Diesel Fuel Consumption in Europe

By
John Christopoulos

Submitted to
Central European University
Department of Economics

In partial fulfilment of the requirements for the degree of Master of Arts

Supervisor: Professor Attila Rátfai

Budapest, Hungary
2008

ABSTRACT

In this study a panel data set consisting of 14 European countries during the period of 1990-2004 is used to conduct an analysis of the factors affecting diesel consumption. This study controls for the fact that the users that are likely to purchase a diesel powered car are the most intensive users. This peculiarity of the diesel fuel market is used in order to produce more realistic elasticity estimates. Other factors such as fuel efficiency and reduced utilization due to more cars per people are used to control for the factors that affect short-run diesel fuel consumption. The results show that after these factors are controlled for the range in which the elasticities fall becomes much smaller and that the elasticities become similar to those found in gasoline demand studies.

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my family for their support throughout my life. Secondly, I would like to thank MOL Plc. for their valuable support. Last but not least, I would like to thank my supervisor, Prof. Ratfai, as well as Prof. Korosi and Prof. Kezdi for their help.

TABLE OF CONTENTS

INTRODUCTION

Over the past years the car market of Europe has been changing considerably, the time when gasoline fuel was used for cars and diesel fuel was used mainly for commercial vehicles such as buses and trucks has long passed. In most European countries there has been an increase in the share of diesel cars[1] in the market and therefore an increase in diesel fuel consumption[2]. This new reality is important because although traditionally it was possible to differentiate fuel demand of households and businesses by the fuel they were using now it is not possible, this leads to problems when estimating demand for fuel because households and businesses have different income and price elasticities. Diesel fuel is being used more and more by the household sector and therefore this new trend plays an important role when estimating its demand. Therefore an analysis of fuel demand has to take into consideration these recent trends in order to appropriately measure these elasticities.

Recent increases in the net prices of crude oil have led to huge increases in the price of automotive fuels that consumers are faced with. The increased demand of diesel fuel that stems mostly from the rising popularity of diesel fueled passenger cars has also contributed to the rise of the price of diesel fuel. Diesel and gasoline fuel go through a different process to be refined and companies cannot adjust their short to medium term production capacities. This implies that companies should adjust their production capacities based on projections of future demand. Therefore it is crucial to correctly identify the factors that affect the demand of diesel fuel.

The increase in the share of diesel cars in most European countries is mostly due to differences in the lower net price of diesel fuel compared to that of gasoline

[1] Cars are defined as vehicles intended for private use by households.
[2] View Graphs 1 & 2 in Appendix on diesel cars and diesel fuel consumption respectively.

fuel, the efficiency of diesel compared to gasoline engines, the lower taxation usually of diesel fuel and the fact that the cost gap of diesel compared to gasoline engines has been closing due to technological advances. On the other hand diesel cars are still more expensive to acquire than gasoline powered cars which leads to only high usage individuals purchasing a diesel and not gasoline powered car, since only these high-usage individuals will reap the benefits of the fuel savings over the extra cost of acquiring a diesel powered vehicle as it is pointed out by Verboven (2002) and Pock (2007). Verboven (2002) claims that there is quality based price discrimination in the car market which effectively segments the high mileage from the low mileage consumers. In Verboven (2002) a comparison of diesel and gasoline annual mileages to those predicted by a demand model, where the only source of heterogeneity between consumers is annual mileage, show that it is a safe assumption to say that annual mileage is the main factor affecting the purchasing decision between a gasoline and diesel powered car. Therefore any econometric analysis estimating the demand of diesel fuel should consider that the users that purchase a diesel and not a gasoline powered car are usually high mileage consumers since otherwise it would be irrational for them to do so as was shown in Verboven (2002).

Literature on diesel consumption is very limited especially if we consider diesel fuel as a household good and not only as a fuel used by businesses for vehicles other than passenger cars such as trucks and buses. In Pock (2007) it is claimed that all previous estimations of gasoline demand per total passenger cars suffer from a bias because they ignore the increasing share of diesel cars in the market. The car market therefore has become a quiet interesting case in the past years and in turn this has affected all previous studies on gasoline and diesel demand. Epsey (1998) conducted a meta-analysis of income and price elasticities when estimating gasoline consumption by reviewing articles published between 1966 and 1997 and reviewed the explanatory variables to which estimations are sensitive to. Epsey's (1998) paper reviews the various techniques that have been used over the years to model gasoline consumption. Theoretically the model under which diesel fuel is estimated should be similar to models used to estimate gasoline consumption since the goods are consumed in an identical way. The only additional fact to be considered in the formulation of a diesel demand model is that the consumers most probable to purchase a diesel instead of a gasoline powered car are high mileage consumers as it was considered in Pock (2007).

In this analysis a diesel consumption model is formulated under a similar basis as the ones that are usually formulated to explain gasoline consumption while some changes are applied in order to make the model better suitable to explain the peculiarities of the diesel fuel market. The main contribution of this analysis is that it takes into consideration the fact that the users purchasing diesel powered vehicles are more intensive users and uses this peculiarity to estimate price and income elasticities that are believed to be more realistic.

The countries included in this analysis are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Spain, Sweden and the United Kingdom. The elasticities obtained in this analysis are similar to those obtained when estimating gasoline fuel consumption after the peculiarities of the diesel consumption market were controlled for. In the two main econometric techniques used it is shown that the range of elasticities in the GMM/IV specification is smaller and that the elasticities of the GMM/IV specification are higher (in absolute value) than those obtained without controlling for the peculiarities of the diesel fuel consumption market. The elasticity estimates are also compared to those usually found in analyses of gasoline fuel consumption for reasons of comparison.

In the next chapter the basic model underlying this study will be presented. In Chapter 2, the dataset used in the analysis will be described. In Chapter 3 the different specifications of the model to be estimated will be shown and the theory underlying each specification will be discussed. In Chapter 4 the output of the estimations are presented and the econometric techniques are discussed. In Chapter 5 the results of this analysis are discussed and explained. In the "Conclusion" the analysis is shortly summarized.

CHAPTER 1: MODEL

In this thesis it is assumed that there are three factors affecting the demand of diesel and those are the number of diesel powered cars, their fuel efficiency and their utilization as was done in Sweeney (1978), Baltagi and Griffin (1983,1997) and Pock (2007).

Obviously the amount of diesel powered cars in the market affect the amount of diesel demanded. The number of diesel cars is affected by the acquisition costs and operating costs. Diesel passenger cars are considered to be of a higher utility because they are more durable and more fuel efficient than gasoline powered passenger cars, however they are more expensive and as was mentioned in the introduction this leads to high usage individuals purchasing diesel instead of gasoline powered cars. Most countries included in this analysis show an increasing amount of diesel powered cars[3]. Epsey (1998) observed in his meta-analysis of gasoline demand models that the specification of consumption either as an aggregate measure, as per capita or as per vehicle did not affect the elasticities of obtained in studies. Therefore diesel fuel consumption per diesel powered vehicle will be estimated in this thesis and the number of diesel powered cars will not be included as an explanatory variable.

The fuel efficiency of cars differs by model and changes with technological progress throughout the years. Therefore I will use as a good proxy for the fuel efficiency of the cars in use in a given country a measure of how old the passenger car stock is. In order for the average age of the car stock to be a good proxy for the fuel efficiency of the stock of cars it must be the case that consumers do not change their preferences over time, however the increasing popularity of Sport Utility Vehicles

[3] View Graph 1 in the Appendix.

(SUV's) is proof of changes in consumer preferences[4]. This measure can approximately tell us the fuel efficiency of a country's stock of cars compared to that of another country. If a country has a younger stock of cars it should be more fuel efficient if preferences towards vehicles remained the same. However because of limited access to data in this study I cannot control for changes in consumer preferences and therefore this measure may prove to be a bad proxy for fuel efficiency.

The third factor affecting consumption, vehicle utilization, tells us how much cars are being used. Consumers can adjust their driving habits due to many reasons such as changes in the price of fuel or in their income. Vehicle utilization may also vary as the number of vehicles per people increase[5].

Summarizing the above three factors, the demand for diesel fuel changes when the stock of diesel powered vehicles change, either through a change in the number of vehicles or through the replacement of older vehicles with newer more or less efficient ones, or when the utilization of the stock of these vehicles changes. Since I will be estimating diesel consumption per diesel powered passenger car I am interested in changes in vehicle utilization and fuel efficiency of the stock of diesel powered passenger cars.

In this study I assume that vehicle utilization is captured by changes in income, the price of fuel and the number of vehicles per driving population. The number of cars per driving population is used to capture reduced utilization due to more cars per drivers, a measure of driving population was used instead of a measure of total population in order to capture for demographic effects as was done in Schmalensee and Stoker (1999) and Pock (2007). The average age of cars will be used as a proxy

[4] For more information on this view the discussion in Pock (2007) and Zachariadis (2006).
[5] Pock (2007) reports that a family with two cars does not drive twice as much as a family owning one car.

for fuel efficiency even though possible changes in consumer preferences can invalidate its ability to be a good proxy for fuel efficiency. Gross Domestic Product will be used to control for general economic activity since we do not have a measure for the number of commercial vehicles.

This study assumes that in each period the number of users purchasing a diesel and not a gasoline powered car depends on the consumer's annual mileage. In Pock (2007) a model is formulated explaining the "break-even" annual mileage however I will not go into such detail. I simply assume that given the acquisition and operating costs of diesel compared to gasoline cars in each period it is only rational for consumers of a given annual mileage to purchase a diesel powered car. When I observe an increase in the relative number of registrations of diesel to gasoline vehicles I interpret it as a decrease in the "break-even" annual mileage for which it is rational to purchase a diesel powered vehicle. A variable measuring the ratio of new diesel passenger car registrations to new gasoline passenger cars therefore will be used to capture the "break-even" mileage in each period and therefore also how intensive the users entering the diesel demand market are.

In this study diesel fuel consumption per diesel powered passenger cars will be estimated using a partial adjustment model in which it is assumed that diesel consumption per diesel powered passenger cars takes some time to adjust to its desired level. Since consumers cannot adjust their stock of vehicles immediately and therefore also their consumption per vehicle the following relationship exists between actual and desired levels of consumption per vehicle, where the desired level is represented by an asterisk (*) and θ is the speed of adjustment:

$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1} = \theta\cdot\left[\log\left(\frac{dieselconsumption}{dieselcars}\right)_t^* - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1}\right] + v_t \quad (1)$$

If we assume that the desired level of diesel fuel consumption per diesel powered car is defined as:

$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t^* = \alpha + \beta_i \cdot x_{i,t} + \eta_t \tag{2}$$

then we can derive the below equation[6]

$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t = \alpha \cdot \theta + (1-\theta) \cdot \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1} + \theta \cdot \beta_i \cdot \log(x_{i,t}) + \varepsilon_t \tag{3}$$

This partial adjustment specification will be the basis of the model that will estimate diesel fuel consumption per diesel powered passenger car. The control variables $(x_{i,t})$ that were mentioned above will be described in more detail in the next chapter and will be added to control for changes in vehicle utilization and the fuel efficiency of the car stock in different model specifications that will be discussed in the "Estimation" chapter.

[6] View Appendix for a step by step derivation.

CHAPTER 2: DATA

In this analysis a panel data set is used that is compiled from two sources, the Statistical Office of the European Communities (EUROSTAT) and the European Automobile Manufacturer's Association (ACEA). The data consists of annual observations of 14 countries (EU-15 excluding Portugal) for the years 1991-2004. The variables that will be used are the total number of diesel powered passenger cars, the total number of passenger cars, consumption of diesel fuel, an index of gross domestic product, income, total population over 19 years old, price of diesel fuel, a measure of average age and a ratio of new registrations of diesel powered passenger cars to gasoline powered passenger cars.

The variables for the number of diesel *(CARSD)* and total number *(CARS)* of passenger cars are in thousands. As it was mentioned in the introduction most countries show an increasing number of diesel powered cars. Diesel consumption *(CONSD)* is an aggregate measure reflecting what is consumed by road transport and is measured in thousands of tones[7]. The variable *(GDP)* is gross domestic product at market prices indexed to 1995. The income variable *(INCOME)* is defined as net national income at market prices and is measured in euros per inhabitant. The total population over the age of 19 *(POP)* will be used as a measure of total drivers. This variable will be used to divide the number of passenger cars in order to capture reduced utilization due to more cars per people since when there are more cars per people vehicle utilization is lower as it was discussed previously. It is used instead of total population since otherwise income elasticities are overstated as was shown in Schmalensee and Stoker (1999. The price of diesel fuel *(PD)* is the price of the automotive fuel at the pump which is the price consumers are directly faced with, the

[7] A more detailed definition of diesel and gasoline fuel is given in the Appendix.

prices were half yearly prices of 1000 litres of the product and an average was constructed to have annual averages of the prices. The average age variable *(AGE)* was created manually since only the number of cars in different age groups was available from EUROSTAT. The ratio of new registrations of diesel powered passenger cars to those of gasoline powered cars *(REG)* will be used to capture the relative fuel consuming intensity of users entering the diesel demand market in any given period, this variable was constructed manually from the percentages of annual diesel registrations and total volume of new registrations that were obtained from the statistics of ACEA.

All of the variables were obtained from EUROSTAT except for the *REG* variable which, as it was mentioned, was constructed with data obtained from ACEA. These variables will be used to estimate the basic partial adjustment model for diesel fuel consumption per diesel powered passenger car as well as its modifications which will be discussed in the next chapter.

CHAPTER 3: MODEL SPECIFICATIONS

The basic model that will be used to capture diesel consumption per diesel powered passenger cars is a partial adjustment model that controls for income, price of fuel and the total number of passenger cars per driving age population, as they were defined in the “Data” section, these control variables are assumed to capture the factors affecting changes in vehicle utilization. It is assumed that consumers cannot adjust their vehicle stock immediately and therefore also their consumption per vehicle. Thus we use a partial adjustment model that also includes three variables that affect car utilization (*INCOME*, *PD*, and *CARS/POP)*. This basic model is given below:

$$\log\left(\frac{CONSD}{CARSD}\right)_{t,i} = \beta_0 + \beta_1 \cdot \log\left(\frac{CONSD}{CARSD}\right)_{t-1,i} + \beta_2 \cdot \log\left(INCOME\right)_{t,i} + \beta_3 \cdot \log\left(PD\right)_{t,i} + \beta_4 \cdot \log\left(\frac{CARS}{POP}\right)_{t,i} + \varepsilon_{t,i} \qquad (4)$$

where the variables are as defined in the data part of the study, t is a time subscript and i is the country subscript.

Another model to be estimated is adding average age of cars to the basic model. This addition is made in order to attempt to capture the technological advances throughout time that improve fuel efficiency. However, as it was already mentioned, in order for the *AGE* variable to be a good proxy for fuel efficiency consumer preferences towards vehicles should remain unchanged throughout time. This assumption for the *AGE* variable does not hold in reality since preferences are very likely to change. Therefore it is possible that the *AGE* variable is a bad proxy for fuel efficiency even though the fuel efficiency of vehicles has been improving throughout the years with advances in technology. The initial model that was formulated above is shown below in a different specification that attempts to control for the change in technology throughout the years in addition to the factors that affect utilization:

$$\log\left(\frac{CONSD}{CARSD}\right)_{t,i} = \beta_0 + \beta_1 \cdot \log\left(\frac{CONSD}{CARSD}\right)_{t-1,i} + \beta_2 \cdot \log(INCOME)_{t,i} + \beta_3 \cdot \log(PD)_{t,i} + \beta_4 \cdot \log\left(\frac{CARS}{POP}\right)_{t,i} + \beta_5 \cdot \log(AGE)_{t,i} + \varepsilon_{t,i} \quad (5)$$

the variables are as defined in the "Data" section, t is a time subscript and i is the country subscript.

The last model specification to be estimated is adding the variable *REG* to the basic model. This addition aims to control for the intensity of the users purchasing diesel cars in each period. As the *REG* variable increases an influx of less-intensive users is expected to enter the diesel fueled car owning population since it is assumed that this increase implies a decrease in the "break-even" annual mileage at which it is rational to purchase a diesel fueled car. The basic model therefore becomes:

$$\log\left(\frac{CONSD}{CARSD}\right)_{t,i} = \beta_0 + \beta_1 \cdot \log\left(\frac{CONSD}{CARSD}\right)_{t-1,i} + \beta_2 \cdot \log(INCOME)_{t,i} + \beta_3 \cdot \log(PD)_{t,i} + \beta_4 \cdot \log\left(\frac{CARS}{POP}\right)_{t,i} + \beta_5 \cdot \log(REG)_{t,i} + \varepsilon_{t,i} \quad (6)$$

the variables are again as defined above, t is a time subscript and i is the country subscript.

All three of the above models will also be estimated with the addition of the control variable *GDP*. The *GDP* variable attempts to capture overall economic activity and thus control for diesel fuel consumption by the commercial sector, that is by vehicles such as trucks and buses.

The equations will be estimated using fixed effects (cross-section fixed effects, period fixed effects and cross-section and period fixed effects), random effects (cross-section random effects and period random effects) and also as a pooled sample assuming exogeneity of the regressors. However, due to the presence of a lagged dependent variable which is not independent of the error term these estimates are inconsistent. A standard solution for this problem is to use Instrumental Variables (IV) with the Generalized Method of Moments (GMM) technique. The IV's that will be used are the exogenous variables as well as a dynamic IV of the lagged dependent variable.

CHAPTER 4: ESTIMATIONS

The basic model of this analysis (equation 4) was estimated with the following estimation techniques: cross-section fixed effects, period fixed effects, cross-section and period fixed effects, cross-section random effects, period random effects and as a pooled sample. In these estimations the White diagonal standard errors were used, which "are robust to observation specific heteroskedasticity in the disturbances, but not to correlation between residuals for different observations" (Eviews Users Guide 5, 2004, pp. 854). Generalized Least Squares cross-section weights were used wherever there were signs of serial correlation and the weighted estimations are shown in the tables. All of the residuals obtained from the estimations showed a close to normal distribution. The residuals show no signs of autocorrelation, more specifically regressions of the residuals on themselves for a number of lags showed insignificant coefficients[8]. Afterwards, due to the presence of the lagged dependent variable as was mentioned in the previous section, the model was then estimated using the GMM/IV technique. In the GMM/IV technique the exogenous regressors were used as Instrumental Variables and the dependent lagged variable in a dynamic form with lags from -2 to -5. The "orthogonal deviations" option was used to remove the individual fixed-effects[9], the White period ("AB n-step") GMM weights were used and the White period standard errors method was used. By conducting Sargan's test of over-identifying restrictions we could not reject the null of the over-identifying restrictions being valid[10]. The results obtained from the estimation of the basic model are as expected in terms of the signs of the coefficients with almost all of the different

[8] View Section A.5 in the Appendix for test results.
[9] For more on this method see Arellano and Bover (1995).
[10] View Section A.5 in the Appendix for test results.

techniques that were used[11]. In Table 1 the estimation output of the random cross section effects and the GMM/IV estimations are shown.

Table 1

Method:	***Cross-Section Random Effects***	***GMM/IV Orthogonal Deviations***
Constant	-1.03**	-
	(0.69)	-
Log(*CONSD/CARSD*)-1	0.98	0.87
	(0.01)	*(0.05)*
Log(*INCOME*)	0.06	0.10
	(0.03)	*(0.02)*
Log(*PD*)	-0.07*	-0.20
	(0.03)	*(0.04)*
Log(*CARS/POP*)	-0.10**	0.06**
	(0.06)	*(0.04)*
Observations	152	138
R-Squared	0.993	0.814
D-W stat	1.80	-
J-Stat	-	11.65
Instrument Rank	-	14

*Insignificant at 5% level,**Insignificant at 10% level

In Table 1 above we see the coefficients of the variables and their standard errors are given in parentheses. The coefficient on *INCOME* is positive which implies that an increase in consumers' income positively affects the consumption of diesel fuel per diesel powered car. The coefficient on the price of diesel fuel is negative as expected, implying that an increase in the price of the fuel will result in lower consumption per diesel car. The coefficient on *CARS/POP* which is the number of total cars per driving age population, as it was defined, is positive and insignificant in the GMM/IV technique. However, in the Random Cross-Section Effects technique it is positive meaning that more cars per population will result in less diesel consumption per diesel cars. From the coefficient on our lagged dependent variable we can get the speed of adjustment. For example since our coefficient is 0.98 in the

[11] The estimation output from all the estimation techniques can be found in the Appendix.

Random Cross-Section Effects specification we can subtract this number from 1 to get $\theta = 1 - \beta_1$. In this particular case $\theta = 0.02$ which can be used to divide our elasticities in order to get the long run elasticities since the ones reported are short run due to the presence of the lagged dependent variable.

The next model that was estimated (equation 5) included the average age of the car stock variable as a proxy for changes in fuel efficiency due to technological advances. As it was mentioned before, in order for the average age of the car stock to be a good proxy consumers preferences towards vehicles must remain unchanged over the years. In Table 2 the results from the cross-section and period fixed effects estimation and the GMM/IV technique are shown[12].

Table 2

Method:	***Cross-Section & Period Fixed Effects***	***GMM/IV Orthogonal Deviations***
Constant	-4.49**	-
	(4.12)	-
Log(*CONSD/CARSD*)-1	0.66	0.82
	(0.07)	*(0.13)*
Log(*INCOME*)	0.45	0.30
	(0.16)	*(0.13)*
Log(*PD*)	-0.50	-0.31
	(0.12)	*(0.07)*
Log(*CARS/POP*)	-0.53**	-0.37**
	(0.39)	*(0.32)*
Log(*AGE*)	-0.04**	0.06**
	(0.22)	*(0.21)*
Observations	113	99
R-Squared	0.995	0.682
D-W stat	2.11	-
J-Stat	-	8.99
Instrument Rank	-	13

*Insignificant at 5% level,**Insignificant at 10% level

[12] The output from all the estimation techniques for this model specification can be found in the Appendix.

In this model the GMM/IV technique is the same as was described above and by conducting Sargan's Test we could not reject that the over-identifying restrictions are valid. As we can see there are insignificant coefficients on the variables of *AGE* and *CARS/POP*. The signs of the elasticities are as were expected in the output shown in Table 2 (except for the coefficient on *AGE*). The results from the different econometric techniques are shown in the Appendix and as we see the *AGE* variable is always very close to zero. Therefore, it turns out that the average age of the car stock is not a good proxy for fuel efficiency of cars since we would have expected it to be negative always and significant, possibly because consumer preferences change.

The final model specification (equation 6) in which the intensity of use of the new entrants into the market is captured by the *REG* variable is estimated using different estimation techniques (cross-section fixed effects, period fixed effects, cross-section and period fixed effects, cross-section random effects, period random effects and as a pooled sample)[13]. Afterwards, this model specification is also estimated using the GMM/IV technique that was described above. The estimation output from the "random cross-section effects", the "random period effects" and the GMM/IV techniques are shown in Table 3.

In this specification the coefficients on the variables we estimated are all of the expected sign as in our basic model and are all significant. The coefficient on the variable *REG* is negative with all the techniques used. As it was already discussed previously an increase in the *REG* variable implies a decrease in the annual mileage at which it is rational to purchase a diesel powered car compared to gasoline powered one. The negative coefficient on the variable *REG* confirms this assumption since it implies a decrease in diesel consumption per diesel powered vehicle as it increases

[13] The estimation output of all the techniques can be found in the Appendix.

and an increase in diesel consumption per diesel powered car as the variable decreases.

Table 3

Method:	***Random Cross Section Effects***	***Random Period Effects***	***GMM/IV Orthogonal Deviations***
Constant	-1.27	-0.62**	-
	(0.57)	*(0.44)*	-
Log(*CONSD/CARSD*)-1	0.91	0.92	0.80
	(0.02)	*(0.02)*	*(0.05)*
Log(*INCOME*)	0.11	0.08	0.17
	(0.02)	*(0.02)*	*(0.03)*
Log(*PD*)	-0.10	-0.13	-0.23
	(0.03)	*(0.02)*	*(0.07)*
Log(*CARS/POP*)	-0.10*	-0.10	0.00
	(0.06)	*(0.04)*	*(0.10)***
Log(*REG*)	-0.05	-0.04	-0.05
	(0.01)	*(0.01)*	*(0.01)*
Observations	151	151	137
R-Squared	0.990	0.994	0.861
D-W stat	2.02	1.82	-
J-Stat	-	-	13.44
Instrument Rank	-	-	14

*Insignificant at 5% level,**Insignificant at 10% level

The variable *GDP* which was supposed to capture general economic activity and therefore control for changes in commercial vehicle consumption of diesel was included in all specifications and all estimation techniques. Although it was insignificant only in the model specification which controlled for fuel efficiency (Equation 5) it was excluded from our results because it was capturing the same effect as the *INCOME* variable, which in turn became insignificant when the *GDP* variable was included. Therefore, from this we can conclude that *GDP* variable does not perform as expected as a proxy for commercial vehicle activity since it captures the income effect.

The model is also estimated without controlling for the peculiarities of the diesel demand market. Basically, it is also estimated in its partial adjustment form with the inclusion of the price of fuel and income control variables. This is done for comparison of the results with a specification that does not take into consideration the additions that this analysis believes help to better understand diesel fuel consumption. The estimation of this specification was done with the same techniques as the above specifications. The results from the cross-section fixed effects, the cross-section and period fixed effects and the GMM/IV techniques are shown in Table 4.

Table 4

Method:	***Cross-Section Fixed Effects***	***Cross-Section & Period Fixed Effects***	***GMM/IV Orthogonal Deviations***
Constant	0.38**	-0.39**	-
	(0.32)	(0.67)	-
Log(*CONSD/CARSD*)-1	0.85	0.77	0.88
	(0.03)	(0.04)	(0.01)
Log(*INCOME*)	0.12	0.29	0.12
	(0.04)	(0.05)	(0.01)
Log(*PD*)	-0.21	-0.32	-0.21
	(0.05)	(0.08)	(0.03)
Observations	152	152	138
R-Squared	0.995	0.993	0.813
D-W stat	1.79	1.96	-
J-Stat	-	-	12.42
Instrument Rank	-	-	14

*Insignificant at 5% level,**Insignificant at 10% level

All of the GMM/IV estimations where tested by conducting Sargan's Test of over-identifying restrictions and in all cases we could not reject the null hypothesis that the over-identifying restrictions are valid. In the non-GMM/IV estimations we conducted tests on the residuals and their distributions are close to normal. After regressing the residuals obtained from these equations on lags of themselves we found that all of the lags included were insignificant. As was mentioned before, because of the inclusion of a lagged dependent variable which is not independent with the error term in the model specifications, the non-GMM/IV estimations are inconsistent. However they were still included in this thesis for reasons of comparison to the GMM/IV estimates. In the next chapter, the elasticities obtained from the all of the above estimations will be discussed.

CHAPTER 5: DISCUSSION

In the previous chapter the different model specifications that were formulated were estimated using various econometric techniques. In this section I will discuss the variations in elasticities between these different techniques as well as the variation with estimations of other studies. Afterwards I will briefly discuss the implications of the elasticities that have been estimated.

After comparing the results of the GMM/IV estimates with those of the ordinary panel estimations the first thing that is noticeable is that the GMM/IV estimates vary a lot less. However it should also be considered that a far greater number of equations were estimated using ordinary panel data methods than were estimated using GMM/IV. In Table 5 the range of the elasticities estimated are shown based on the two econometric methods used.

Table 5

	Ordinary Panel Variation	***GMM/ IV Variation***	***Total Variation***
CONSD/CARSD -1	*0.66 – 1.00*	*0.80 – 0.88*	*0.66 – 1.00*
INCOME	*-0.07 – 0.45*	*0.12 – 0.30*	*-0.07 – 0.45*
PD	*-0.50 - -0.06*	*-0.31 - -0.20*	*-0.50 - -0.06*
CARS/POP	*-0.53 – 0.22*	*-0.37 – 0.06*	*-0.53 – 0.22*
AGE	*-0.07 – 0.01*	*0.06*	*-0.07 – 0.06*
REG	*-0.09 - -0.04*	*-0.05*	*-0.09 - -0.04*

The coefficients estimated on the lagged dependent variable range from 0.66 to 1 and therefore the speed of adjustment ranges from 0 – 0.34, which implies that the long run elasticities obtained in this study are at least three times larger than the short run elasticities. The income elasticities estimated are within the expected range except for an instance in which we obtained a negative elasticity, otherwise it is within 0.02 and 0.45. Price of fuel elasticities are always negative, ranging from -0.50 to -0.06 as expected. The elasticity of the *CARS/POP* variable differed the most in our estimations, ranging from -0.53 to 0.22. Although it was within a reasonable (negative) range in most of the estimations, it was estimated to be positive in a total of three estimations which in turn led to the huge variation observed. The elasticities obtained for the *AGE* variable were mostly insignificant; this could be a result of the average age of the car stock being a bad proxy for fuel efficiency since controlling for changes in consumer preferences was not possible. The elasticities of the *REG* variable were always negative in all model specifications and with all econometric techniques used. Therefore, as a larger number of diesel cars enter the market, diesel consumption per vehicle decreases which is in line with our assumption that the "break-even" mileage at which it is rational to purchase a diesel powered car decreases as this variable increases.

The elasticities obtained in this study show much less variation than those obtained by Pock (2007) for income in terms of the GMM/IV specification. In terms of the elasticity of price, Pock's results show much less variation than the results obtained in this analysis using the ordinary panel data techniques. However in terms of the GMM/IV specification the variation is quite similar. Pock's (2007) results also show great variation for the speed of adjustment when using the GMM/IV specification whereas this study's varied by vary little (0.80 – 0.88).

If Epsey's (1998) survey of the literature on gasoline demand is used as a comparison for the diesel demand elasticities obtained in this study we can see that they differ very little. In terms of the short run price elasticity, the range of our elasticity estimates is in the same range as most of the studies surveyed and analyzed by Epsey (1998). On the other hand, the short run income elasticities obtained in this analysis are a bit lower than the ones observed by Epsey (1998) but not in an uncommon range at all. The control variables that were added to the basic model performed fairly well. More specifically, the *CARS/POP* variable was significant in all but the GMM/IV specifications, the *AGE* variable was mostly insignificant for reasons that were discussed, the *REG* variable was significant and its negative coefficient is of a considerable size.

CONCLUSION

In this thesis diesel fuel consumption was analyzed using three different model specifications and with the use of various econometric techniques. The main contribution in this analysis is that it takes into consideration some peculiarities of the diesel fuel market and attempts to produce more realistic elasticity estimates in this way. It was found that using the average age of the car stock as a proxy for fuel efficiency did not produce significant estimates, probably because of changes in consumer preferences that this analysis did not control for. It was found that using the ratio of new registrations of diesel powered vehicles to gasoline powered vehicles it was possible to control for the influx of less intensive users into the diesel demand market. As this variable increased it was interpreted as a reduction in the "break-even" annual mileage at which it is rational to purchase a diesel instead of a gasoline powered car. Reduced vehicle utilization due to more cars per people proved insignificant with the GMM/IV technique and the inclusion of a variable to control for general economic activity in order to capture increased activity of the commercial vehicle sector proved to not be significant.

Overall, the elasticities that have been obtained in this analysis fall within reasonable ranges and do not vary a lot in the different model specifications and with the different econometric techniques used. The fact that they are in the same range as those usually obtained in gasoline demand studies confirms the assumption that both gasoline and diesel fuel are similar goods that are consumed in a nearly identical way. The smaller range in elasticities that was observed in the GMM/IV techniques, especially for the price and income elasticities, shows that this technique is more suitable for estimating diesel fuel consumption.

Results such as the ones obtained in this study, as well as the introduction of peculiarities of the diesel fuel market in studies, should be taken into consideration in order for oil companies to correctly project future demand. This will allow them to adjust their production capacities in such a way that they will correctly anticipate future demand and not cause price increases due to limited supply.

APPENDIX

A.1 Definitions of Variables

Gasoline is defined as motor spirit on EUROSTAT. It "covers leaded and unleaded motor spirit, aviation spirit and car spirit. It consists of a mixture of light hydrocarbons distilling between 30°C and 215°C. It is used as fuel for spark ignition engines. It may include additives, oxygenates and octane enhancers, including lead components such as TEL (Tetraethyl lead) and TML (Tetramethyl lead)" (EUROSTAT, 2008).

Diesel is defined as Gas/Diesel Oil on EUROSTAT. It "is primarily a medium distillate distilling between 180°C and 380°C. It comprises transport diesel (as road diesel for diesel compression ignition usually of low sulphur content) and heating and other gas oil (light heating oil for industrial and commercial purposes, marine diesel and diesel used in rail traffic and other gas oil including heavy gas oils distilling between 380°C and 540°C used as petrochemical feedstocks). Also blending components are included" (EUROSTAT, 2008).

A.2 Derivation

If diesel fuel consumption is defined as:

$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t^* = \alpha + \beta_i \cdot x_{i,t} + \eta_t \qquad \text{(a)}$$

and the following relationship exists between actual and desired levels of consumption per vehicle, where the desired level is represented by an asterisk (*) and θ is the speed of adjustment:

$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1} = \theta \cdot \left[\log\left(\frac{dieselconsumption}{dieselcars}\right)_t^* - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1}\right] + \nu_t \qquad \text{(b)}$$

Then by plugging equation (a) into (b) we get:

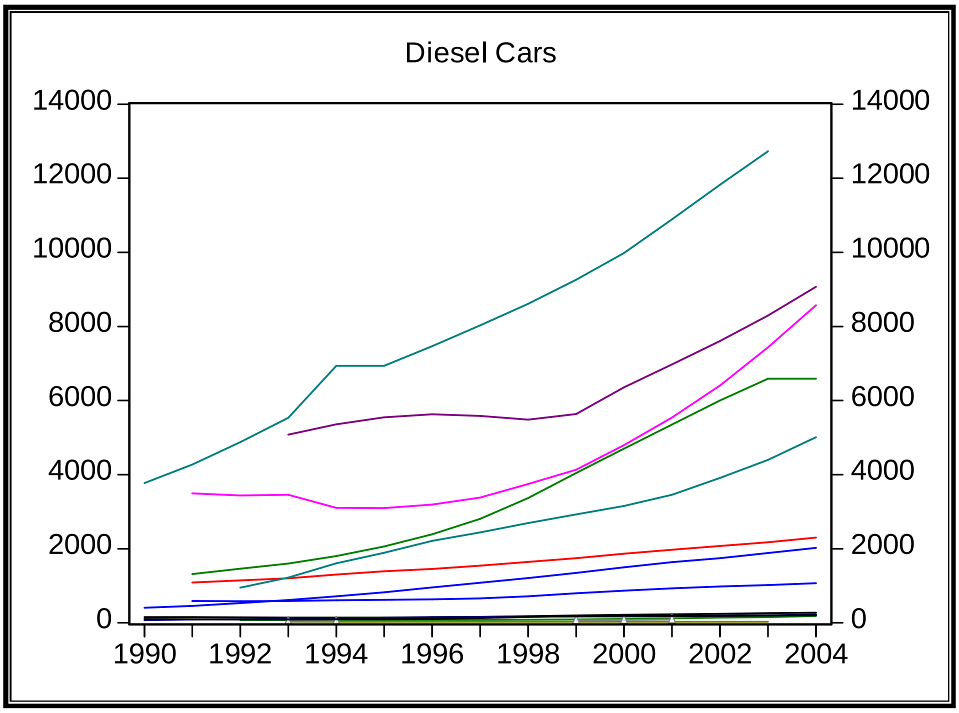
$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1} = \theta \cdot \left[(\alpha + \beta \cdot x_t + \eta_t) - \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1}\right] + \nu_t \qquad \text{(c)}$$

Rearranging the terms in (c) we get:

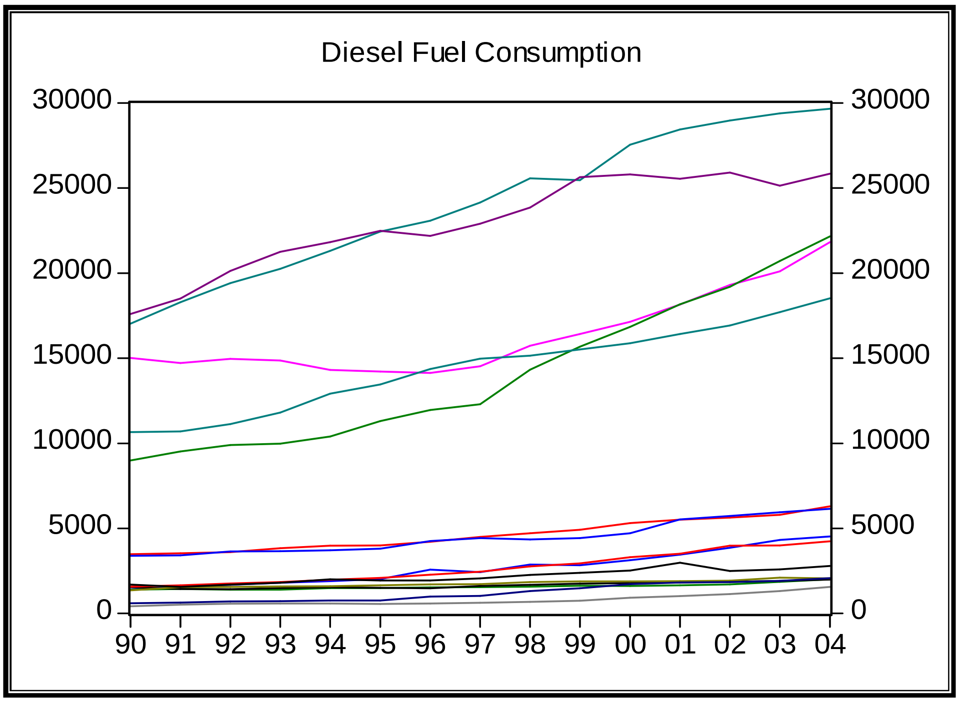
$$\log\left(\frac{dieselconsumption}{dieselcars}\right)_t = \alpha \cdot \theta + (1-\theta) \cdot \log\left(\frac{dieselconsumption}{dieselcars}\right)_{t-1} + \theta \cdot \beta_i \cdot \log(x_{i,t}) + \varepsilon_t \qquad \text{(d)}$$

A.3 Graphs

Graph 1: Amount of Diesel fueled cars.



Graph 2: Diesel Fuel Consumption



A.4 Tables

Table A.1: 1st Specification

Method:	*Cross-Section Fixed Effects*	*Period Fixed Effects*	*Cross-Section & Period Fixed Effects*	*Random Cross Section Effects*	*Random Period Effects*	*Pooled*
Constant	1.54	-0.32	1.99	-1.03	-0.65	-0.49
	(1.26)	(0.57)	(1.67)	(0.69)	(0.53)	(0.49)
Log(*CONSD/CARSD*)-1	0.84	0.98	0.75	0.98	0.98	0.99
	(0.03)	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)
Log(*INCOME*)	0.09	0.03	0.22	0.06	0.04	0.03
	(0.05)	(0.02)	(0.07)	(0.03)	(0.02)	(0.02)
Log(*PD*)	-0.22	-0.11	-0.31	-0.07	-0.07	-0.07
	(0.05)	(0.03)	(0.08)	(0.03)	(0.03)	(0.02)
Log(*CARS/POP*)	0.09	-0.09	0.22	-0.10	-0.09	-0.08
	(0.10)	(0.05)	(0.13)	(0.06)	(0.05)	(0.04)
Observations	152	152	152	152	152	152
R-Squared	0.999	0.993	0.995	0.989	0.993	0.996
D-W stat	1.80	1.63	1.97	1.80	1.65	1.48

Table A.2: 2nd Specification

Method:	*Cross-Section Fixed Effects*	*Period Fixed Effects*	*Cross-Section & Period Fixed Effects*	*Random Cross Section Effects*	*Random Period Effects*	*Pooled*
Constant	-1.84	-1.57	-4.49	-1.47	-1.19	-1.57
	(4.25)	(0.50)	(4.12)	(1.10)	(0.97)	(0.50)
Log(*CONSD/CARSD*)-1	0.75	0.97	0.66	0.97	0.98	0.97
	(0.06)	(0.01)	(0.07)	(0.02)	(0.01)	(0.01)
Log(*INCOME*)	0.24	0.10	0.45	0.10	0.08	0.10
	(0.15)	(0.01)	(0.16)	(0.04)	(0.04)	(0.01)
Log(*PD*)	-0.29	-0.10	-0.50	-0.09	-0.09	-0.10
	(0.08)	(0.02)	(0.12)	(0.04)	(0.03)	(0.02)
Log(*CARS/POP*)	-0.24	-0.16	-0.53	-0.15	-0.14	-0.16
	(0.38)	(0.04)	(0.39)	(0.10)	(0.09)	(0.04)
Log(*AGE*)	-0.07	0.01	-0.04	-0.02	-0.04	0.01
	(0.20)	(0.04)	(0.22)	(0.07)	(0.06)	(0.04)
Observations	113	113	113	113	113	113
R-Squared	0.993	0.998	0.995	0.989	0.991	0.998
D-W stat	2.11	1.68	2.11	1.90	1.83	1.68

Table A.3: 3rd Specification

Method:	*Cross-Section Fixed Effects*	*Period Fixed Effects*	*Cross-Section & Period Fixed Effects*	*Random Cross Section Effects*	*Random Period Effects*	*Pooled*
Constant	-0.81	0.23	0.09	-1.27	-0.62	-0.78
	(1.35)	(0.43)	(1.83)	(0.57)	(0.44)	(0.36)
Log(*CONSD/CARSD*)-1	0.79	0.88	0.73	0.91	0.92	0.91
	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.01)
Log(*INCOME*)	0.17	-0.07	0.16	0.11	0.08	0.06
	(0.05)	(0.02)	(0.06)	(0.02)	(0.02)	(0.01)
Log(*PD*)	-0.20	-0.26	-0.39	-0.10	-0.13	-0.10
	(0.04)	(0.04)	(0.08)	(0.03)	(0.02)	(0.02)
Log(*CARS/POP*)	-0.08	-0.12	-0.15	-0.10	-0.10	-0.11
	(0.11)	(0.04)	(0.18)	(0.06)	(0.04)	(0.03)
Log(*REG*)	-0.06	-0.07	-0.09	-0.05	-0.04	-0.04
	(0.01)	(0.01)	0.01	(0.01)	(0.01)	(0.01)
Observations	151	151	151	151	151	151
R-Squared	0.998	0.995	0.997	0.990	0.994	0.996
D-W stat	2.18	2.06	2.47	2.02	1.82	1.63

Table A.4: 4th Specification

Method:	*Cross-Section Fixed Effects*	*Period Fixed Effects*	*Cross-Section & Period Fixed Effects*	*Random Cross Section Effects*	*Random Period Effects*	*Pooled*
Constant	0.38	0.59	-0.39	0.02	0.23	0.40
	(0.32)	(0.25)	(0.67)	(0.24)	(0.21)	(0.18)
Log(*CONSD/CARSD*)-1	0.85	1.00	0.77	1.00	1.00	1.00
	(0.03)	(0.06)	(0.04)	(0.01)	(0.01)	(0.01)
Log(*INCOME*)	0.12	0.01	0.29	0.03	0.02	0.02
	(0.04)	(0.02)	(0.05)	(0.02)	(0.02)	(0.02)
Log(*PD*)	-0.21	-0.10	-0.32	-0.06	-0.07	-0.09
	(0.05)	(0.03)	(0.08)	(0.03)	(0.03)	(0.02)
Observations	152	152	152	152	152	152
R-Squared	0.995	0.998	0.993	0.996	0.988	0.993
D-W stat	1.79	1.61	1.96	1.79	1.63	1.45

A.5 Test Results

Table A.5: Distribution of residuals, autocorrelation and over-identification tests.

Specification	***1***	***2***	***3***		***4***	
Skewness	*0.56*	*0.54*	*0.65*	*0.71*	*0.38*	*0.47*
Kurtosis	*4.60*	*4.49*	*4.16*	*4.64*	*3.98*	*4.76*
Insignificant Lags	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*No*
Sargan's Test	χ^2 *(11.65,10)=0.30*	χ^2 *(9,8)=0.34*	χ^2 *(13.44,9)=0.14*		-	

REFERENCES

Arellano, M., and Bover, O. (1995). *Another Look at the Instrumental Variables Estimation of Error components Models*, Journal of Econometrics, 68, 29–51.

ACEA (2008), www.acea.be , (accessed last at 6/6/2008).

Baltagi, B. H. and Griffin, J. M. (1983), *Gasoline demand in the OECD: An application of pooling and testing procedures*. European Economic Review, 22, 117-137.

Baltagi, B. H. and Griffin, J. M. (1997). *Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline*. Journal of Econometrics, 77, 303-327.

Epsey, M. (1998) *Gasoline demand revisited: An international meta-analysis of elasticities*. Energy Economics 20_1998.273]295

EUROSTAT (2008), www.ec.europa.eu/eurostat , (accessed last at 6/6/2008).

Eviews Users Guide 5, 2004, pp. 854

Pock, M. (2007) *Gasoline and Diesel Demand in Europe: New Insights.* Reihe Ökonomie / Economics Series 202.

Schmalensee, R., Stoker, T. M. (1999*): Household gasoline demand in the United States*. Econometrica, 67, 3: 412-418.

Sweeny, J. (1978). *The demand for gasoline in the United States: a vintage capital model. In: Workshop on energy and supply and demand*, International Energy Agency, Paris, 240-277.

Verboven, F. (2002). *Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars*. RAND Journal of Economics Vol. 33. No. 2, Summer 2002, pp.275-297

Zachariadis, T. (2006): *On the baseline evolution of automobile fuel economy in Europe*. Energy Policy, 34: 1773-1785.