



universität  
wien

# DIPLOMARBEIT

Titel der Diplomarbeit

"A comparative study on gasoline demand, using  
cointegration and vector error correction."

Verfasser

Norbert Ernst

angestrebter akademischer Grad

Magister der Sozial- und  
Wirtschaftswissenschaften (Mag.rer.soc.oec.)

Wien, im März 2010

Studienkennzahl lt. Studienblatt:

A140

Studienrichtung lt. Studienblatt:

Diplomstudium Volkswirtschaft

Betreuer:

Univ. Prof. Dr. Robert Kunst

## *ACKNOWLEDGEMENTS*

I want to thank several persons, without their help the completion of this work would have been impossible.

First of all my family, who gave me through all the years the necessary backup. I am also very grateful to Univ. Prof. Dr. Robert Kunst, for taking the task of assisting me in my work and giving me outstanding support. Furthermore i want to mention MMag. Markus Pock, who generously put his data set from Pock (2007) at my disposal.

I would also like to thank Thomas Kemetmüller, Matthias Schnetzer, Andreas Wolf, Gerald Stampfel, Mario Müller, Willi Waitz, Claudia Kemetmüller, Benedikt Sagl, Sebastian Wojta, Martin Ludwig, Khalid Shoumariyeh and Sarah Ott for their friendship and moral support.

And last but not least Miriam Van den Nest for her inspiration.

# Contents

|          |                                                  |           |
|----------|--------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                              | <b>5</b>  |
| <b>2</b> | <b>Cointegration and Vector error correction</b> | <b>8</b>  |
| 2.1      | Concept of cointegration . . . . .               | 9         |
| 2.2      | Definitions and Explanations . . . . .           | 9         |
| 2.3      | Vector error correction . . . . .                | 12        |
| 2.4      | Two step estimator . . . . .                     | 13        |
| 2.5      | Maximum likelihood estimator . . . . .           | 14        |
| 2.6      | A tale of two Econometricians . . . . .          | 15        |
| <b>3</b> | <b>Gasoline elasticities in the U.S.A.</b>       | <b>20</b> |
| 3.1      | Model specification . . . . .                    | 20        |
| 3.2      | Data . . . . .                                   | 22        |
| 3.3      | Estimations & Results . . . . .                  | 22        |
| <b>4</b> | <b>Gasoline elasticities in Europe</b>           | <b>31</b> |
| 4.1      | European Data . . . . .                          | 33        |
| 4.2      | Model Estimation . . . . .                       | 34        |
| <b>5</b> | <b>Conclusion</b>                                | <b>44</b> |

# List of Figures

|     |                                                           |    |
|-----|-----------------------------------------------------------|----|
| 3.1 | Real Gasoline Price . . . . .                             | 21 |
| 3.2 | Long-term equilibrium relationship, 1975 - 1981 . . . . . | 23 |
| 3.3 | Long-term equilibrium relationship, 2001 - 2007 . . . . . | 28 |
| 3.4 | VEC 1975 - 1981 . . . . .                                 | 29 |
| 3.5 | VEC 2001 - 2007 . . . . .                                 | 30 |
| 4.1 | Total Gasoline Demand . . . . .                           | 32 |
| 4.2 | Gasoline demand U.S.A. . . . .                            | 38 |
| 4.3 | Gasoline elasticity Germany . . . . .                     | 39 |
| 4.4 | Gasoline elasticity UK . . . . .                          | 40 |
| 4.5 | VEC Germany . . . . .                                     | 41 |
| 4.6 | VEC UK . . . . .                                          | 42 |
| 4.7 | VEC Germany . . . . .                                     | 43 |

# List of Tables

|     |                                        |    |
|-----|----------------------------------------|----|
| 2.1 | Results of estimations (2.23) - (2.26) | 17 |
| 2.2 | Results of estimations (2.27) - (2.30) | 18 |
| 3.1 | Cointegration coefficients             | 25 |
| 3.2 | VEC coefficients                       | 25 |
| 3.3 | Johansen Cointegration coefficients    | 26 |
| 3.4 | Johansen Error correction coefficients | 26 |
| 4.1 | Orders of integration                  | 35 |
| 4.2 | 95% confidence interval                | 36 |
| 4.3 | Results of the unit root test          | 36 |
| 4.4 | Maximum likelihood estimates           | 38 |

# Chapter 1

## Introduction

After the oil price shocks in the 1970's and beginning 1980's, a lot of empirical research on consumer behavior and demand elasticities has been done. Espey (1998) provides a detailed analysis of past studies on gasoline demand. Since the recent shocks in the oil price this topic becomes more important in economic debates. On one hand for ecological reasons. The debate of global warming seems to become more important for policy makers and the coefficient of demand elasticity is crucial when it comes to the decision of taxation. But gasoline is also a good which is used by many people in every day life. Bearing in mind the extent of the media reports during the price peaks in the past years, the price of gasoline seems to be of vital interest for a large part of the population.

In this paper we investigate whether there has been a change in gasoline demand elasticities of the U.S.A., and if gasoline elasticities vary between different European countries. Austria, Germany and the UK. Although many studies exist on this topic, a comparison of them is difficult because the dataset and the estimated models vary extensively. We use a coherent dataset for the whole period 1975 - 2007 for the US, and 1990 - 2004<sup>1</sup> for the European countries.

---

<sup>1</sup>As will be explained later in more detail we use a dataset supplied by Pock (2007)

Analyzing more than 100 studies and distinguishing between ten different categories of models, Dahl and Sterner (1991) conclude that income and price are the most important factors influencing gasoline demand. We build on these findings and construct our models accordingly.

As it will turn out during our investigation the time series we use conform to integration of first or higher order, therefore the method of cointegration is used to produce valid results

A special feature of this method is, that the result of a cointegration equation is a long-term equilibrium relationship and the implied vector error correction model gives the speed of adjustment to this equilibrium in case of a shock. So this method provides us with the possibility to model both long run and short run effects.

In the first section of this work we give an introduction to the method of cointegration and present the two estimators which are used during this study. Additionally we show difficulties and possible traps one has to deal with, when working with non-stationary data.

In the second part we investigate if the gasoline elasticities for the U.S.A. have changed or not. We chose the time periods 1975 - 1981 and 2001 - 2007. Both periods contain similar shocks and the distance between them is large enough that a change could be expected.

In the third part of the study we compare national differences between Austria, Germany and the UK. Different policies regarding public transport and taxation on the different kind of fuels are expected to produce different responses for each country. As Pock (2007) showed, a major point when analyzing gasoline demand in Europe is the spread of diesel powered cars. We stick to our model presented above and do not include variables to capture the effect of diesel and look if it is still possible to receive reasonable results. Due to data problems we are bound in this part of our investigation to the period 1990 - 2004, so the price turbulences which peaked in the summer 2008 could not be captured.

In both, the American and the European parts, we estimate the long-term cointegration relationship and the according VEC model. For this purpose we use from Engle and Granger (1987) the Engle-Granger two step estimator (EG-2) and the maximum likelihood (ML) method developed by Johansen (1988) and compare the outcomes. The EG-2 estimator is consistent and easy to implement, but as Banerjee et al. (1993) argue inefficient and therefore the ML is more commonly used.

Finally we summarize our results, discuss strengths and weaknesses of this study and give an outlook for possible future studies.

## Chapter 2

# Cointegration and Vector error correction

In time series analysis, the non-stationarity of the observed time series is a crucial problem. The standard proof of consistency of OLS shows that with increasing sample size the sample moments settle down to their population values, see Banerjee et al. (1993). But this is only true in case of stationarity. When estimating a regression using non stationary data, the investigator might receive a significant correlation between two independent time series. This phenomenon was described by Yule (1926) as *spurious regression*. Transforming the time series to make them stationary, for example by differencing, is one solution, but one has to accept a loss of information about the long run relationship. Another possibility to tackle this problem is using cointegration.

We want to give an overview about the method of cointegration and vector error correction. First we give an explanation of the concept of cointegration and vector error correction. Afterwards we give a short introduction to the Engle-Granger two step procedure and the ML-estimator from Johansen (1988). Finally in section 2.6 we give an example from Stock and Watson (1988) to show the difficulties when working with integrated data. Read-

ers who are already familiar with these techniques can skip this section and proceed with the estimations and the results in section 3.

## 2.1 Concept of cointegration

First we want to give a general idea about cointegration before continuing with the definitions and application. As already mentioned above, using cointegration techniques gives us the insight about long run equilibrium relationships of variables of interest. Here, equilibrium is not meant as Quandt (1978) defined it, where supply equals demand and following market clearing exists. But as Banerjee et al. (1993) summarize, an equilibrium state is defined as one in which there is no inherent tendency to change.

A intuitive explanation is presented by Engle and Granger (1991), who explain the concept of cointegration using the example of the regional price difference between tomato prices in north and south California. If the difference between the two regions is getting too large, it would be possible to make profits by purchasing tomatoes in one region, transporting and selling them in the other one. It is expected that market mechanism increase prices in the region where they are low and decrease in the other one. If the difference is small enough no profits can be made because of risk and transport cost. So if a shock occurs that has the effect that the two prices drift away, there will be a tendency that pulls them back together. Even if the prices rise in the long term, they are expected to move together. The ratio to which the two prices are tending, is called equilibrium.

## 2.2 Definitions and Explanations

We will now give some formal definitions for cointegration and will then show how this method can be used in practice, starting with Engle and Granger (1991)

*"If  $x_t, y_t$  are  $I(1)$  but there exists a linear combination*

$$z_t = m + ax_t + by_t$$

*which is both  $I(0)$  and has a zero mean, then  $x_t, y_t$  are said to be cointegrated"*

Or, another definition from Banerjee et al. (1993)

*"The components of the vector  $x_t$  are said to be co-integrated of order  $d, b$ , denoted  $x_t CI(d, b)$ , if (i)  $x_t$  is  $I(d)$ <sup>1</sup> and (ii) there exists a non-zero vector  $\alpha$  such that  $\alpha'x_t I(d-b)$ ,  $d \geq b > 0$ . The vector  $\alpha$  is called the co-integration vector."*

In this case, there is also spoken of co-integration, although the result of  $\alpha'x_t$  is not stationary. Nevertheless the definition above is correct, co-integration is only useful if  $\alpha'x_t$  is a stationary time series. Otherwise the problem of spurious regression still exists.

We will now go ahead with presenting an example from Banerjee et al. (1993)<sup>2</sup> to illustrate the definitions presented above. Consider the two time series  $x_t$  and  $y_t$

$$x_t + \beta y_t = u_t \tag{2.1}$$

$$x_t + \alpha y_t = \epsilon_t \tag{2.2}$$

with

$$u_t = u_{t-1} + \epsilon_{1t} \tag{2.3}$$

$$\epsilon = \rho\epsilon_{t-1} + \epsilon_{2t} \tag{2.4}$$

---

<sup>1</sup>Where  $I(d)$  describes a time series which gets stationary after taking differences  $d$  times.

<sup>2</sup>We follow the example from Banerjee et al. (1993) very closely, who base themselves on Engle and Granger (1987)

where  $|\rho| < 1$ , and  $(\epsilon_{1t}, \epsilon_{2t})'$  are iid error terms.

$$E(\epsilon_{1t}) = E(\epsilon_{2t}) = 0 \quad (2.5)$$

$$var(\epsilon_{1t}) = \sigma_{11}; \quad var(\epsilon_{2t}) = \sigma_{22}; \quad cov(\epsilon_{1t}, \epsilon_{2t}) = \sigma_{12} \quad (2.6)$$

Solving (2.1) and (2.2) for  $x_t$  and  $y_t$  with  $\alpha \neq \beta$  gives

$$x_t = \alpha(\alpha - \beta)^{-1}u_t - \beta(\alpha - \beta)^{-1}\epsilon_t \quad (2.7)$$

$$y_t = -(\alpha - \beta)^{-1}u_t + (\alpha - \beta)^{-1}\epsilon_t \quad (2.8)$$

The linear dependence of  $x_t$  and  $y_t$  on  $u_t$ , which is a random walk, makes them I(1) variables. But since  $\epsilon_t$  is stationary,  $x_t + \alpha y_t$  is I(0).  $[1:\alpha]$  is the cointegration vector and  $x + \alpha y$  is the long-run equilibrium relationship.

Although it is possible to use OLS to estimate the correlation of non-stationary time series, in case a cointegration relationship, as described above, exists, the interpretation of the estimation output, specially concerning the coefficient and the reported standard error is different.

In a usual regression like

$$y_t = c + \beta x_t + u_t \quad (2.9)$$

the coefficient  $\beta$  is interpreted as the effect of a change in  $x_t$  on  $y_t$ . The characteristics of a cointegration estimation however are described by Johansen (2006).

*"A coefficient in an identified cointegration relation can be interpreted as the effect of a long-run change to one variable on another, keeping all others fixed. The difference with the usual interpretation of a regression coefficient is that because the relation is a long-run relation, that is, a relation between*

*long-run values, the counterfactual experiment should involve a long-run change in the variables.”*

There is a additional effect on the reported standard errors. The coefficients from a cointegration equation are *super-consistent*. That means, relying on Stock and Watson (1988), that they converge at a rate of  $T$  to their true values instead of the root of  $T$  as in an OLS regression. As a consequence the true standard errors are smaller than they are reported in the estimation output. This makes it difficult to determine the true confidence interval. This question is crucial for our work, since we want to figure out if certain coefficients, in our case the ones for price and disposable income, have changed or not.<sup>3</sup>

## 2.3 Vector error correction

Since a cointegration equation only gives a static relationship, error correction models are useful to capture the short-run dynamics related to the long-run relationship. Or as Banerjee et al. (1993) explain.

*“Error-correction mechanism (ECMs) are intended to provide a way of combining the advantages of modeling both levels and differences. In an error-correction model the dynamics of both short-run (changes) and long-run (levels) adjustment processes are modeled simultaneously.”*

Following Engle and Granger (1987), the idea is that a proportion of the disequilibrium from a period is corrected in the following one. As Banerjee et al. (1993) explain, if there exists an equilibrium relationship as

$$y^* = \theta x^* \tag{2.10}$$

---

<sup>3</sup>Hughes et al. (2006) solve this question comparing the F-statistics from the different estimations. Their method, however, only admits assessing whether all coefficients have changed but cannot tell which of them. Since the influence of the seasonal dummies are very strong, distinction between the F-statistics could capture a change in the seasonal consumer behavior and not imply an increase in dependence of gasoline.

the error correction term is

$$y_t - \theta x_t \quad (2.11)$$

or in case the equilibrium relationship is not known but estimated

$$y_t - \hat{\theta} x_t \quad (2.12)$$

The error correction term can also take the form

$$y_t - \sum_{j=1}^p \theta_j x_{jt} \quad (2.13)$$

A formal definition is given by Engle and Granger (1987) in terms of the backshift operator  $B$

"A vector time series  $x_t$  has an error correction representation if it can be expressed as:

$$A(B)(1 - B)x_t = -\gamma z_{t-1} + u_t$$

where  $u_t$  is a stationary multivariate disturbance, with  $A(0)=I$ ,  $A(1)$  has all elements finite,  $z_\gamma = \alpha' x_\gamma$ , and  $\gamma \neq 0$ .

The relationship between cointegration and error correction model was first stated by Granger (1981) and later defined in the *Granger Representation Theorem*, which states that every cointegrated time series can be expressed as an error correction model<sup>4</sup>.

## 2.4 Two step estimator

Engle and Granger (1987) introduced a two-step estimator for models using

---

<sup>4</sup>For a more detailed explanation or the mathematic proof, see Engle and Granger (1987)

cointegrated time series. Since this method is used in the empirical part of this work we give a short overview of this technique.

In the first step the coefficients of the cointegration equation are estimated and in the second step, they are used in the error correction form to determine the short-run dynamics. Engle and Granger (1987) show that in both steps only OLS is required and all parameters are consistently estimated.

If  $y_t$  and  $x_t$  are cointegrated, the cointegration vector  $\alpha$  can be estimated through the regression

$$y_t = \hat{\alpha}x_t + v_t \quad (2.14)$$

Banerjee et al. (1993) argue that  $v_t$  contains all the omitted dynamics. In the second step the dynamics are modeled.

$$\Delta y_t = \hat{\theta}v_{t-1} + \hat{\gamma}_1\Delta y_{t-1} + \dots + \hat{\gamma}_j\Delta y_{t-j} + \hat{\rho}_1\Delta x_{t-1} + \dots + \hat{\rho}_k\Delta x_{t-k} \quad (2.15)$$

The coefficient  $\hat{\theta}$  states the proportion of how much of the disequilibrium in the period  $t - 1$  is adjusted in  $t$ . The estimates of the parameters in the first step converge to their probability limits at rate  $T$ , while in the second step with the usual asymptotic rate  $T^{1/2}$ .

## 2.5 Maximum likelihood estimator

Although the estimates are consistent, Banerjee et al. (1993) argue that the EG-2 is inefficient and therefore suggest dynamic estimation methods. One of the most commonly used is the maximum likelihood estimator introduced by Johansen (1988). The ML delivers better results than the EG-2 because it takes the error structure of the process into account when modeling the long-term relationship. We just sketch the idea of the maximum likelihood estimator. For a deeper understanding see Johansen (1988, 1991, 2006).

Starting with a VAR of the form

$$X_t = \mu + \sum_{j=1}^p \eta_j X_{t-j} + \epsilon \quad (2.16)$$

where  $X_t$  is a  $k$ -vector of  $I(1)$  variables and  $\epsilon$  are iid Gaussian errors. This VAR can be rewritten in the form.

$$\Delta X_t = \mu + \Pi X_{t-1} + \sum_{j=1}^{p-1} \Gamma \Delta X_{t-j} + \epsilon_t \quad (2.17)$$

where

$$\Pi = \alpha \beta' \quad (2.18)$$

$\alpha$  and  $\beta$  are  $p \times r$  matrices, and the linear combinations  $\beta' X_t$  are stationary. The space spanned by  $\beta$  which Johansen (1988) defines as the cointegration space, is the space spanned by the rows of the matrix  $\Pi$ .

While the Engle-Granger 2 step estimator can only estimate one cointegration relation, Johansen (2006) presents a ML-estimator which estimates the rank of the  $\Pi$  matrix, using the eigenvectors and the eigenvalues, and thus the number of cointegration relations. This restriction is then used to estimate the cointegration relationship.

## 2.6 A tale of two Econometricians

We now want to reproduce an example from Stock and Watson (1988) to illustrate the dangers involved when ignoring the order of integration of time series. In this example two econometricians study the relation between aggregate real per capita consumption  $C_t$ , aggregate disposable income  $Y_t$  and the price index  $P_t$ .

The according processes are generated by Stock and Watson (1988), who rely on the theories of the consumption function of Friedman (1957), which

are unknown to the econometricians<sup>5</sup>.

$$Y_t = Y_t^P + Y_t^S \quad (2.19)$$

$$Y_t^P = Y_{t-1}^P + u_t \quad (2.20)$$

$$C_t = Y_t^P \quad (2.21)$$

$$P_t = P_{t-1} + v_t \quad (2.22)$$

The disposable income consists of two parts, the permanent  $Y_t^P$  and the transitory  $Y_t^S$  components, where the first one is assumed to follow a random walk, while the second is an independently and identically distributed random variable.

According to Friedman's permanent income hypothesis, consumers spend precisely the permanent part of their disposable income.

The price level is a random walk with unforecastable changes and mean zero. But it does not confuse consumers since real disposable income and consumption are independent of the price level.  $u_t$  and  $v_t$  are mutually independent and normally distributed with mean zero and unit variance.

The first econometrician estimates the following equations:

$$c_t = \alpha_1 + \beta_1 p_t \quad (2.23)$$

$$c_t = \alpha_2 + \beta_2 t \quad (2.24)$$

---

<sup>5</sup>Using a pseudo random number generator Stock and Watson (1988) produced time series with 150 observations according to the equations (2.19)-(2.22)

$$\Delta c_t = \alpha_3 + \beta_3 \Delta y_t \quad (2.25)$$

$$\Delta c_t = \beta_4 y_{t-1} \quad (2.26)$$

(2.23) to check the influence of the price level on consumption (money illusion), whether consumption has a linear trend (2.24), or not. (2.25) for the marginal propensity to consume and (2.26) to test the permanent income hypothesis<sup>6</sup>

| Coefficient | Value | t-statistic |
|-------------|-------|-------------|
| $\beta_1$   | 0.4   | 5.12        |
| $\beta_2$   | 0.69  | 16.9        |
| $\beta_3$   | 0.28  | 8.06        |
| $\beta_4$   | 0.41  | -2.15       |

Table 2.1: Results of estimations (2.23) - (2.26)

The results are astonishing.  $\beta_1$  and  $\beta_2$  are both strongly significant. The marginal propensity is significant but rather small. So the econometrician receives significant results and concludes that beside a very small marginal propensity to consume, consumers have money illusions and consumption contains a linear time trend. As clear as the results of the estimations might appear, they are all wrong.

The first regression is the classical case of a spurious regression.  $C_t$  and  $P_t$  are both uncorrelated random walks. In the second equation the econometrician tries to explain a random walk by a deterministic trend, and therefore according to Banerjee et al. (1993) is also spurious. The coefficient of the third equation is biased downwards, because disposable income measures the

---

<sup>6</sup>Equation (2.26) is not, like the others, taken from Stock and Watson (1988) but from Banerjee et al. (1993) to illustrate, as will turn out later, the example of an unbalanced regression.

change in permanent income with error. (2.26) is an example for an unbalanced regression, since it tries to explain an I(1) variable through an I(0), which leads to biased t-statistics and misleading inferences about the significance of the coefficient.

The second econometrician estimates the following regressions:

$$c_t = \gamma_1 + \delta_1 y_t \quad (2.27)$$

$$c_t = \gamma_2 + \delta_2 c_{t-1} + \delta_3 c_{t-2} \quad (2.28)$$

$$c_t = \gamma_3 + \delta_4 c_{t-1} + \delta_5 y_{t-1} \quad (2.29)$$

$$c_t = \gamma_4 + \delta_6 c_{t-1} + \delta_7 p_{t-1} + \delta_8 \Delta p_{t-1} \quad (2.30)$$

(2.27) for the marginal propensity to consume, (2.28) to test whether consumption follows a random walk and (2.29) and (2.30) to find further variables for a potential forecast.

The results for the time series produced by Stock and Watson (1988) are presented in table 2.6.

| <b>Coefficient</b> | <b>Value</b> | <b>t-statistic</b> |
|--------------------|--------------|--------------------|
| $\delta_1$         | 0.94         | -2.74              |
| $\delta_2$         | 0.97         | 11.7               |
| $\delta_3$         | -0.01        | -0.13              |
| $\delta_4$         | 1.03         | 14.3               |
| $\delta_5$         | -0.07        | -0.97              |
| $\delta_6$         | 0.95         | 45.0               |
| $\delta_7$         | 0.004        | 0.17               |
| $\delta_8$         | 0.06         | 0.87               |

Table 2.2: Results of estimations (2.27) - (2.30)

The marginal propensity to consume is large but less than one. When

testing the hypothesis that consumption follows a random walk, she finds only the first lag significant and receives no different result when replacing the second lag of consumption through the first of income. Lags of the price deliver no better forecasting model as well.

Following Stock and Watson (1988) it can be said that the results of our second econometrician are largely correct. Consumption and income are cointegrated so (2.27) produces super-consistent estimators. At estimation (2.28) Stock and Watson (1988) suggest that it can be rewritten as

$$c_t = \gamma_2 + (\delta_2 + \delta_3)c_{t-1} - \delta_3(c_{t-1} - c_{t-2}) \quad (2.31)$$

As a result  $\delta_3$  is a coefficient of a stationary variable. The same strategy is suggested for (2.29), which gives

$$c_t = \gamma_3 + (\delta_4 + \delta_5)c_{t-1} - \delta_5(c_{t-1} - y_{t-1}) \quad (2.32)$$

Here  $\delta_5$  is a coefficient on a stationary time series because income and consumption ( $y_{t-1} - c_{t-1}$ ) are cointegrated. At regression (2.30)  $\Delta P_{t-1}$  is a mean zero random variable, so following Stock and Watson (1988) the estimator is consistent.  $P_{t-1}$  is not cointegrated with any other variable and therefore the coefficient can not be written as coefficient on a mean zero stationary regressor. Hence the estimator has a non-standard-distribution and the usual critical values do not apply.

Final remarks are delivered by Banerjee et al. (1993) *"The moral of the econometricians' story is the need to keep track of the orders of integration on both sides of the regression equation, which usually means incorporating dynamics; models that have restrictive dynamic structures are relatively likely to give misleading inferences simply for reasons of inconsistency of orders of integration."*

## Chapter 3

# Gasoline elasticities in the U.S.A.

In this section we investigate if the gasoline elasticities in the U.S.A. change from the period 1975-1981 to 2001-2007.

### 3.1 Model specification

We choose a common model, where per capita gasoline demand ( $G$ ), is a function of per capita real disposable income ( $Y$ ) and the real price ( $P$ ).

$$G = f(Y, P) \tag{3.1}$$

This kind of model has been used on various studies before, see Dahl and Sterner (1991). This is favorable for our purpose, since we don't want to investigate on new types of models, but check if the elasticity and consequently the dependence on gasoline has changed or not. For our empirical estimation we express (3.1) as:

$$\ln(G_t) = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln Y_t + m_{it} + \epsilon_t \tag{3.2}$$

All the variables have the meaning as described above, but  $\epsilon_t$  is an error term and  $m_{it}$  is a monthly seasonal dummy built in to capture the seasonal

effects present in the consumption of gasoline. Using this model, we rely on Hughes et al. (2006) who used this model before.<sup>1</sup> The result of this estimation gives the long run relationship of  $G$ ,  $Y$  and  $P$ . We estimate the same model for the periods 1975 - 1981 and 2001 - 2007. As can be seen in figure 3.1, both contain similar price shocks. Afterwards we estimate, accordingly

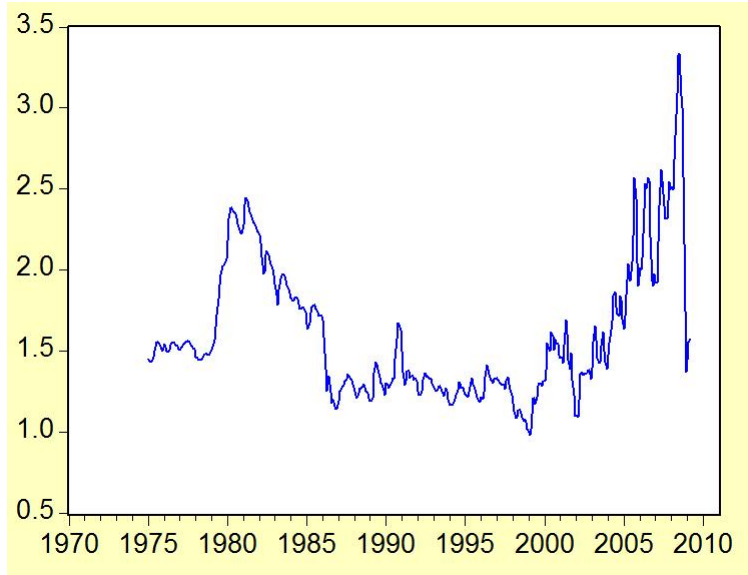


Figure 3.1: Real Gasoline Price

to the cointegration relationships, the vector error correction model which gives us the speed of adjustment to the long-run equilibrium. For a deeper explanation of cointegration and vector error correction techniques see Section 2. The different coefficients can then be compared to find out, whether there has been a shift in gasoline elasticities, or not.

---

<sup>1</sup>Hughes et al. (2006) used this model for similar purposes but had a different approach concerning the dynamic structure and the way to compare the models.

## 3.2 Data

We use a monthly dataset, coherent from January 1975 to December 2007. Gasoline consumption consists of US domestic production plus imports minus exports and stock changes measured in total barrels. This time series is made available by the Energy Information Administration (EIA) under the category *Product Supplied*.

The stated price stands for U.S. city average price per gallon for unleaded regular gasoline, adjusted with GDP deflator base 2000<sup>2</sup>. Source, U.S. Bureau of labour statistics.

Disposable income is given as monthly seasonal data adjusted at annual rates, corrected with the GDP deflator base 2000, from the Bureau of Economic Analysis. Gasoline consumption and disposable income are both used as per capita, and therefore are divided by the total US population from the US bureau of census.

## 3.3 Estimations & Results

By the estimation of our equations we follow closely the suggestions of Engle and Granger (1987) and the theoretical advices from Banerjee et al. (1993).

First we take a look at our time series of interest and check whether they are stationary or not. The ADF unit root test showed, that all time series are not stationary at levels, but are stationary after taking the first differences. Afterwards the cointegration regression was estimated for both periods. From January 1975 to December 1981 and 2001 to 2007 for the same months, see Figure 3.2 and Figure 3.3. We decided to include all seasonal dummies into the equation, although some of them are insignificant, to make sure that most of the seasonal effects are captured. Since the number of observations is 84, the loss of degrees of freedom seems acceptable. The same approach is

---

<sup>2</sup>For the year 1975, the price for leaded gasoline was used, due to a lack of data

also used later at the error correction model.

| Dependent Variable: LOG(G) |             |                       |             |        |
|----------------------------|-------------|-----------------------|-------------|--------|
| Method: Least Squares      |             |                       |             |        |
| Date: 10/16/09 Time: 17:12 |             |                       |             |        |
| Sample: 1975M01 1981M12    |             |                       |             |        |
| Included observations: 84  |             |                       |             |        |
| Variable                   | Coefficient | Std. Error            | t-Statistic | Prob.  |
| LOG(Y)                     | 0.611488    | 0.097687              | 6.259641    | 0.0000 |
| LOG(P)                     | -0.361822   | 0.024227              | -14.93490   | 0.0000 |
| M1                         | -0.069682   | 0.014801              | -4.707926   | 0.0000 |
| M10                        | 0.000804    | 0.014745              | 0.054520    | 0.9567 |
| M11                        | -0.058798   | 0.014738              | -3.989578   | 0.0002 |
| M2                         | -0.136524   | 0.014821              | -9.211295   | 0.0000 |
| M3                         | -0.022670   | 0.014798              | -1.531917   | 0.1301 |
| M4                         | -0.020248   | 0.014808              | -1.367430   | 0.1759 |
| M5                         | 0.012853    | 0.014762              | 0.870649    | 0.3869 |
| M6                         | 0.024933    | 0.014818              | 1.682617    | 0.0969 |
| M7                         | 0.036570    | 0.014794              | 2.471982    | 0.0159 |
| M8                         | 0.039186    | 0.014778              | 2.651567    | 0.0099 |
| M9                         | -0.025487   | 0.014772              | -1.725395   | 0.0889 |
| C                          | -4.345151   | 0.706004              | -6.154566   | 0.0000 |
| R-squared                  | 0.880953    | Mean dependent var    | -0.087766   |        |
| Adjusted R-squared         | 0.858845    | S.D. dependent var    | 0.073383    |        |
| S.E. of regression         | 0.027570    | Akaike info criterion | -4.193135   |        |
| Sum squared resid          | 0.053209    | Schwarz criterion     | -3.787999   |        |
| Log likelihood             | 190.1117    | F-statistic           | 39.84655    |        |
| Durbin-Watson stat         | 1.457376    | Prob(F-statistic)     | 0.000000    |        |

Figure 3.2: Long-term equilibrium relationship, 1975 - 1981

Before we continue to analyze the result, it is necessary to examine whether a cointegration relationship exists, or not. Therefore the residuals of the estimations have to be checked on stationarity. The reported t-values of the ADF-unit root test are *-3.802* for the first and *-5.706* for the second period. As Banerjee et al. (1993) explain, the critical values for the confidence interval stated by the ADF-Unit root test are invalid and must not be used. MacKinnon (1991) gives an alternative procedure to calculate correct values instead. Following this method we receive a 5% critical value of *-3.41*. Therefore a cointegration relationship exists and the estimation output gives the corresponding coefficients of the long run equilibrium relationship. The result is very satisfying, because all variables have the expected sign and with

the exception of some seasonal dummies, are strongly significant. The size of the coefficients seem, from an economic point of view, also reasonable.

We proceed with estimating the vector error correction model. Taking the error term of the cointegration equation

$$\ln(G_t) = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln Y_t + m_{it} + \epsilon_t \quad (3.3)$$

and using it as a regressor of  $\Delta G$ , together with the first four differences of price and income to get:

$$\Delta G = \gamma_1 \epsilon_{(t-1)} + \gamma_2 \Delta Y_{(t-1)} + \dots + \gamma_5 \Delta Y_{(t-4)} + \gamma_6 \Delta P_{(t-1)} + \dots + \gamma_9 \Delta P_{(t-4)} + \theta_i m_{it} + u_t \quad (3.4)$$

In the estimation output,  $\epsilon$  of equation (3.3) will be denoted as *EC75\_81* for the first period from 1975 - 1981, and *EC01\_07* for the later one. Here again all insignificant variables except the seasonal dummies have been excluded. See Figures 3.4 and 3.5.

The *EC* coefficient gives the speed of adjustment to the long-run equilibrium in case of a deviation. For the years 75 - 81 this means that 91% of the deviation from the equilibrium will be adjusted in the next month. For the years 2001 - 2007 it's 60%. They are highly significant and have the expected sign, but from an economic point of view the numbers seem to be very high.

The objective of this work is to explore whether the dependence on gasoline has increased in the last 30 years. For this purpose we compare the coefficients of the estimations of the two time periods and check if they have changed, or not. This task turns out to be more complex than expected. As described in Section 2.2, the coefficients are super-consistent and therefore the reported standard errors and t-Statistics are not correct. Although it is not possible to compute the correct confidence interval, we do know that the true one is smaller. And therefore our procedure to determine the change in the coefficients is to calculate confidence intervals for each factor for both

periods, using the standard errors from the estimation output. If they overlap, it is not possible to make a reliable statement, but otherwise we can conclude that there has been a significant change in the regarding variable.

| 95% Confidence interval |             |             |             |             |
|-------------------------|-------------|-------------|-------------|-------------|
| Coefficient             | 1975 - 1981 |             | 2001 - 2007 |             |
|                         | Lower limit | Upper Limit | Lower Limit | Upper Limit |
| Y                       | 0.416       | 0.806       | 0.259       | 0.574       |
| P                       | -0.410      | -0.313      | -0.057      | -0.013      |

Table 3.1: Cointegration coefficients

As can be seen in figures 3.2 & 3.3 the coefficient of  $Y$  changed from  $0.611$  to  $0.417$  implying that the influence of the disposable income on gasoline consumption decreased, but since the confidence interval is overlapping, see table 3.1 it is not possible to talk of a significant change.

But concerning the price on the other hand the decrease from  $-0.361$  to  $-0.035$  is significant.

To compare the VEC-coefficients we proceed the same way, but since super-consistency is here not the case, the estimated confidence interval is correct.

| 95% Confidenc interval |             |             |             |             |
|------------------------|-------------|-------------|-------------|-------------|
| Coefficient            | 1975 - 1981 |             | 2001 - 2007 |             |
|                        | Lower limit | Upper Limit | Lower Limit | Upper Limit |
| EC                     | -0.97       | -0.52       | -0.81       | -0.4        |

Table 3.2: VEC coefficients

The error correction term decreased from  $-0.919$  to  $-0.603$  which implies that consumers need longer to adjust their behavior and return to the equilibrium state. Due to the overlapping confidence intervals, this change can not be considered as significant.

Now, we repeat the estimation but using the ML estimator instead. At this approach the three time series demand, income and price are the endoge-

nous variables and the seasonal dummies are used as exogenous variables. The lag length has been selected by using the Akaike and Schwarz Information Criterion. The Johansen cointegration test rejects the null hypothesis of no cointegration relationship, but does not reject the null for one. The results of the cointegration coefficients for the ML method, normalizing  $g = 1$ , can be seen in table 3.3 and the error correction terms in table 3.4. Standard errors are in () and t-statistics in [ ].

| Coefficient | Period                    |                           |
|-------------|---------------------------|---------------------------|
|             | 1975 - 1981               | 2001 - 2007               |
| Income      | 0.41<br>(0.15)<br>[-2.69] | 0.36<br>(0.19)<br>[-1.93] |
| Price       | -0.37<br>(0.04)<br>[9.94] | -0.04<br>(0.03)<br>[1.45] |

Table 3.3: Johansen Cointegration coefficients

| Coefficient | Period                     |                            |
|-------------|----------------------------|----------------------------|
|             | 1975 - 1981                | 2001 - 2007                |
| Demand      | -0.52<br>(0.11)<br>[-4.64] | -0.36<br>(0.05)<br>[-6.34] |
| Income      | -0.08<br>(0.04)<br>[-1.99] | -0.5<br>(0.05)<br>[-0.96]  |
| Price       | -0.08<br>(0.06)<br>[-1.21] | 0.69<br>(0.32)<br>[2.16]   |

Table 3.4: Johansen Error correction coefficients

The outcome of the ML estimates confirm the results of the EG-2 estimator. The decline of the coefficient for income is not significant, but for the price it is. Furthermore are the coefficients for price from the two estimator

almost the same and the error correction terms for price and income change significantly.

Taking into account the empirical results of the first part of this study, there exists statistical evidence that dependency on gasoline increased during the last 30 years. The most striking point is the change in the coefficient for the price in the long-run equilibrium relationship.

Dependent Variable: LOG(G)  
Method: Least Squares  
Date: 10/16/09 Time: 17:10  
Sample: 2001M01 2007M12  
Included observations: 84

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| LOG(Y)             | 0.417063    | 0.078595              | 5.306498    | 0.0000 |
| LOG(P)             | -0.035450   | 0.010861              | -3.263845   | 0.0017 |
| M1                 | -0.045586   | 0.006103              | -7.469794   | 0.0000 |
| M11                | -0.035386   | 0.006132              | -5.770598   | 0.0000 |
| M2                 | -0.125612   | 0.006119              | -20.52789   | 0.0000 |
| M3                 | -0.010600   | 0.006196              | -1.710692   | 0.0916 |
| M4                 | -0.029514   | 0.006379              | -4.626533   | 0.0000 |
| M5                 | 0.022810    | 0.006486              | 3.517038    | 0.0008 |
| M7                 | 0.040265    | 0.006254              | 6.438465    | 0.0000 |
| M8                 | 0.042636    | 0.006270              | 6.800603    | 0.0000 |
| M9                 | -0.039482   | 0.006291              | -6.275789   | 0.0000 |
| C                  | -3.324654   | 0.616698              | -5.391053   | 0.0000 |
| M6                 | -0.002157   | 0.006388              | -0.337665   | 0.7366 |
| M10                | 0.001620    | 0.006211              | 0.260885    | 0.7949 |
| R-squared          | 0.947056    | Mean dependent var    | -0.066263   |        |
| Adjusted R-squared | 0.937223    | S.D. dependent var    | 0.045263    |        |
| S.E. of regression | 0.011341    | Akaike info criterion | -5.969799   |        |
| Sum squared resid  | 0.009003    | Schwarz criterion     | -5.564663   |        |
| Log likelihood     | 264.7316    | F-statistic           | 96.31924    |        |
| Durbin-Watson stat | 1.115290    | Prob(F-statistic)     | 0.000000    |        |

Figure 3.3: Long-term equilibrium relationship, 2001 - 2007

Dependent Variable: D(LOG(G))  
Method: Least Squares  
Date: 11/24/09 Time: 22:24  
Sample (adjusted): 1975M02 1981M12  
Included observations: 83 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| EC75_81(-1)        | -0.747242   | 0.113654              | -6.574705   | 0.0000 |
| M1                 | -0.132038   | 0.014550              | -9.074892   | 0.0000 |
| M10                | -0.028606   | 0.013976              | -2.046781   | 0.0444 |
| M11                | -0.116360   | 0.013976              | -8.325617   | 0.0000 |
| M2                 | -0.133065   | 0.013976              | -9.520900   | 0.0000 |
| M3                 | 0.053136    | 0.013976              | 3.801908    | 0.0003 |
| M4                 | -0.060154   | 0.013976              | -4.304016   | 0.0001 |
| M5                 | -0.025387   | 0.013976              | -1.816427   | 0.0736 |
| M6                 | -0.055825   | 0.013976              | -3.994341   | 0.0002 |
| M7                 | -0.049501   | 0.013976              | -3.541792   | 0.0007 |
| M8                 | -0.056836   | 0.013976              | -4.066629   | 0.0001 |
| M9                 | -0.123819   | 0.013976              | -8.859304   | 0.0000 |
| C                  | 0.060060    | 0.009883              | 6.077371    | 0.0000 |
| R-squared          | 0.854719    | Mean dependent var    | 7.51E-05    |        |
| Adjusted R-squared | 0.829814    | S.D. dependent var    | 0.063381    |        |
| S.E. of regression | 0.026147    | Akaike info criterion | -4.307260   |        |
| Sum squared resid  | 0.047856    | Schwarz criterion     | -3.928406   |        |
| Log likelihood     | 191.7513    | F-statistic           | 34.31882    |        |
| Durbin-Watson stat | 2.079271    | Prob(F-statistic)     | 0.000000    |        |

Figure 3.4: VEC 1975 - 1981

Dependent Variable: D(LOG(G))  
Method: Least Squares  
Date: 11/24/09 Time: 22:25  
Sample (adjusted): 2001M02 2007M12  
Included observations: 83 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| EC01_07(-1)        | -0.603695   | 0.102738              | -5.876051   | 0.0000 |
| D(LOG(P(-2)))      | 0.035441    | 0.019676              | 1.801281    | 0.0760 |
| M1                 | -0.084476   | 0.005302              | -15.93426   | 0.0000 |
| M2                 | -0.120010   | 0.005083              | -23.61151   | 0.0000 |
| M3                 | 0.071891    | 0.005199              | 13.82735    | 0.0000 |
| M4                 | -0.063231   | 0.005264              | -12.01278   | 0.0000 |
| M5                 | 0.008263    | 0.005457              | 1.514131    | 0.1346 |
| M6                 | -0.067412   | 0.005567              | -12.11022   | 0.0000 |
| M7                 | 0.002550    | 0.005303              | 0.480795    | 0.6322 |
| M8                 | -0.038270   | 0.005102              | -7.500769   | 0.0000 |
| M9                 | -0.122662   | 0.005114              | -23.98732   | 0.0000 |
| M10                | 0.000815    | 0.005156              | 0.158126    | 0.8748 |
| M11                | -0.077064   | 0.005219              | -14.76675   | 0.0000 |
| C                  | 0.041013    | 0.003689              | 11.11844    | 0.0000 |
| R-squared          | 0.977479    | Mean dependent var    | 0.000814    |        |
| Adjusted R-squared | 0.973236    | S.D. dependent var    | 0.058112    |        |
| S.E. of regression | 0.009507    | Akaike info criterion | -6.320975   |        |
| Sum squared resid  | 0.006236    | Schwarz criterion     | -5.912978   |        |
| Log likelihood     | 276.3205    | F-statistic           | 230.3717    |        |
| Durbin-Watson stat | 2.037805    | Prob(F-statistic)     | 0.000000    |        |

Figure 3.5: VEC 2001 - 2007

## Chapter 4

# Gasoline elasticities in Europe

In the second part of the study we compare the gasoline elasticity of three European countries, Austria, Germany and Britain. Using the same model as in the first part, with real price and income per capita as the chosen factors influencing per capita gasoline demand.

$$G = f(Y, P)$$

The approach will be very similar as in the section on the U.S. After checking the degree of integration of the time series of the different countries, we try to estimate the cointegration relation and the according vector error correction model of each country. The comparison will be done the same way as in the preceding section. The investigated time period is 1990 - 2004.

The goal is to detect national differences in the response to changes in the gasoline price. Since the economic framework conditions, for example the extent of public transports, tolls and taxation of the different kinds of fuel varies in each country, different elasticities would not be surprising.

Investigating on gasoline demand in Europe seems to be more problematic than in the U.S. Leaving aside the national differences mentioned above, one reason for this might be, that the usage of diesel powered cars is more widespread. In the United States the proportion of sold diesel to gasoline rose

just two percentage points during the period 1990 - 2004<sup>1</sup>. In Austria for example the share of diesel powered passenger cars went up from 13.7 % in 1990 to 49.2 % in 2004, see Pock (2007). One reason for this development is the technological progress like direct injection and turbo charges which improved the driveability of diesel cars. Another reason is the taxation of gasoline and diesel in European countries, which leads to an on average 18 cents per litre higher gasoline price<sup>2</sup>. By contrast, in the United States diesel is still more expensive than gasoline. This leads to a decreasing total gasoline consumption in Europe, see figure 4.1, while in the US total and per capita gasoline consumption is still increasing, see figure 4.2.

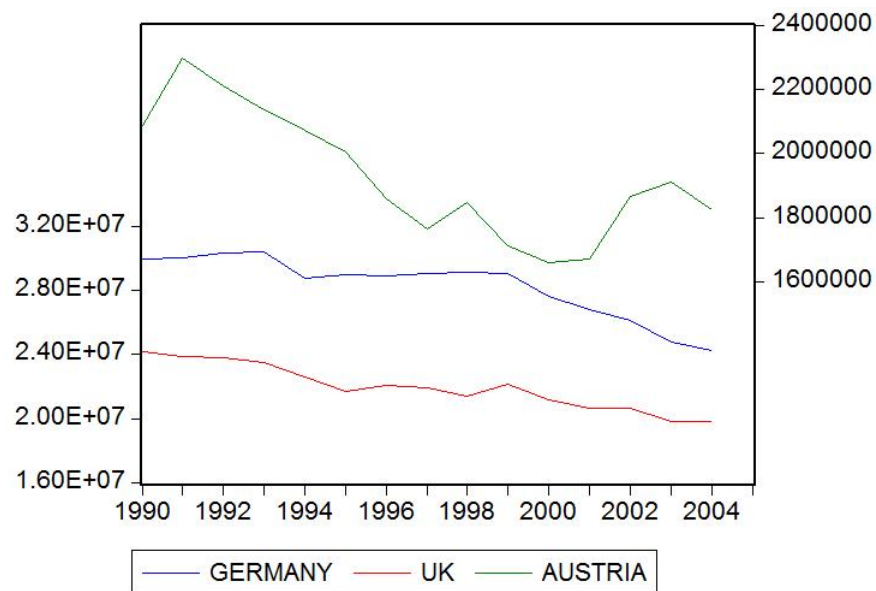


Figure 4.1: Total Gasoline Demand

<sup>1</sup>We compare *Distillate Fuel Oil* and *Finished Motor Gasoline* of the category *Products Supplied* provided by the Energy Information Agency. Although *Distillate Fuel Oil* includes products which are used not only for automobiles and trucks, but also for railroad locomotives and agricultural machinery, the comparison seems appropriate

<sup>2</sup>The average is calculated for all EU countries for the period 1990 - 2004. Here again we rely on Pock (2007).

In earlier studies Baltagi and Griffin (1983) estimated different models for 18 European countries and found that for the period 1960 - 1978 the income effect is insignificant for the majority of the countries and the price of gasoline is statistically not significant for seven out of 18 countries. When using pooling estimators, instead of OLS, better results have been achieved. Pock (2007) built on the findings of Baltagi and Griffin (1983) and added the shift to diesel powered cars in his estimations. He concluded that, *"car owners react to increasing fuel prices by gradually replacing their gasoline powered cars with diesel powered ones"* Pock (2007). This makes sense, since the acquisition costs of diesel cars are higher than for gasoline ones. Switching to diesel cars is only rational beyond a certain price level<sup>3</sup>. This implies, that models which do not include the shift to diesel powered cars and therefore the rise in diesel demand, deliver overstated gasoline demand elasticities. Keeping these findings in mind we proceed and watch, if the use of cointegration and vector error correction techniques can tackle the problems mentioned above.

## 4.1 European Data

After starting to collect European data, one problem appeared immediately. While for American data the institutions offer coherent time series, for Europe this is not the case. Data available from Eurostat is insufficient for our purpose and the statistics offered by the national institutions are hardly comparable with each other. Fortunately Markus Pock made his data set from Pock (2007), in which he completed or substituted the eurostat data when necessary from national statistics agencies, ministries and automobile associations, available for our work.<sup>4</sup> This however limits our estimations to the

---

<sup>3</sup>The exact break even point is given by Pock (2007) as the total life-cycle cost  $A + \sum_{t=1}^T (1+r)^{-t}(p_t q_t + c_t)$  with acquisition costs A, operating life expectancy T, retail fuel price p, annual fuel consumption q, interest rate r and non fuel operating costs c.

<sup>4</sup>Without this contribution the second part of our study could not have been completed. So we want express at this point our most profound thanks.

time period 1990 - 2004. Therefore the response to the oil price shock which lasted to 2008 and had his biggest jumps after 2005 can not be captured.

Another difficulty concerning this dataset, since the time series are yearly data, is the small number of observations. Whereas the US dataset contained 84 data points for each time period, which left us even after including 11 seasonal dummies with a comfortable large number of degrees of freedom, for our European estimations there are only 15 observations.

As in the previous section we use per capita gasoline consumption real income and the real price.

Where fuel consumption is measured in tonnes per year. Real income per capita is measured as PPP adjusted real GDP in US\$ per total population and the prices are CPI-adjusted retail prices in EURO per liter.

All time series used in our estimations are in logarithmic form.

## 4.2 Model Estimation

We start the estimations determining the order of integration of our time series using an augmented-Dickey-Fuller unit root test.

Only two time series are stationary, see table 4.1. So the usage of standard OLS estimations can be excluded. But in order that cointegration techniques can be applied the degree of integration of for each country must be the same. This is clearly not the case for Austria and Germany.

The null hypothesis of a unit root in the German real per capita GDP data can be rejected with a t-statistic of  $-3.77^5$ . So the decision is very clear. But at least a cointegration relationship between the price and demand may exist. For Austria however this is not the case. While the null hypothesis can not be rejected for the real retail price with a t-statistic of  $-2.87^6$ , the income gets only stationary after differentiating three times. For the UK the

---

<sup>5</sup>The corresponding p-value is  $0.015$

<sup>6</sup>Which gives a p-value of  $0.07$

| Country & time series | Order of Integration |
|-----------------------|----------------------|
| <b>Austria</b>        |                      |
| Gasoline demand       | I(1)                 |
| Price                 | I(0)                 |
| Income                | I(3)                 |
| <b>Germany</b>        |                      |
| Gasoline demand       | I(1)                 |
| Price                 | I(1)                 |
| Income                | I(0)                 |
| <b>UK</b>             |                      |
| Gasoline demand       | I(1)                 |
| Price                 | I(1)                 |
| Income                | I(1)                 |

Table 4.1: Orders of integration

results are more pleasant. All three times series are integrated of order one.

As an outcome of the unit root test, we have to exclude Austria from our further investigations, and the following comparison between Germany and UK excludes income. Although estimating gasoline elasticities for UK including income would be valid, since the aim of this study is to compare different countries there is no point in doing so. Recalling the findings of Baltagi and Griffin (1983) a negligible influence of income on the European gasoline demand is not surprising.

Therefore our estimated regression for the longterm equilibrium relationship is<sup>7</sup>

$$\ln(G_t) = \beta_0 + \beta_1 \ln P_t + \epsilon_t \quad (4.1)$$

The outcome can be seen in figures 4.3 and 4.4. The coefficient of the price is significant in both regressions and has the expected sign. The size is

---

<sup>7</sup>Since we use yearly data, the seasonal dummies have also been excluded.

at  $-0.7$  for Germany twice as much as for the UK with  $-0.3$ , which suggests a higher dependence on gasoline for the latter country. To verify this statement, we calculate lower and upper limits for a 95 % confidence interval for each country.

| <b>Price</b> | Lower limit | Upper Limit |
|--------------|-------------|-------------|
| Germany      | -0.92       | -0.55       |
| UK           | -0.49       | -0.13       |

Table 4.2: 95% confidence interval

As can be seen in table 4.2, the upper limit for Germany of  $-0.55$  and the lower limit for Great Britain of  $-0.49$  are not overlapping. So the difference of the elasticities can be seen as significant. As in section 3.3, the coefficients of a cointegration relationship are superconsistent, therefore the true confidence interval is smaller as the one given in table 4.2.

Before this outcome can be considered as certain, the error terms of the equilibrium relationships have to be checked on stationarity. We compare the critical values from an augmented Dickey-Fuller unit root test with the adjusted ones from MacKinnon (1991).

| <b>Country</b> | <b>aDF unit root test</b> | <b>5% Critical values</b> |
|----------------|---------------------------|---------------------------|
| Germany        | -2.9                      | -1.97                     |
| UK             | -3.05                     | -1.97                     |

Table 4.3: Results of the unit root test

As can be seen in table 4.2 the null hypothesis of a unit root in the error terms can clearly be rejected. Therefore the regressions in figures 4.3 and 4.4 are said to be long-term equilibrium relationships, the statements made above are valid and we can proceed by estimating the vector error correction model.

The error term  $\epsilon$  of the equation (4.1) is taken together with first two lags of the change in price as regressors of  $\Delta G$

$$\Delta G = \gamma_1 \epsilon_{t-1} + \gamma_2 \Delta P_{t-1} + \gamma_3 \Delta P_{t-2} \quad (4.2)$$

Following the general to specific method insignificant variables have been excluded. The outcome is presented in figures 4.5 and 4.6. For the German model none of the lags on the price are significant. After eliminating all insignificant price terms, the coefficient of the error correction term is highly insignificant. Regarding the R-squared, the explanatory power of this model is very low. Therefore we estimate the error correction model, using  $\Delta P$  as the dependent variable, see figure 4.7. Here the coefficient of the error correction term is significant and confirms the finding that a cointegration relationship exists.

The coefficient of the error correction term for the UK is significant and has with  $-0.34$  the expected sign and a reasonable size.

Now we repeat again the estimations using the maximum likelihood method. The Johansen cointegration test confirms the cointegration relationship for the UK. But for Germany he fails to reject the null hypothesis of no cointegration for both, the three and the two dimension model. So the EG-2 might have produced incorrect results concerning Germany. We therefore just estimate the cointegration relationship for the UK and compare it with the result of the EG-2 estimator. See the results in table 4.4. Standard errors are in (), t-statistics in [ ]

In contrast to the results for the United States, for Europe the outcome of the ML-Estimator contradicts the ones of the EG-2 estimator. While for the UK, the difference of the coefficients  $-0.3$  and  $-0.42$  is not significant, for Germany the case looks different. While the EG-2 estimator confirms the existence of a cointegration relationship, the ML-estimator rejects it. As Banerjee et al. (1993) explains the Engle-Granger 2 step procedure is strongly biased in small samples, so the results for Germany, based on the Johansen procedure, have to be taken more seriously.

| Coefficient               | United Kingdom           |                            |
|---------------------------|--------------------------|----------------------------|
|                           | Demand                   | Price                      |
| Cointegration coefficient |                          | -0.42<br>(0.11)<br>[-3.8]  |
| EC-coefficient            | 0.14<br>(0.08)<br>[1.63] | -0.82<br>(1.67)<br>[-4.84] |

Table 4.4: Maximum likelihood estimates

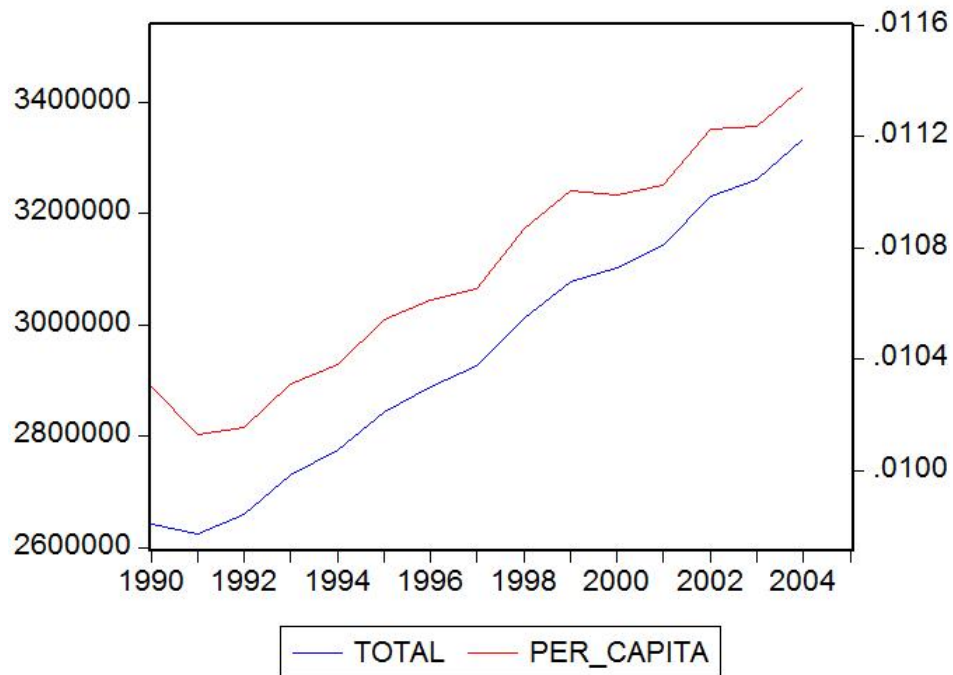


Figure 4.2: Gasoline demand U.S.A.

Dependent Variable: DE\_G  
Method: Least Squares  
Date: 11/09/09 Time: 17:34  
Sample (adjusted): 1990 2004  
Included observations: 15 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| DE_P               | -0.736457   | 0.091573              | -8.042324   | 0.0000 |
| C                  | -1.198680   | 0.018817              | -63.70026   | 0.0000 |
| R-squared          | 0.832645    | Mean dependent var    | -1.064270   |        |
| Adjusted R-squared | 0.819771    | S.D. dependent var    | 0.078891    |        |
| S.E. of regression | 0.033492    | Akaike info criterion | -3.831457   |        |
| Sum squared resid  | 0.014582    | Schwarz criterion     | -3.737050   |        |
| Log likelihood     | 30.73593    | F-statistic           | 64.67898    |        |
| Durbin-Watson stat | 1.269241    | Prob(F-statistic)     | 0.000002    |        |

Figure 4.3: Gasoline elasticity Germany

Dependent Variable: UK\_G  
Method: Least Squares  
Date: 11/09/09 Time: 17:34  
Sample (adjusted): 1990 2004  
Included observations: 15 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| UK_P               | -0.309195   | 0.090982              | -3.398431   | 0.0048 |
| C                  | -1.042049   | 0.023540              | -44.26736   | 0.0000 |
| R-squared          | 0.470454    | Mean dependent var    | -0.980860   |        |
| Adjusted R-squared | 0.429720    | S.D. dependent var    | 0.077770    |        |
| S.E. of regression | 0.058730    | Akaike info criterion | -2.708171   |        |
| Sum squared resid  | 0.044840    | Schwarz criterion     | -2.613765   |        |
| Log likelihood     | 22.31128    | F-statistic           | 11.54934    |        |
| Durbin-Watson stat | 0.437368    | Prob(F-statistic)     | 0.004757    |        |

Figure 4.4: Gasoline elasticity UK

Dependent Variable: D(DE\_G)  
Method: Least Squares  
Date: 11/19/09 Time: 17:59  
Sample (adjusted): 1991 2004  
Included observations: 14 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| EC_DE(-1)          | -0.010336   | 0.246959              | -0.041852   | 0.9673 |
| C                  | -0.017601   | 0.006524              | -2.698104   | 0.0194 |
| R-squared          | 0.000146    | Mean dependent var    | -0.017653   |        |
| Adjusted R-squared | -0.083175   | S.D. dependent var    | 0.023037    |        |
| S.E. of regression | 0.023976    | Akaike info criterion | -4.491984   |        |
| Sum squared resid  | 0.006898    | Schwarz criterion     | -4.400690   |        |
| Log likelihood     | 33.44389    | F-statistic           | 0.001752    |        |
| Durbin-Watson stat | 1.751210    | Prob(F-statistic)     | 0.967305    |        |

Figure 4.5: VEC Germany

Dependent Variable: D(UK\_G)  
Method: Least Squares  
Date: 11/19/09 Time: 18:00  
Sample (adjusted): 1993 2004  
Included observations: 12 after adjustments

| Variable           | Coefficient | Std. Error            | t-Statistic | Prob.  |
|--------------------|-------------|-----------------------|-------------|--------|
| EC_UK(-1)          | -0.344593   | 0.106232              | -3.243769   | 0.0101 |
| D(UK_P(-2))        | 0.165128    | 0.043351              | 3.809057    | 0.0042 |
| C                  | -0.024174   | 0.004693              | -5.150646   | 0.0006 |
| R-squared          | 0.678224    | Mean dependent var    | -0.018472   |        |
| Adjusted R-squared | 0.606718    | S.D. dependent var    | 0.024831    |        |
| S.E. of regression | 0.015572    | Akaike info criterion | -5.274362   |        |
| Sum squared resid  | 0.002182    | Schwarz criterion     | -5.153135   |        |
| Log likelihood     | 34.64617    | F-statistic           | 9.484881    |        |
| Durbin-Watson stat | 2.514630    | Prob(F-statistic)     | 0.006081    |        |

Figure 4.6: VEC UK

| Dependent Variable: D(DE_P)                 |             |                       |             |        |
|---------------------------------------------|-------------|-----------------------|-------------|--------|
| Method: Least Squares                       |             |                       |             |        |
| Date: 02/22/10 Time: 17:47                  |             |                       |             |        |
| Sample (adjusted): 1991 2004                |             |                       |             |        |
| Included observations: 14 after adjustments |             |                       |             |        |
| Variable                                    | Coefficient | Std. Error            | t-Statistic | Prob.  |
| C                                           | 0.028587    | 0.014273              | 2.002873    | 0.0683 |
| EC_DE(-1)                                   | -1.190296   | 0.540324              | -2.202931   | 0.0479 |
| R-squared                                   | 0.287957    | Mean dependent var    | 0.022689    |        |
| Adjusted R-squared                          | 0.228620    | S.D. dependent var    | 0.059726    |        |
| S.E. of regression                          | 0.052457    | Akaike info criterion | -2.926094   |        |
| Sum squared resid                           | 0.033020    | Schwarz criterion     | -2.834800   |        |
| Log likelihood                              | 22.48266    | F-statistic           | 4.852904    |        |
| Durbin-Watson stat                          | 1.917291    | Prob(F-statistic)     | 0.047884    |        |

Figure 4.7: VEC Germany

## Chapter 5

## Conclusion

In the past there have been many studies on gasoline demand elasticities, sometimes with contradictory results. For example Espey (1998) argues that the periodicity of the data does not affect the elasticities in general, while Dahl and Sterner (1991) come to the result that models using monthly or quarterly data maybe more unreliable than those who use yearly data. A common finding of these two meta-analyse is that income and price are the main factors influencing gasoline demand. Dahl and Sterner (1991) describe models which do not include both variables even as misspecified. Taking these recommendations we built our model correspondingly. Since the aim of this study has been to compare different elasticities regarding time and country we stick with this rather simple model.

For the U.S.A. the estimations for the long-run equilibrium relationship show a decisive change from 1975 - 1981 to 2001 - 2006, which mainly confirms the findings of Hughes et al. (2006). The influence of income declined, concerning the EG-2 estimator, from  $0.61$  to  $0.41$ , but the change for the coefficient of income can not be considered as significant. The coefficient of price changed significantly from  $-0.36$  to  $-0.03$ . One possible interpretation of these figures is that dependence on gasoline has increased and that for drivers for example it is more difficult to switch to alternative ways of

transport.

The corresponding vector error correction models state the speed of adjustment to the equilibrium relationship. It changed from  $-0.91$  to  $-0.61$  but because of the overlapping confidence intervals this change cannot be considered as significant.

Another point that should be noted is that coefficients for the error correction term are very high. For the period 1975 - 1981 this would mean that 91% of the deviation from the equilibrium are adjusted after one month.

The repetition of the estimations, using the ML-estimator developed by Johansen (2006) confirmed the results widely. The coefficients for income changed from  $0.41$  to  $0.36$ , here again the change is not significant. The coefficients for price of  $-0.37$  to  $-0.04$  are almost identical to those obtained by the EG-2 method.

Despite some difficulties, with showing that elasticities have changed, the first part of the study has achieved its goal.

The part for European demand delivers less substantial results. We started again with the EG-2 estimator. Austria had to be excluded due to not matching orders of integration. The remaining comparison between Germany and UK did not contain income as an explanatory variable for the same reason. The long-run relationships for these countries show significant differences. While for Germany the coefficient of the price is  $-0.73$  for the UK it is  $-0.3$ . Therefore it can be said, concerning the EG-2 method, that German consumers react stronger to price changes than British consumers.

The VECM for the UK delivers with an ec-term of  $-0.34$  a reasonable result, but as in the US part a lagged change of price appears with an unexpected sign. For Germany the error correction term is insignificant when taking demand as the dependent variable, but significant for the price.

The estimations from the ML-estimator deliver a price elasticity for the UK of  $-0.42$ , so not significantly different from the results from the EG-2 estimator. The outcome for Germany however is contradictory. The EG-2

approach confirms the existence of a cointegration relationship, while the Johansen cointegration test rejects it. This result does not catch us completely by surprise since Banerjee et al. (1993) explains that the EG-2 estimator tends to be strongly biased in small samples. Our study seem to confirm this finding. For the U.S.-section, where we investigated on time series with 84 observations, the results delivered by the Engle-Granger 2 step estimator are very close to the ones from the ML-method. But for the Europe section, where only 15 data points have been available per time series, the results are contradictory.

The findings for Europe must be described as less meaningful. This is most likely for two reasons. First, the data available for this study was restrictive concerning the time period and therefore the degrees of freedom in our estimations are very low. Additionally some time series behaved unexpectedly. For example the real retail price for Austria turned out to be stationary.

The second reason might be a possible misspecification of our European models. At the beginning of the European section we stated the findings of Pock (2007), that when estimating models for gasoline demand in Europe, one should include a variable that captures the increasing spread in diesel powered cars. Not doing so could have led to misleading results.

Summarizing it can be said that this study delivers some answers concerning gasoline demand, but opens up questions for possible future studies. To research why elasticities changed in the U.S.A. or why price, income and demand are not cointegrated for Austria and Germany might be of some interest. But therefore more empirical research is necessary.

# Bibliography

- B. Baltagi and J. Griffin. *European Economic Review*, volume 22, pages 117–137. North Holland, 1983.
- A. Banerjee, J. Dolado, J. W. Galbraith, and D. F. Hendry. *Co-Integration, Errorcorrection, and the econometric analysis of non-stationary data*. Oxford University Press, 1993.
- C. Dahl and T. Sterner. Analysing gasoline demand elasticities: a survey. *Energy Economics*, 13:203–210, 1991. ISSN 3.
- R. F. Engle and C. W. J. Granger. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55:251–276, 1987.
- R. F. Engle and C. W. J. Granger. *Long-run economic relationships readings in cointegration*. Oxford University Press, 1991.
- M. Espey. Gasoline demand revisited: an international meta-analysis of elasticities. *Energy Economics*, 20:273–295, 1998.
- M. Friedman. *A Theory of the Consumption Function*. Princeton University Press, 1957.
- C. W. J. Granger. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, pages 121–130, 1981.

- J. E. Hughes, C. R. Knittel, and D. Sperling. Evidence of a shift in the short-run price elasticity of gasoline demand. 2006.
- S. Johansen. *Statistical Analysis of Cointegration Vectors*, pages 131–152. Oxford University Press, 1988.
- S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580, 1991.
- S. Johansen. *Handbook of Econometrics*, volume 1, pages 540–577. Palgrave Macmillan, 2006.
- J. G. MacKinnon. *Critical Values for Co-Integration Tests*, pages 267–267. Oxford University Press, 1991.
- M. Pock. Gasoline and diesel demand in europe: New insights. *Economic Series*, 202:Institute for Advanced Studies, 2007.
- R. E. Quandt. Tests of the equilibrium vs disequilibrium hypotheses. *International Economic Review*, 19:435–452, 1978.
- J. H. Stock and M. W. Watson. Variable trends in economic time series. *The Journal of Economic Perspectives*, 2(3):147–174, 1988.
- G. U. Yule. Why do we sometimes get nonsense correlations between time series?, a study in sampling and the nature of time series. *Journal of the Royal Statistical Society*, 89(1):1–63, 1926.

## **Eidesstattliche Erklärung**

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Wien, im März 2010

*Unterschrift*

(Norbert Ernst)

# CURRICULUM VITAE

## Personal Data

Name: Norbert Ernst  
Birth Date: August 1983  
Birth Place: Vienna  
Nationality: Austria  
Adress: Görgengasse 2/2/4  
1190 Wien, Austria  
Telephone: +43 650 26 159 48  
email: norberternst@gmx.at

## Education

1989 - 1993 Elemantary School Krottenbachstrasse  
1993 - 1997 BRG XIX  
1997 - 2002 HAK Floridsdorf  
since 2003 Diplomstudium Volkswirtschaft  
Universität Wien

## **Abstract**

The recent spike in the price of gasoline has shown its economic importance. The reaction of consumer behavior to turbulences in the gasoline price has important implications for policy makers. Not just for ecological reasons, but also the economic dependence on gasoline seems to be of some interest. There have been many studies on the elasticities of gasoline demand, but most of them focus on the period of the late 1970's. In this paper, we investigate if there has been a change for gasoline elasticities. We compare the period 1975-1981 with the recent period 2001-2007 for the U.S., using the method of cointegration and vector error correction to investigate the long run relationship and the speed of adjustment. In the second part we investigate if the gasoline elasticities differ between countries, comparing Austria, Germany and Britain.

### **Abstract**

Die jüngsten Preissprünge haben die ökonomische Bedeutung von Benzin einmal mehr aufgezeigt. Die Reaktion der Konsumenten auf solche Turbulenzen hat wichtige Implikationen für die Wirtschaftspolitik. Nicht nur aus ökologischen Gründen, sondern auch die ökonomische Abhängigkeit von Benzin scheint von Interesse. Es gibt eine ganze Fülle an Studien die sich mit Benzinnachfrage beschäftigen. Die meisten aber betrachten die Periode Ende der 70er Jahre. In dieser Arbeit untersuchen wir, ob es eine Veränderung bezüglich der Benzinpreiselastizität gegeben hat. Wir vergleichen die Perioden 1975 - 1981 und 2001 - 2007 für die U.S.A.. Im Zweiten Teil dieser Arbeit untersuchen wir Unterschiede zwischen drei europäischen Ländern Österreich, Deutschland und Groß Britannien. Die hierzu angewandte Methode ist Kointegration und Vektor-Fehler-Korrektur.