Ecological Transition.

4. Estimation

4.1. Baseline results

We carry out the estimation of Eq. (2) by using the OLS and IV procedures. The corresponding results are shown in Table 3.[14] In both estimation procedures we have applied the covariance matrix estimation proposed by Driscoll and Kraay (1998), which yields a robust covariance matrix estimation with heteroskedasticity and very general forms of temporal and cross-sectional dependence when, as in our case, the time dimension is large. The reason for this is that, by applying the groupwise heteroskedasticity test of Greene (2000), the serial correlation test of Wooldridge (2010), and the cross-sectional dependence statistic (CD) proposed by Pesaran (2004),[15] these problems have been significantly revealed.

While the efficiency of the OLS estimator is slightly higher, the IV procedure will be more appropriate if there is endogeneity. For this last procedure, we use a city-specific population as an instrumental variable of the variation in the number of retailers, in line with other research on the issue, such as Campbell and Hopenhayn (2005), Clemenz and Gugler (2006), and Sen and Townley (2010).[16] Second, taking advantage of panel data information (e.g., Evans et al., 1993; Reed, 2015), we also employ the six-month lagged values for brand class identity ($class_{m,it}$) and for the number of local competitors ($ln(N_{m,it})$). We are aware that the reliability of the IV procedure depends on the use of appropriate instruments, which should be exogenous and sufficiently correlated with the endogenous regressors. A set of diagnostic tests presented in the bottom panel of Table 3 indicates that the instruments used are adequate. Specifically, from the Hansen J test, we cannot reject the null hypothesis that instruments are uncorrelated with the error term at the standard signification levels suggesting then that the instruments employed are exogenous, while the Kleibergen-Paap rk Wald F statistic is well above

[14] Note that the relative effect corresponding to the class of stations dropped in the Eq. (2) is given by minus the sum of the coefficients associated with the rest of classes (i.e., $class_{others} = -class_{Repsol} - class_{Cepsa} - class_{low\,cost}$), while their corresponding standard error has been obtained through the delta method.

[15] Specific values for these tests are, respectively, $\chi^2(7961) = 72 \times 10^{14}$, $F(1, 7944) = 240,403.2$, and $CD = 56,836.1$.

[16] Data on the city-specific population has been drawn from the Spanish National Statistics Institute.

the conventional rule of thumb ($F > 10$) proposed by Staiger and Stock (1997), indicating a sufficiently strong correlation between our instruments and the regressors that are presumed to be endogenous. Moreover, the validity of our set of instruments also depends on whether our resulting model is identified. Rejection of the null hypothesis of the model's under-identification in the Kleibergen-Paap rk LM test suggests that this is the case (Kleibergen and Paap, 2006). Finally, the Durbin-Wu-Hausman test suggests that the OLS estimates might be inconsistent, probably due to some degree of simultaneity between prices and sellers.

Therefore, taking the above mentioned diagnostic tests into account, here we focus on the IV estimates presented in Table 3, despite the apparent similarities with those obtained from OLS. The first thing that draws our attention is that, with a high level of confidence, the effects exerted by the presence of each class of seller are not equal to each other. Regardless of the number of stations and the proportion of each class in the local market, the relative effects of stations belonging to the dominant networks are significantly positive. Specifically, the presence of a Repsol station increases prices in its surrounding competitors by 0.058% above the average effect of the stations. In the case of a Cepsa station, the results show that the impact runs in the same direction. This last type of station relatively increases the prices of surrounding competitors by 0.097%. These results contrast with those corresponding to "low-cost" stations. In this last case, competitors' prices after the entry of one of these stations relatively decrease by 0.194%. As we know, the rest of the stations have been captured by the variable $class_{others}$. Because this last group is quite heterogeneous and numerous, it is not surprising that their corresponding coefficient is not significantly different from the average effect at the conventional levels of significance.

The signs of the corresponding coefficients associated with the variables that help us to control the number and composition of competitors across local markets are what we expected. They are clearly consistent with both the prices fixed by each class of stations and, according to the coefficients just discussed, with their relative effect on the neighboring sellers. Specifically, a new competitor belonging to the dominant network will cause an increase in the competitors' average price that is consistent with their higher price as well as the relatively positive impact on the prices of their nearby sellers. The opposite situation would occur with the entry of stations belonging to the "low-cost" brands or even with stations within the general classification of other brands.

Finally, note that the patterns exhibited by those stations associated with the dominant network could be consistent with the theoretical models that contemplate the existence of captive consumers who are unwilling to substitute certain well-positioned brands (e.g., Rosenthal, 1980; Chen and Riordan, 2008). The greater presence of stations associated with Repsol and Cepsa, as well as the higher capacity for expenditure on advertising of these two companies, could positively contribute to enhance the brand values perceived by consumers, which would to some extent account for the results obtained concerning the dominant sellers.

Table 3: Baseline results

	OLS		IV	
$class_{Repsol}$	0.083***	(0.030)	0.058*	(0.033)
$class_{Cepsa}$	0.111***	(0.037)	0.097**	(0.039)
$class_{low\,cost}$	-0.244***	(0.055)	-0.194***	(0.060)
$class_{others}$	0.051	(0.054)	0.039	(0.058)
$ln(N_{Repsol})$	0.896***	(0.055)	0.926***	(0.059)
$ln(N_{Cepsa})$	0.626***	(0.049)	0.674***	(0.054)
$ln(N_{low\,cost})$	-2.479***	(0.089)	-2.378***	(0.099)
$ln(N_{others})$	-1.024***	(0.048)	-0.975***	(0.051)
R^2	0.979		0.980	
Time observations	2,070		1,888	
Total observations	11,964,618		10,482,118	
Hansen J			1.244 [0.265]	
Kleibergen-Paap rk Wald F			11,000	
Kleibergen-Paap rk LM			243.991 [0.000]	
Durbin-Wu-Hausman			5,839.66 [0.000]	
$H_0 : \beta_1 = ... = \beta_M$	20.60 [0.000]		11.10 [0.011]	

Dependent variable is the average price of surrounding competitors. All regressions include a constant term, and dummy variables for time and ZIP codes. Estimated coefficients and standard errors are multiplied by 10^2. Standard errors for $class_{others}$ coefficients are obtained by using the delta method. Driscoll-Kraay's standard errors with 6 lags in the autocorrelation structure are presented in parenthesis. We use ***, **, and * to denote statistical significance at the 1, 5 and 10% levels, respectively. P-values are in brackets.

4.2. Robustness check

In this section, we are first interested in knowing whether, by using the IV procedure as it has been defined so far, the results presented in Table 3 remain robust to the introduction of some dynamics in our model, and also to changing competitors that are considered relevant. More concretely, on the one hand, we extend the Eq. (2) by including the lagged dependent variable as an additional regressor.[17] In this case, long-run impacts can also be calculated in order to facilitate the comparison with those coefficients obtained in the baseline analysis. That is, by assuming that ρ is the coefficient of the lagged dependent variable, we will multiply the parameters capturing the short-run impact by $1/(1-\rho)$. On the other hand, we alternatively delineate local markets within a 2.5-minute driving distance of the geographical location of each station. A reduction of time with respect to the period used in the baseline analysis (which is in accordance with Perdiguero and Borrell (2019)) seems more in line with several previous works. In fact, this involves a travel distance of about 2 and 5 kilometers in urban and interurban roads with speed limits between 50 and 120 km/h, respectively. It is interesting to note that in papers such as Barron et al. (2004), Hosken et al. (2008) markets are defined in a 2.4-km (1.5-mile) radius around the station, and in both Hastings (2004) and Bernardo (2018) only a 1.6-km (1-mile) radius was considered. Therefore, new variables for both the average competitors' price ($ln(\overline{P}_{it})$) and the number of each station class are constructed according to the new size of the isochrone. So, on this occasion we have approximately half of the observations of the baseline analysis because many stations have been excluded as they are now assumed to act as spatial monopolies.

The estimates from considering a dynamic model and smaller local markets are firstly displayed in Table 4. In both cases, diagnostic tests confirm the model's identification and the suitability of the IV estimator over OLS. In the dynamic model, the coefficient associated with the lagged dependent variable is positive and statistically significant (at a 1% level), which suggests a remarkable persistence of retail prices over

[17]It is well known that standard panel estimators with a lagged dependent variable can yield biased coefficients when the time dimension (T) is small and the cross-sectional dimension (N) tends to infinity. However, this bias is expected to be negligible in our analysis given the reasonably large number of time periods. See Nickell (1981), Kiviet (1995), and Judson and Owen (1999).

time. After calculating the long-run coefficients, we can see that they are quite similar to those provided by the previous baseline analysis. Additionally, after reducing the size of local markets, our findings remain quite robust. Namely, once controlled for the number of sellers and their composition, the relative effects exerted by the presence of each class of station are not statistically equal to each other. Particularly, we find that the presence of stations associated with Repsol and Cepsa relatively increases the prices fixed by its surrounding competitors, while "low-cost" brands cause the opposite effect.

Finally, we ask ourselves how robust results would be in response to a different choice of instrumental variables in the estimation process. Specifically, we have decided to use as instruments the six-month lagged values for the brand class identity ($class_{m,it}$) and the number of local competitors $ln(N_{m,it})$), without considering the city-specific population. The results obtained from this exactly identified case are presented at the last columns of Table 4.[18] Interestingly, the relative impacts of each class of station are not statistically equal to each other. Once again, the signs of these impacts are consistent with those obtained in the baseline analysis.

[18]In this case, note that we cannot test the exogeneity assumption of instruments.

Table 4: Robustness check based on IV method

	(I) Lagged dependent variable model				(II) 2.5-min isochrones of driving time		(III) Only lagged values as IV	
	Short-run impacts		Long-run impacts					
$class_{Repsol}$	0.002*	(0.001)	0.053*	(0.032)	0.186**	(0.081)	0.057*	(0.032)
$class_{Cepsa}$	0.003**	(0.001)	0.092**	(0.040)	0.212**	(0.094)	0.097**	(0.039)
$class_{low\,cost}$	-0.006***	(0.002)	-0.185***	(0.062)	-0.564***	(0.141)	-0.194***	(0.060)
$class_{others}$	0.001	(0.001)	0.041	(0.036)	0.166**	(0.071)	0.039	(0.030)
$ln(N_{Repsol})$	0.028***	(0.006)	0.918***	(0.187)	1.534***	(0.353)	0.925***	(0.060)
$ln(N_{Cepsa})$	0.021***	(0.003)	0.667***	(0.127)	1.507***	(0.395)	0.673***	(0.054)
$ln(N_{low\,cost})$	-0.073***	(0.005)	-2.358***	(0.128)	-7.163***	(0.490)	-2.378***	(0.099)
$ln(N_{others})$	-0.030***	(0.004)	-0.970***	(0.154)	-2.055***	(0.290)	-0.975***	(0.051)
$ln(\overline{P}_{it-1})$	96.914***	(0.170)						
R^2	0.998				0.970		0.980	
Time observations	1,888				1,888		1,888	
Total observations	10,475,863				5,040,683		10,482,118	
Hansen J	1.103 [0.294]				0.204 [0.651]			
Kleibergen-Paap rk Wald F	8,263.292				42.488		12,000	
Kleibergen-Paap rk LM	242.311 [0.000]				127.858 [0.000]		243.933 [0.000]	
Durbin-Wu-Hausman	1,232.10 [0.000]				41,997.09 [0.000]		330.86 [0.000]	
$H_0 : \beta_1 = ... = \beta_M$	9.40 [0.009]				16.18 [0.001]		111.10 [0.011]	

Dependent variable is the average price of surrounding competitors. All regressions include a constant term, and dummy variables for time and ZIP codes. Estimated coefficients and standard errors are multiplied by 10^2. Standard errors for $class_{others}$ coefficients and long-run impacts are obtained by using the delta method. Driscoll-Kraay's standard errors with 6 lags in the autocorrelation structure are presented in parenthesis. We use ***, **, and * to denote statistical significance at the 1, 5 and 10% levels, respectively. P-values are in brackets.

5. Conclusions and policy implications

Despite the policy efforts made in Spain to implement measures aimed at increasing competition in the retail fuel market, the country is among those that regularly have the highest pre-tax prices in the EU. Market structure, behavior, and performance indicators suggest that even with the intensive entry of new sellers, the retail distribution of transport fuel in this country is far from competitive. To be able to introduce adequate measures to increase effective competition, we first need to know which sellers notably contribute to it and which sellers do not. With this purpose in mind, we have built a parsimonious model which helps to reveal the ability of certain types of stations to change their near competitors' prices. On the one hand, we have explored the differential impact derived from the presence of those stations associated with vertically integrated companies that have a dominant presence in the retail market (i.e., Repsol and Cepsa) and, on the other hand, from those stations that advertise themselves as "low-cost" brands.

Our main empirical analysis is based on more than ten million observations corresponding to the variables of interest, which were collected daily between late 2010 and mid-2016. Namely, we refer to data on diesel prices, brands, and the geographical location of the stations operating on the Spanish peninsula. Prices are subsequently expressed excluding the corresponding general, special, and regional fuel oil taxes. In addition, we exploit the Open Source Routing Machine service to precisely delineate reasonable local markets within a 5-minute driving distance of the geographical location of each gas station. After controlling for heterogeneity in the number of each type of competitor, and by using standard errors that are robust to heteroskedasticity, temporal and cross-sectional dependence, we focus our attention on estimates obtained by the IV procedure to prevent simultaneity bias.

The empirical findings suggest that each type of gas station contributes differently to competitors' prices in local markets. On the one hand, the entry of stations which belong to the network of any of the two dominant operators is less favorable to decreasing competitors' prices than the entry of any other type of gas station. In fact, our results indicate that their presence would cause competitors' prices to relatively increase in comparison with the average impact that would result from the presence of a representative gas station. On the other hand, those stations advertised as "low cost" significantly impact in the opposite direction. In other words, this last group of sellers does not only fix prices generally below the average in consistence with their label, but also produces relative downward effects on the prices of local competitors. The importance of the effect exerted by "low-cost" stations gives us an idea of the aggressive competition they cause. We have further verified that these outcomes are robust to the introduction of some dynamics in the model, redefinition of the size of the relevant geographic markets, and the use of alternative instrumental variables.

With the aim of evaluating the implications for future policies in the sector, it seems advisable to take into account the foreseeable context that will have to be faced in the coming decades. By this we are referring to the expected contraction of demand for fossil fuels as a result of decarbonization policies, which will end up causing a reduction in the number of stations. This is what can be expected, at the very least, from the draft of the Spanish Integrated Energy and Climate Plan (2021-2030). As reported in this plan, the need for fossil fuels for motor vehicles at the end of the next decade will have substantially decreased due to changes in means of transport as well as to the introduction of electric vehicles. Specifically, it estimates that 16% of all vehicles running on Spanish roads will be electric or will use advanced biofuels by 2030, which contrasts with the less than 1% of this type of vehicles currently being driven in Spain. Moreover, according to a recent communication released by the European Commission (COM/2018/773 final), the objective is that zero emissions from the entire fleet of vehicles will be achieved by 2050. Consequently, given the foreseeable scenario of a contraction in oil demand, it seems appropriate to initially promote the exit of stations belonging to the two dominant firms in order to reduce pre-tax prices. The effect on price levels would then be determined through two channels. On the one hand, lower prices would be obtained since these types of stations are among the most expensive in the industry. And on the other hand, the effect on competitors' prices would also be relatively favorable to reducing pre-tax prices.

Even though the network expansion of these two dominant firms has been constrained by current legislation in certain provinces (Law 8/2015), this measure does not appear to be sufficient. The relocation of these types of sellers to local areas with a higher density of competitors could even cause undesirable effects on price levels. One possible recommendation is to prevent the creation and the renewal of flagging contracts with the above mentioned dominant companies. In fact, an important part of the stations operating in their networks are owned by dealers. Those stations that do not renew their contracts could, however, remain in the market as an independent brand as long as the demand allows them to do so. After all, we must bear in mind that the efficiency gains of the remaining type of stations would come from both their lower prices and, if they adopt a "low-cost" strategy, their higher gains on the prices of the surrounding competitors.

Finally, it should be underlined that any gain in pre-tax price reduction does not seem advisable to be transferred to consumers. At least this could be justified from an environmental point of view. Namely, the literature has highlighted that any reduction in petroleum product prices for consumers would have significant undesirable implications on the country's carbon emissions (Balaguer and Cantavella, 2016). Also it should be kept in mind that a decrease in fuel prices is likely to discourage the incorporation of new fuel-saving vehicles to the fleet (e.g. Rivers and Schaufele, 2017). Therefore, it seems more appropriate that any market efficiency improvement derived from lower prices should be transferred to other productive sectors by exploiting the tax capacity gains on each unit sold. This is particularly feasible in Spain, where the current tax rates on fuel oil are lower than the EU average.

Acknowledgements

We thank the members of the Energy Policy Research Group at the University of Cambridge (UK), where the project of this work began during a research stay carried out by J. Balaguer in spring 2016, and the School of Business and Economics at the University of Loughborough (UK) for their helpful comments and suggestions during a research stay undertaken by J. Ripollés in winter 2019. We also appreciate the useful comments of two anonymous referees. We hold ourselves solely responsible for any errors. Financial support from the Spanish Ministry of Science, Innovation and Universities (RTI2018-100983-B-I00), and Universitat Jaume I (UJI-B2019-53) are gratefully acknowledged.

References

Albalate, D. and Perdiguero, J. (2015). Entry regulation asymmetries and petrol competition in a mixed motorway network. *Journal of Transport Economics and Policy (JTEP)*, 49(4):603–625.

Bachmeier, L. J. and Griffin, J. M. (2003). New evidence on asymmetric gasoline price responses. *Review of Economics and Statistics*, 85(3):772–776.

Bacon, R. W. (1991). Rockets and feathers: the asymmetric speed of adjustment of uk retail gasoline prices to cost changes. *Energy economics*, 13(3):211–218.

Balaguer, J. and Cantavella, M. (2016). Estimating the environmental kuznets curve for spain by considering fuel oil prices (1874–2011). *Ecological Indicators*, 60:853–859.

Balaguer, J. and Pernías, J. C. (2013). Relationship between spatial agglomeration and hotel prices. Evidence from business and tourism consumers. *Tourism Management*, 36:391–400.

Balaguer, J. and Ripollés, J. (2016). Asymmetric fuel price responses under heterogeneity. *Energy Economics*, 54:281–290.

Balaguer, J. and Ripollés, J. (2018a). The dynamics pattern of price dispersion in retail fuel markets. *Energy Economics*, 74:546–564.

Balaguer, J. and Ripollés, J. (2018b). Revisiting the importance of border effect in subnational regions. Evidence from a quasi-experimental design. *Papers in Regional Science*, 97(4):1113–1130.

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.

Bello, A., Contin-Pilart, I., and Palacios, M. (2018). Pricing and margins in the retail automotive fuel market: Empirical evidence from spain. *Energy Journal*, 39.

Bernardo, V. (2018). The effect of entry restrictions on price: evidence from the retail gasoline market. *Journal of Regulatory Economics*, 53(1):75–99.

Bernardo, V., González, J. L. J., and Perdiguero, J. (2014). Competencia y medidas liberalizadoras en el sector minorista de hidrocarburos: análisis del efecto sobre los precios de la entrada de nuevos operadores en españa. *Cuadernos Económicos de ICE*, (88):127–144.

Campbell, J. R. and Hopenhayn, H. A. (2005). Market size matters. *The Journal of Industrial Economics*, 53(1):1–25.

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

Chen, Y. and Riordan, M. H. (2008). Price-increasing competition. *The RAND Journal of Economics*, 39(4):1042–1058.

Clemenz, G. and Gugler, K. (2006). Locational choice and price competition: some empirical results for the austrian retail gasoline market. *Empirical Economics*, 31(2):291.

CNMC (2009). Primer informe anual de supervision del mercado de hidrocarburos liquidos, Julio 2009. pages 1–50. Comision Nacional de Mercados y Competencia.

CNMC (2018). Supervision de la distribucion de carburantes en estaciones de servicio, Diciembre 2018. pages 1–28. Comision Nacional de Mercados y Competencia.

Coloma, G. (2002). The effect of the repsol-ypf merger on the argentine gasoline market. *Review of Industrial Organization*, 21(4):399–418.

Cooper, T. E. and Jones, J. T. (2007). Asymmetric competition on commuter routes: the case of gasoline pricing. *Southern Economic Journal*, pages 483–504.

Deltas, G. (2008). Retail gasoline price dynamics and local market power. *The Journal of Industrial Economics*, 56(3):613–628.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80(4):549–560.

Eckert, A. (2013). Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys*, 27(1):140–166.

Evans, W. N., Froeb, L. M., and Werden, G. J. (1993). Endogeneity in the concentration–price relationship: Causes, consequences, and cures. *The Journal of Industrial Economics*, pages 431–438.

García, J. P. (2010). Dynamic pricing in the spanish gasoline market: A tacit collusion equilibrium. *Energy Policy*, 38(4):1931–1937.

Greene, W. H. (2000). Econometric analysis 4th edition. *International edition, New Jersey: Prentice Hall*, pages 201–215.

Hastings, J. S. (2004). Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california. *American Economic Review*, 94(1):317–328.

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.

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 102(5):2147–82.

Im, K. S., Pesaran, M. H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.

Jiménez, J. L. and Perdiguero, J. (2014). Información, presión institucional y estrategías empresariales: el efecto lunes en el mercado español de gasolina. *ICE, Revista de Economía*, (876).

Judson, R. A. and Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters*, 65(1):9–15.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, 68(1):53–78.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Kvasnička, M., Staněk, R., and Krčál, O. (2018). Is the retail gasoline market local or national? *Journal of Industry, Competition and Trade*, 18(1):47–58.

Kwoka, J. and Shumilkina, E. (2010). The price effect of eliminating potential competition: Evidence from an airline merger. *The journal of Industrial Economics*, 58(4):767–793.

Lach, S. and Moraga-González, J. L. (2017). Asymmetric price effects of competition. *The Journal of Industrial Economics*, 65(4):767–803.

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

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.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Noel, M. (2009). Do retail gasoline prices respond asymmetrically to cost shocks? the influence of edgeworth cycles. *The RAND Journal of Economics*, 40(3):582–595.

Noel, M. D. (2007). Edgeworth price cycles, cost-based pricing, and sticky pricing in retail gasoline markets. *The Review of Economics and Statistics*, 89(2):324–334.

Nowakowski, M. and Karasiewicz, G. (2016). Market structure and price-cost margins in european retail gasoline industry. *Journal of Management and Business Administration*, 24(3):105–124.

Pennerstorfer, D. (2009). Spatial price competition in retail gasoline markets: evidence from austria. *The Annals of Regional Science*, 43(1):133–158.

Perdiguero, J. and Borrell, J. R. (2018). Driving competition in local markets with near-perfect substitutes: an application on the spanish retail gasoline market. *Empirical Economics*, pages 1–20.

Perdiguero, J. and Borrell, J. R. (2019). Driving competition in local markets with near-perfect substitutes: an application on the spanish retail gasoline market. *Empirical Economics*, 57(1):345–364.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels.

Reed, W. R. (2015). On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics*, 77(6):897–905.

Rivers, N. and Schaufele, B. (2017). Gasoline price and new vehicle fuel efficiency: Evidence from canada. *Energy Economics*, 68:454–465.

Rosenthal, R. W. (1980). A model in which an increase in the number of sellers leads to a higher price. *Econometrica: Journal of the Econometric Society*, pages 1575–1579.

Sen, A. (2005). Does increasing the market share of smaller firms result in lower prices? empirical evidence from the canadian retail gasoline industry. *Review of Industrial Organization*, 26(4):371–389.

Sen, A. and Townley, P. G. (2010). Estimating the impacts of outlet rationalization on retail prices, industry concentration, and sales: empirical evidence from canadian gasoline markets. *Journal of Economics & Management Strategy*, 19(3):605–633.

Shepard, A. (1991). Price discrimination and retail configuration. *Journal of Political Economy*, 99(1):30–53.

Simpson, J. and Taylor, C. (2008). Do gasoline mergers affect consumer prices? the marathon ashland petroleum and ultramar diamond shamrock transaction. *The Journal of Law and Economics*, 51(1):135–152.

Staiger, D. and Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65:557–586.

Tappata, M. and Yan, J. (2017). Competition in retail petrol markets. *Australian Economic Papers*, 56(3):201–218.

Taylor, C. and Fischer, J. (2003). A review of west coast gasoline pricing and the impact of regulations. *International Journal of the Economics of Business*, 10(2):225–243.

Van Meerbeeck, W. (2003). Competition and local market conditions on the belgian retail gasoline market. *De Economist*, 151(4):369–388.

Vita, M. G. (2000). Regulatory restrictions on vertical integration and control: The competitive impact of gasoline divorcement policies. *Journal of Regulatory Economics*, 18(3):217–233.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Zimmerman, P. R. (2012). The competitive impact of hypermarket retailers on gasoline prices. *The Journal of Law and Economics*, 55(1):27–41.

Appendix A

Table A1: Stations per million inhabitants (Inh.) and vehicles

	Spain		Germany		France		UK	
	Inh.	Vehicles	Inh.	Vehicles	Inh.	Vehicles	Inh.	Vehicles
2006	198	343	184	308	214	372	162	284
2017	247	349	175	227	167	285	128	223
% variation	24.75	1.75	-4.89	-26.30	-21.96	-23.39	-20.99	-21.48

The number of stations has been taken from National Oil Industry Associations (referring to December 31 of each year); the population was obtained from Eurostat; and the vehicles in use were collected from the Spanish Dirección General de Tráfico, the German Federal Motor Transport Authority (Kraftfahrt-Bundesamt), the French Comité des Constructeurs Français d'Automobiles, and the Society of Motor Manufacturers and Traders in UK.

Appendix B

Table B1: Regional tax rates on diesel fuel (expressed in euros/liter)

In force since:	Jan 1 2006	Jul 10 2010	Jan 1 2012	Jan 10 2012	Mar 1 2012	Apr 1 2012	May 1 2012	Jun 2 2012	Jul 1 2012	Jun 23 2012	Nov 1 2012	Jan 1 2013	Jan 1 2014	Jan 1 2015	Apr 12 2015	Jan 1 2016
Andalusia	0.000	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.048	0.048	0.048	0.048	0.048	0.048	0.048
Aragon	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Asturias	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.040	0.040	0.040	0.040	0.040
Basque Com.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cantabria	0	0	0	0	0	0	0	0.048	0.048	0.048	0.048	0.048	0.024	0	0	0
C-La Mancha	0.024	0.024	0.024	0.024	0.024	0.024	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048
C-Leon	0	0	0	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.016	0.016	0
Catalonia	0.024	0.024	0.024	0.024	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048
Extremadura	0	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.048	0.048	0.048	0.048	0.0384	0.0384	0.0384	0.0384
Galicia	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.048	0.048	0.048	0.048
La Rioja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Madrid	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017
Murcia	0	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.048	0.048	0.048	0.048	0.048	0.048
Navarre	0	0	0	0	0	0	0	0	0	0	0	0	0.024	0.024	0.024	0.024
Valencian Com.	0.012	0.012	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048

Authors' elaboration based on data from the Spanish Tax Agency.

Appendix C

Table C1: Results of Im-Pesaran-Shin panel unit root test

$ln(\overline{P}_{it})$	$ln(N_{Repsol})$	$ln(N_{Cepsa})$	$ln(N_{low\,cost})$	$ln(N_{others})$
-290*** [0.000]	-220*** [0.000]	-43.921*** [0.000]	-41.213*** [0.000]	-430*** [0.000]

Cross-sectional dependence is controlled by subtracting cross-sectional means from the observed data (demeaned data), while serial autocorrelation is taken into account by considering the optimal lag length according to the Akaike Information Criteria. Superscripts ***, **, and * denote rejection of the null hypothesis (all the panels contain a unit root) at the 1, 5 and 10% levels, respectively. P-values are in brackets.