202

Reihe Ökonomie

Economics Series

Gasoline and Diesel Demand in Europe:

New Insights

Markus Pock

202

Reihe Ökonomie

Economics Series

Gasoline and Diesel Demand in Europe:
New Insights

Markus Pock

January 2007

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Markus Pock
Department of Economics and Finance, HealthEcon
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
☎: +43/1/599 91-210
fax: +43/1/599 91-555
email: pock@ihs.ac.at

Founded in 1963 by two prominent Austrians living in exile – the sociologist Paul F. Lazarsfeld and the economist Oskar Morgenstern – with the financial support from the Ford Foundation, the Austrian Federal Ministry of Education and the City of Vienna, the Institute for Advanced Studies (IHS) is the first institution for postgraduate education and research in economics and the social sciences in Austria. The **Economics Series** presents research done at the Department of Economics and Finance and aims to share "work in progress" in a timely way before formal publication. As usual, authors bear full responsibility for the content of their contributions.

Das Institut für Höhere Studien (IHS) wurde im Jahr 1963 von zwei prominenten Exilösterreichern – dem Soziologen Paul F. Lazarsfeld und dem Ökonomen Oskar Morgenstern – mit Hilfe der Ford-Stiftung, des Österreichischen Bundesministeriums für Unterricht und der Stadt Wien gegründet und ist somit die erste nachuniversitäre Lehr- und Forschungsstätte für die Sozial- und Wirtschafts-wissenschaften in Österreich. Die **Reihe Ökonomie** bietet Einblick in die Forschungsarbeit der Abteilung für Ökonomie und Finanzwirtschaft und verfolgt das Ziel, abteilungsinterne Diskussionsbeiträge einer breiteren fachinternen Öffentlichkeit zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den Autoren und Autorinnen.

Abstract

This study utilizes a panel data set from 14 European countries over the period 1990-2004 to estimate a dynamic model specification for gasoline and diesel demand. Previous studies estimating gasoline consumption per total passenger cars ignore the recent increase in the number of diesel cars in most European countries leading to biased elasticity estimates. We apply several common dynamic panel estimators to our small sample. Results show that specifications neglecting the share of diesel cars overestimate short-run income, price and car ownership elasticities. It appears that the results of standard pooled estimators are more reliable than common IV/GMM estimators applied to our small data set.

Keywords
Dynamic panel data, gasoline demand, error components, omitted variable

JEL Classification
C23, Q41

Comments
This research has been initiated during the doctoral seminar held by Robert Kunst at the University of Vienna and could be finished at IHS with the help of sabbatical funding of the Austrian National Bank, OeNB. I am grateful to Robert Kunst for his comments and advices and to Angela Parker and Harald Hutter for their editorial support.

Contents

1 Introduction **1**

2 Rising diesel car share in Europe **2**

3 Model specification **6**

3.1 Description 6

3.2 Data 10

4 Estimators **11**

5 Discussion **13**

5.1 Omitted variable bias 13

5.2 Gasoline demand 14

5.3 Diesel demand 19

6 Summary and conclusions **20**

Appendix **22**

A.1 Data description 22

A.2 Omitted variable bias 23

References **27**

1 Introduction

Petroleum products are major ingredients of a growing economy in developed countries. Gasoline and diesel fuel in particular have become essential for private and economic mobility. With growing energy and fuel consumption, however, the dependency on these scarce resources has become striking, as oil crises in the past and recent oil price developments indicate. In addition, environmental externalities caused by rising fuel consumption, such as emissions of carbon dioxide, nitric oxides, and carcinogenic airborne fine particulates from diesel engines are becoming a major concern. Hence, policy-makers have an interest in how expected increases in income and fuel prices would affect fuel consumption and automobile use over time.

There is a vast literature on estimating gasoline demand at the aggregate level as well as at micro level[1]. Many of the previous studies used gasoline consumption per total passenger cars or per capita as the dependent variable and included total passenger cars per capita as the explanatory variable[2], implicitly assuming a diminishing share of diesel-powered cars. For some countries, like the USA, for example, this holds true, but especially in the Western European countries an extensive increase in the number of registered diesel cars can be observed at the expense of the number of gasoline-powered cars. Thus, one could expect the estimated coefficients of the equations that explain gasoline consumption using total passenger cars instead of the number of gasoline-powered cars to be biased due to an omitted variable. Accordingly, the income, price and car ownership elasticities should be overestimated in dynamic demand models[3]. The higher the diesel car share, the more severe the bias should be.

The objective of this paper is to investigate a dynamic fuel consumption equation adopted from the flow adjustment model by Houthakker and Taylor (1970), which has been extensively applied in the literature[4], but with new data accounting for the omitted variable problem just mentioned. The balanced panel data set comprises time series of 14 European countries over the period 1990 till 2004 for gasoline and diesel consumption, the number of gasoline and diesel-powered passenger cars, gross domestic product per capita, and CPI-adjusted gasoline and diesel retail prices. The panel data stem from EUROSTAT and were supplemented by data from national statistics institutions.

[1] For surveys see Bohi and Zimmermann (1984), Dahl (1986), Dahl and Sterner (1991), and Espey (1998) for a meta-analysis.

[2] See Baltagi and Griffin (1983, 1999) using a panel over 18 OECD countries from 1960 to 1978 and 1960 to 1990, respectively.

[3] In the appendix we give a some intuition of the bias direction.

[4] See for instance Houthakker et al. (1974), Sweeny (1978), Baltagi and Griffin (1983,1997), and Baltagi et al. (2003).

To our knowledge, there exists no other study using such specific time series for estimation. In a recent pooled study with regional French data, Baltagi et al. (2003) tried to take account of the observed effect of a shift towards diesel-fueled cars by means of a constructed petrol price index based on gasoline and diesel fuel prices, but leaving gasoline consumption per car unadjusted. Our paper contributes to the literature of fuel demand by producing separate estimates for unbiased income and price elasticities with respect to the omitted variable of diesel-powered cars in a dynamic gasoline and diesel demand equation.

Secondly, we compare the estimates of several common dynamic panel data estimators along the lines of previous studies done by Baltagi and Griffin (1983, 1999), and Baltagi et al. (2003). Due to the short time dimension of the panel, we refrained from comparing forecast performance. As recent simulation studies revealed[5], the bias-corrected Within estimator (Kiviet (1995)) and IV/GMM estimators according to Andersen and Hsiao (1982), Arellano and Bond (1991), and Blundell and Bond (1998) are supposed to be a practical device which remedies the estimation bias in small dynamic panels.

In the next section, we give a short overview of the basic facts concerning the evolution of diesel-fueled cars in Europe. Section 3 describes the model specification, the data set and the applied pooled estimators. In Section 4, results of the estimations are given and discussed. Section 5 summarizes and concludes.

2 Rising diesel car share in Europe

Diesel engines are up to 30% more efficient in thermodynamic terms than Otto motors due to higher combustion temperature. Compared on the basis of identical engine capacity, a diesel motor has less engine power, higher weight and higher production costs than the gasoline motor. From an environmental point of view, a diesel engine emits approx. 30% less CO_2 than its gasoline counterpart, thus diesel technology is seen as a major device for meeting the Kyoto commitments (EC-ACEA (2003)). The reverse side of the coin is the probably harmful diesel exhaust[6], which poses a serious air pollution

[5] See for instance Kiviet (1995), Judson and Owen (1999), and Bun and Kiviet (2003).

[6] In a growing number of scientific studies, diesel exhaust has been linked to respiratory diseases, heart disease, cancer and premature deaths (EPA (2002), WHO (2003)). Diesel exhaust is a complex mixture of gases and particles. NO_x, CO, SO_x and particulate matter consisting of polycyclic aromatic hydrocarbons (PAH), in particular, are suspected of causing health problems. According to Künzli et al. (2000), air pollution causes 6% of overall mortality or 40.000 deaths per year in Switzerland, Austria and France, with half of these deaths attributed to motor traffic emissions. Pope et al. (2002) found evidence of carcinogenicity of fine particles ($PM_{2.5}$) from diesel exhaust.

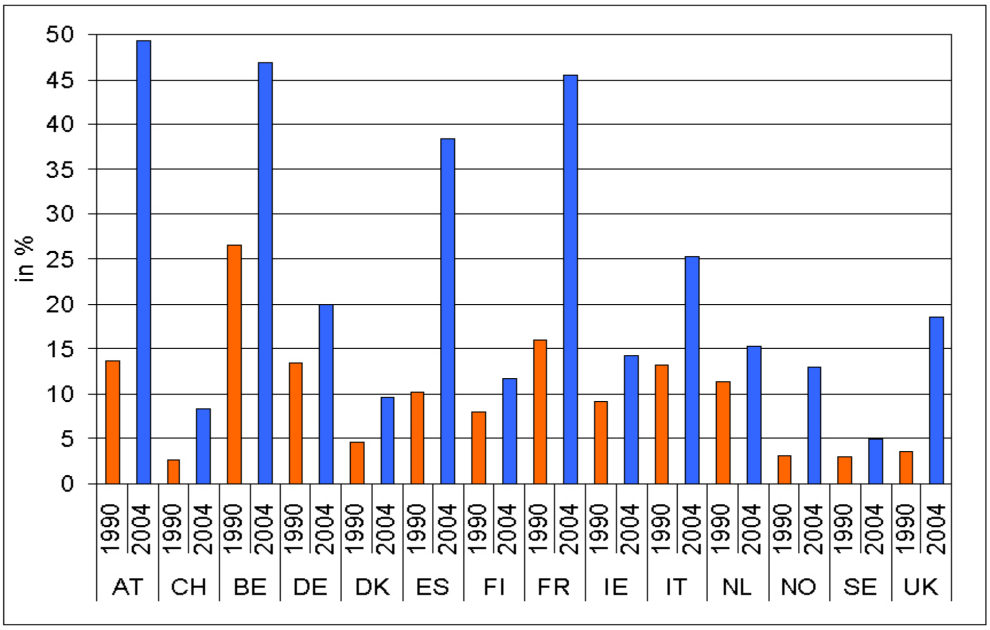

Figure 1: Diesel share of total passenger cars in percent for 14 European countries in 1990 and 2004

problem despite technological advances in emission control.

Although diesel technology has been available for over 80 years now, the penetration of the European passenger car markets did not start until the end of the 1980s, when technological innovations like direct injection and turbo chargers have increased the torque and power of diesel motors and thus improved the comfort and driveability of diesel cars. In addition to high oil prices, fuel tax differentiation favouring diesel in most Western European countries has enhanced the growth of the diesel car share during the last fifteen years. For instance, the share in France amounted to 4.7% in 1980 compared to 16%, 35.6% and 45.5% in 1990, 2000 and 2004, respectively. Austria, Belgium and Spain have also experienced a strong rise in the diesel share of their passenger car fleet, namely from 13.7%, 26.6% and 10.2% in 1990 to 49.2%, 46.9% and 38.4% in 2004, respectively (see Fig.1).

The observed shift to diesel cars in most European countries stems from the fact that diesel fuel is more efficient in economic terms (approx. 2 litres per 100 km) and cheaper than gasoline due to lower taxes in most European countries, thereby offsetting the higher purchase costs of diesel cars and the slightly higher production costs of diesel fuel. Figure 2 shows the retail fuel price differences in five European countries. In

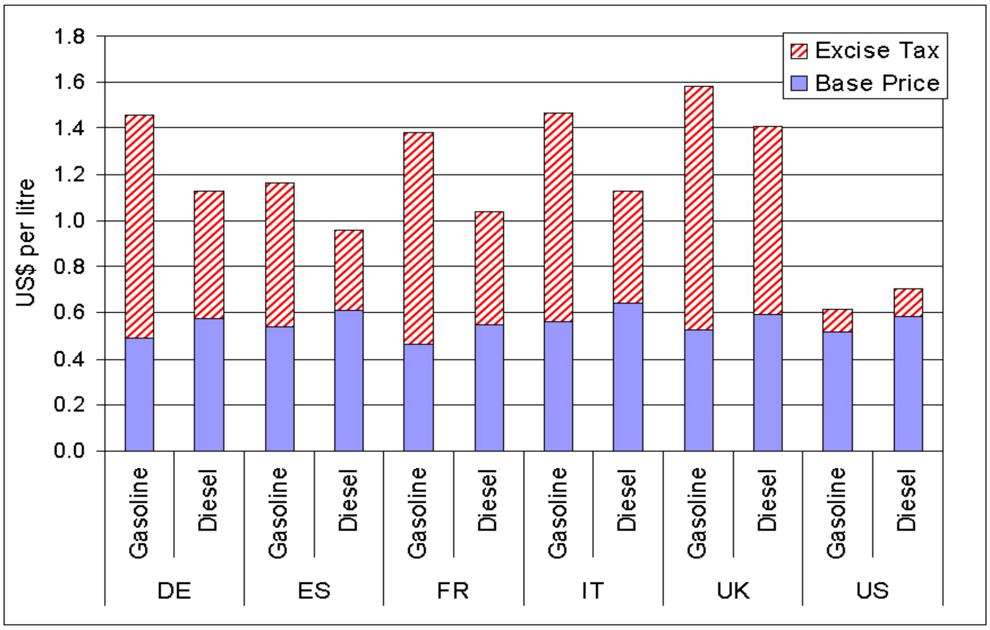

Figure 2: Fuel retail prices of five European countries and USA, Nov 2005 (source: International Energy Agency)

contrast to the USA, where stringent regulations for motor car emissions as well as low and equal fuel taxing de facto excluded diesel cars from the market, European policies focused on fuel economy and reduction of CO_2 emissions (EC-ACEA (2003)). Also, lobbying by the commercial transport sector contributed to the lower diesel taxes. The average of the difference between gasoline and diesel retail prices from 1990 till 2004 over all EU-countries[7] in our data amounts to 18 cents per litre. The relative diesel price is an important factor in the recently observed strong growth in diesel cars in most European countries, a trend which has probably been amplified by the recent oil price increase. The correlation between the average price of diesel relative to gasoline and the growth in the diesel car share from 1990 to 2004 amounts to 0.5 in our data set.

However, the fuel economy of diesel engines combined with the lower diesel retail price is not the whole story. The consumer's decision to purchase a certain type of a car is also influenced by the latter's acquisition costs and non-fuel running costs[8]. The

[7] Except Switzerland, because it is the only country considered with higher retail prices for diesel than for gasoline.

[8] Driving comfort, which encompasses greater car safety, engine power, air conditioning, GPS, etc., is also an important issue. The rising demand for these amenities has led to an increased market share of bigger, heavier and faster diesel vehicles (SUV - sports utility vehicles) in Europe. Together with higher traffic density leading to uneconomic driving habits, this has considerably reduced the

manufacturing costs for diesel cars are higher than those for gasoline cars, resulting in approx. 5% higher net purchasing costs for cars with comparable engine power (Proost and Mayeres (2000)). Taxation of car purchases varies between European countries. In car producing countries, such as France, Germany, Italy and the United Kingdom, the normal value-added tax is applied with other registration fees being of minor importance. Other fiscal policies target the level of fuel consumption by taxing vehicle price, weight, engine capacity or power (Orfeuil (2001)). The progressivity of such tax schemes favours gasoline cars leading to a smaller diesel car share in countries like Denmark, Finland, Ireland, the Netherlands and Norway (see Fig.1).

Non-fuel running costs comprise for instance the annual tax on car ownership, insurance and maintenance costs, and road user charges. The tax on car ownership is relevant to the consumer's decision between a gasoline and a diesel car, because in most European countries it is indexed to the engine's power or capacity, thus favouring gasoline cars (Orfeuil (2001)). In Finland, gasoline-powered vehicles are exempted from the annual motor vehicle tax. A switch to CO_2-based taxation systems is currently being discussed in European countries. In contrast, by 2006 Sweden had introduced a CO_2-related annual road tax on passenger cars, which has immediately increased the diesel share of new car registrations (ACEA (2006)).

Passenger car owners who have above-average annual driving distances are thus more inclined to choose diesel over gasoline cars[9]. The money saved from lower fuel costs and higher efficiency makes up for the more expensive upfront cost of the vehicle. Formally, the total life-cycle costs of a car can be written as $A + \Sigma_{t=1}^{T}(1+r)^{-t}(p_t q_t + c_t)$, with the acquisition costs A, operating life expectancy T, retail fuel price p, annual fuel consumption q, interest rate r and non-fuel operating costs c. By assuming constant kilometrage km, constant gasoline and diesel consumption[10] of 8.8 and 7 liters per 100 km, respectively, and a constant price mark-up ω to equal diesel price to gasoline price one derives the kilometrage break-even, where a rational individual opts for the diesel car given equivalent engine power and vehicle size:

$$\overline{km} = \frac{100(f^{-1}\Delta A + \Delta c)}{(8.8\omega - 7)p_d}, \qquad f = \frac{(1+r)^T - 1}{r(1+r)^T} \tag{1}$$

gains in fuel economy obtained due to technology improvements and the observed *dieselisation* of European car fleets (see Zachariadis (2006) and the references therein). As a side effect, the discrepancy between official driving cycles measuring fuel efficiency and real-world driving records has increased during the last decade.

9 In 2000, for instance, Austrian households drove an average of 12.000 and 16.000 km with their gasoline- and diesel-powered cars, respectively (Statistics Austria (2001)).

10 See Statistics Austria (2001).

with the total discounting factor f, the difference in acquisition costs, ΔA, and annual non-fuel running costs, Δc, of a diesel and gasoline car, respectively. The current diesel price p_d is assumed to grow in line with the inflation rate, hence r is the net interest rate. Due to the higher efficiency of diesel in terms of distance driven per liter, the gasoline retail price could be approx. 80% of the diesel price and one would still observe the process of substitution by diesel-powered cars, *ceteris paribus*. The kilometrage break-even $\overline{km}$ decreases with increasing diesel prices and net interest rate as well as a lower spread of the acquisition and annual running costs. Accordingly, lower diesel car shares can be found in those countries (see Fig.1) where a relatively higher tax has been imposed on the purchase and ownership of diesel cars (Denmark, Finland, Netherlands, Norway, Sweden) or where the difference in retail fuel prices is relatively small (Ireland, Switzerland, United Kingdom)[11].

3 Model specification

3.1 Description

According to Sweeny (1978), and Baltagi and Griffin (1983, 1997), fuel consumption is calculated on the basis of a vehicle's utilization, its fuel efficiency, i.e. fuel consumption per distance driven, and the stock of cars in use:

$$Gasoline\ consumption = \frac{\sharp\ km}{\sharp\ cars} \cdot \frac{Gasoline\ consumption}{\sharp\ km} \cdot \sharp\ cars. \tag{2}$$

Data limitations require a representation in which gasoline consumption per passenger car GAS is explained by variables reflecting utilization and efficiency, leading to the following log-linear demand equation:

$$(GAS)^* = \alpha(Y)^\gamma(PG)^\beta(CAR)^\delta \tag{3}$$

Estimation of the gasoline demand per car is thus based on real income per capita Y, real gasoline price PG and stock of passenger cars per driver CAR. All these variables are supposed to influence vehicle utilization. CAR aims to capture reduced utilization caused by the rising number of cars per household[12] and - by using the number of

[11] Proost and Mayeres (2000) found the break-even annual kilometrage for medium capacity diesel cars in 10 EU countries to be highest in Finland and lowest in France.

[12] A family owning two cars does not drive twice the distance of a one-car family. The average annual kilometrage of the Austrian households, for instance, amounts to 14.153 km for the first and 9.942 km for the second gasoline-powered car in 2000 (Statistics Austria (2001)).

drivers instead of total population as the denominator - avoid demographic effects[13]. Fuel efficiency, on the other hand, is not observable on an aggregated data level and over the chosen time horizon for all countries. It is determined by the technical characteristics of a car's engine, driving habits, traffic density, geographical conditions, the vehicle's weight, usage of air conditioning, etc. Since the vehicle stock has a gestation period which itself is expected to vary with economic growth, fuel efficiency can be expressed by distributed lags of the economic variables (see Baltagi and Griffin (1983)). Including the lagged dependent variable as regressor instead mirrors this distributed lag specification (see Baltagi et al. (2003)).

From a model viewpoint, lagging gasoline consumption per car is justified by adopting a flow adjustment model according to Houthakker et al. (1974). Utilization is adapted to desired utilization via a habit-persistence mechanism. Applied to our case, this means that adjustment of realized gasoline consumption GAS to the desired level of gasoline consumption $(GAS)^*$ over time is assumed to follow a first-order process:

$$\frac{GAS_t}{GAS_{t-1}} = \left(\frac{GAS_t^*}{GAS_{t-1}}\right)^{\theta}, \qquad 0 < \theta < 1 \tag{4}$$

where θ is the adjustment coefficient.

After plugging equation (3) into (4), log linearizing, and pooling the data for country i and time t, one obtains the frequently used dynamic demand equation for gasoline:

$$\begin{aligned} \ln GAS_{i,t} &= \theta \ln \alpha + (1-\theta) \ln GAS_{i,t-1} + \theta\beta \ln Y_{i,t} + \theta\gamma \ln PG_{i,t} + \\ &\quad + \theta\delta \ln CAR_{i,t} + u_{i,t} \end{aligned} \tag{5}$$

where the disturbance term $u_{i,t}$ is specified as a two-way error component model:

$$u_{i,t} = \mu_i + \lambda_t + \epsilon_{it}, \quad i = 1, \ldots, I, \quad t = 1, \ldots, T \tag{6}$$

with country-specific effect μ_i, time-specific effect λ_t and white noise $\epsilon_{i,t}$. The individual and time effects can be modeled as fixed or random. The specification of the error term can be generalized to be correlated over the cross-sections and/or over time.

Under formulation (5), the short-run elasticities of gasoline demand per car with

[13] Schmalensee and Stoker (1999) pointed out that using the number of licensed drivers instead of population heavily reduces estimated income elasticity with US data. They conclude that many studies ignoring demographic changes overstate income elasticities. As a proxy for the number of licensed drivers we took population aged between 18 and 69 (see the data appendix). The Within estimator applied to our panel set gives approx. 20% higher income and price elasticities when using total population instead, confirming the above findings.

respect to per capita income, real price and total cars per driver are $\theta\beta$, $\theta\gamma$ and $\theta\delta$, respectively. The corresponding long-run responses are given by β, γ and δ, with the speed of adjustment to the long-run equilibrium $(1-\theta)$. For the short-run and long-run transformed elasticities of total gasoline demand relative to per capita income, real price and total car fleet the results are $(1-\theta\beta)$, $(1-\theta\gamma)$, $(1-\theta\delta)$, and $(1-\beta)$, $(1-\gamma)$, $(1-\delta)$, respectively.

The short-run effect of gasoline price works primarily through adjustment of car utilization, while in the long-run consumers adapt their car fleet to long-run changes in gasoline prices. In the event of gasoline price increases there will not only be a shift to more efficient gasoline-fueled cars but also one to diesel-fueled cars, as long as diesel fuel and diesel car ownership costs are relatively low. Looking at the recent developments in the composition of the passenger car stock, exactly the latter has happened in most European countries (see Fig.1). Gasoline-powered cars with higher utilization rates will be replaced first. The underlying shift towards diesel cars over time can be thought of as a kind of selection mechanism. Those individuals with higher car utilization are more likely to switch to diesel cars than others. Because the acquisition costs of diesel cars are higher than for gasoline-powered cars, there will be a certain economic threshold for engine substitution depending on car utilization, diesel prices and non-fuel operating costs (see the discussion in section 2). Gradually, the less intensive car users are left with the gasoline-powered car fleet, which certainly decreases gasoline consumption per gasoline car.

The demand equation defined in (5) is afflicted with two shortcomings. First, it ignores the just mentioned indirect effect of the steadily increasing number of diesel cars on per car gasoline consumption. On the other hand, equation (5) relates gasoline consumption to total passenger cars and uses total cars per driver as an explanatory variable, even though diesel-fueled vehicles are operated by diesel and not gasoline fuel. As it is shown in the appendix, this specification of the gasoline demand equation neglects variables. Hence, the elasticity estimates do suffer from an omitted variabel bias leading to overestimation of the elasticity estimates.

In order to correct these defects of the common gasoline demand equation, we restate equation (5) by using gasoline consumption per gasoline-powered car $GASG$ as the dependent variable and the number of gasoline-powered passenger cars per driver $CARG$ as a regressor variable. In addition, the regressor variable diesel-powered passenger cars per driver $CARD$ is appended to get:

$$\begin{aligned}\ln GASG_{i,t} &= \theta \ln \alpha + (1-\theta) \ln GASG_{i,t-1} + \theta\beta \ln Y_{i,t} + \theta\gamma \ln PG_{i,t} + \\ &+ \theta\delta \ln CARG_{i,t} + \theta\phi \ln CARD_{i,t} + u_{i,t} \qquad (7)\end{aligned}$$

where the coefficient ϕ measures the long-run elasticity of diesel car substitution. An F-test of the joint significance of included variables rejects the null of the redundancy of the $CARG$ variable in both one-way and two-way fixed or random effects models.

Along the same lines, the diesel consumption demand equation is derived by means of an error component model:

$$\begin{aligned}\ln DIESD_{i,t} &= \theta \ln \alpha + (1-\theta) \ln DIESD_{i,t-1} + \theta\beta \ln Y_{i,t} + \theta\gamma \ln PD_{i,t} + \\ &+ \theta\delta \ln CARG_{i,t} + \theta\phi \ln CARD_{i,t} + u_{i,t} \qquad (8)\end{aligned}$$

Here, the dependent variable is diesel consumption per diesel-fueled car ($DIESD$). The exogenous regressors are income per capita Y, diesel price PD, and numbers of gasoline and diesel-fueled passenger cars per drivers, $CARG$ and $CARD$, respectively. The $CARG$ variable is included to capture the opposite substitution effect away from gasoline towards diesel cars. Again, an F-test clearly rejects the null of the redundancy of the $CARG$ variable.

Besides diesel-powered passenger cars there are other road vehicles that consume diesel fuel, such as trucks and buses for transporting goods and persons. These could also be expected to influence diesel consumption. However, the added variables truck per diesel car or truck per capita proved to be insignificant, being in line with the results obtained by Baltagi and Griffin (1983). One explanation for this result is that the variable truck per capita, in particular, captures the economic activity that is already modeled by the variable income per capita Y.

In order to decide between a one-way or two-way and fixed or random effects model for the postulated fuel demand equations, the following tests were conducted (see Baltagi (2005)). First, a Hausman-type specification test rejects the null of *no systematic difference between the Within and GLS coefficient estimates* in both the one- and two-way models in each demand equation[14], supporting a fixed effects model. Next, we conducted a likelihood ratio and F-test to distinguish between the one- and two-way

[14] The $\chi(5)$-statistic amounts to 44.3 and 34.3 respectively in the one-way and two-way error component model of the gasoline demand equation given in (5), and 53.3 and 53.2 respectively in the diesel demand equation.

model. Both statistics clearly point to the individual fixed effects model[15]. Hence, we decided to use the one-way fixed effects model when applying the Within and related estimators described in the following section[16]. Correspondingly, when applying standard GLS to equations (7) and (8) for reasons of comparison, only country-specific effects were considered.

Further, in both fuel demand equations with fixed effects a Wald test reveals cross-sectional heteroskedasticity. In addition, several tests for serial correlation indicate first-order autocorrelation of the residuals after estimating fixed and random effects models. We thus add to the selection of estimators a feasible GLS, allowing for country-specific AR(1) autocorrelation, cross-sectional correlation and heteroskedasticity (see section 3).

3.2 Data

The annual data set comprises 14 European countries ranging from 1990 till 2004: Austria (AT), Belgium (BE), Switzerland (CH), Germany (DE), Denmark (DK), Spain (ES), Finland (FI), France (FR), Ireland (IE), Italy (IT), the Netherlands (NL), Norway (NO), Sweden (SE) and the United Kingdom (UK). The basic data were obtained from EUROSTAT. The panel was then revised and supplemented by data from national statistical institutions[17]. Fuel consumption is measured in tonnes per year. Income per capita is calculated as PPP-adjusted real GDP in US$ per total population and real fuel prices as CPI-adjusted retail prices in EURO per liter[18]. In order to account for demographic changes, the stock of car series are divided by the population aged 18 to 69 rather than by total population as a proxy for adult drivers.

Table 1 summarizes the extent of inter- and intra-country data variation for the variables. There is more between than within variation in the data, suggesting a pooled estimation procedure. The common pooled estimators assume homogeneity of the parameters across the countries and time, i.e. the parameters are assumed to be constant for

[15] The unrestricted model included country and time-specific effects, while the restricted version excluded the time effects. For the gasoline and diesel demand equation the tests yield $F(13,164) = 1.3$ and $\chi(13) = 18.7$, and $F(13,164) = 0.9$ and $\chi(13) = 14.0$ respectively, being below the corresponding distribution with the degrees of freedom given in brackets. If the individual effects are additionally restricted, the statistics clearly reject redundancy of the effects in both equations.

[16] At the first glance, it seems counterintuitive not to take time effects into account in the data generating process given the strong shift towards diesel-powered cars in most countries over time. Indeed, if we drop the $CARDA$ variable from the gasoline demand equation, we obtain significant test statistics advocating the incorporation of time effects. We conclude that the country-specific variable $CARDA$ captures these time trends better than country-unspecific time dummies.

[17] For details see the data appendix.

[18] The conversion factor between tonnes and liters of a specific fuel type, which is implied in then regression of fuel consumption in tonnes on fuel price per liter, will be captured by the regression constant.

Table 1: Analysis of Variances of the variables in the data set (I=14, T=15, 1990-2004)

	ln $GASG$	ln $DIESD$	ln Y	ln PG	ln PD	ln $CARG$	ln $CARD$
Overall	9.279	84.804	6.009	5.145	11.097	6.493	144.156
Between	95.6%	86.4%	56.4%	63.9%	65.6%	87.2%	79.7%
Within	38.4%	16.9%	43.6%	36.1%	34.4%	12.8%	20.3%

all i, t in equations (5), (7) and (8), enabling poolability of country data. However, this assumption was doubted and alternative estimators proposed[19]. Indeed, homogeneity is rejected in our data set[20]. Nevertheless, given the dominance of between variation in the independent variables, one might expect heterogeneous estimators to perform less favorably than homogeneous estimators. A comparison study concerning forecast performance of homogeneous and heterogeneous estimators was conducted by Baltagi and Griffin (1997) and Baltagi et al. (2003) using a similar data set to ours. They showed that heterogeneous estimators could not outperform homogeneous ones. In this study we therefore concentrate on homogeneous estimators, which are briefly described in the next section.

4 Estimators

The following standard pooled estimators are applied to our data set, assuming exogeneity of the regressors: pooled **OLS**, ignoring any effects, the Within estimator **LSDV**, which allows for individual fixed effects, and **GLS**, where the country-specific effects are assumed to be random. In addition, a feasible GLS estimator **GLS-HC**, allowing for heteroskedasticity with cross-sectional correlation and a panel-specific AR(1) error structure, is estimated with individual fixed effects, having a richer error structure than the latter GLS estimator.

Since our model is dynamic with individual country effects, the lagged dependent variable is correlated with the error term and thus leads to inconsistent estimates of OLS, Within[21] and GLS. As a remedy, the literature builds on the instrumental variable estimation method providing consistent but not necessarily efficient estimates of the model parameters. The instruments are the exogenous variables and their lagged

[19] See Maddala et al. (1994,1997) for their shrinkage estimator, and Pesaran and Smith (1995).

[20] Applied to equation (7), a Chow-test for the equality of slope coefficients across countries and time with varying intercepts yields an F-value of 4.04, which is distributed as $F(78, 126)$. The null hypothesis is clearly rejected. However, standard poolability tests are known to overreject.

[21] The well-known Nickell bias is given in the appendix.

values. In our study we report the **Within-2SLS**, which transforms the data around the country means, and the first-difference 2SLS estimator (**FD-2SLS**) proposed by Anderson and Hsiao (1982). The latter estimator eliminates unobserved country heterogeneity by first-differencing, but this leads to autocorrelation with the error term. In order to preserve consistency, the second lag of the dependent variable in levels is used as an instrument for the differenced lagged dependent variable. Further, following Arellano and Bond (1991), we used their one-step GMM estimator (**AB1**). The endogenous variable in first differences is instrumented here with suitable lags of its own levels and differenced exogenous regressors. Because it incorporates more orthogonality conditions, the Arellano-Bond estimator is more efficient than FD-2SLS. Blundell and Bond (1998) state that with persistent data differenced IV and GMM estimators suffer from small sample bias due to weak internal instruments. As a solution they suggest a system GMM estimator (**sys-GMM**) with first-differenced instruments for the equation in levels and instruments in levels for the differenced equation. In their simulation study, the small sample properties of this estimator seem to be preferable in comparison to other IV and GMM estimators. However, with increasing number of regressors the moment conditions get close to the number of observations in small samples, as in our case. Too many instruments produce over-fitting of the instrumented variable and the resulting estimates are biased toward those of the OLS (see for instance Baltagi (2005), p.153). This is exactly what we observed for our sample. We therefore used a subset of the instrument matrices[22] in order to avoid this small sample bias.

A weakness of all IV, GMM and system GMM estimators is that their desirable properties only hold asymptotic for large N. Thus, in samples with a small number of cross-sectional units, as in our case, the estimates can be biased and inefficient (see Bun and Kiviet (2006)). An alternative based on the bias-corrected Within estimator (**LSDVc**) has recently been used in the econometric literature. The Within estimator, although inconsistent, has a smaller variance compared to IV and GMM estimators. By correcting for the Nickell (1981) bias via approximation terms developed in Kiviet (1995), one obtains an estimator with favorable properties in small samples. Monte Carlo studies done by Kiviet (1995), Judson and Owen (1999), and Bun and Kiviet (2003) demonstrate that the bias-corrected Within estimator (LSDVc) indeed often outperforms IV and GMM estimators in samples with small N, T. In our study, we applied the corrected Within estimator using an approximation term up to the order of T^{-1} (see for instance Bun and Kiviet (2003), formula B1). Initial values of the true coefficients were obtained

[22] For the implementation in STATA see Roodman (2005).

Table 2: Elasticity estimates of gasoline demand equation without diesel (5)

	Short-run				Long-run		
	GAS_{t-1}	Y_t	PG_t	CAR_t	Y_t	PG_t	CAR_t
Exogenous regressors							
OLS	1.003	0.003	0.015	-0.061	-	-	-
	(84.47)	(0.14)	(0.59)	(2.87)	-	-	-
Within	0.788	0.257	-0.107	-0.696	1.209	-0.502	-3.277
	(17.11)	(4.03)	(3.49)	(5.78)	(4.27)	(4.13)	(6.45)
LSDVc	0.822	0.231	-0.097	-0.620	1.296	-0.543	-3.471
	$(24.25)^b$	$(3.88)^b$	$(3.27)^b$	$(6.01)^b$	$(3.72)^b$	$(3.29)^b$	$(6.22)^b$
GLS	0.980	0.0174	-0.005	-0.112	0.876	-0.236	-5.666
	(49.43)	(0.63)	(0.17)	(3.21)	(0.66)	(0.17)	(1.21)
GLS-HC	0.750	0.270	-0.122	-0.81	1.077	-0.489	-3.257
	(401.68)	(97.89)	(101.58)	(171.88)	(89.29)	(64.28)	(204.35)
Endogenous regressors							
Within-2SLS	0.573	0.407	-0.171	-1.142	0.953	-0.400	-2.674
	(6.66)	(4.42)	(4.07)	(5.59)	(5.82)	(5.45)	(9.89)
FD-2SLS	0.922	0.471	-0.207	-0.934	6.056	-2.664	-12.011
	(2.44)	(2.67)	(4.55)	(2.34)	(0.20)	(0.20)	(0.22)
FD-GMM	0.736	0.416	-0.143	-1.007	1.576	-0.542	-3.819
	(7.46)	(4.41)	(3.82)	(4.97)	(2.36)	(2.50)	(3.62)
sys-GMM	1.138	0.005	-0.109	0.116	-	-	-
	(13.51)	(0.80)	(1.62)	(1.07)	-	-	-

Numbers in parentheses denote t-statistics from panel robust standard errors
^{b}Bootstrapped

via the consistent FD-GMM estimator[23].

We computed a total of 9 estimators for each fuel demand equation given in (5), (7) and (8). Anticipating the results, the estimated coefficients reveal the expected signs: positive income, negative price, and negative car-ownership effects. The results are discussed in more detail in the next section.

5 Discussion

5.1 Omitted variable bias

In order to quantify the bias caused by omitting diesel cars, we in addition estimate the gasoline demand equation (5) where the total car series is used instead of the correct gasoline-powered cars series (see Tab.2). As outlined in the appendix, this is expected to lead to overestimated coefficients in absolute terms compared to the gasoline demand

[23] Bruno (2005) implemented the routine *xtLSDVc* in STATA.

specification (7) because it ignores the increasing share of diesel-powered cars in Europe over the last ten years (see Fig.1). The resulting bias is due to an omitted variable bias.

Under the assumption that equation (7) comprehensively describes the major factors determining gasoline demand, we see from comparing Table 2 and 3 that there is clear overestimation of the short- and long-run elasticities in the misspecified gasoline specification (5) with few exceptions. For illustration, the Within estimator of the misspecified equation gives short-run estimates for the lagged dependent variable GAS_{t-1}, income Y_t, gasoline price PG_t, and total cars per driver CAR_t of 0.788, 0.257, -0.107 and -0.696 respectively, whereas these of the correctly specified equation (7) yield 0.705, 0.075, -0.106, and for the gasoline-fueled cars per driver series $CARG_t$ -0.228. Income and car ownership in particular are heavily overestimated in the short-run as well as in the long-run. In contrast, the difference in the short-run price elasticities is negligible. Generally speaking, the higher the share of diesel cars in the investigated countries the worse the overestimation of income, price and car ownership elasticities.

It may be of interest whether our estimates of the gasoline demand equation (5), which is misspecified with respect to omitting diesel cars, correspond to previous studies neglecting the growing share of diesel-fueled passenger cars, too. Baltagi and Griffin (1997) used a dynamic panel model with lagged income and car stock as additional regressors from 1960 to 1990 consisting of 18 OECD countries with 13 European countries. Best accordance was achieved by the estimates of the Within estimator. The short-run estimates for lagged gasoline consumption per car, income, price and car ownership were 0.87, 0.39, -0.11 and -0.74 respectively, which are largely in line with our results (see Table 1). Taking into account the fact that in some European countries like France the diesel car share was already substantial before 1990, we conclude that the estimates of the elasticities obtained by Baltagi and Griffin (1997) are to some extent overestimated. The same should hold true for the results obtained by Baltagi et al. (2003) using a panel from 21 French regions over the period 1973-1998.

5.2 Gasoline demand

We now turn to the parameter estimates of the gasoline demand in equation (7), including the diesel car series. The task is to compare the results for the different applied estimators in order to retrieve information about their usefulness in small samples.

The pooled OLS estimator is included for reasons of comparison. It yields the highest coefficient for the lagged dependent variable, which is in line with Baltagi and Griffin (1997) and Baltagi et al. (2003) because it is biased due to the omitted country-specific effects. The same holds true for the random effects estimator GLS assuming uncorrelated effects which was rejected by a Hausman-type test. Here we find a coefficient of 0.918. Both OLS and GLS yield insignificant income and price effects. In addition, the short-run elasticities of car ownership seem to be underestimated compared to the other estimators, and the long-run elasticities overestimated due to the higher dynamic coefficient.

The Within estimator cancels out the country-specific effects and their possible correlation with the explanatory variables. It is therefore a widely applied estimator. All coefficients are reasonable in magnitude as well as sign and are significant[24] except for the coefficient of income. The coefficient of lagged gasoline consumption shows pronounced habit persistence yielding a long-run response of 3.4 times the short-run elasticities[25]. The short-run price elasticity of 0.106 is quite low, which corresponds to previous studies. In contrast, the long-run response to gasoline price increase is much more inelastic than previously stated[26]. Further, the short-run coefficient of $CARG_t$ is below previous results. It may not capture the shift towards diesel-powered cars that was observed during the last few years in most European countries. Hence, a one percent increase in the number of gasoline-powered cars per driver will lower gasoline demand per gasoline car by only 0.23 percent, but in the long-run it will decrease by 0.77 percent. Since the stock of gasoline-fueled cars enters the equation both as a dependent and an independent variable, one can calculate the long-run transformed elasticity of gasoline consumption with respect to the gasoline car fleet $(1 + \delta)$, i.e. when a household purchases a second gasoline car *ceteris paribus*, total gasoline consumption will increase by approx. 23%, which is only half of the findings of Dahl and Sterner (1991) and Baltagi and Griffin (1997)[27]. Next, the coefficients of the variable diesel-powered passenger cars per driver, $CARD$, yield -0.05 and -0.17 for short-run and long-run responses. As mentioned earlier, the negative elasticity may stem from a kind of selection mechanism. Consumers and firms in most European countries have switched their car fleets to cheaper diesel-fueled cars over the last few years. The higher the car utilization, the more likely is the switch to a diesel auto, leaving more low-level users with gasoline-powered cars.

Due to serial correlation found in the residuals of standard models, the GLS-HC estimator introduced in the previous section is expected to perform well. The estimates are

[24] However, one has to take into account the fact that the size of the t-statistics is distorted for all considered estimators.

[25] This corresponds to the findings of Bohi and Zimmermann (1984) and Dahl and Sterner (1991), i.e. approx. 3.3, but not to those of Baltagi and Griffin (1997), 7.7, and Baltagi et al. (2003), 4.7. One explanation could be that in studies covering older periods the share of diesel-powered cars was negligible even in European countries, whereas in the latter studies the omitted variable bias became substantial.

[26] Except for the results of Baltagi et al. (2003), which were in our range.

[27] Baltagi et al. (2003), see Table 2 there, find a decrease of approx. 13 percent, which is obviously due to their omitting the diesel-powered car series.

highly significant and correspond in magnitude to those of the Within estimator. The coefficient of the lagged dependent variable, 0.61, is somewhat smaller than the Within and LSDVc estimates but in line with the consistent FD-GMM estimates, though the latter also suffer from a negative small sample bias. Owing to its very small standard errors, this GLS estimator accounting for a flexible error structure seems to be appropriate for our panel model, confirming the findings of Baltagi and Griffin (1997). As Bun and Kiviet (2006) demonstrate, the small sample bias of the standard errors of a feasible GLS estimator is in line with that of other estimators considered.

In order to account for the endogeneity problem of the lagged gasoline consumption variable $GASG_{t-1}$, four common instrumental variable estimators were applied. The Within-2SLS estimator yields a low estimate of the lagged dependent variable, whereas the short-run elasticities reveal some overestimation compared to the standard Within counterpart[28]. The same holds true for the Anderson-Hsiao estimator FD-2SLS, where the dynamic coefficient is indeed negative and insignificant.

In contrast, the one-step Arellano-Bond estimator (FD-GMM) yields reasonable estimates. The coefficient of the lagged gasoline demand, 0.676, is somewhat lower than that obtained with the LSDV counterparts, whereas the estimated short-run elasticities of income and car stock are higher. This leads to higher long-run elasticities than the LSDV and GLS-HC counterparts but remains reasonable[29].

Turning to the Blundell-Bond estimator (sys-GMM), we observed a high dynamic coefficient when applying the full set of moment conditions, which quickly increases with the number of regressors. In small samples this may cause an overfit of the instrumented variable, biasing the estimates towards that of OLS. We therefore restricted the instrument set, which decreased the estimate of the dynamic coefficient to a reasonable value, 0.797, however the standard errors of the estimated elasticities still remained high, leading to insignificant estimates.

Alternatively, one can exploit the efficiency property of the Within estimator by using a bias-corrected version. Because the FD-2SLS estimator revealed unreasonable estimation results, we used the FD-GMM estimator for the initial values in the LSDVc approximation procedure. Comparing LSDVc with the Within estimator in Table 3, one can see that Nickell bias correction leads to a higher coefficient estimate of the lagged dependent variable and to somewhat lower short-run elasticity estimates in absolute values, resulting in higher long-run elasticities. The estimates are realistic and in line

[28] Similarly, Baltagi and Griffin (1997) found that 2SLS pooled estimates of this coefficient are generally lower than pooled OLS, GLS or LSDV.

[29] Except the long-run elasticity of $CARG$, -1.109, which yields a negative transformed long-run elasticity of gasoline consumption with respect to gasoline-powered cars per driver.

Table 3: Elasticity estimates of gasoline demand equation (7)

	Short-run					Long-run			
	$GASG_{t-1}$	Y_t	PG_t	$CARG_t$	$CARD_t$	Y_t	PG_t	$CARG_t$	$CARD_t$
Exogenous regressors									
OLS	0.932	0.036	-0.029	-0.062	-0.019	0.524	-0.416	-0.902	-0.275
	(53.00)	(1.46)	(1.31)	(3.48)	(4.24)	(1.58)	(1.35)	(3.77)	(5.18)
Within	0.705	0.075	-0.106	-0.228	-0.049	0.254	-0.360	-0.774	-0.166
	(12.33)	(1.45)	(3.63)	(2.75)	(3.64)	(1.59)	(4.22)	(3.07)	(4.59)
LSDVc	0.780	0.065	-0.090	-0.208	-0.037	0.295	-0.408	-0.943	-0.169
	$(15.28)^b$	$(1.49)^b$	$(2.96)^b$	$(3.93)^b$	$(2.77)^b$	$(1.43)^b$	$(2.88)^b$	$(3.21)^b$	$(3.01)^b$
GLS	0.918	0.040	-0.031	-0.075	-0.022	0.489	-0.370	-0.911	-0.270
	(43.86)	(1.51)	(1.36)	(3.33)	(4.37)	(1.69)	(1.41)	(3.84)	(5.48)
GLS-HC	0.611	0.081	-0.122	-0.250	-0.025	0.209	-0.314	-0.641	-0.180
	(96.71)	(9.55)	(68.22)	(47.50)	(2.08)	(9.26)	(65.90)	(47.48)	(27.54)
Endogenous regressors									
Within-2SLS	0.331	0.111	-0.190	-0.271	-0.105	0.166	-0.84	-0.405	-0.156
	(2.29)	(1.70)	(4.20)	(2.52)	(4.19)	(1.80)	(6.20)	(2.64)	(8.22)
FD-2SLS	-0.210	0.235	-0.169	-0.625	-0.206	-	-	-	-
	(0.28)	(1.57)	(4.28)	(1.80)	(2.00)	-	-	-	-
FD-GMM	0.676	0.155	-0.102	-0.360	-0.073	0.477	-0.316	-1.109	-0.226
	(8.52)	(1.30)	(3.32)	(1.70)	(2.28)	(1.57)	(3.00)	(2.36)	(4.02)
sys-GMM	0.797	-0.020	-0.024	-0.089	-0.044	-0.099	-0.120	-0.436	-0.217
	(3.32)	(1.06)	(0.72)	(1.05)	(0.94)	(3.87)	(0.43)	(2.78)	(6.82)

Numbers in parenthereses denote t-statistics from panel robust standard errors

^{b}Bootstrapped

Table 4: Elasticity estimates of diesel demand equation (8)

	Short-run					Long-run			
	$DIESD_{t-1}$	Y_t	PD_t	$CARG_t$	$CARD_t$	Y_t	PD_t	$CARG_t$	$CARD_t$
Exogenous regressors									
OLS	0.859	0.139	-0.142	-0.091	-0.122	0.988	-1.005	-0.645	-0.861
	(30.03)	(3.65)	(4.95)	(2.80)	(5.48)	(4.26)	(4.28)	(3.04)	(19.22)
Within	0.511	0.670	-0.132	-0.347	-0.404	1.368	-0.271	-0.709	-0.825
	(8.86)	(7.12)	(4.04)	(4.04)	(8.92)	(8.05)	(3.92)	(3.76)	(19.37)
LSDVc	0.561	0.631	-0.128	-0.349	-0.370	1.438	-0.291	-0.794	-0.842
	$(11.24)^b$	$(7.66)^b$	$(3.30)^b$	$(3.84)^b$	$(10.03)^b$	$(7.32)^b$	$(3.29)^b$	$(3.50)^b$	$(15.12)^b$
GLS	0.754	0.288	-0.148	-0.180	-0.199	1.172	-0.604	-0.733	-0.811
	(18.22)	(4.99)	(4.42)	(3.84)	(6.42)	(6.27)	(4.11)	(3.78)	(20.78)
GLS-HC	0.409	0.769	-0.129	-0.378	-0.464	1.303	-0.218	-0.640	-0.785
	(97.81)	(102.68)	(82.86)	(45.30)	(179.42)	(85.86)	(58.95)	(38.67)	(419.45)
Endogenous regressors									
Within-2SLS	0.094	1.061	-0.186	-0.422	-0.696	1.171	-0.205	-0.465	-0.768
	(0.97)	(8.41)	(4.82)	(3.16)	(9.59)	(10.69)	(5.15)	(3.19)	(25.17)
FD-2SLS	1.723	0.022	-0.013	-0.450	0.298	-	-	-	-
	(0.35)	(0.01)	(0.09)	(0.70)	(0.10)	-	-	-	-
FD-GMM	0.316	0.879	-0.111	-0.367	-0.566	1.284	-0.162	-0.536	-0.827
	(2.37)	(5.07)	(2.96)	(1.64)	(5.87)	(6.65)	(2.10)	(1.76)	(14.58)
sys-GMM	0.952	-0.012	-0.081	-0.028	-0.046	-0.237	-1668	-0.580	-0.953
	(25.27)	(3.48)	(3.00)	(0.79)	(1.60)	(1.40)	(1.26)	(0.79)	(4.55)

[30]umbers in parenthereses denote t-statistics from panel robust standard errors

^{b}Bootstrapped

with Within and GLS coefficient estimates with one exception: the long-run elasticity of gasoline car ownership amounts here to -0.94, which results in an implausibly low transformed elasticity of 0.06, substantially lower than the uncorrected counterpart.

Recapitulating, the IV/GMM estimators considered in this small sample dynamic panel model produce somewhat unstable estimates. The one-step Arellano-Bond estimator is the only one of the four IV/GMM estimators that seems to perform well in our setting. Our results show that good candidates for estimating small dynamic panel data sets are the Within estimator, its bias-corrected version LSDVc, GLS with heteroscedastic, cross-sectional correlation and/or AR(1) error structure, and the Arellano-Bond estimator FD-GMM.

5.3 Diesel demand

Next, we apply the set of standard panel estimators to the above-specified diesel demand in equation (8) (see Table 4). Again, the estimates of the IV/GMM estimators vary considerably. The dynamic coefficient estimate of the Within-2SLS estimator tends towards zero, that of the system GMM estimator, like OLS, towards one, and the FD-2SLS estimate lies outside the unit circle but is insignificant. Only the Arellano-Bond estimator provides elasticities in the range of the preferable Within, LSDVc and GLS-HC estimates. However, the estimated dynamic coefficient, 0.32, appears to be unrealistically low, therefore we compare the estimates of the diesel and gasoline demand equations on the basis of the Within estimator.

First, the estimated coefficient of the lagged diesel consumption variable amounts to 0.511, which entails a lower habit persistence than for the gasoline demand equation. The long-run responses are only twice the size of the short-run elasticities. Whereas the estimated income elasticities of the gasoline demand equation are low and mostly insignificant, the estimates given in Table 4 suggest a strong dependence of diesel consumption on income. Because the variable Y is measured as GDP per capita, income can be interpreted as economic income or activity, so Y captures the effect of the number of diesel-powered trucks, buses and other heavy-load vehicles on diesel demand[31].

The short-run diesel price elasticity is higher than that of the gasoline price but still quite inelastic. This makes sense if we assume a higher share of diesel cars in the stock of vehicles with intensive utilization due to economic activities. Firms are expected to react more economically to a fuel price increase than consumers with low car usage. In contrast, the long-run response to gasoline price increases is more elastic in the gasoline

[31] This could explain why the inclusion of the *trucks per capita* series affords no improvement in estimating the diesel demand equation. See also Baltagi and Griffin (1983), footnote 5.

demand equation (see Table 3) than the response to diesel price increases as defined in the diesel demand equation (8), being in accordance with economic behaviour. In the long-run, high-usage consumers and firms shift their car fleets towards lower cost engines with increasing fuel prices. After switching, the users react to changes in income and diesel price with more elastic short-run and more inelastic long-run responses than those for the gasoline demand specification.

6 Summary and conclusions

The share of diesel-fueled passenger cars in the total car stock has continuously increased during the last decade especially in Europe. Previous studies on gasoline demand have not accounted for this fact by relating gasoline consumption to the total number of cars. This paper argues that estimates based on such a misspecified demand equation are biased owing to the related omitted variable problem. We thus expect estimates of income, price and car ownership elasticities reported in recent studies covering gasoline demand to be overstated. In particular, the elasticity of gasoline consumption with respect to car ownership is overestimated when no distinction is made between gasoline and diesel-powered cars.

To our small 14x15 panel data set we applied 9 common dynamic panel estimators. Three of the four IV/GMM estimators considered conveyed unreliable estimates. The quite good performance of the Andersen-Hsiao (1982) estimator in small samples[32] could not be affirmed by our study. The Blundell-Bond (1998) estimator suffers from a too large set of instruments relative to the sample size. Only the one-step Arellano-Bond (1991) estimator seems to perform well in our small sample setting. Our results show that good candidates for estimating small dynamic panel data sets with a dynamic coefficient close to one are the standard Within estimator, its bias-corrected version LSDVc, GLS with heteroscedastic, cross-sectional correlation and/or AR(1) error structure, and the Arellano-Bond estimator FD-GMM.

The qualitative results of the estimation of the gasoline demand equation correspond to those obtained by previous studies like Baltagi and Griffin (1983, 1997) and Baltagi et al. (2003). The income elasticity is positive, the price effect is negative, and the effect of increased car ownership on gasoline consumption is negative. Comparing the Within estimates of these studies, however, the coefficient estimates in this paper are found to be somewhat lower in absolut terms, depending on the specified model and the included countries, as well as on time period. This is accredited to the omitted variable bias

[32] See for instance Kiviet (1995) and Judson and Owen (1999).

effect, from which the estimate of the gasoline-fueled cars per driver variable especially suffers. We interpret its low value together with an inelastic short-run gasoline price response as such that car owners react to increasing fuel prices by gradually replacing their gasoline-powered cars with diesel-powered ones. In most European countries such a shift towards diesel cars can be observed since the 1990s. When fuel supply tightens, the income and price elasticities of diesel consumption should therefore display relatively more elastic responses than in the gasoline case. Exactly this can be observed in the estimation results.

Acknowledgement:
I would like to thank the following institutions for kindly supplying me with data: ARBÖ - Auto-, Motor- und Radfahrerbund Österreichs (AT); ÖAMTC - Österreichische Automobil-, Motorrad- und Touring Club (AT); FVMI-Fachverband der Mineralölindustrie Österreichs (AT); DESTATIS - Statistisches Bundesamt Deutschland (DE); DBI - De Danske Bilimportører (Danish Car Importers Association); DGT - Dirección General de Tráfico, Servicio de Estadística (ES); Tilastokeskus - Statistics Finland; CCFA - Comité des Constructeurs Français d'Automobiles (FR); ACI - Automobile Club d'Italia; ANFIA - Associazione Nazionale fra Industria Automobilistiche (IT); CBS - Centraal Bureau voor de Statistiek (NL); SSB - Statistik Sentralbyrå (NO); SCB - Statistiska Centralbyrån (SE); DfT - Department for Transport (UK).

A Appendix

A.1 Data description

The data framework originates from EUROSTAT. We completed or substituted individual items of data whenever necessary and available from national statistics agencies and ministries as well as local automobile associations. Due to data limitations, Greece and Portugal were excluded from the panel set. The final balanced panel comprises data for the period 1990 till 2004 from 14 European countries: Austria (AT), Belgium (BE), Switzerland (CH), Germany (DE), Denmark (DK), Spain (ES), Finland (FI), France (FR), Ireland (IE), Italy (IT), the Netherlands (NL), Norway (NO), Sweden (SE) and the United Kingdom (UK).

The stock of passenger cars is given as an annual average or for a certain point during the year, mainly per end of December. The official data of the car series for Germany suffer from breaks due to the German reunification in 1991 and a switch in the reference date in 2001. We reconstructed our time series out of data from DESTATIS and EUROSTAT, smoothing out the switch and updating the series backwards from 1991 using the growth rates of the original series.

As a proxy for the average annual number of drivers within a country we used end-of-year numbers from EUROSTAT population data for the 18 to 69 age group. The use of the number of drivers instead of the country's total population avoids demographic effects.

Annual data series for purchasing power parity (PPP) in US dollars and real GDP adjusted by PPP, basis 2000, for each investigated country in US dollars were taken from the OECD Economic Outlook No. 76 and Annual National Accounts.

For annual gasoline and diesel consumption we used EUROSTAT-data on final energy consumption of road traffic in thousand tonnes per year.

EUROSTAT also provided data on gasoline and diesel retail prices per 100 liters inclusive of taxes, averaged per year. In order to account for the gradual displacement of leaded gasoline fuel by unleaded (the first showing a price mark-up) in most European countries during the early 1990s, we constructed the gasoline price series by means of a stepwise changing weighted average of the two fuel sorts from 1990 to 1994. By the end of 1999, leaded fuel had been withdrawn in most European countries.

The relative fuel prices were calculated with the help of national consumer price indices (CPI), basis 1995, retrieved from AMECO, the annual macro-economic database of the European Commission's Directorate General for Economic and Financial Affairs (DG ECFIN). Note that we do not adjust relative fuel prices for PPP, as we presume that consumer decisions are affected by local prices rather than by those that prevail in a reference country.

Out of this we constructed the series for estimation: gasoline consumption per total number of cars in logarithm $\ln(GAS)$, gasoline consumption per gasoline-powered car in logarithm $\ln(GASG)$, diesel consumption per diesel-powered car in logarithm $\ln(GASD)$, PPP-adjusted real output per capita in US$ in logarithm $\ln(Y)$, number of total cars per capita in logarithm $\ln(CAR)$, number of gasoline-powered cars per capita in logarithm $\ln(CARG)$, number of diesel-powered cars per capita in logarithm $\ln(CARD)$, the relative price of gasoline to other goods $\ln(PG)$, and the relative price of diesel to other goods $\ln(PD)$.

A.2 Omitted variable bias

Here, we illustrate the omitted variable bias effect of the Within estimator due to the incorrect use of the number of total passenger cars C instead of gasoline-powered cars CG in the variables gasoline consumption per gasoline-powered car $GASG$ and gasoline cars per adult $CARG$ in the gasoline demand specification given in equation (7).

First, we define $D_t = C_t/CG_t$ capturing the growth in the numbers of diesel-fueled passenger cars over time as the ratio between total number of cars and number of gasoline passenger cars. In other words, D is the inverse of the share of gasoline-powered cars. Substituting CG for C/D in the variables $GASG$ and $CARG$ in equation (7) and rearranging the terms, one gets:

$$\begin{aligned} \ln GAS_{i,t} &= \alpha \ln GAS_{i,t-1} + \beta \ln Y_{i,t} + \gamma \ln PG_{i,t} + \delta \ln CAR_{i,t} + \mu_i + \epsilon_{i,t} + \\ &\quad \phi \ln CARD_{i,t} + \alpha \ln D_{i,t-1} - (1+\delta) \ln D_{i,t} \end{aligned} \tag{9}$$

where the variable GAS is gasoline consumption per total cars. Comparing equation (5) with (7) one notes that the term in the second line of equation (9) captures the effect of the omitted variables $CARD_t$ and D, hidden in the dependent variable as well as in the two regressors GAS_{t-1} and CAR, total cars per driver. Note that (9) differs from (7) by a slight change in the coefficient notation.

Further, applying within transformation in order to eliminate the individual fixed effects, stacking the observations over time and across countries, and collating the logarithmized exogenous variables into the NTxK matrix $\tilde{X} = [\widetilde{Y} \vdots \widetilde{PG} \vdots \widetilde{CAR}]$ yields:

$$\tilde{y} = \tilde{W}\boldsymbol{\kappa} + \tilde{\epsilon} + \phi \widetilde{CARD} + (\alpha L - 1 - \delta)\tilde{D} \tag{10}$$

with the coefficient vector $\boldsymbol{\kappa} = (\alpha, \beta, \gamma, \delta)'$, the lag-operator L, the NTx1 matrices $\tilde{y} = (I_N \bigotimes (I_T - i_T i_T'/T) \ln GAS$, $\tilde{\epsilon}$ and $\tilde{D}$, and the NTx$(K+1)$ matrix $\tilde{W} = [\tilde{y}_{-1} \vdots \tilde{X}]$, respectively.

Applying OLS to the above equation without the last term yields the Within or least-square-dummy-variable estimator LSDV:

$$\begin{aligned} \hat{\boldsymbol{\kappa}}_{\text{LSDV}} &= (\tilde{W}'\tilde{W})^{-1}\tilde{W}'\tilde{y} \\ &= \boldsymbol{\kappa} + (\tilde{W}'\tilde{W})^{-1}\tilde{W}'\tilde{\epsilon} + (\tilde{W}'\tilde{W})^{-1}\tilde{W}' \left[\phi \widetilde{CARD} + (\alpha L - \delta - 1)\tilde{D}\right] \end{aligned} \tag{11}$$

with the true parameters $\boldsymbol{\kappa}$. Equation (11) implies that the LSDV estimator is biased due to the so-called Nickell bias given by the second term on the RHS, and in addition through the correlation of the omitted variables $\widetilde{CARD}$ and $\tilde{D}$ with the regressor matrix $\tilde{W}$ represented by the third term.

Nickell (1981) examines the bias of the LSDV estimator with exogenous regressors for $N \to \infty$ and finite T:

$$\text{plim}_{N\to\infty}(\hat{\alpha} - \alpha) = (\text{plim}_{N\to\infty}\frac{1}{N}\tilde{y}_{-1}' M_{\tilde{X}} \tilde{y}_{-1})^{-1} \text{plim}_{N\to\infty}\frac{1}{N}\tilde{y}_{-1}'\tilde{\epsilon} \tag{12}$$

and

$$\text{plim}_{N\to\infty}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = -\text{plim}_{N\to\infty}[(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y}_{-1}]\text{plim}_{N\to\infty}(\hat{\alpha} - \alpha) \tag{13}$$

with the projection matrix $M_{\tilde{X}} = I - \tilde{X}(\tilde{X}'\tilde{X}')^{-1}\tilde{X}'$ and $\boldsymbol{\beta} = (\beta, \gamma, \delta)'$. The bias of α given in equation (12) is of order $O(T^{-1})$ and negative for positive values of α. For the AR(1) panel model the simplest bias approximation is given by[33]

$$-\frac{1+\alpha}{T-1}, \tag{14}$$

Hence, the absolute value of the bias in estimating α increases with α and gets quite large as $\alpha \to 1$. The presence of exogenous regressors influences the bias given by equation (12) and (13). Phillips and Sul (2004) proved that the inconsistency of the parameter estimates is decreased in absolute values when exogenous variables are present. This result is in contrast to Nickell (1981). Also, the bias on $\hat{\boldsymbol{\beta}}$ depends on the relationship between the transformed exogenous variables and $\tilde{y}_{-1}$. If there is positive (negative) correlation between the projected variables, then equation (13) indicates that its coefficient will be upward (downward) biased. For our data set, this means that the Nickell-bias

[33] See Nickell (1981), formula (19). With $\alpha = 0.9$ and $T = 15$ this amounts to -0.136.

induces the estimated income, price and car ownership elasticities in the demand models of equations (7) and (8) to be over-, under- and over-estimated, respectively, in absolute terms.

Kiviet (1995) derived approximation formula for the bias of the LSDV estimator with strictly exogenous regressors but when both N and T are small. Subtracting the approximated bias from the LSDV estimate one arrives at the corrected LSDV estimator which performed well in simulation studies[34]. This estimator has also been applied to our data set (see section 5).

The following discusses the effects of neglecting the diesel car series on the parameter estimates, as it is stated in the text. Extending the third term on the right hand side of equation (11), the omitted variable bias of the estimates in the specific gasoline demand equation (9) is given by:

$$\phi(\tilde{W}'\tilde{W})^{-1}\tilde{W}'\widetilde{CARD} + \alpha(\tilde{W}'\tilde{W})^{-1}\tilde{W}'\tilde{D}_{-1} - (1+\delta)(\tilde{W}'\tilde{W})^{-1}\tilde{W}'\tilde{D} \tag{15}$$

To size up the direction of the expected bias of the estimated coefficients in our specific data set we have to evaluate the covariances between the omitted variables and the corresponding transformed regressor. For instance, in equation (15) the first element of the last term is equivalent to $\text{cov}(M_{\tilde{X}}\tilde{y}_{-1}, M_{\tilde{X}}\tilde{D})/\text{var}(M_{\tilde{X}}\tilde{y}_{-1}) = -0.54$, with the projection matrix $M_{\tilde{X}}$ defined above. Because the true value of δ probably lies around -0.2, the bias of the estimated coefficient α, which is due to the omission of $\tilde{D}$ in the demand equation, is positive. Calculating the empirical counterparts of all expressions in (15) along the same lines and assuming true values for $\alpha = 0.8$ and $\phi = -0.05$, we are able to evaluate the direction of the omitted variabel bias for each estimated coefficient. Depending on the resulting sign this bias augments or reduces the Nickell bias given by equation (12) and (13).

The omitted variable bias of the lagged dependent variable coefficient $\hat{\alpha}$ is small and positive, and thus moderately mitigates the negative Nickell bias given in equation (14). The omitted variable effect by itself results in an overestimation of $\hat{\alpha}$. In the case of the coefficient estimate of the gasoline price variable, the low and negative omitted variable bias only marginally reduces the Nickell bias and hence can be ignored. In contrast, the omitted variable bias causes a considerable overestimation in absolute terms of the negative coefficient of the variable cars per driver $\widetilde{CAR}$ and to a lesser extent that of the positive coefficient of the income variable $\tilde{Y}$. Consequently, this augments the negative and positive Nickell bias of $\hat{\delta}$ and $\hat{\beta}$, respectively.

[34] See for instance Kiviet (1995), Judson and Owen (1999), and Bun and Kiviet (2003).

Alternatively, the extent and direction of the omitted variable bias and the small sample bias in dynamic panels can be read off from the Within and corrected LSDV coefficient estimates of the gasoline demand equations (5) and (7). Given that the LSDVc estimator fully corrects for the Nickell bias, the latter can be extracted from the difference between the Within and LSDVc estimates of the correct specified gasoline demand equation (7). For instance, the difference in the estimates of the lagged dependent variable yields -0.075 (see Table 3), which is nearly half of the Nickell bias calculated by the approximation formula in equation (14). The pure omitted variable bias is obtained by taking the differences between the LSDVc estimates of the equation (5) and (7). For the coefficients α, β, γ and δ one obtains 0.042, 0.166, -0.007 and -0.412, respectively (see Table 2 and 3). Thus, by not taking into account the recent strong development of the diesel-fueled cars in the common dynamic gasoline demand equation, the resulting omitted variable bias causes an overestimation in absolute terms of all coefficient estimates. The elasticity of the car fleet per driver is in particular affected, whereas that of the gasoline price is negligible. Further, the omitted variable bias mitigates the Nickell bias in the case of the lagged dependent variable. In contrast, the overall bias of the car fleet elasticity is considerably aggravated by its omitted variable bias.

References

[1] ACEA (2006): European Union Economic Report, June 2006. European Automobile Manufacturers Association, Brussels. http://www.acea.be/ASB20/axidownloads20s.nsf/CategorizedOverviewACEA?OpenForm&Language=English&cat1=2&cat2=9

[2] Anderson, T.W., Hsiao, C. (1982): Formulation and estimation of dynamic models using panel data. *Journal of Econometrics, 18*, 47-82.

[3] Arellano, M., Bond, S. (1991): Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277-297.

[4] Baltagi, B.H. (2005): *Econometric analysis of panel data*, John Wiley & Sons Ltd, Chichester.

[5] Baltagi, B.H., Bresson, G., Griffin, J.M., Pirotte, A. (2003): Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption. *Empirical Economics, 28*, 795-811.

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

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

[8] Blundell, R., Bond, S. (1998): Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115-143.

[9] Bohi, D.R., Zimmermann, M. (1984): An update on econometric studies of energy demand. *Annual Review of Energy, 9*, 105-154.

[10] Bruno, G.S.F. (2005): Estimation and inference in dynamic unbalanced panel data models with a small number of individuals. *CESPRI WP n.165*, Universit Bocconi-CESPRI, Milan.

[11] Bun, M.J.G., Kiviet, J.F. (2001): The accuracy of inference in small samples of dynamic panel data models: simulation evidence and empirical results. *Tinbergen Institute discussion paper*, TI 2001-006/4.

[12] Bun, M.J.G., Kiviet, J.F. (2003): On the diminishing returns of higher-order terms in asymptotic expansions of bias. *Economics Letters, 79*, 145-152.

[13] Bun, M.J.G., Kiviet, J.F. (2006): The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of Econometrics, 132*, 409-444.

[14] Dahl, C.A. (1986): Gasoline demand survey. *The Energy Journal, 7*, 67-82.

[15] Dahl, C.A., Sterner, T. (1991): Analyzing gasoline demand elasticities: A survey. *Energy Economics, 11*, 203-210.

[16] EC-ACEA (2003): Monitoring of ACEAs commitment on CO_2 emission reduction from passenger cars (2002). Joint Report of the European Automobile Manufacturers Association and the Commission Services, Brussels. http://www.acea.be/ASB20/axidownloads20s.nsf/CategorizedOverviewACEA?OpenForm&Language=English&cat1=2&cat2=7&cat3=15

[17] EPA (2002): Health assessment document for diesel engine exhaust. Prepared by the National Center for Environmental Assessment. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC, EPA/600/8-90/057F. http://epa.gov/ttn/atw/dieselfinal.pdf

[18] Houthakker, H.S., Taylor, D. (1970): *Consumer demand in the United States, analyses and projections*. Harvard University Press, Cambridge, Massachusetts.

[19] Houthakker, H.S., Verleger, P.K., Sheehan, D. (1974): Dynamic demand for gasoline and residential electricity. *American Journal of Agricultural Economics, 56*, 412-418.

[20] Judson, R.A., Owen, A.L. (1999): Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters, 65*, 9-15.

[21] Kiviet, J.F. (1995): On bias, inconsistency and efficiency of some estimators in dynamic panel data models. *Journal of Econometrics, 68*, 53-78.

[22] Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, Jr.F., Puybonnieux-Texier, V., Quénel, P., Schneider, J., Seethaler, R., Vergnaud, J.C., Sommer, H. (2000): Public-health impact of outdoor and traffic-related air pollution: a European assessment. *The Lancet, 356*,9232: 795-801.

[23] Maddala, G.S., Srivastava, V.K., Li, H. (1994): Shrinkage estimators for the estimation of short-run and long-run parameters from panel data models. Ohio State University, Columbus, Working Paper.

[24] Maddala, G.S., Trost, R.P., Li, H., Joutz, F. (1997): Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators. *Journal of Business and Economic Statistics, 15*, 90-100.

[25] Nickell, S. (1981): Biases in dynamic models with fixed effects. *Econometrica, 49*, 1417-1426.

[26] Orfeuil, J. (2001): Tax regimes on cars and CO_2 emissions in European countries. OEIL - Observatoire de l'Économie et des Institutions Locales, Université de Paris XII, France. http://www.univ-paris12.fr/servlet/com.univ.utils.LectureFichierJoint?CODE=1136791632612&LANGUE=0

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

[28] Phillips, P.C.B., Sul, D. (2004): Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Cowles Foundation for Research in Economics, Yale University, discussion paper No.1438.

[29] Pope, C.A., Burnett, R.T., Thun, M.J, Calle, E.E., Krewski, D., Ito, K., Thurston, G.D. (2002): Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association 287*,9: 1132-1141.

[30] Proost, S., Mayeres, I. (2001): Should diesel cars in Europe be discouraged?. FEEM Working Paper No. 1.2001. http://ssrn.com/abstract=268488

[31] Roodman, D. (2005): XTABOND2: Stata module to extend xtabond dynamic panel data estimator. Center for Global Development, Washington. http://econpapers.repec.org/software/bocbocode/s435901.htm

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

[33] Statistics Austria (2001): Energieeinsatz der Haushalte. *Statistische Nachrichten 10/2001*: 789-794.

[34] 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.

[35] WHO (2003): Health aspects of air pollution with particulate matter, ozone, and nitrogen dioxide. World Health Organization, Bonn, Germany. http://www.euro.who.int/document/e79097.pdf

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

Author: Markus Pock

Title: Gasoline and Diesel Demand in Europe: New Insights

Reihe Ökonomie / Economics Series 202

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 59991-555 • http://www.ihs.ac.at

ISSN: 1605-7996