GASOLINE AND DIESEL CONSUMPTION FOR ROAD TRANSPORT IN SPAIN: A DYNAMIC PANEL DATA APPROACH

Rosa M. González-Marrero[1]
Departamento de Análisis Económico, IUDR (ULL)
Cátedra Economía de las Infraestructuras (FEDEA-ABERTIS)
Rosa M. Lorenzo-Alegría
Departamento de Análisis Económico, IUDR (ULL)
Gustavo A. Marrero[2]
Departamento de Análisis Económico (ULL), FEDEA, ICAE and CAERP
Programa de Investigación de Energía y Cambio Climático (FOCUS-ABENGOA-FEDEA)

ABSTRACT:

This paper studies the factors explaining per capita fuel consumption for road transport in Spain, distinguishing between diesel and gasoline consumption. The main contribution of the paper is to specify an empirical fuel consumption model in a dynamic panel data (DPD) framework, and then to properly apply estimation techniques, based on the system GMM procedure of Arellano and Bover (1995) and Blundell and Bond (1998). We find that alternative and more traditional estimation procedures (pooling-OLS, the within group, first difference GMM), which are shown to generate bias estimates, produce important differences that may even change policy recommendations. We find that most explanatory variables are significant in explaining the evolution of gasoline consumption in Spain, while diesel consumption is found to be independent of most of these factors. This finding highlights the necessity to estimate a different model for gasoline than for diesel. The intensive dieselization process that has taken place in Spain over the last decade has resulted in diesel consumption being exposed to factors - i.e., regulatory - which are not of a strictly economic nature.

JEL: R41, O13, O56

KEY WORDS: Fuel consumption, road transport, fixed-effect model.

---

[1] Correspondence to: Rosa M. González-Marrero, Departamento Análisis Económico, Universidad de La Laguna, Campus de Guajara, La Laguna, 38071 Santa Cruz de Tenerife; Fax: 922 317204, Tfno: 922317113, rmglzmar@ull.es
[2] Gustavo A. Marrero acknowledges financial support of FOCUS-ABENGOA through Programa de Investigación de Energía y Cambio Climático FEDEA-ABENGOA and the Ministry of Education and Science of Spain through the SEJ2006-14354 program.

1. Introduction

The transport sector in Spain represented almost 40% of final energy consumption in 2006, one of the highest within EU27 countries. This sector generates about 25% of total CO2 emissions, with road transport contributing the most to said emissions. Having such a large road transport sector poses a serious roadblock for Spain to reaching the goals set by the Kyoto Protocol and the recently proposed 20/20/20 plan. Hence, energy transport policies and environmental policies should be intimately related. In order to implement effective policy measures to abate fuel consumption, it is crucial to properly characterize the relationships between energy consumption and the factors that explain this consumption.

This paper studies the factors explaining per capita fuel consumption for road transport in Spain. In addition to the breakdown by Spanish regions, which allows us to use a panel data approach, another relevant aspect of this research is the distinction made between gasoline and diesel consumption. Polemis (2006), Zervas (2006) and Labandeira and López-Nicolás (2002), among others, have already emphasized the importance of making this distinction, which is especially relevant for the case of Spain, since it is probably, along with France, one of the countries in which the *dieselization* process has been the most significant in the last decade.[3] This process has led to a very uneven distribution in the consumption of gasoline and diesel, implying that the conclusions derived from an analysis of overall energy consumption could be misleading. Hence, we estimate a *gasoline model* and a *diesel model* and compare their results, which is a first contribution of the paper.

There is an extensive literature that characterizes the consumption of energy in road transport. For example, see Schipper et al. (1992) and Johansson and Schipper (1997) for

[3] For instance, the diesel to gasoline consumption ratio in road transport in Spain, which was 1.71 in 1998, had risen to 3.78 by 2006.

OECD countries, Mazzarino (2000) for Italy, Kwon (2005) for the United Kingdom, Polemis (2006) for Greece, Tapio et al. (2007) for the EU-15, Zervas (2006) for Ireland, Alvaes and Bueno (2003) for Brazil, Samimi (1995) for Australia, Nicol (2003) for Canada and the United States, Ramanathan (1999) for India, Koshal et al. (2007) for Japan, and Belhaj (2002) for Morocco, among others. And yet, despite the increasing demand for energy consumption in Spain and its active *dieselization* process, few empirical studies have been conducted on the Spanish case. Some exceptions are the works of Labeaga and López-Nicolás (1997) and Labandeira and López-Nicolás (2002), which estimate the demand for automotive fuel, though they mainly focus on analyzing the effects of taxes on overall consumption.

The main contribution of this paper is to write the empirical fuel consumption model in a dynamic panel data (DPD) framework, and then properly apply estimation techniques. A DPD approach is shown to have important advantages with respect to a traditional static or time series analysis. First of all, energy consumption is dynamic by nature [Johansson and Schipper (1997)]. This might be due, for example, to the persistence of fuel usage habits, requiring that a dynamic model be specified. Secondly, a DPD approach allows for working with the entire data panel and for specifying unobserved or omitted fixed effects to estimate the relevant parameters in the model [Hsiao (1986)].

With regards to the estimation procedure, we use the one-step system GMM estimator proposed by Arellano and Bover (1995) and developed by Blundell and Bond (1998), which allows for endogeneity, measurement error and omitted variables problems.[4] In order to discuss the importance of considering this estimation method, we follow Blundell et al. (2000), and compare the system GMM estimates with respect to alternative, more traditional methods –the within groups, pooling-OLS, the first difference GMM of Arellano and Bond

[4] In the growth literature, Forbes (2000), Shioji (2001), Levine et al. (2000) and Bond et al. (2001), among others, use the one-step system GMM estimator that we consider in this paper.

(1991). There are few exceptions in the energy literature that seriously consider the weakness of traditional methods in estimating DPD models. For example, Halkos (2003), Gang (2004) and Metcalf (2008) address the endogeneity problem and use the first difference GMM estimator, but this method does not consider the *weak* instruments problem of this procedure when time series are persistent [Blundell and Bond (1998)], which is the case for aggregate emissions and energy time series. Huang et al. (2008), which revisits the causal relationship between energy consumption and GDP, is an exception that properly addresses both the endogeneity and the weak instruments problems and considers a system GMM approach.

The endogenous variable is per capita gasoline and diesel consumption . In addition to the dynamic fuel consumption term (the lagged level of fuel), we consider as explanatory variables the real prices of gasoline and diesel, the per capita GDP, the fleet of gasoline and diesel per capita (the motorization rate by fuel type) and the total fleet divided by the total kilometers of road as a proxy for the saturation of the road network.

We find that most explanatory variables are significant – and with the appropriate signs - in explaining the evolution of gasoline consumption in Spain, while diesel consumption is found to be independent of most of these factors. The poorer adjustment of the diesel model could be due to the intensive dieselization process that has taken place in Spain over the last decade, which has resulted in diesel consumption being exposed to factors - i.e., regulatory - which are not of a strictly economic nature. Moreover, we find that the alternative estimation procedures produce important differences that may even change policy recommendations, thus highlighting the need to carry out further research in this field.

This paper is structured as follows. Section 2 presents the DPD fuel consumption model and briefly comments on the system GMM estimation approach. Section 3 describes the data used in the analysis. Section 4 estimates the gasoline and diesel model and discusses the main results. Finally, Section 5 presents the main conclusions.

2. A Dynamic Panel Data Model for Fuel Consumption

In this section we present a DPD model for fuel consumption.[5] We consider two alternative models, one for gasoline consumption (*the gasoline model*) and another for diesel consumption (*the diesel model*). In its general specification, fuel consumption is explained by lagged levels of fuel and additional explanatory variables,

$$y_{it} = \alpha_i + \beta y_{it-1} + \lambda' X_{it} + \varepsilon_{i,t} \tag{1}$$

where y_{it} is the log of fuel consumption of region i at time t; X_{it} is a set of K variables, which are dependent on each region and time and can affect fuel consumption, such as fuel prices, regional GDP, the road network, etc.. This set of explanatory variables is discussed in more details in the next section; α_i considers those fixed factors which are time-invariant and inherent to each region, and they are not observed or not included in the model, such as geographical, social or local policy regional aspects or initial energy efficiency use;[6] finally, ε_{it} encompasses effects of a random nature which are not considered in the model, and it is assumed to have a standard error component structure:

A1: $E[\varepsilon_{it}] = 0;\ E[\alpha_i \varepsilon_{it}] = 0; E[\varepsilon_{it}\varepsilon_{is}] = 0,\ i = 1,..., N; t = 1,..., N \ and \ s \neq t$ **.**

We also consider a common assumption in DPD models [Arellano and Bond (1991)], which is that y_{1i} is predetermined,

A2: $E[y_{i1}\varepsilon_{it}] = 0,\ for \ i = 1,..., N \ and \ t = 2,..., T$

[5] Given the close relationship between energy and income, the dynamic specification is similar to that used in the convergence-growth literature [Barro and Sala-i-Martin (1995), among many others]. Álvarez et al. (2005) and Marrero (2009) have adapted this dynamic approach to economic-pollution models.

[6] Fixed effects, such as differences in the initial energy efficiency use, would be omitted in a standard OLS pool regression, resulting in bias estimates (i.e., the β estimates is upward bias.) See Anderson and Hsiao (1982) and Hsiao (1986) for more details about this point.

The endogenous variable y_{it} is measured in Kilo-tones per habitant: it is gasoline consumption divided by population – *GASO* - for *the gasoline model,* while it is diesel consumption divided by population - *DISL* - for *the diesel model.* The dynamic term of fuel consumption (y_{it-1}), denoted by $GASO_{t-1}$ and $DISL_{t-1}$, control for convergence across states. Indeed, the interpretation of equation (1) depends on the level of β. A β smaller than one is consistent with conditional convergence, which means that regions relatively close to their steady-state per capita fuel consumption levels will experience a slowdown in their consumption growth. In this case, α_i and all explanatory variables affect to the steady-state the fuel consumption of region *i* is converging to. On the other hand, if β is greater than one, there is no convergence effect and α_i and all regressors would measure differences in steady-state energy consumption growth rates. Estimated β will be lower than one in all cases, hence we will focus on the conditional convergence interpretation.

Traditional procedures for estimating a DPD model like (1) (i.e., fixed or random effects methods or pooling-OLS) are known to be unsuitable [Anderson and Hsiao (1982); Hsiao (1986)]. Holtz-Eakin et al. (1988) and Arellano and Bond (1991) propose an alternative approach, where first differences in the regression equation are taken to remove unobserved time-invariant country specific effects and then particular moment conditions for lagged variables are exploited to find a set of instruments and construct a GMM-based estimator. Their GMM approach (GMM-DIF) allows us to handle endogeneity, measurement errors and omitted variables problems. However, the GMM-DIF approach shows important bias problems in small sample when variables are persistent, which is the case of economic and energy macroeconomic variables. Under these circumstances, the instruments used in the GMM-DIF estimator have proven to be *weak* and the first difference estimator is poorly behaved. Arellano and Bover (1995) and Blundell and Bond (1998) propose an alternative

GMM procedure which might overcome the *weak* instruments problem. This procedure estimates a system of equations in both first-differences and levels, where the instruments in the level equations are lagged first differences of the variables. In this paper we use the one-step system GMM estimator. In contrast to the two-step version, the one-step GMM estimator has standard errors that are asymptotically robust to heteroskedasticity and have been found to be more reliable for finite sample inference.[7]

In addition to the dynamic term, the other explanatory variables assumed in (1) are among those traditionally considered as indicators for characterizing the behavior of the road transportation sector [Eltony (1993), Bentzen (1994), Kirby et al. (2000), Alves and Bueno (2003), Polemis (2006)]**.** All variables are measured in logs and are taken for each Spanish region between 1998 and 2006. These variables are the following: *GDPpc* is the per capita Gross Domestic Product,[8] *P.GASO* and *P.DISL* are the average real prices of gasoline and diesel, respectively;[9] *FLEET.GASO* and *FLEET.DISL* are the existing fleet of gasoline and gasoil vehicles per capita (motorization rate), respectively, at the end of the period; *SAT* is the total number of vehicles divided by total kilometers of road at the end of the period, which is a proxy of the saturation of the road network.

Taking specification (1) and the variables defined above as our starting point, we then estimate the following models, the first for gasoline and the second for diesel:

$$\begin{aligned}&\ln(GASO_{i,t}) = \alpha_i + \beta \ln(GASO_{i,t-1}) + \lambda_1 \ln(GDPpc_{i,t}) + \lambda_2 \ln(SAT_{i,t}) + \\ &\lambda_3 \ln(P.GASO_{i,t}) + \lambda_4 \ln(P.DISL_{i,t}) + \lambda_5 \ln(FLEET.GASO_{i,t}) + \lambda_6 \ln(FLEET.DISL_{i,t}) + \varepsilon_{i,t}\,,\end{aligned} \tag{2}$$

[7] See Blundell and Bond (1998), Blundell et al. (2000) and Bond (2002), among others. See the technical appendix for more details about this point.

[8] Regional GDP is built on the basis of regional physical, economic indicators, such as retail sales, industrial production index, car sales, overnight stay of tourists, consumption of cement, etc, which are highly related to the domestic level of activity in each Region. De la Fuente (2002), among others, has used the Spanish regional GDP dataset to study the source of convergence in Spain.

[9] Petrol prices are retail prices and deflated by each regional CPI .For more details about how petrol prices are determined in Spain, see Perdiguero (2006).

$$\ln(DISL_{i,t}) = \alpha_i + \beta \ln(DISL_{i,t-1}) + \lambda_1 \ln(GDPpc_{i,t}) + \lambda_2 \ln(SAT_{i,t}) + \\ \lambda_3 \ln(P.GASO_{i,t}) + \lambda_4 \ln(P.DISL_{i,t}) + \lambda_5 \ln(FLEET.GASO_{i,t}) + \lambda_6 \ln(FLEET.DISL_{i,t}) + \varepsilon_{i,t}, \quad (3)$$

3. Fuel consumption and transport data for Spain: 1998-2006

In this section we briefly describe the data used in our analysis. Table 1 shows annual growth rates between 1998 and 2006 for each variable at a regional level.

TABLE 1: LISTS THE AVERAGE ANNUAL VARIATION RATES FOR ALL VARIABLES FOR THE PERIOD 1998-2006

REGION	consumption gasoline pc	consumption diesel pc	gasoline real price	diesel real price	per capita GDP	gasoline fleet pc	diesel fleet pc	total fleet/ road network
Andalusia	-3.44	6.19	2.76	5.09	2.97	-1.48	11.47	4.84
Aragón	-3.90	6.25	2.63	4.78	2.79	-1.77	9.64	2.14
Asturias	-3.44	5.51	3.01	5.18	2.99	-1.36	8.23	2.39
Cantabria	-3.78	6.67	2.93	5.03	3.07	-1.30	9.46	4.28
Castilla y León	-3.33	5.91	2.72	4.90	3.26	-1.11	9.78	2.94
Castilla La Mancha	-4.76	5.87	3.03	5.23	2.14	-1.71	10.43	5.40
Catalonia	-5.16	4.11	2.71	4.85	2.03	-2.53	8.33	2.85
Valencia	-4.33	5.13	3.01	5.05	1.64	-2.33	7.85	3.72
Extremadura	-2.89	6.96	3.03	5.06	3.75	-0.82	12.01	3.94
Galicia	-3.24	3.74	3.00	4.99	2.99	-1.32	7.94	3.18
Madrid	-6.72	6.42	2.91	3.86	2.13	-4.15	10.63	3.19
Murcia	-4.62	6.58	2.29	4.54	1.98	-2.42	8.94	4.82
Navarre	-3.33	6.04	2.64	4.60	2.65	-2.62	7.38	2.74
Basque Country	-3.99	6.28	2.49	4.59	3.26	-2.01	7.62	2.86
La Rioja	-4.18	4.07	2.36	4.43	1.72	-2.66	8.08	3.26
SPAIN	**-4.07**	**5.72**	**2.77**	**4.81**	**2.63**	**-1.97**	**9.19**	**3.50**

Per capita gasoline consumption fell between 1998 and 2006 in all Spanish regions at an average rate of 4.1%, while per capita diesel consumption increased for each region at a higher average rate of 5.7%. During this period, gasoline and diesel real prices increased by 2.8% and 4.8% in Spain, respectively. Although significant differences were noted in their growth rates in the time dimension, differences within regions are very small. The large time-volatility of fuel prices resulted from Spain's enormous (nearly 100%) dependence on foreign oil, on the important fluctuations in the euro/dollar exchange rate, and on changes in fuel taxes and their repercussions on the final price.[10]

The data on per capita GDP growth rate showed an annual increase of 2.6% nationally, varying between 1.6% in Valencia and 3.8% in Extremadura. Overall, per-capita GDP growth showed a notable regularity among the different regions, as evidenced by the

[10] For example, in 2001 the price in dollars of a barrel of Brent crude fell 14%, while in 2004 and 2005 it rose by 33% and 42%, respectively.

generalized slowdown between 2001 and 2003-2004 and the subsequent recovery until 2006 for most of the regions.

The motorization rates by type of fuel (*gasoline fleet pc* and *diesel fleet pc*) show a trend similar to those of their associated fuel consumption series. For most regions, it decreased for gasoline vehicles (2.0% in Spain), while it increased for diesel (9.2% in Spain). As is the case with fuel consumption, the dieselization process has changed the vehicle fleet composition in Spain: the ratio of diesel to gasoline vehicles, which was 43% in 1998, had risen to 105% by 2006. However, the intensity of this substitution process varied greatly depending on the Region. Thus, Madrid is the region that experienced the most significant substitution process from gasoline to diesel vehicles, with a decline in the motorization rate for gasoline vehicles of 4.2% and an increase for diesel of 10.6%. Other regions show a different pattern. For example, in Extremadura, the motorization rate for gasoline vehicles remained relatively stable, while that of diesel experienced the largest increase, within all Spanish regions, of 12% per year.

As for the number of vehicles versus road network kilometers (the saturation level of road), it increased in all regions, showing the highest average annual growth rate in Castilla La Mancha (5.4%) and the lowest in Aragon (2.1%).

4. Gasoline and Diesel Model Results

The goal of this section is threefold. First, we emphasize the importance of considering an appropriate quantitative approach when estimating a dynamic fuel consumption model; second, we point out the differences between the gasoline and the diesel model results; and third, we show the main determinants of per capita gasoline and diesel consumption in Spain.

The estimation procedure employs the one-step GMM estimator proposed by Arellano and Bover (1995) and developed by Blundell and Bond (1998), with heteroskedasticity-consistent asymptotic standard errors.[11] We first find evidence supporting the good properties of the system GMM estimates. Following Blundell et al. (2000), we compare the results of alternative methods for the gasoline and the diesel model: the OLS pooling estimates (OLS-POOL), the Within Group estimates (WG), the first-difference GMM approach (GMM-DIF) of Arellano and Bond (1991) and the system GMM method (GMM-SYS). Tables 2 and 3 show the results for the gasoline and diesel model, respectively, for all these alternative methods. The p-value of the t significance test associated with each parameter is shown. We also show standard specification tests for each model. First, notice that the Haussman test rejects the null hypothesis of random effects at any standard level of significance. For any GMM-based estimates, we show the *m1* and the *m2* tests and conclude that moment conditions underlying GMM estimates seem to be robustly supported.[12]

[11] For a given cross-sectional sample size, the use of too many instruments in models with endogenous regressors may result in seriously biased estimates (Álvarez and Arellano, 2003; Arellano and Bond, 1998). Hence, even when computing speed is not an issue, these authors recommend not using the entire series history as instruments. We use instruments up to t-3. Including more lags does not change results significantly.

[12] The most frequently used tests to validate the assumptions underlying GMM methods are the *m1*, *m2* and Sargan tests. The *m1* and *m2* tests are based on the standardized average residuals autocovariance, which are asymptotically N(0,1) distributed under the null hypothesis of no autocorrelation. The Sargan test, in contrast, is distributed chi-squared with degrees of freedom equal to the number of moment restrictions minus the number of parameters, estimated under the null hypothesis that moment conditions are valid. However, the Sargan test is less meaningful since it requires that the error terms be independently and identically distributed, which is not expected in our case. Hence, we will consider primarily the *m1* and *m2* tests.

Table 2: Estimates of the gasoline DPD model

[redacted]

Note: '*WG*' is Within Groups estimation, *OLS-POOL* is OLS applied to the entire pool of data. For GMM estimates, we take as instruments the lagged levels of *y* and the endogenous regressors dated t-2 and earlier and the pre-determined regressors dates t-1 and earlier. We use the lagged difference of *y* and all regressors dated *t-1* as additional instruments in the system GMM estimation. For the DIF-GMM and SYS-GMM, we report their one-step estimations. The null of the Haussman test is the existence of random effects. The null of the m1 and m2 test is the absence of first- and second-order serial correlation between regressors and residuals, respectively. Number of regressors: 8; number of cross sections: 15 (all Spanish regions except Ceuta and Melilla, Balears and Canary islands); number of time periods: 9 (1998-2006); number of time periods adjusted for GMM-DIF and GMM-SYS: 6 (2001-2006).

Table 3: Estimates of the Diesel DPD model

[redacted]

Note: See Note on Table 2.

For each model, we compare OLS-POOL, WG, GMM-DIF and GMM-SYS estimates. Based on the results shown in Tables 2 and 3, OLS-POOL seems to give an upward-biased estimate of the β coefficient (0.853 for the gasoline and 0.998 for the diesel model), while WG appears to give a downward-biased estimate of this coefficient (0.387 for the gasoline and 0.495 for the diesel model). Using GMM-DIF, the β coefficient is barely lower than the WG estimates, suggesting the possibility of important finite sample bias due to the weak instruments problem [Blundell and Bond (1998)]. This comparison also highlights how the estimated coefficients of the remained regressors, which are our main interest, differ among the alternative procedures. Hence, using a method resulting in bias estimates (the OLS-POOL, WG or the GMM-DIF) might lead to misleading conclusions. For example, the coefficients associated with the network saturation variable in the gasoline and diesel model are not significant under the WG and GMM-DIF estimates, while it is significant under the GMM-SYS procedure; for the diesel model, the per capita GDP variable is significant under the WG, while it is not under the GMM-SYS and OLS-POOL; the magnitude of the estimated price-elasticities are smaller under the GMM-SYS than under the WG estimates in the diesel model.

In summary, this comparison suggests that the WG estimates are severely biased, that there exists a problem with weak instruments and hence that the GMM-DIF is biased similarly to WG, and that the GMM-SYS approach is a convenient way to overcome the weak instruments problem. This conclusion is an important contribution of the paper, and not always properly considered in the related literature. We will focus our attention on the one-step GMM-SYS estimates from now on.

Comparing the results of the gasoline and diesel models, we find important differences in the magnitude and significance of the variables. The results allow us to conclude that the estimates of the coefficients associated with the explanatory variables are less significant in the diesel model than in the gasoline model. The poorer adjustment of the diesel model could be due to the dieselization process that has taken place in Spain over the last decade, which has resulted in diesel consumption being exposed to other factors which are not of a strictly economic nature.

The parameters estimated for the $DISL_{t-1}$ variable (specific to the diesel model) and the $GASO_{t-1}$ (specific to the gasoline model) are positive and less than one at the 1% level of significance. The estimate is 0.867 for the diesel model and 0.558 for the gasoline model. Hence, the evidence for conditional convergence is significant in both cases, though it is greater for the gasoline case. The estimates indicate that the rate of convergence for the per capita fuel consumption ratio, conditioned to its long-term equilibrium levels in each region, is about 13% for diesel consumption and about 44% for gasoline.

A common result in both models is that per capita GDP is not significant in explaining per capita fuel consumption. From Section 3, we showed how per capita fuel consumption and GDP experienced important increases from 1998 to 2006. However, it seems that the former evolved independently of per capita GDP at a regional level. Behind this result lies the fact that regions with different per capita GDP levels shared similar fuel consumption patterns. For example, this is the case of the Basque Country, with a large per capita GDP, Castilla and Leon, with intermediate per capita GDP, and Extremadura, with one of the smallest per capita GDPs in Spain; however, they experienced a similar increase in per capita fuel consumption (between 3.6% and 3.9%).[13]

From a conceptual and methodological standpoint, the above result is important. Since per capita GDP can be interpreted as a proxy for personal income, our result suggests a negligible fuel-income elasticity, at least at current income and fuel consumption levels. This finding is in contradiction with most results in the related literature. For example, Dahl and Sterner (1991) showed that short-term income elasticity on gasoline demand varied between 0.30 and 0.52 in the different studies they considered.[14] However, notice that the GDP elasticity under the WG estimate was significant and about 0.44 for diesel, which is indeed consistent with the Dahl-Sterner range; this, however, is the result of a bias estimate. With this example, we are not claiming that Dahl-Sterner estimates are wrong. In fact, differences between their estimates and ours may only be due to differences in the sample used. We are just stressing the importance of considering an appropriate estimation approach for handling fuel consumption models, because, otherwise, results can lead to misleading conclusions.

Regarding the real price of fuel, the GMM-SYS procedure estimate for its elasticity is negative and significant for the case of gasoline, though its magnitude is well below one (-0.29). This result confirms the evidence that the elasticity of the demand price for gasoline is

[13] Nevertheless, we are aware that the short time dimension of the data can also influence this result.

[14] More recently, studies such as that by Koshal et al. (2007) gave values of 0.29. For a detailed review, see Graham and Glaister (2002) and Goodwin et al. (2004).

low in the short term, as verified by, among others, Kayser (2000) with data for the United States. Results at the international level place the price-elasticity in the -0.2 and -0.3 range [Dahl and Sterner (1991)]. Contrary to what occurred for the per capita GDP variable, our estimations are now consistent with results in the literature. This result indicates that fuel demand is highly inelastic, at least at current price levels. Moreover, it supports a result commonly discussed in the literature: fuel taxes are convenient for increasing fiscal revenues, but they are not effective enough to reduce fuel consumption [Kirby et al. (2002)].

In addition, the real price of diesel is significant in explaining short-term changes in per capita gasoline consumption, although its parameter is small (0.21). This last result is also consistent with the transport literature due to the strictness that exists in substituting types of vehicles in the short term [Polemis (2006)]. We should emphasize that the recent intensive switch from gasoline to diesel vehicles has been basically due to regulatory reasons (dieselization) rather than to a change in the price of the alternative fuel. For the case of the diesel consumption model, neither its own price nor the price of gasoline is significant. This result is one of the most important differences between the gasoline and the diesel model estimates. This finding also highlights the need to consider different models for gasoline and diesel consumption.

The remaining variables are specific to the road transport sector and include relevant aspects that can affect fuel consumption. The per capita diesel and gasoline fleet variables show the motorization rate for each type of fuel vehicle. For the gasoline model, the coefficient of the per capita gasoline fleet variable is highly positive and significant (0.64), while the coefficient of the per capita diesel fleet is negative but much smaller in magnitude (-0.08). Regarding these variables, results for the diesel model are controversial.

We find that the per capita diesel fleet variable is non-significant in explaining diesel consumption. A feasible explanation of this result is that the lower price of diesel fuel and its higher efficiency have led to a more intensive use of existing diesel vehicles. Hence, an important part of the increase in diesel consumption is not directly related with the stock of diesel vehicles, i.e., the *rebound effect* [Schipper et al (2002)]. On the other hand, results show that the per capita gasoline fleet variable is positive and significant, although its coefficient is much smaller than that associated with the gasoline model. [15]

The coefficients of the measure of the degree of saturation of the road network (the ratio between total fleet and road network) are negative and significant in both models. Moreover, their estimates are similar: -0.048 for diesel and -0.059 for gasoline. The fact that estimated coefficients are similar in both models is a clear indication that road saturation affects both diesel and gasoline vehicles in a similar way. This result suggests that a reduction in road congestion promotes mobility, which may induce an increment in per capita fuel consumption.[16] Needless to say, the way to reduce energy consumption for transportation is not to saturate the road network artificially; the solution should involve increasing the usage of more sustainable modes of transportation (public and non-motorized transport), improving and enlarging the road network while at the same time restricting and penalizing private mobility, and promoting a more efficient use of cars through other measures such as car-pooling. However, this important issue needs to be studied in more detail within a different framework, which goes beyond the scope of this paper.

5. Final Remarks

This paper has estimated a DPD model for fuel consumption in order to characterize the main determinants of gasoline and diesel consumption for road transport in Spain. This information is necessary in order to implement a proper transport policy and to forecast fuel

[15] Further investigation on this important topic would constitute a prominent extension of this paper.

[16] As noted by Goodwin (1996), improving the infrastructure has an induced effect on the demand for transport. Moreover, Cervero and Hansen (2002) provided empirical evidence of the existence of a direct relationship between investing in roads and the demand for transport, namely that an expansion of infrastructure generates demand for transport, which in turn induces the creation of infrastructure.

consumption. We used panel data for fifteen Spanish Regions from 1998 to 2006. As explanatory variables, we considered real GDP and fuel prices, which are the most commonly used in the related literature, as well as other relevant and novel variables, such as the motorization rate and the congestion of the road network.

In this paper we use the one-step system GMM approach of Arellano and Bover (1995) and Blundell and Bond (1998), which has been shown to solve many of the problems that arise in traditional panel data procedures. When compared with the system GMM results, we found that traditional panel data estimation procedures [the within-group estimates, OLS-pooling or the first difference GMM approach of Arellano and Bond (1991)] might exhibit significantly biased estimates, which might even change policy recommendations. Our results emphasize the need to revisit DPD fuel consumption results obtained with traditional procedures, and show the relevance of considering a suitable estimation method.

For the sample used, we found that most explanatory variables are significant in explaining the evolution of gasoline consumption in Spain, while diesel consumption was found to be independent of most of these factors. The intensive dieselization process that has taken place in Spain over the last decade has may resulted in diesel consumption being exposed to factors - i.e., regulatory - which are not of a strictly economic nature. These conspicuous differences between the results for the gasoline and diesel models imply that the conclusions derived from an analysis of overall fuel consumption could be misleading. Moreover, it emphasizes the necessity to investigate the determinants of diesel consumption using a different model than that used for gasoline consumption, which is a prominent extension of this paper.

Our estimates confirm that the elasticity of the demand price for fuel consumption is low – even negligible for diesel - in the short term, which supports the view that the policy of taxing fuel has little effect on reducing fuel consumption. Our results are also consistent with the evidence of small cross price elasticities for gasoline and diesel consumption. This result indicates that gasoline and diesel are imperfect substitutes in the short run. Per capita GDP is not significant in explaining per capita gasoline and diesel consumption. This result suggests negligible fuel-income elasticity, which is in contradiction with most results in the related literature. We show that this fact might be due to the usual practice of traditional and unsuitable estimation methods in DPD fuel consumption models.

Finally, an important finding of this work is the negative and significant relationship between the degree of saturation of the road network and both types of per capita fuel consumption. This result shows that reducing road network saturation – i.e., by increasing the road network -, could promote mobility and a higher transport demand ("*induced travel demand*"), which can favour higher levels of fuel consumption.

The implication of our results to transport policy is clear. In order to reduce per capita fuel consumption, without negatively affecting growth and welfare, the simultaneous application of different measures must be implemented. These should involve improving and enlarging the road network, increasing the usage of more sustainable modes of transportation (public and non-motorized transport), promoting a more efficient use of cars, and, at the same time, restricting and penalizing the use of private vehicles.

REFERENCES

- Alvaes D., Bueno R. Short-run, long-run and cross elasticities of gasoline demand in Brazil. Energy Economics 2003, 25, 191-199.
- Álvarez J., Arellano M. The time series and cross-section asymptotics of dynamic panel data estimators. Econometrica 2003, 71(4), 1121-1159.
- Álvarez F., Marrero G. A., Puch L. A. Air pollution and the Macroeconomy across European Countries. FEDEA 2005 Working Paper 2005-26.
- Anderson T.W., Hsiao C. Formulation and estimation of dynamic models using panel data. Journal of Econometrics 1982, 18, 47-82.
- Arellano M., Bover O. Another look at the instrumental-variable estimation of error-components models. Journal of Econometrics 1995, 68, pp. 29–52.
- Arellano M., Bond S. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies 1991, 58, 277-297.
- Barro R. J., Sala-i-Martin X. Economic Growth, Advanced Series in Economics, McGraw-Hill, 1995.
- Belhaj M. Vehicle and fuel demand in Morocco. Energy Policy 2002, 30, 1163-1171.
- Bentzen J. An Empirical analysis of gasoline demand in Denmark using cointegration techniques. Energy Economics 1994, 16, 139-143.
- Blundell R.W., Bond S.R., Windmeijer F. Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator. In: Baltagi, B., Editor, 2000. Nonstationary Panels, Panel Cointegration, and Dynamic PanelsAdvances in Econometrics Vol. 15, JAI Press, Elsevier Science, Amsterdam, pp. 53–91
- Blundell R., Bond S. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics 1998, 87, 115-143.
- Bond S. Dynamic panel data models: a guide to microdata methods and practice. Portuguese Economic Journal 2002, 1, 141-162.
- Bond S., Hoeffler A., Temple J. GMM Estimation of Empirical Growth Models. Economics Papers 2001-W21, Economics Group, Nuffield College, University of Oxford.
- Cervero R., Hansen M. Induced travel demand and induced road investment. A Simultaneous equation analysis. Journal of Transport Economic and Policy 2002, 36(3), 469-490.
- Dahl C., Sterner T. Analyzing gasoline demand elasticities: a survey. Energy Economics 1991, 13, 203-310.
- De la Fuente A. On the source of convergence: a close look at the spanish regions. European Economic Review 2002, 46 (3), 569-599.
- Dorand H. E., Schmidt P. GMM estimator with improved finite sample properties using principal components of the weighting matrix, with an application to the dynamic panel data models. Journal of Econometrics 2006, 133(1), 387-409.

- Eltony M. The demand for gasoline in the GCC: an aapplication of pooling and testing procedures. Energy Economics 1993, 18 (3), 203-209.
- Forbes K. A reassessment of the relationship between inequality and growth. American Economic Review 2000, 90(4), 869-887.
- Gang L. Estimating Energy Demand Elasticities for OECD Countries. A Dynamic Panel Data Approach. Research Department of Statistics Norway 2004, Dicussion Paper 373.
- Goodwin P. Empirical evidence on induced traffic. Transportation 1996, 23(1), 35-54.
- Goodwin P., Dargay J., Hanley M. Elasticities of road traffic and fuel consumption with respect to price and income: A review. Transport Review 2004, 24 (3), 275-292.
- Grahan D. J., Glaister S. The demand for automobile fuel: a survey of elasticities. Journal of Transport Economics and Policy 2002, 36, 1-26.
- Halkos G.E. Environmental Kuznets Curve for sulphur: evidence using GMM estimation and random coefficient panel data models. Environment and Development Economics 2003, 8, 581-601.
- Holtz-Eakin D., Newey W., Rosen H. S. Estimating vector autoregressions with panel data. Econometrica 1998, 56(6), 1371-1395.
- Hsiao C. Analysis of panel data. Econometric Society monographs 11, Cambridge. University Press, 1986.
- Huang B.N., Hwang M. J., Yang C. W. Causal relationship between energy consumption and GDP growth revisited: A dynamic panel data approach. Ecological Economics 2008, 67 (1), 41-54.
- Johansson O., Schipper L. Measuring long-run automobile fuel demand: separate estimations of vehicle stock, mean fuel intensity, and mean annual driving distance. Journal of Transport Economic and Policy 1996, 31(3), 277-292.
- Kayser H. A. Gasoline demand and car choice estimating gasoline demand using household information. Energy Economics 2000, 22, 331-348.
- Kirby H. R., Hutton B., McQuaid R. W., Raeside R., Zhang X. Modelling the effects of transport policy levers on fuel efficiency and national fuel consumption, Transportation Research Part D 2000, 5, 265-282.
- Koshal R. K., Manjulika K., Yuko Y., Sasuke M., Keizo Y. Demand for gasoline in Japan. International Journal of Transport Economics 2007, 34, 351-367.
- Kwon T-H. The determinants of the changes in car fuel efficiency in Great Britain (1978-2000). Energy Policy 2005, 2, 261-275.
- Labeaga J. M., López-Nicolás A. A study of petrol consumption using Spanish panel data. Applied Economics 1997, 29, 795-802.
- Labandeira X., López-Nicolás A. La imposición de los carburantes de automoción en España; algunas observaciones teóricas y empíricas. Hacienda Pública Española 2002, 160-1, 177-210.
- Levine R., Loayza N., Beck T. Financial intermediation and growth: causality and causes. Journal of Monetary Economics 2000, 46, 31-77.

- Marrero G.A. Greenhouse gases emissions, growth and the energy mix in Europe: a dynamic panel data approach. Working Paper 2009-16, FEDEA.
- Mazzarino M. The economics of the greenhouse effect: evaluating the climate change impact due to the transport sector in Italy. Energy Policy 2000, 28, 957-966.
- Metcalf G. An empirical analysis of energy intensity and its determinants at the state level. The Energy Journal 2008, 29, 1-26.
- Nicol C. J. Elasticities of demand for gasoline in Canada and the United States. Energy Economics 2003, 25, 201-214.
- Perdiguero G. J. Dinámica de precios en el mercado español de gasolina: un equilibrio de colusión tácita, Documento de Trabajo de Funcas 2006, nº 253.
- Polemis M. L. Empirical assessment of the determinants of road energy demand in Greece. Energy Economics 2006, 28, 385-403.
- Ramanathan R. Short and long-run elasticities of gasoline demand in India: an empirical analysis using cointegration techniques. Energy Economics 1999, 21, 321-330.
- Samimi R. Road transport energy demand in Australia. Energy Economics 1995, 17, 329-339.
- Schipper L., Steiner R., Duerr P., An F., Strom S. Energy use in passenger transport in OCDE countries: Changes since 1970. Transportation 1992, 19, 25-42.
- Schipper, L., C. Marie-Lilliu y L.Fulton, Diesels in Europe. Analysis of Characteristics, usage patterns, energy savings and CO_2 emission implications, Journal of Transport Economics and Policy 2002, 36(2), 305-340.
- Shioji E. Public capital and economic growth: a convergence approach. Journal of Economic Growth 2001, 6, 205-227.
- Tapio P., Banister D., Luukkanen J., Vehmas J., Willamo R. Energy and transport in comparison: Immaterialisation, dematerialisation and decarbonisation in the EU15 between 1970 and 2000. Energy Policy 2007, 35, 433-451.
- Windmeijer F. A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 2005, 126, 1, 25-51.
- Zervas E. CO_2 benefit from the increasing percentage of diesel passenger cars. Case of Ireland. Energy Policy 2006, 34, 2848-2857.

Appendix: System GMM Estimation of DPD Models

The fixed effect treatment leads to the well known within group estimator (WG) [Hsiao (1986)], which has been applied to multiple frameworks. However, the within transformation in a panel dynamic model implies a correlation of order 1/T between the lagged dependent term y_{it-1} and the error ε_{it}, which leads to biased estimates [Anderson and Hsiao (1981); Hsiao (1986)]. In addition, a fuel consumption equation such as (1) suffers from endogeneity and, maybe, from measurement errors problems. For instance, gasoline and diesel prices are jointly determined with gasoline and diesel consumption. The WG method neither properly handle these problems.

Holtz-Eakin et al. (1988) and Arellano and Bond (1991), among others, point out these problems and propose a GMM-based estimation approach. The current response of these authors is to first difference the model equation, remove the fixed effect term and then use the following orthogonally conditions, which, under assumptions A1 and A2 (see Section 2 of the paper), are valid for the first difference model:

$$E[y_{it-s}\Delta\varepsilon_{it}] = 0,\ t = 3,..., T \ and \ \ 2 \le s \le t-1,\ for\ i = 1,..., N, \qquad \textbf{(3)}$$

Regressors in the gasoline and diesel models are either endogenous (prices, GDP and registrations) or pre-determined (the road network and the vehicle fleets ratios).[17] Assuming a similar condition to A2 but for the regressors in X,

A3: $E[x_{i1}\varepsilon_{it}] = 0,\ for\ \ i = 1,..., N\ and\ t = 2,..., T$ **,**

we have additional 0.5(T-1)(T-2) moment conditions,

[17] In the case of exogenous regressors, additional moment conditions are available. See Arellano and Bond (1991) for more detail about this point.

$$E\left[x_{it-s}\Delta\varepsilon_{it}\right]=0,\ t=3,\ldots,T \text{ and } 2\le s\le t-1, \text{ for } i=1,\ldots,N, \tag{4.a}$$

for each endogenous regressor, and another (T-1)(T-2) moment conditions,

$$E\left[x_{it-s}\Delta\varepsilon_{it}\right]=0,\ t=3,\ldots,T \text{ and } 1\le s\le t-1, \text{ for } i=1,\ldots,N, \tag{4.b}$$

for each pre-determined regressor.

For the case of K=1 and endogenous regressor,[18] we have a total of *Nd=(T-1)(T-2)* moment conditions. Conditions in (3) and (4.a) can be written more compactly as

$$E\left[Z_i^{'}\Delta\varepsilon_i\right]=0,\ i=1,\ldots,N, \tag{5}$$

where Z_{iDIF} is a *(T-2)xNd* matrix, given by

$$Z_{iDIF}=\begin{pmatrix} y_{i1}\,x_{i1} & 0 & \ldots & & 0 \\ 0 & y_{i1}\,y_{i2}\,x_{i1}\,x_{i2} & \ldots & & \ldots \\ . & \ldots & \ldots & & \ldots \\ . & \ldots & \ldots & y_{i1}\ldots y_{iT-3}\,x_{i1}\ldots x_{iT-3} & 0 \\ 0 & 0 & \ldots & 0 & y_{i1}\ldots y_{iT-2}\,x_{i1}\ldots x_{iT-2} \end{pmatrix} \tag{6}$$

These are the moment conditions exploited by the standard first-difference GMM estimator (GMM-DIF).

However, the GMM-DIF estimator has been found to have large finite sample bias and poor precision when the set of instruments is *weak* [Blundell and Bond (1998).], which is the case of our fuel consumption model. To deal with this problem, Arellano and Bover (1995) and Blundell and Bond (1998) assume additional conditions to **A1**, **A2** and **A3**,

[18] For ease of exposition, we restrict notation to the case of only one endogenous regressor (i.e., K=1). The extension of the general case is straighforward.

A4: $E[\eta_i \Delta y_{i2}] = 0,\ i = 1,\ldots, N$,

A5: $E[\eta_i \Delta x_{i2}] = 0,\ i = 1,\ldots, N$

which allows the use of other 2·(T-2) moment conditions for a model in levels,

$$E[u_{it} \Delta y_{it-1}] = 0,\ t = 3,\ldots, T\ , \qquad \textbf{(7)}$$

$$E[u_{it} \Delta x_{it-1}] = 0,\ t = 3,\ldots, T \qquad \textbf{(8)}$$

which stay informative even for high persistent time series. Their proposal consists in a stacked system of all *(T-2)* equations in first differences and all *(T-2)* equations in levels for *t=3,4,…,T*, and combine restrictions (3), (4), (7) and (8) to form a linear system GMM estimator (GMM-SYS) based on the following instrument matrices:

$$Z_i = \begin{pmatrix} Z_{iDIF} & 0 \\ 0 & Z_{iSYS} \end{pmatrix}, \qquad \textbf{(9)}$$

with Z_{iDIF} given by (6) and Z_{iSYS} by

$$Z_{iSYS} = \begin{pmatrix} \Delta y_{i2}\ \Delta x_{i2} & 0 & \ldots & & 0 \\ 0 & \Delta\ y_{i3}\ \Delta x_{i3} & \ldots & & \ldots \\ . & \ldots & \ldots & & \ldots \\ . & \ldots & \ldots & \Delta y_{iT-2}\ \Delta x_{iT-2} & 0 \\ 0 & 0 & \ldots & 0 & \Delta y_{iT-1}\ \Delta x_{iT-1} \end{pmatrix};$$

Monte Carlo analysis has shown that using GMM-SYS greatly reduces the finite sample bias and improves the precision of the estimator in presence of weak instruments.[19] The linear GMM estimator is given by $\left(\tilde{X}'ZH_N Z'\tilde{X}\right)^{-1}\left(\tilde{X}'ZH_N Z'\tilde{Y}\right)$, where, for the GMM-DIF,

$$\tilde{X}=\begin{bmatrix}\Delta X_1\\ \ldots\\ \Delta X_N\end{bmatrix};\tilde{Y}=\begin{bmatrix}\Delta Y_1\\ \ldots\\ \Delta Y_N\end{bmatrix};Z=\begin{bmatrix}Z_{1,DIF}\\ \ldots\\ Z_{N,DIF}\end{bmatrix},$$

while for the GMM-SYS case,

$$\tilde{X}=\begin{bmatrix}\Delta X_1\\ \ldots\\ \Delta X_N\\ X_1\\ \ldots\\ X_N\end{bmatrix};\tilde{Y}=\begin{bmatrix}\Delta Y_1\\ \ldots\\ \Delta Y_N\\ Y_1\\ \ldots\\ Y_N\end{bmatrix};Z=\begin{bmatrix}Z_1\\ \ldots\\ Z_N\end{bmatrix}.$$

For each GMM-FIF and GMM-SYS case, two different choices of H_N result in two different GMM estimators. The one-step estimator sets

$$H_{N,GMM\,1}=\left(\frac{1}{N}\sum_{i=1}^{N}Z_i^{'}HZ_i\right)^{-1},$$

where the H matrix is a (T-2) square matrix with 2's on the main diagonal, -1 on the first off-diagonals and zeros elsewhere. The two-step GMM estimator uses

$$H_{N,GMM\,2}=\left(\frac{1}{N}\sum_{i=1}^{N}Z_i^{'}\Delta\hat{u}_i\Delta\hat{u}_i^{'}Z_i\right)^{-1},$$

[19] Indeed, Blundell and Bond (1998) and Bond et al. (2001) shows that an optimal combination of differenced and level equations allow us to calculate a GMM estimator using the full set of linear moment conditions implied by assumptions A1-A5.

where estimated residuals are from a consistent one-step estimator (i.e., the one-step), which is an asymptotically efficient GMM estimator.

Under spherical disturbances, GMM1 and GMM2 are equivalent in the first-difference model. Otherwise, GMM2 is more efficient. However, Monte Carlo studies have shown that the efficiency gains of the two-step estimator are generally small. It also has the problem of converging to its asymptotic distribution relatively slowly. Hence, in finite samples, its variance-covariance matrix can be seriously biased. Moreover, for the case where the total number of instruments is large relative to the cross-section dimension of the panel, there may be computational problems in calculating the two-step estimates and serious estimation errors may arise [Arellano and Bond (1998); Doran and Schmidt (2006)]. With this in mind, most empirical works with a relatively small cross-section dimension report results of the one-step GMM estimator, which has standard errors that are asymptotically robust to heteroskedasticity and have been found to be more reliable for finite sample inference [Blundell and Bond (1998), Blundell et al. (2000); Windmeijer (2005); Bond (2002)]. This is the strategy considered in this paper.

There exist some tests to validate the assumptions underlying GMM methods. The standard approach for testing the validity of the moment conditions in GMM estimation is the Sargan test of overidentifying restrictions and the m2 second-order serial correlation test [Arellano and Bond (1991)]. Under the null hypothesis that moment conditions are valid, the Sargan test is distributed chi-squared with degrees of freedom equal to the number of moment restrictions minus the number of parameters estimated. The m2 test is normally distributed under the null hypothesis of the absence of second-order serial correlation between regressors and residuals.