Exploring the life of fuel price responses in retail markets. The effect of cross-sectional aggregation

Jacint Balaguer [*a] and Jordi Ripollés [†a]

[a]Department of Economics, Universitat Jaume I, Spain.

Abstract

Empirical studies on vertical price transmission in retail fuel markets commonly use average group data of petrol stations. In this paper a simulation approach is employed to illustrate that, in this case, the persistence of price responses tends to be overestimated. To explore the real extent of the problem, we apply the mean group (MG) and the mean group with common correlated effects (MG-CCE) estimators to individual data from petrol stations. When heterogeneity in the pricing dynamics is captured by MG and MG-CCE estimators, persistence of retail price responses becomes considerably lower than the typical OLS estimations from average group data would suggest.

Keywords: Fuel price responses; Cross-sectional aggregation; Dynamic persistence; Overestimation.
JEL classification: C51; C23; L71; Q41.

Acknowledgements
Financial support from the European Union FEDER founds and the Spanish Ministry of Economy and Competitiveness (ECO2014-58975-P) is gratefully acknowledged.

[*]Tel.: 0034 964387161. E-mail address: jacint.balaguer@eco.uji.es.
[†]Tel.: 0034 964388609. E-mail address: jripolle@eco.uji.es. (corresponding author)

1 Introduction

Since the seminal paper by Bacon (1991) there has been an increasing interest in knowing the speed of fuel price responses to input price shocks (e.g. Borenstein et al., 1997; Polemis, 2012; Contín-Pilart et al., 2009; Remer, 2015; Asane-Otoo and Schneider, 2015). Much of the effort in this field has been devoted to testing possible asymmetries in the speed of response to upward and downward shocks,[1] while little attention has been paid to knowing to what extent the degree of aggregation underlying the data commonly used could affect these estimates. However, it is well known that aggregation over time or over individuals in dynamic models, like those used in this research area, could imply substantial bias.

A reduction of information from temporal aggregation could preclude model specification from properly capturing the intertemporal lag distribution of the real phenomena (Geweke, 1978). This is a type of omitted variables problem that has been revealed to be clearly significant in the estimation of the fuel price adjustment towards long-run equilibrium through Error Correction Models (ECMs). For instance, in the early paper by Bachmeier and Griffin (2003),[2] the coefficient of the adjustment from weekly data is about five times higher than that obtained from daily data. Thus, the use of daily frequencies is recommended insofar as they are available to researchers (Bachmeier and Griffin, 2003). On the other hand, several theoretical papers have demonstrated that cross-sectional aggregation to build average group data could cause a reduction in the estimated speed of responses and, therefore, an overestimation of the dynamics of the aggregate phenomena studied. This sort of bias, which is critically dependent on the behavioural heterogeneity of the individuals involved in the analysis (e.g. Pesaran, 2003; Stoker, 1993), is revealed to be especially significant under application of ECMs (Lippi, 1988). Given

[1] For a complete meta-analysis at this regard see, for example, Perdiguero-García (2013).

[2] The aim of this paper is to re-examine, by using an ECM, the previous results from Borenstein et al. (1997) based on weekly data.

that the behaviour of petrol stations could be quite heterogeneous (Haucap et al., 2017), undermining the consequences of cross-sectional aggregation by appealing to the representative-agent paradigm could be, in this case, an unsuitable strategy which deserves special attention as highlighted by Faber (2015).[3]

The present paper aims to conduct a comprehensive investigation of the extent to which estimates from average group data of petrol stations could be overstating the life of retail fuel price responses. From a literature search we can see that, in spite of the potential bias commented above, most of the evidence on fuel price responses is based on data at weekly or lower frequencies which, in turn, are aggregated to build data averaged by country. With the purpose of trying to isolate the possible impact of cross-sectional aggregation on the previous analyses, we will focus on those works that include estimates of the adjustment towards long-run equilibrium based on daily data.[4] These papers offer us results from different degrees of cross-sectional aggregation, which can be usefully exploited for a preliminary exploration of the importance of the problem.[5]

First, we can consider two papers that employ data at country level. With the aim of illustrating the length of the dynamics in each case, let us now to use a synthetic measure. Concretely, we are going to pay attention to the average number of days needed to adjust 95% of the retail price deviations from the long-run equilibrium after an input price shock from the wholesale-refined fuel market.[6] Thus, in

[3] In this paper it is indicated that the pooled estimation, where parameters are restricted to be equal across individuals, is not suitable to analyze the market as a whole. Then, the author advocates the use of a separate analysis for each individual station.

[4] To our knowledge, this is the maximum degree of data disaggregation available in this research area. Probably because of the difficulty in obtaining daily information on retail fuel prices, few papers with this temporal disaggregation have been published to date.

[5] Some papers that employ daily data, but do not involve data aggregation across petrol stations, are not included. This is, for example, the case studied by Bachmeier and Griffin (2003), where shocks from crude oil prices to wholesale petrol prices are analysed and, therefore, the behavioural heterogeneity of sellers operating in the retail market is not a relevant issue.

[6] This can easily be derived from the estimated coefficients of speed of adjustment towards the long-run equilibrium provided in each of the papers. The period necessary to adjust 95% of

the work by Al-Gudhea et al. (2007), which provides estimates for USA, around 464 days are required to fill this percentage of the gap after a shock. The results from Balaguer and Ripollés (2012) concerning Spain would imply that, following a shock in international fuel markets, about 251 days are needed to close this percentage of the gap. Considering the surprising number of days required in both cases, we might suspect that aggregation across petrol stations could be causing a relevant overestimation of the life of price responses, at least, when it is performed at country level.

Second, we can take into account the paper by Bettendorf et al. (2009). In this case, retail prices are aggregated for the Shell brand in the Dutch petrol market. Authors indicated that competing oil companies in the Netherlands could easily monitor and follow the pricing policy of this company with the largest market share (30%), although it is recognized that differences in pricing strategies may exist between the firms operating in that market. From the results of the estimations of this leader company, we obtain that the period needed to adjust 95% of the retail price deviations from the equilibrium is only 27 days. Hence, evidence may suggest that data disaggregation by brand (which presumably involves relatively homogeneous sellers) could be contributing significantly to reduce overestimation in the life of price responses.

Third, the recent papers by Remer (2015) and Balaguer and Ripollés (2016) are based on data from individual petrol stations with the usual purpose of testing asymmetries in the response of prices. In the first of the two cases, the retail price observations belong to petrol stations from several states in USA (New Jersey, Maryland, Virginia, Philadelphia and Washington). From the estimates, we find that adjustments would need about 43 days until 95% of the gap was closed. From

the deviations is obtained from the natural logarithm of 0.05 divided by the value of the speed adjustment coefficient. When the adjustment is split in order to distinguish the effect of upward and downward shocks, for the sake of simplicity, we will consider the average coefficient.

the second of the papers we can observe the period required to close this percentage of the gap for the two greater metropolitan areas in Spain. Specifically, we are referring to Madrid and Barcelona, for which 25 and 27 days were required, respectively. It is clear that the lengths of price responses in the last three papers contrasted with those where data are aggregated at country level, which suggests the importance of the problem when different operating brands with singular pricing strategies are disregarded. At this point, we further ask ourselves whether aggregation at brand level could be considered a reasonable empirical strategy to prevent an important part of the overestimation of the length of the dynamics.

The rest of the paper is organized as follows. In Section 2, we start by describing the baseline model. Following, we will discuss a set of results obtained by simulations to draw attention to the problem statement. In Section 3, we present an application on Spanish diesel prices that have been collected daily to prevent, as much as possible, temporal aggregation bias. Then, we will compare the empirical results from heterogeneous panel estimates with those derived from the common OLS estimates based on cross-sectionally aggregated data. First, we will perform this comparison using the entire sample. Thus, we try to examine to what extent aggregation across a large diversity of brands, which probably hides a high degree of behavioural heterogeneity, affects the accuracy of the estimation results. Second, we further focus on a sub-sample of the leader's petrol stations. The purpose will be to know whether the extent of the problem would persists through this empirical strategy. In Section 4, we test the robustness of the results. Finally, in Section 5, we present our concluding remarks.

2 The analytical framework

2.1 Model specification

It is well known that, in the case of cointegration between non-stationary time series, implementation of an ECM specification can be a particularly useful strategy to describe dynamics between variables (Engle and Granger, 1987). This empirical strategy is mostly being adopted by studies on vertical price transmission in fuel markets. Here, we adopt this approach by introducing the possible heterogeneity of the behaviour of each of the firms ($i = 1, 2, \dots, N$):

$$\Delta p_{it} = \sum_{m=1}^{M} \beta_{im} \Delta p_{it-m} + \sum_{j=0}^{J} \delta_{ij} \Delta wp_{t-j} + \theta_i (p_{it-1} - \alpha_i - \phi_i wp_{t-1}) + \epsilon_{it} \qquad (1)$$

where Δ is the first-differences operator, p_{it} is the retail price of the i-th firm at time t ($t = 1, 2, \dots, T$), and wp_t is the corresponding wholesale fuel price at time t, which is common for all operating firms. The speed of adjustment towards the level of equilibrium is captured by θ_i, characterizing the long-run dynamics of the model. Lastly, ϵ_{it} is a random disturbance term, which is assumed to be *iid*.

Second, we also consider a restricted version of Eq. (1) as follows:

$$\Delta \bar{p}_{.t} = \sum_{m=1}^{M} \beta_m \Delta \bar{p}_{.t-m} + \sum_{j=0}^{J} \delta_j \Delta wp_{t-j} + \theta (\bar{p}_{.t-1} - \alpha - \phi wp_{t-1}) + \epsilon_t \qquad (2)$$

where $\bar{p}_{.t} = \frac{1}{N} \sum_{i=1}^{N} p_{it}$, and the subscript i in coefficients and the random disturbance term is disregarded. Then, the specification from Eq. (2) implies some degree of cross-sectional aggregation as is usual in this research area.

2.2 Simulation

With the aim of illustrating the problem statement, let us obtain artificial series of retail fuel prices based on Eq. (1). For this purpose we consider, as the wholesale price series (wp_t), spot prices of refined diesel fuel. Specifically, we consider the wholesale spot prices at Amsterdam-Rotterdam-Antwerp (Euros/litre), which represent the principal and direct raw material cost in the European retail markets for diesel. These prices are provided by Platts of the McGraw-Hill Company and correspond to a sample period that ranges from 10 June 2010 to 25 November 2012.[7] A preliminary analysis of time series indicates that it follows an I(1) process. Namely, in accordance with the Augmented Dickey-Fuller test (ADF) proposed in Dickey and Fuller (1979), wholesale price series are non-stationary in levels (ADF$=-1.59$) but their first differences can be considered stationary (ADF$=-20.32$).

Cointegration between retail price and wholesale price series would require the coefficient of the speed of adjustment towards equilibrium in Eq. (1) to be negative ($\theta_i < 0$). In order to facilitate our simulation process, let us assume that the coefficient (θ_i) is uniformly distributed on a plausible bounded interval $[-1, 0]$. Therefore, we allow for heterogeneity in the speed at which firms adjust towards long-run equilibrium. Moreover, for the sake of simplicity, the remaining coefficients in Eq. (1) are fixed (in accordance with the values shown in the caption of Table 1). Finally, we assume that ϵ_{it} is a random disturbance term, which is distributed normally with mean zero and constant variance (of 0.05).

[Please insert Table 1 about here]

Taking into account our wholesale fuel prices (wp_{it}) and the above assumptions, in accordance with Eq. (1) let us first obtain 1,000 simulated panels of retail prices (p_{it}) corresponding to each of the uniformly distributed coefficients that have been

[7] Missing values resulting from closure of the spot markets at weekends and in holidays have been filled in with prices from the day before.

generated (θ_i). Each of the simulated panels are thus composed of $N = 300$ and $T = 900$. Subsequently, in each panel, we can aggregate retail prices across the N firms to estimate the speed of adjustment regardless of the behavioural heterogeneity, in accordance with the restricted ECM from Eq. (2). In Table 1 we show the estimated speed of adjustment (averaged for the 1,000 simulated panels) and the corresponding period required to close 95% of the gap. The estimated coefficient is, in absolute terms, lower than the mean of the uniformly distributed coefficients (θ_i) in the interval $[-1, 0]$. Consequently, when aggregate retail prices are employed, the period required to close this gap overestimates the mean period for individuals. Specifically, the expected life of price responses in terms of 95% of the gap is overestimated by about 2 days.

We have considered it interesting to replicate the above simulation, but now allowing for a lower range of heterogeneity behaviour of individuals regarding the coefficient of speed (θ_i). With this purpose we define the speed of adjustment as uniformly distributed in a shorter range $[-0.8, -0.2]$, which involves the same mean speed coefficient as before. As we can see in Table 1, the estimated coefficient of speed from aggregate data is, once again, lower in absolute terms and the life of price adjustment is consequently oversized. However, note that the difference between the mean of the true individual coefficients and the estimated coefficient is lower than when more heterogeneity is allowed. Thus, the expected life of price responses is also overestimated but to a lesser extent than before.

3 An econometric exploration

Now we present an application based on real prices fixed by petrol stations. We first estimate the dynamic heterogeneous panel represented by Eq. (1) by adopting the econometric approach proposed by Pesaran and Smith (1995), which is widely known as the mean group (MG) estimator. When large time panels are available, the MG estimator has proved to be a suitable procedure to prevent inconsistent and highly misleading estimates of the coefficients from aggregate data if behaviour differs across individuals. In this case individual micro-relations should be estimated separately by OLS and, then, the mean of the estimated coefficients and their standard errors calculated explicitly. In this section we will further apply the MG common correlated effects (MG-CCE) procedure proposed by Pesaran (2006). It is an extension of the MG procedure that provides consistent estimates under serial correlation in errors and various forms of cross-sectional dependence among individuals (Pesaran and Tosetti, 2011). We must take into account that this extension may be especially useful in our case where the pricing behaviour of each i-firm can be driven by unobserved factors that are common to their neighbouring firms, such as greater (less) demand derived from the higher (lower) level of traffic in some locations. In practice, the correlated effects will be captured by the lagged cross-sectional average of the prices fixed by all competitors operating at each moment of time within a certain radius.[8] Finally, estimates of Eq. (1) by MG and MG-CCE will be subsequently compared with those obtained by OLS from aggregated data. In this last case, we will consider the restricted ECM from Eq. (2).

Our retail prices are now diesel prices (Euros/litre) collected daily at midnight from the website of the Spanish Ministry of Industry, Energy and Tourism (http://geoportal.mityc.es/) throughout the period from 10 June 2010 to 25 November 2012.[9] The sample includes prices from 590 petrol stations involving a large number of operating brands as well as a relevant group of independent sellers in Spain.[10] It is obvious that it presumably contains petrol stations with notably

[8]A similar application, in which the average behaviour of neighbouring individuals is taken into account, can be seen in Eberhardt and Teal (2013).

[9]Petrol stations that, for reasons such as closure due to repair work or holidays, did not provide prices for any part of the period have not been considered.

[10]We consider the stations operating throughout the entire period located in the metropolitan areas of Madrid, Barcelona and Valencia. Concretely, our sample covers the following cities:

different pricing strategies. Therefore, to explore the possible reduction in the overestimation of the length of the dynamics from aggregation at brand level, we have also carried out fieldwork to identify those petrol stations belonging to the leading brand in the sector (Repsol). We then alternatively use the sub-sample resulting from considering the prices fixed by the 226 petrol stations that operate under this brand. In short, we employ a full sample and a sub-sample of retail prices consisting in two panels of 531,000 and 203,400 retail price observations, respectively. All these prices are expressed net of taxes following the information by the Spanish Ministry of Economy's Tax Office. To implement the MG-CCE procedure we will need the precise geographical location of the petrol stations. To this end, we extracted the coordinates (longitude and latitude) from the website of the Spanish Ministry of Industry, Energy and Tourism.

Before applying the ECM represented by Eq. (1), we analysed the time series properties to ensure that there is cointegration between the retail and the wholesale price series. First, the analysis of time series indicates that they follow an I(1) process. Specifically, the Breitung and Das (2005) panel unit root test, which is robust to the presence of cross-sectional dependence,[11] was employed on the set of

Alacuás, Albal, Albalat dels Sorells, Alboraya, Albuixech, Alcácer, Alcalá de Henares, Alcobendas, Alcorcón, Aldaya, Alfafar, Alfara del Patriarca, Almácera, Badalona, Badia del Vallès, Barberà del Vallès, Barcelona, Begues, Benetúser, Beniparrell, Boadilla del Monte, Bonrepós y Mirambell, Brunete, Burjasot, Castellbisbal, Castelldefels, Catarroja, Cerdanyola del Vallès, Cervelló, Chirivella, Colmenar Viejo, Corbera de Llobregat, Cornellà de Llobregat, Coslada, Cuart de Poblet, El Papiol, El Prat de Llobregat, El Puig, Emperador, Esplugues de Llobregat, Foyos, Fuenlabrada, Gavà, Getafe, Godella, Humanes de Madrid, L'Hospitalet de Llobregat, La Palma de Cervelló, Las Rozas, Leganés, Lugar Nuevo de la Corona, Madrid, Majadahonda, Manises, Masalfasar, Masamagrell, Masanasa, Mejorada del Campo, Meliana, Mislata, Molins de Rei, Moncada, Montcada i Reixac, Montgat, Móstoles, Museros, Paiporta, Pallejà, Paracuellos de Jarama, Parla, Paterna, Picaña, Picasent, Pinto, Pozuelo de Alarcón, Puebla de Farnals, Puzol, Rafelbuñol, Ripollet, Rivas-Vaciamadrid, Rocafort, San Antonio de Benagéber, San Fernando de Henares, San Sebastián de los Reyes, Sant Adrià de Besòs, Sant Andreu de la Barca, Sant Boi de Llobregat, Sant Climent de Llobregat, Sant Cugat del Vallès, Sant Feliu de Llobregat, Sant Joan Despí, Sant Just Desvern, Sant Vicenç dels Horts, Santa Coloma de Cervelló, Santa Coloma de Gramenet, Sedaví, Silla, Tabernes Blanques, Tiana, Torrejón de Ardoz, Torrelles de Llobregat, Torrente, Tres Cantos, Valencia, Velilla de San Antonio, Viladecans, Villanueva de la Cañada, Villanueva del Pardillo, Villaviciosa de Odón, and Vinalesa.

[11] In accordance with the LM-statistic of Breusch and Pagan (1980), we can reject cross-sectional independence among the retail prices (with p-values virtually equal to zero).

individual retail prices, while the ADF unit root test was applied to cross-sectionally aggregated retail prices. Results are presented in Table A1 of the Appendix A. Additionally, prices of the corresponding raw material are the wholesale spot prices of refined diesel, whose stationary analysis has been performed in the previous subsection. Second, we chose the bootstrap approach of Westerlund and Edgerton (2007) to test for cointegration between individual retail prices and wholesale prices, as it allows for heterogeneity and is robust to very general forms of cross-sectional interdependence. While two of the statistics proposed are group-mean tests (G_τ and G_α), under the alternative hypothesis that at least one cross-sectional unit is cointegrated, the other two statistics are panel tests (P_τ and P_α), with the alternative hypothesis that the whole panel is cointegrated. We obtain that the set of individual retail price series and wholesale price series are cointegrated. Finally, the ADF test also indicates cointegration between aggregate retail prices and wholesale prices.

We can now estimate Eq. (1) for both the full and the sub-sample. To know whether behavioural heterogeneity should be considered in our estimation process, we employ a Hausman-type statistical test. As we can see in Table 2, in all cases we overwhelmingly reject the null hypothesis of behavioural homogeneity across the N individuals. Thus, we can think that MG and MG-CCE should provide more accurate estimates, whereas the OLS to aggregate data across firms, where the assumption of a representative agent is implicitly imposed, should be taken with caution. At this point we would ask ourselves whether the results from the heterogeneous-type panel estimator are substantially different from those obtained from aggregate data. In Table 2 we also present the results to answer this question and, based on the speed coefficients (θ), in Figure 1 we provide a general comparison among the estimated adjustments resulting from the MG and MG-CCE procedures in Eq. (1), and OLS in Eq. (2).

[Please insert Table 2 about here]

[Please insert Figure 1 about here]

Although all the estimated coefficients of the speed are significantly negative (indicating that retail fuel prices return towards their equilibrium after price shocks), the size of the estimated coefficients from MG and MG-CCE clearly differ from those obtained from OLS based on aggregate data. For the full sample, the magnitude of the estimated speed of adjustment towards long-run equilibrium is consistently much higher in the first two cases. The magnitudes of the estimated speed of adjustment to equilibrium by taking into account the MG and MG-CCE procedures are about two times higher than those based on OLS for aggregated prices. The importance of the difference can be illustrated in the Figure 1(a) in terms of the life of retail price responses. For example, as can be seen, retail prices require 27 and 25 days to adjust the 95% of the gap according to the MG and MG-CCE estimates based on disaggregated data, while it becomes 58 days when heterogeneity is neglected in the OLS estimates based on aggregate data.

To explore to what extent there is a reduction in the overestimation of the length of the dynamics from aggregation only at brand level, we can focus our attention on the sub-sample of stations belonging to Repsol. In this case, the magnitudes of the estimated speed of adjustment from the MG and MG-CCE procedures is only about 23% higher than that based on OLS for aggregated prices. As we can see in the Figure 1(b), the number of days required to adjust 95% of the gap after a wholesale-refined fuel price shock are 30 and 29 from the MG and MG-CCE procedures, respectively. These results obtained taking into account heterogeneity barely diverge from those based on aggregate retail prices at brand level, which are about 37 days. This would be consistent with the fact that behaviour is more homogeneous among petrol stations within a brand than if all the brands and independent stations are included the analysis. It therefore seems that the overestimation of the length of the dynamics would be substantially reduced when aggregation is performed at brand level.

4 Robustness check

To test the robustness of our empirical results,[12] we alternatively consider crude oil prices as the origin of shocks, in line with many other works in the literature (e.g. Borenstein et al., 1997; Grasso and Manera, 2007). Specifically, we employ the wholesale prices for Europe Brent crude oil (Euros/litre), which were obtained from the US Energy Information Administration. A preliminary analysis indicates that the crude oil price series follow an I(1) process. That is, according to the ADF test, the series of crude oil price are non-stationary in levels (ADF=-1.83), but stationary in first differences (ADF=-22.31). Moreover, as can be seen in Table A1 of the Appendix A, the results of the Westerlund and Edgerton (2007) tests suggest that there is cointegration between the set of individual retail prices and the crude oil price series, while the outcome of the ADF test also indicates cointegration between aggregate retail price and crude oil price series.

The main results can be seen in Table 3. Moreover, in Figure 2 we also provide a graphical representation based on the estimated coefficients of the speed of adjustment (θ). As expected, estimates now indicate the need for more days to reach 95% of the transmission in relation to previous results concerning shocks from the wholesale refined fuel market. The comparison between MG or MG-CCE estimates and those obtained with OLS show us the robustness of the outcome from considering the crude oil prices alternatively. Thus, for example, the magnitude of the estimated speed of adjustment to equilibrium by taking into account the MG-CCE

[12]Our empirical results are also robust to the decomposition of estimated coefficients in accordance with upward and downward input price shocks. Results from decomposition are available from the authors upon request.

procedure is about two and a half times (one and a half times) higher than that based on aggregated prices for the full sample (sub-sample for Repsol). Hence, the expected number of days required to close the percentage of the gap with equilibrium is almost three times (one and a half times) lower than that based on aggregated prices for the full sample (sub-sample for Repsol).

[Please insert Table 3 about here]

[Please insert Figure 2 about here]

5 Concluding remarks

The literature on vertical price transmission in fuel markets typically uses ECM specifications to measure the speed at which the retail prices adjust towards the long-run equilibrium. These papers are generally based upon aggregate data across petrol stations, which might lead to bias in the estimation of the speed. We first address the problem statement by discussing a set of simulated results from wholesale refined fuel prices. Our simulation illustrated that application of OLS by using aggregate data and, therefore, excluding potential behaviour heterogeneity may cause an overestimation of the life of adjustment towards the long-run equilibrium. It has been further illustrated how the problem is reduced if behaviour among firms involved in the analysis is more similar.

To explore whether the period of price responses might be frequently overestimated, we exploit a dataset of real retail prices for diesel. These prices refer to a sample of petrol stations operating in Spain. They have been collected daily to prevent, as much as possible, temporal aggregation bias. In this way we have tried to isolate the possible impact on the estimated speed derived from aggregation over a large number of brands and independent stations included in the sample. Indeed, application of a Hausman-type statistical test has suggested that the behaviour of the operating sellers is significantly heterogeneous. To account for this, we have applied the MG and MG-CCE procedures on individual data following Pesaran and Smith (1995) and Pesaran (2006). Empirical evidence from both heterogeneous panel procedures consistently suggests that the speed at which real prices return to the long-run equilibrium after a shock is substantially greater than that estimated by the OLS from aggregate data. That is, the life of the retail price responses seems to be notably oversized when the presence of heterogeneity is not taken into account by the dynamic model. Aggregation when analysis is performed for a sub-sample of petrol stations belonging to an specific company (i.e. Repsol) seems rather less problematic.

Thus, our findings can be seen as complementary to the recommendation put forward earlier by Bachmeier and Griffin (2003). That is, besides the use of high frequencies in this study area, we further advice on the importance of using disaggregated data for petrol stations as far they are available. Evidence from empirical researches adopting the approach advised here could be useful to better know when price changes in international fuel markets would be transmitted to local markets. It could be particularly interesting to improve the cost-saving strategies of retailers in the fuel sector. Thus, by monitoring the prices of crude oil or refined fuel in international markets, they can take advantage of a substantial temporary drop in the wholesale prices by purchasing from distributors in the precise moment in which the drop is fully transmitted. Moreover, it is well known that fuel is a very important consumer product in industrialized countries and, therefore, the evolution of its retail prices plays an essential role in the control of inflation. Hence, for monetary authorities it can be of great interest to have accurate information about when a substantial oil price shock will be transmitted to the local market. In fact, monetary measures introduced to moderate consumer price responses would probably be implemented too late if they are based on the estimates from aggregate data. Finally, knowing more accurately the life of price responses could be also helpful in order to design adaptive fiscal policies, consisting in timely modifications of fuel taxes, oriented at reducing the retail price volatility of fuel products.

References

Al-Gudhea, S., Kenc, T., and Dibooglu, S. (2007). Do retail gasoline prices rise more readily than they fall?: A threshold cointegration approach. *Journal of Economics and Business*, 59(6):560–574.

Asane-Otoo, E. and Schneider, J. (2015). Retail fuel price adjustment in germany: A threshold cointegration approach. *Energy Policy*, 78:1–10.

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 Ripollés, J. (2012). Testing for price response asymmetries in the spanish fuel market. new evidence from daily data. *Energy Economics*, 34(6):2066–2071.

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

Bettendorf, L., Van der Geest, S. A., and Kuper, G. H. (2009). Do daily retail gasoline prices adjust asymmetrically? *Journal of Applied Statistics*, 36(4):385–397.

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

Breitung, J. and Das, S. (2005). Panel unit root tests under cross-sectional dependence. *Statistica Neerlandica*, 59(4):414–433.

Contín-Pilart, I., Correljé, A. F., and Palacios, M. B. (2009). Competition, regulation, and pricing behaviour in the spanish retail gasoline market. *Energy Policy*, 37(1):219–228.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431.

Eberhardt, M. and Teal, F. (2013). No mangoes in the tundra: spatial heterogeneity in agricultural productivity analysis. *Oxford Bulletin of Economics and Statistics*, 75(6):914–939.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55(2):251–276.

Faber, R. P. (2015). More new evidence on asymmetric gasoline price responses. *The Energy Journal*, 36(3):287–307.

Geweke, J. (1978). Temporal aggregation in the multiple regression model. *Econometrica*, 46(3):643–661.

Grasso, M. and Manera, M. (2007). Asymmetric error correction models for the oil–gasoline price relationship. *Energy Policy*, 35(1):156–177.

Haucap, J., Heimeshoff, U., and Siekmann, M. (2017). Fuel prices and station heterogeneity on retail gasoline markets. *The Energy Journal*, 38(6):81–103.

Lippi, M. (1988). On the dynamic shape of aggregated error correction models. *Journal of Economic Dynamics and Control*, 12(2):561–585.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11(6):601–618.

Perdiguero-García, J. (2013). Symmetric or asymmetric oil prices? a meta-analysis approach. *Energy Policy*, 57:389–397.

Pesaran, M. H. (2003). Aggregation of linear dynamic models: an application to life-cycle consumption models under habit formation. *Economic Modelling*, 20(2):383–415.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. H. and Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics*, 161(2):182–202.

Polemis, M. L. (2012). Competition and price asymmetries in the greek oil sector: an empirical analysis on gasoline market. *Empirical Economics*, 43(2):789–817.

Remer, M. (2015). An empirical investigation of the determinants of asymmetric pricing. *International Journal of Industrial Organization*, 42:46–56.

Stoker, T. M. (1993). Empirical approaches to the problem of aggregation over individuals. *Journal of Economic Literature*, 31(4):1827–1874.

Westerlund, J. and Edgerton, D. L. (2007). A panel bootstrap cointegration test. *Economics Letters*, 97(3):185–190.

Appendix A

[Please insert Table A1 about here]

Table 1: Results from simulation

	$\theta_i \sim [-1, 0]$	$\theta_i \sim [-0.8, -0.2]$
Mean of uniformly distributed coefficients	-0.500	-0.500
Days to close 95% of the gap	5.99	5.99
OLS estimator on aggregate retail prices		
Number of lags (M)	1	1
Number of lags (J)	0	0
$\frac{\sum_{r=1}^{R} \theta_{OLS_r}}{R}$	-0.390^{***}	-0.476^{***}
	(0.007)	(0.005)
Days to close 95% of the gap	7.67	6.29

We assume that $\beta_{im} = 0$ for $m \geq 1$, $\delta_{ij} = 0.5$ if $j = 0$ and 0 otherwise, $\alpha_i = 0$, and $\phi_i = 1$. θ_{OLS_r} is the estimated coefficient from the panel r. The total number of simulated panels is $R = 1,000$. Standard errors are reported in parentheses, and *** indicates significance at the 1% level.
Source: Own elaboration.

Table 2: Results based on real response of retail fuel prices (to refined fuel prices)

	Full sample (all brands)	Sub-sample (Repsol)
MG estimator on individual retail prices, Eq. (1)		
Number of lags (M)	7	7
Number of lags (J)	8	8
θ_{MG}	-0.113^{***} (0.001)	-0.100^{***} (0.002)
Days to close 95% of the gap	26.51	29.96
Obs. (N x T)	531,000	203,400
Individuals (N)	590	226
Hausman test	36,454.00 [0.000]	1,230.58 [0.000]
MG-CCE estimator on individual retail prices, Eq. (1)		
Number of lags (M)	7	7
Number of lags (J)	8	8
θ_{MG-CCE}	-0.122^{***} (0.005)	-0.102^{***} (0.005)
Days to close 95% of the gap	24.56	29.37
Obs. (N x T)	531,000	203,400
Individuals (N)	590	226
Hausman test	92,408.94 [0.000]	5,177.49 [0.000]
OLS estimator on aggregate retail prices, Eq. (2)		
Number of lags (M)	7	7
Number of lags (J)	8	8
θ_{OLS}	-0.052^{***} (0.012)	-0.082^{***} (0.019)
Days to close 95% of the gap	57.61	36.53
Obs. (T)	900	900

Akaike Information Criteria were employed in order to select the optimal lag length. In addition to the variables listed in Eq. (1) and Eq. (2), all regressions include in the long-term relationship a time trend and a set of daily dummies that control for the possible effects of seasonal changes in demand for each day of the week. The MG-CCE incorporates, as an explanatory variable, the lagged cross-sectional average of the retail prices fixed by all the competitors operating at each moment of time within a 200-metre radius. Standard errors are reported in parentheses and p-values are presented in brackets. We use *** to indicate significance at the 1% level.
Source: Own elaboration.

Table 3: Results based on real response of retail fuel prices (to crude oil prices)

	Full sample (all brands)	Sub-sample (Repsol)
MG estimator on individual retail prices, Eq. (1)		
Number of lags (M)	8	7
Number of lags (J)	8	7
θ_{MG}	-0.068^{***} (0.001)	-0.067^{***} (0.001)
Days to close 95% of the gap	44.05	44.71
Obs. (N x T)	$531,000$	$203,400$
Individuals (N)	590	226
Hausman test	$11,321.51$ [0.000]	882.71 [0.000]
MG-CCE estimator on individual retail prices, Eq. (1)		
Number of lags (M)	8	7
Number of lags (J)	8	7
θ_{MG-CCE}	-0.080^{***} (0.005)	-0.068^{***} (0.005)
Days to close 95% of the gap	$37,45$	44.05
Obs. (N x T)	$531,000$	$203,400$
Individuals (N)	590	226
Hausman test	$82,304.23$ [0.000]	681.49 [0.000]
OLS estimator on aggregate retail prices, Eq. (2)		
Number of lags (M)	8	7
Number of lags (J)	8	8
θ_{OLS}	-0.031^{***} (0.008)	-0.047^{***} (0.013)
Days to close 95% of the gap	96.64	63.74
Obs. (T)	900	900

Akaike Information Criteria were employed in order to select the optimal lag length. In addition to the variables listed in Eq. (1) and Eq. (2), all regressions include in the long-term relationship a time trend and a set of daily dummies that control for the possible effects of seasonal changes in demand for each day of the week. The MG-CCE incorporates, as an explanatory variable, the lagged cross-sectional average of the retail prices fixed by all the competitors operating at each moment of time within a 200-metre radius. Standard errors are reported in parentheses and p-values are presented in brackets. We use *** to indicate significance at the 1% level.
Source: Own elaboration.

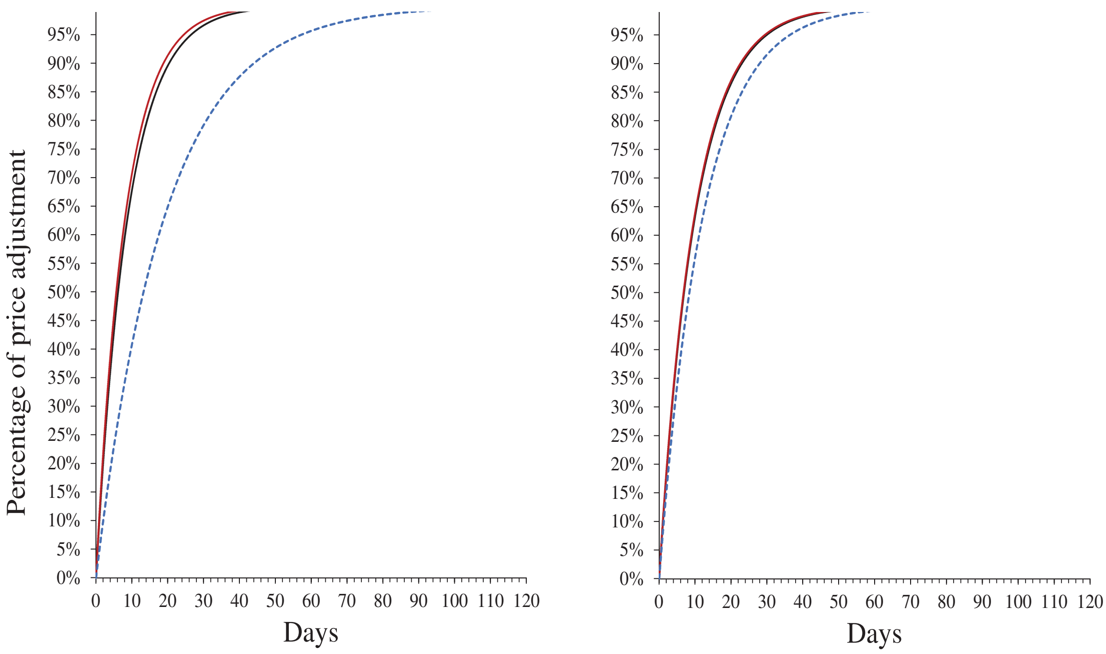
Table A1: Unit root and cointegration tests on real prices

	Full sample (all brands)		Sub-sample (Repsol)	
	Levels	First diff.	Levels	First diff.
Unit root test				
Breitung-Das test on individual retail prices	0.006	-29.013***	-0.041	-17.739***
ADF test on aggregate retail prices	-1.748	-5.681***	-1.565	-5.935***
Cointegration between wholesale-refined fuel prices and retail prices				
Westerlund tests on individual retail prices				
G_τ	-6.239***		-5.897***	
G_α	-102.855***		-105.218***	
P_τ	-162.114***		-162.756***	
P_α	-110.663***		-117.814***	
ADF test on aggregate retail prices	-3.682***		-3.653***	
Cointegration between crude oil prices and retail prices				
Westerlund tests on individual retail prices				
G_τ	-4.789***		-4.275***	
G_α	-60.185***		-51.696***	
P_τ	-114.794***		-72.537***	
P_α	-58.582***		-59.491***	
ADF test on aggregate retail prices	-3.941***		-3.531***	

We use *** to denote the rejection of the null hypotheses (unit root and non-cointegration) at the 1% significance level. The lag order for the Augmented Dickey-Fuller, Breitung-Das and Westerlund tests are obtained by using the Akaike Information Criteria. Critical values for the Augmented Dickey-Fuller and Breitung-Das tests are based on MacKinnon (1996) and Breitung and Das (2005), respectively. The Westerlund tests employ bootstrapped robust critical values based on 500 replications, where the Bartlett kernel bandwidth is set according to the $4(T/100)^{2/9} \approx 7$ rule.
Source: Own elaboration.

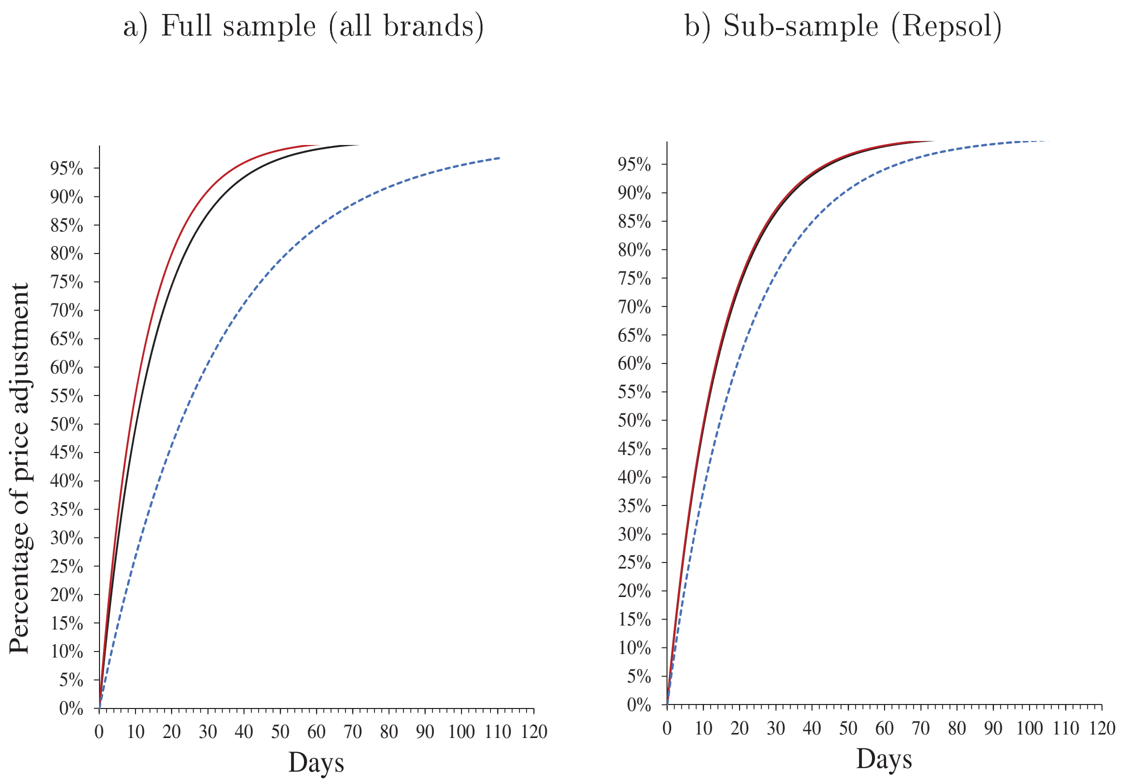
Figure 1. Retail price adjustment after a change in wholesale-refined fuel price

a) Full sample (all brands) b) Sub-sample (Repsol)



The black and red continuous lines represent the results from Eq. (1) by using the MG and MG-CCE procedures, respectively. The dashed blue line denotes those obtained from Eq. (2) by using the OLS procedure on aggregated data.
Source: Own elaboration based on the estimated results of Table 2.

Figure 2. Retail price adjustment after a change in crude oil price



The black and red continuous lines represent the results from Eq. (1) by using the MG and MG-CCE, respectively. The dashed blue line denotes those obtained from Eq. (2) by using the OLS procedure on aggregated data.
Source: Own elaboration based on the estimated results of Table 3.