Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Income Elasticity of Gasoline Demand: A Meta-Analysis

Author: **Ondřej Kokeš**

Supervisor: **PhDr. Tomáš Havránek**

Academic Year: **2011/2012**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature. The author also declares that he has not used this thesis to acquire another academic degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 13, 2012

Signature

Acknowledgments

I would like to thank my supervisor PhDr. Tomáš Havránek for the thesis topic and invaluable comments throughout the year. I am also extremely grateful to prof. Carol Dahl, not only for the published papers on gasoline demand, but also for publishing her dataset that was used for this meta-analysis.

Abstract

In this thesis I summarize previous studies estimating income elasticity of gasoline demand, analyze the models employed, comment on the evolution of econometric tools used, and finally perform a meta-analysis. This thesis is the first survey on gasoline income elasticity that takes into account publication bias. It also distinguishes between models including car stock information in estimation. I estimate the underlying short-run elasticity to be 0.1, long-run with car stock 0.234, and long-run without car stock 0.644. These results, on average, point to less income-elastic demand for gasoline than what previous surveys found.

JEL Classification	Q41, C52, C81, C83
Keywords	gasoline demand, income elasticity, meta-analysis, publication bias

Author's e-mail	`ondrej.kokes@gmail.com`
Supervisor's e-mail	`tomas.havranek@ies-prague.org`

Abstrakt

V této práci se věnuji výzkumným pracem na téma důchodové elasticity poptávky po benzínu, popisuji jednotlivé modely, shrnuji vývoj užitých ekonometrických metod a na závěr provádím meta-analýzu nasbíraných dat. Tato práce je první analýzou důchodové elasticity, která bere v potaz publikační selektivitu. V modelech odděluje odhady ze studií, které využívaly informaci o počtu automobilů na daném území. Finální odhad důchodové elasticity je 0,0999 pro krátké období, 0,234 pro dlouhé období ve vzorku studií využívající informaci o počtu automobilů a 0,644 pro dlouhé období s modely bez této informace. Díky odhadnutí publikační selektivity dostáváme odhady nižší než ty, které se dříve objevovaly v literatuře.

Klasifikace JEL	Q41, C52, C81, C83
Klíčová slova	poptávka po benzínu, důchodová elasticita, meta-analýza, publikační selektivita

E-mail autora	`ondrej.kokes@gmail.com`
E-mail vedoucího práce	`tomas.havranek@ies-prague.org`

Contents

List of Tables

List of Figures

Acronyms

ARDL Auto Regressive Distributed Lag

ECM Error Correction Model

EIA U.S. Energy Information Administration

GCC Gulf Cooperation Council

GDP Gross Domestic Product

iid independent and identically distributed

OECD Organisation for Economic Co-operation and Development

OPEC Organization of Petroleum Exporting Countries

RMSE Root Mean Square Error

UECM Unrestricted Error Correction Model

Bachelor Thesis Proposal

Author	Ondřej Kokeš
Supervisor	PhDr. Tomáš Havránek
Proposed topic	Income Elasticity of Gasoline Demand: A Meta-Analysis

Topic characteristics I plan to examine the income elasticity of gasoline demand using meta-analysis. The topic was chosen as the oil market has become volatile once again, so same as during the oil shocks in the 1970s, it is now more important to thoroughly study the income and price effects on consumers. As the demand is increasing all over the world, the explanation of elasticities is beneficial for both the oil companies and the governments in order to adjust their policies.

Meta-analysis as a method does not rely on a new research, but leverages the work and data from previous publications. As a toolset for the thesis it is well suited as it differs from a mere description of results of other studies by using econometric tools and data visualisation to uncover the differences in results, publication bias, and other problems often encountered.

Hypotheses

1. The differences in results of various studies can be explained by varying models, estimation methods, data characteristics, and publication bias.
2. Previous meta-analyses do not account for publication bias in the studies examined. Therefore the population value of income elasticity may not estimated correctly.
3. Non-US (or non-OECD) data differ significantly from the rest of the world and should not be pooled together with the rest of the world

Methodology I will use meta-analytical tools used in most recent MRAs, specifically funnel plot and forest plot visualisation, and for the MRA itself I will use most probably a mixed-effects model.

Outline

1. Meta-analysis in economics
   (a) History of the method, it's usage in research today
   (b) It's advantages and disadvantages
   (c) The tools used – models, plots
2. Income elasticity of gasoline demand
   (a) Reasons of gathering the data, consumer behaviour
   (b) Its importance in the past (e.g. oil shocks, demand boost of the 1990s)
   (c) Explanaion of the differences in research results
3. MRA itself
   (a) Overall description, data used, comments on the previous MRAs and elasticity surveys
   (b) Summary of the data, basic characteristics, plots
   (c) Models used, consideration of short- and long-run elasticities
   (d) Summary of hypotheses
   (e) Results and conclusion

Core bibliography

1. Dahl, C. & T. Sterner (1991): "Analysing gasoline demand elasticities: a survey." *Energy Economics* **13(3)**: pp. 203–210.
2. Espey, M. (1998): "Gasoline demand revisited: an international meta-analysis of elasticities." *Energy Economics* **20(3)**: pp. 273–295.
3. Stanley, T.D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.
4. Stanley, T.D. & H. Doucouliagos (2010): "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economic Surveys* **24(1)**: pp. 170-191.
5. Stanley, T.D. & S.B.. Jarrell (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161-170.

Introduction

Modeling consumer behavior when it comes to gasoline demand offers crucial information to various agents, be it governments, producers, or local distributors. Even though hundreds of studies have been performed, results vary broadly and this heterogeneity needs to be modeled as well. Meta-analysis, the approach of statistical summary of research results, has greatly evolved over time and offers a powerful tool to understanding why these studies differ even when they share many common factors.

As both surveys and meta-analyses gather data from previously published studies and look for underlying values, any errors or omissions in the data may affect such practice. One of the caveats may be asymmetry of the sample as the estimated values are sometimes not spread around the population value, but are rather skewed in one way. The problem is when researchers manipulate their data or models in order to make their results more "publishable." If they fail to do so, they may withdraw their paper altogether. These practices affect the outcomes of their studies and need to be accounted for and meta-analysis offers several methods of treating for it. This thesis is the first one in terms of income elasticity of gasoline demand that accounts for this publication bias.

We start with an overview of gasoline demand since the oil shocks of the 1970s in Chapter 1. Chapter 2 is devoted to estimation techniques themselves and their evolution over time. Brief overview of current meta-analytical framework is in Chapter 3 and finally the meta-analysis itself is carried out in Chapter 4.

Chapter 1

Income elasticity of gasoline demand

Elasticity is a basic measure of consumer response to a change of price of a commodity, disposable income, or any other factor of interest. For gasoline, the most commonly estimated elasticities are those of price and income. Even though hundreds of studies have been carried out over the past several decades, the discrepancy between individual results does call for clarification. This variability of the studied elasticity estimates is the driving force behind meta-analyses and literature surveys.

1.1 Usage

The purpose of estimating gasoline demand elasticities is to understand consumer behavior in time. For the results to be comparable, two periods have been established – short and long run. Short-run estimate describes the adjustment during the first year since the change of a variable in question. Other inputs like car stock are assumed to be fixed during this time, not affecting the elasticity. The long run period is not precisely defined and the explanation varies based on the data and models used. For example, the auto regressive distributed lag model assumes effect in infinite periods, on the other hand the polynomial lag model allows for a finite lag structure.

One of the major uses of these estimates is forecasting, predicting future demand based on expected inputs – prices, income, car stock, and others. Various agents are interested in the research results, including producers, exporters, importers, or distributors. Knowing how consumers may react to parameter changes helps to plan their actions. In this regard, price elasticity is more employed as the end-user price is easier to manipulate through taxes, quotas, or tariffs.

Renewed interest in gasoline demand estimation also stems from the ongoing debate over global warming that is closely connected to emission of greenhouse gases. Knowing the consumer response may help governments achieve e.g. the goals set in the Kyoto protocol. For more accurate models of specific income groups, micro-level data are usually employed (Kayser 2000; Nicol 2003).

An example of possible policy adjustment can be found in Alves & Bueno (2003). They discover Brazilian demand to be price inelastic both in the short and long run, so they point out taxing gasoline might be a good source of revenue. Given the widespread fuel switching campaigns, cross-product elasticity is vital as well. Low willingness of consumers to switch to alternative sources of energy has to be accounted for when planning e.g. a gradual move towards an alternative source. Thorough overview of recent studies on fuel switching is in Dahl (2012).

1.2 Decomposition

The distinction of short- and long-run adjustment allows us to observe how consumer adaptation is laid out in time. Graham & Glaister (2002) find that long-run income elasticity is usually two to four times larger than the short-run. This can be seen in our data as well, with density peaking around three. This suggests that 25 to 50% of the total adjustment occurs in the first year since the change.

Another aspect is the real impact caused by changes of income. Unlike the effect of changing prices where people alter their driving habits, usually driving less as prices surge, income effect translates primarily into the change of car stock (Dahl 1982; Dahl & Sterner 1991; Graham & Glaister 2002). Dahl decomposes it even further, attributing majority of the effect to the change of car size rather than the number of vehicles. These findings suggest that omitting to account for vehicle stock, long-run income elasticity estimates may be severely overestimated (Dahl & Sterner 1991). This will be treated in our meta-analysis.

1.3 Oil consumption trends

As Ostro & Naroff (1980) note, there is a high level of dependence on automobiles in terms of commuting in the US. Data from 1974 show that 77% of commuters do so in an automobile and great majority of them, 85%, are drivers. These trips account for 42% of all vehicle-miles traveled. Current census data (U.S. Census Bureau 2009) show an even higher saturation with 86% commuters using personal vehicles, 88% of those driving themselves.

The US is an example of a highly saturated market where aggregate demand for energy does not increase as rapidly as in, for example, developing countries. If we pool together members of the Organisation for Economic Co-operation and Development (OECD) that are generally high-developed nations, we can see their decreasing importance on the oil market. While there is an overall growing trend in both oil and light distillates consumption,[1] the widening gap between the total value and OECD share signals growing importance of non-OECD markets. The world share in oil consumption of OECD in the middle of the 1960s was 75%, down to 53% in 2010.

Absolute consumption of oil is rendered in Figure 1.1(a). The growing difference between OECD share and total world consumption is further confirmed in Figure 1.1(b) where the percentage share of OECD on the world consumption is depicted, both in terms of oil and light distillates. More data including year-over-year comparisons are detailed in Appendix A.

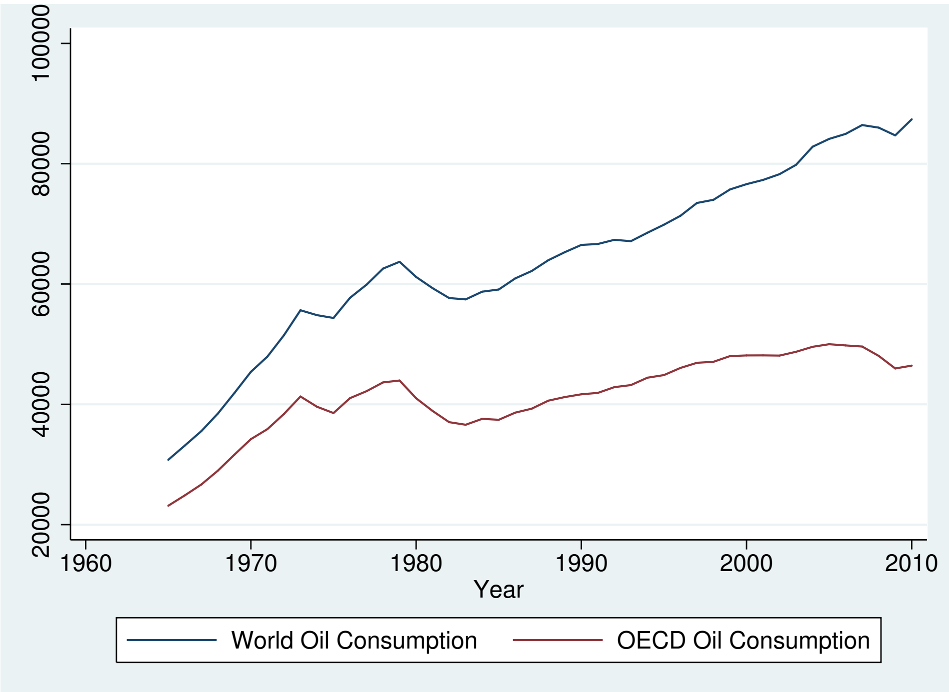
This rather significant development of a previously stable market did rekindle researchers' interest in elasticity estimation. The trend can be clearly seen in Figure 1.2, peaking in the early 1980s, after the second oil shock and general stability. The second peak in the kernel plot indicates the revival of estimation due to new techniques, cointegration in particular. This will be discussed later in Chapter 2.

1.4 Oil prices, shocks, and general volatility

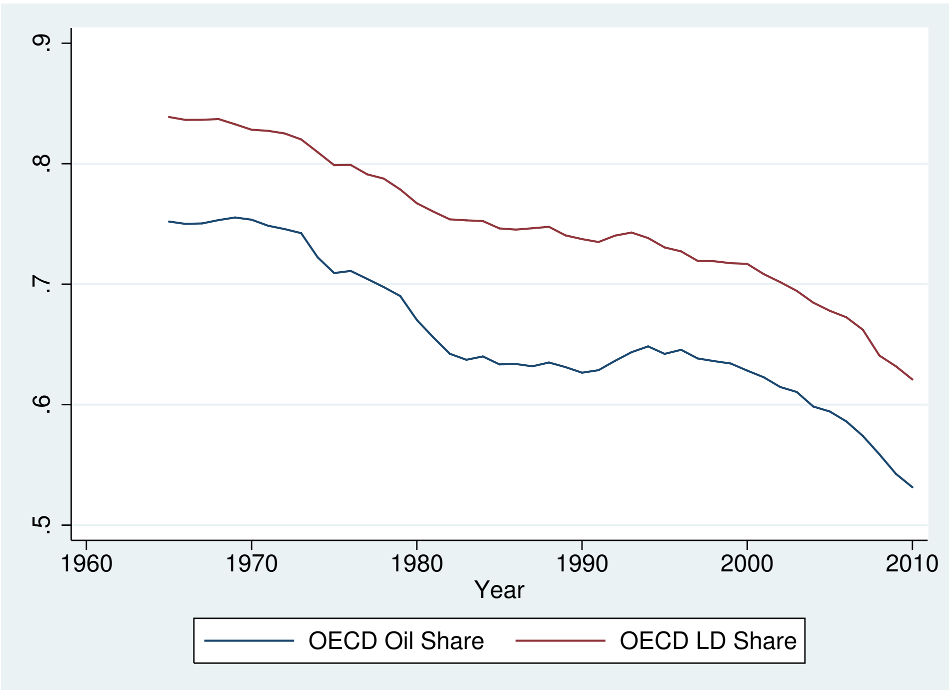
Up until the 1970s oil prices remained fairly stable, slightly increasing in nominal terms,[2] stable or decreasing in real terms. While Figure 1.3 only depicts gasoline and not other derivatives or oil itself, the situation for these is very sim-

[1] In this case "light distillates" refers to aviation and motor gasoline and light distillate feedstock.

[2] Average year-over-year change is 2.2% for the period 1945-1969.



(a) World and OECD shares in oil consumption. Thousands of barrels per day.



(b) OECD share of world consumption of oil and light distillates

Figure 1.1: Trends in oil consumption. Source: BP (2011)

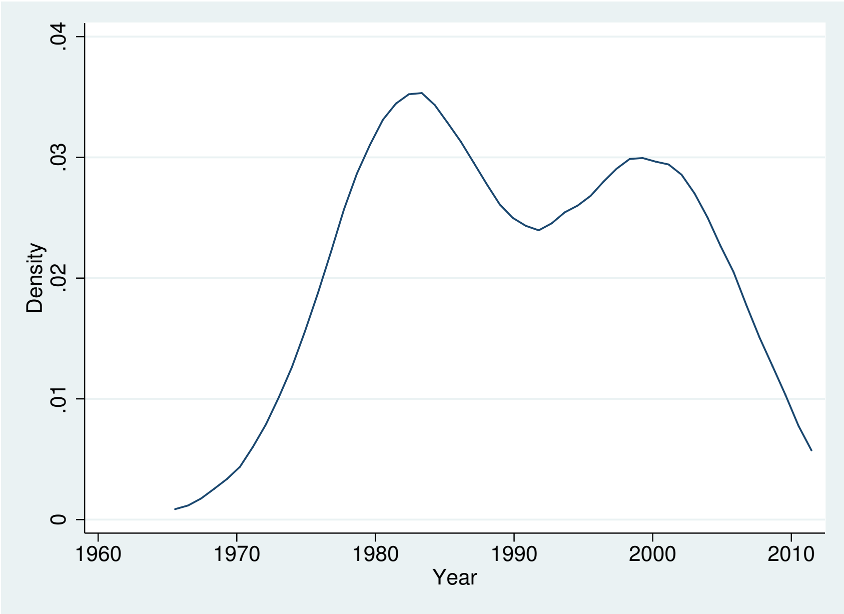


Figure 1.2: Frequency of studies on elasticity of gasoline demand. From a dataset further discussed in Section 4.2.

ilar as the prices are highly correlated between each other.[3] This is mainly due to the ratio of crude oil in these derivatives, namely gasoline and diesel, their crude oil content is 80% and 68% respectively (EIA 2012), so price movements are necessarily similar.

This steady price level of oil products was disrupted by oil shocks of the 1970s, first the Organization of Petroleum Exporting Countries (OPEC) oil embargo of 1973, lasting until March 1974, then the Iranian revolution of 1979, both events causing a price surge. Increasing surplus during the first half of the 1980s, the so-called oil glut, caused a steady decrease of oil prices. Consequently, in December 1985, OPEC members agreed to abandon production quotas for its members, resulting in an even deeper plunge of oil prices. All of these price shifts can be seen in Figure 1.3. Further comments on both the shocks and the early post-shock period can be found in Williams & Mount (1987).

While these unfortunate incidents are regarded negatively, there are two positive aspects to mention:

- Low volatility of oil/gasoline price during the preceding decades did make estimation of elasticities more difficult as robust estimators rely on more

[3] Correlations exceed 0.95 in all cases. Correlation matrices were calculated for oil, gasoline, diesel, and heating oil. All of them on annual, quarterly, and monthly data.

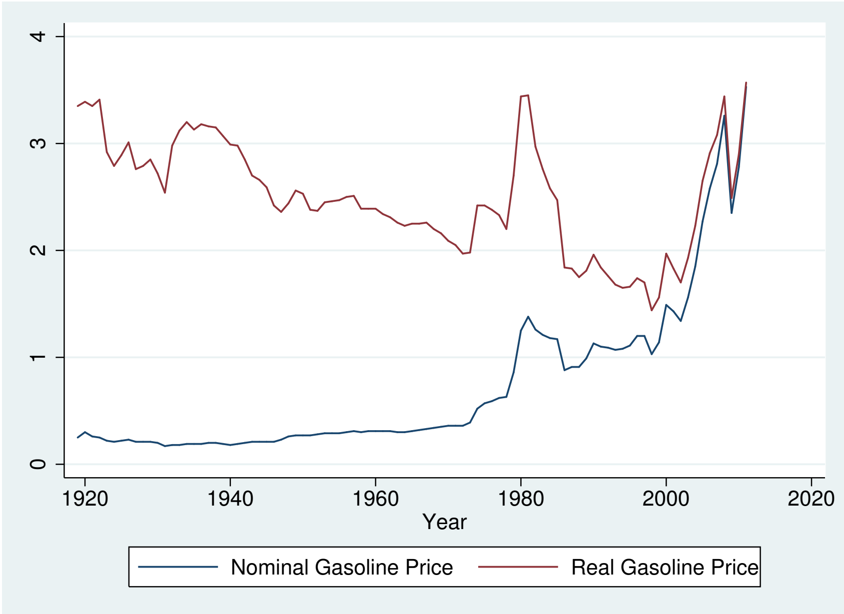


Figure 1.3: Difference in gasoline price trends. Prices are year averages in US dollars per barrel. Real prices are adjusted to the levels of February 2012. Source: EIA (2012)

fluctuating values. The sharp increase of prices in the 1970s allowed researchers to more precisely determine consumer adjustment not only to price changes.

- Producer welfare is not often taken into account. Al-faris (1997) notes that the increased prices and following stabilization not only increased revenues for oil producers, specifically Gulf Cooperation Council (GCC) countries,[4] but also allowed them to further diversify their economies, invest into infrastructure, and increase their standard of living.

Unlike the sudden price increases of the 1970s, the first decade of the 21st century was met with a gradual increase of prices, often referred to as the third oil crisis. Within five years, between 2003 and 2008, oil prices increased five fold in real terms, compared to a tripling in both previous crises. Possible causes, common features, and further analysis is to be found in Kesicki (2010).

[4]GCC countries are known for their oil and natural gas production. The council consists of Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and United Arab Emirates.

Chapter 2

Estimation

The last three decades of the 20th century offer a significant development of econometric methods. We have a new understanding of time series and their estimation, but we can go back to the roots and combine old methods with the new. This chapter offers an overview of the development in the field and also covers contemporary tools used in gasoline demand estimation. Good understanding of the methods helps us in our subsequent analysis of the individual results.

Energy demand estimation does exhibit some unique qualities that do not allow us to treat it as other consumer products. The main point is that people do not demand it directly, they demand transportation where gasoline is an input, so demand for gasoline is often called derived demand. While it is a non-durable good, this dependence on durable goods makes the estimation more difficult. For example as people demand certain amounts of travel, their gasoline consumption is dependent of the efficiency and price of vehicles.

Over the last 40 years, several approaches have been suggested, but consensus has not been reached yet. An overview including recent development follows. All econometric models are assumed to have homoskedastic disturbances with zero mean unless stated otherwise.

2.1 Basic models

All the models used throughout the years have one thing in common – gasoline demand is modeled as a function of its price and real income per capita. If a study does not contain one of these factors, it is generally considered misspecified (Dahl & Sterner 1991) and not considered by researchers performing meta-analyses and literature reviews. Other regressors may include automobile stock, average car efficiency, or prices of other inputs.

The response variable, gasoline demand, is not usually taken as an aggregate value. Some researchers use gasoline demand per driver or, if data are unavailable, they proxy the number of drivers by the adult population. Others use gasoline demand per vehicle.

The main difference between the two types of models presented is in the way they handle the dynamics, how the gasoline demand adjustment is laid out in time.

2.1.1 Static models

The static model does not estimate the short-run adjustment, it only models the overall response in the long run. Dahl (2012) suggests that results from static models should be treated as intermediate run as they yield lower estimates than dynamic models do. This will be further discussed in Subsection 4.2.2.

From here on now, G represents gasoline demanded, Y per capita income, P real prices, and Z_k other relevant explanatory variables.

$$\log G_t = \alpha + \beta_1 \log P_t + \beta_2 \log Y_t + \sum_{k=1}^{K} \beta_{k+2} Z_{kt} + u_t \tag{2.1}$$

2.1.2 Dynamic models

A different model, sometimes referred to as Auto Regressive Distributed Lag (ARDL), described early by Kennedy (1974) or Houthakker *et al.* (1974) does count with certain dynamics, with different consumer adaptation in the short and long run. The function of demand is assumed to be in log-linear form

$$G^* = f_2(P, Y) = \alpha Y^{\beta} P^{\gamma}. \tag{2.2}$$

Given that the desired level may not match the actual demand, there is an adjustment in time toward the ideal demand level.

$$\frac{G_t}{G_{t-1}} = \left(\frac{G_t^*}{G_{t-1}}\right)^{1-\lambda} \tag{2.3}$$

Substituting (2.2) into (2.3), taking the logarithm of both sides of the equation and adding a disturbance term, we arrive at

$$\log G_t = \log \alpha + (1-\lambda)\beta \log Y_t + (1-\lambda)\gamma \log P_t + \lambda \log G_{t-1} + u_t. \quad (2.4)$$

The coefficients of $\log Y_t$ and $\log P_t$ in Equation 2.4 are the short-run estimates for income and price, respectively. Dividing them by $1-\lambda$, thus obtaining β and γ, we get the long-run estimates. This elegant combination of both short- and long-run elasticities within one equation has made this model very popular. This simplified model assumes identical lag structure for the explanatory variables, so lags of the regressors are sometimes included. Thorough overview along with separating into groups can be found in Dahl & Sterner (1991).

There is one aspect of this model that does not bode well for meta-analyses where standard errors are considered. The problem is that while the short-run elasticity is estimated directly including its standard error, the situation is more complicated for the long-run estimate. Since it is computed as a ratio of two regression estimates, its standard error estimate is not outright clear. The so-called delta method can be used to approximate the desired measure of precision, but that is rarely done in practice, most researchers leave the long-term estimate without a standard error. For usage of the delta method in practice see Engsted & Bentzen (1997), the study includes a direct comparison with cointegration methods.

If a meta-analysis is done including the treatment of publication bias, the lack of standard errors is troublesome. In most cases, it is not possible to estimate the standard error from the information reported in studies, because the delta method requires the correlation between the two variables as well. Not even the biased standard error estimates would help us as the role of standard errors in meta-analysis regressions is to understand behavior of researchers when faced with a certain significance of the estimated variables. We may not use such a variable as explanatory if the study author never estimated it himself. One possible solution to this problem is devised in Subsection 4.2.2.

2.2 Cointegration and ECM

As econometric research progressed, new caveats of both dynamic and static models surfaced. As Granger & Newbold (1974) point out, when two unit root processes are regressed on one another, their estimated relationship and

R^2 may or may not be correct, because the problem of a spurious regression may arise. However, when the two variables are found to be cointegrated, in the sense coined by Engle & Granger (1987), the problem no longer arises as cointegration seeks to find a long-term relationship between the variables, despite their unit roots.

2.2.1 Unit roots and cointegration

Unit root testing was devised to check whether a variable is highly dependent on its first lag, its previous value. The test consists of finding out whether a process follows AR(1) where the lagged depended variable has the parameter equal to unity. After rearrangement, the test comes down to testing whether ϕ in (2.5) is significant, the depicted model is the augmented version of the Dickey-Fuller test. The original test without the summation is proposed in Dickey & Fuller (1979).

$$\Delta x_t = \alpha + \beta t + \phi x_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta x_{t-i} + \varepsilon_t \tag{2.5}$$

Given non-standard distributions, alternative critical values were suggested for different sample sizes, number of lags in the augmented test, and for whether the drift and trend terms were included or not. In addition to the Dickey-Fuller test, several other testing methods are summarized in Engsted & Bentzen (1997).

Test of unit roots in gasoline consumption, prices, and real income have been performed in various papers and the null hypothesis of a unit root could not be rejected using standard significance levels. See Alves & Bueno (2003), Akinboade *et al.* (2008) or Bentzen & Engsted (2001) for details.

If a unit root is not rejected, a check for cointegration can be carried out.We are looking for a stationary linear combination of these unit root processes. The coefficients resulting from this then depict their long-run relationships. In our case, we can estimate the following model

$$\log G_t = \alpha + \beta_1 \log Y_t + \beta_2 \log P_t + u_t. \tag{2.6}$$

If we rewrite this equation, leaving u_t on one side, all other variables on the other, we can see that should the linear combination of our three variables be stationary, the disturbances need to be stationary as well. Similarly to unit root testing, the null hypothesis of a unit root needs to be rejected in order for the disturbances to be stationary. As the disturbances are naturally unavailable, residuals are tested. Also as there are more variables in question, different critical values are used, see Engle & Yoo (1987).

This whole derivation in fact validates the long-run estimation using a static model, but only under strict assumptions of stationarity of the disturbances and non-stationarity of the other variables. As most studies did not carry out these tests, the validity of their results can be questioned.

A question arises: What if the variables are unit root processes, but the null hypothesis of a unit root of the disturbance cannot be rejected? Either there is not a long-run relationship between the variables in question or we failed to include a non-stationary variable in our cointegration test, thus leaving it in the error term. Several studies that failed to find a cointegration relationship in energy demand research are mentioned in Engsted & Bentzen (1997). Possible explanations are included, authors attribute this to misspecification.

There are several econometric properties of cointegration that should be mentioned:

- We cannot use standard t- and F-tests as the distributions resulting from cointegration are non-standard.
- Consistency of the long-run elasticity estimates is not affected by missing variables, as long as they are stationary. So if all the non-stationary variables are included in the regression, cointegration will yield consistent estimates.
- The simple OLS approach does have several drawbacks, including a bias in small samples (Banerjee *et al.* 1986). Techniques alleviating it have been devised, see Engsted & Bentzen (1997) for a thorough overview.

2.2.2 Error Correction Model

Cointegration alone does not describe the short-run adjustment, so Engle & Granger (1987) devise the Error Correction Model (ECM) for this. The rationale behind ECM is that whenever the consumer is not in equilibrium, that is the residual resulting from (2.6) is non-zero, he will try to get back to equilibrium in the following period. This adjustment towards equilibrium will allow us to estimate the short-run elasticities. ECM is modeled as follows:

$$\Delta \log G_t = \alpha + \sum_{i=0}^{m} \beta_{1i} \Delta \log Y_{t-i} + \sum_{i=0}^{n} \beta_{2i} \Delta \log P_{t-i} + \sum_{i=1}^{s} \beta_{3i} \Delta \log G_{t-i} + \gamma \hat{u}_{t-1} + \varepsilon_t, \tag{2.7}$$

where m, n, and s are selected so that ε_t is white noise, and $\hat{u}_{t-1}$ are the residuals from the cointegration equation, (2.6) in our case. Thanks to the fact that all G_t, Y_t, and P_t are I(1), their first differences are stationary, I(0), the lagged residuals from (2.6) are stationary as well, as we tested earlier. So the whole model involves only stationary variables and its disturbances are independent and identically distributed (iid), white noise.

In this setting, the first differenced lags of response variables in question depict the short-run elasticity.

2.2.3 Dynamic model revival

While it might seem that both the static and dynamic models were dominated by cointegration and ECM frameworks, it is not entirely so. The whole introduction of unit roots and cointegration certainly clarified the long-term adjustment behavior of consumers, but the small sample properties and the invalidity of standard errors are not favorable qualities. In the work of Pesaran & Shin (1998) a crucial thought is explored: Given our understanding of I(1) processes and cointegration, how does the ARDL model perform? They go on to prove the consistency of both the short- and long-run estimates when using ARDL, given that the underlying variables are I(1) and cointegrated. Also the resulting t- and F-statistics are valid, so hypothesis testing can be carried out.

Although cointegration and ECM provide a seemingly easy and straightforward approach, it is not without flaws and recent research by Pesaran & Shin suggests that ARDL models should not be abandoned just yet.

2.3 Bounds approach to cointegration

Even though the two step cointegration approach has become the leading method for gasoline demand estimation, recent papers use a different technique that alleviates some of the flaws and limitations described in the previous section. Unrestricted Error Correction Model (UECM) proposed by Pesaran *et al.* (2001) does not require the underlying variables to be non-stationary, does not suffer from severe small sample bias, but still remains consistent.

Allowing for both stationary and non-stationary variables in our models does not only ease one assumption, but also widens the possibilities for our variables. The model once again combines both short- and long-run elasticities

$$\Delta \log G_t = \alpha + \sum_{i=0}^{m} \beta_{1i} \Delta \log Y_{t-i} + \sum_{i=0}^{n} \beta_{2i} \Delta \log P_{t-i} + \sum_{i=1}^{s} \beta_{3i} \Delta \log G_{t-i} + \\ + \gamma_1 \log P_{t-1} + \gamma_2 \log Y_{t-1} + \gamma_3 \log G_{t-1} + u_t \tag{2.8}$$

Test of cointegration relationship is carried out using an F-test with the null $\gamma_1 = \gamma_2 = \gamma_3 = 0$, stating there is no long-run relationship. Since the F-statistic was found to be non-standard, Pesaran *et al.* suggest upper and lower bounds for the test, rejecting the null if the upper bound is exceeded, failing to reject if the statistic does not exceed the lower bound. The test is inconclusive in case the F-statistic lies between these two values.

The long-run elasticities are computed as ratios $-\gamma_1/\gamma_3$ and $-\gamma_2/\gamma_3$ for price and income, respectively. Given this indirect inference, their standard errors need to be computed additionally, similarly to the approach suggested in Subsection 2.1.2.

In gasoline demand estimation this UECM model is used by Akinboade *et al.* (2008) or Sa'ad (2009).

2.4 Data selection and pooling

Great majority of studies use annual data, but that of course means only several dozen observations are available at best. Given the asymptotic properties, small sample performance, inability to use many explanatory variables, and other issues with modern econometric tools, researches try to gather more data. There are two means of extending a dataset and thus generally improving the resulting estimation: pooling and using micro-level data.

2.4.1 Pooling

The topic of pooling and comparing homogeneous and heterogeneous models is extensively covered in Baltagi & Griffin (1997). There are two factors they use to judge the quality of estimators. First is the plausibility of the results given previous research and rationale, the other is their forecast quality. Leaving the last 10 years as out-of-sample control group, they order the estimators by their Root Mean Square Error (RMSE). They end up favoring the homogeneous estimators, partly thanks to the fact that their sample are 18 OECD countries that do not exhibit too great inter-country differences.

One caveat of international pooling may be a certain incompatibility of the data. Wheaton (1982) points out two specific problems. First is the difference in standards as he finds some countries to be reporting fuel efficiency differently than the rest, forcing him to create separate models. Second problem may arise with the difference in currencies. Wheaton points to Kravis *et al.* (1978) who constructed a cross national Gross Domestic Product (GDP) deflator for these inter-country comparison purposes.

2.4.2 Micro-level data

Using micro-level data allows researchers to investigate various subgroups within individual countries, separating them by income levels, regions, occupation, marital status, ... These studies aim to estimate the heterogeneity other researchers neglect as they consider countries to be homogeneous. The downside of this method is the availability and extent of the data. Micro-level information is expensive to obtain and is usually gathered using surveys with frequency far lower than one year, frequency used in most of the studies utilizing aggregate data. See Greening (1995), Archibald & Gillingham (1981), or Nicol (2003) for more details on individual studies and Graham & Glaister (2002) for an overview.[1]

2.5 Other remarks

2.5.1 Symmetry of estimators

The problem of asymmetry of price elasticities has surfaced not only in gasoline consumption research. The objective is that people are generally more sensitive to a price increase than to its decrease, while usual estimation is done with assumed symmetry. Summary of research results on this topic in the area of gasoline demand can be found in Dahl (2012). General practice is to augment the demand model so that it differentiates between new price

[1] The section "Micro-level Data: Individual and Household Demand Studies" and the summary in Table 6.

maxima, price cuts, and sub-maximum price recoveries. One of the studies using US data reports significantly larger adjustment when a new maximum is reached, specifically 30% higher elasticity than the symmetric estimate. Price cut responses are merely half of the symmetric ones.

There are only handful of studies including this decomposition, often reporting mixed results, so further research into this topic is necessary to have more reliable results on this subject.

2.5.2 Cointegration and panel data

Using the traditional static and dynamic models, researchers have explored not only single country datasets, but have also dealt with heterogeneous groups of samples, employing techniques like fixed/random effects and the like. This approach gets more difficult in unit root testing and cointegration. The augmented framework has been used in several studies, but still the great majority involves only one country at a time. See Baltagi & Kao (2001) for a comprehensive overview of econometric tools for panel data treatment in cointegration.

2.5.3 Influence of econometrics tools selection

Some of the discrepancy between estimates is attributed to the choice of econometric methods. This aspect is usually covered in overviews and meta-analyses (Espey 1998; Havranek & Irsova 2010). Another approach to corroborate this can be found in Baltagi & Griffin (1997). On the same dataset, various methods are employed in order to find out how influential their selection can be. They find great discrepancy in long-run price estimates, ranging from -0.24 to -1.42, and short-term income elasticity, estimated between -0.65 and -0.92.

Apart from the influence of data selection, most notably the frequency of observations, this volatility is of concern as well.

2.5.4 Developing countries

Gasoline demand in low income economies generally differs from the rest of the world. On one hand, car stock is usually much lower, on the other hand, economic growth often surges. Studies covering these countries frequently yield high income elasticities and Storchmann (2005) suggests this is due to their high marginal rate of consumption. Consensus has not been reached as some studies indicate this does not apply universally.

Another differentiating factor is data availability. Even if researchers use the most advanced econometric techniques, missing information may severely hamper their estimation. For example car stock, as discussed earlier, was shown to be a crucial moderator variable as its omission may lead to severe overestimation of the income effect.

These factors will be controlled for in our meta-analysis.

Chapter 3

Meta-analysis – current methodology

Given the amount of research dedicated to specific economic problems, surveys of the results often get published, gasoline demand is no exception. While these surveys are important, their subjective approach and lack of rigorous assessment is of concern. Meta-analysis tackles the problem of research result variance using econometric methods.

Meta-analysis itself has been long used in science, first coined by Glass (1976), later spread in the area of economics following the publication of Stanley & Jarrell (1989). Only the toolset to be used in Chapter 4 is presented here, for a complete overview of contemporary methods see Nelson & Kennedy (2009) and Stanley *et al.* (2006).

The idea behind meta-analysis is to explore factors that influence research results. After gathering as many studies on the same topic as possible, various information about each work are collected. These include the sample size, econometric methods used for estimation, data used, specification, year published, sometimes even sex of the researcher.

Such an approach aims to be more objective about the inference than traditional survey methods. However, it is still subjective in a way since the underlying data are collected and models are constructed by the researcher and their adjustment may affect the outcome.

3.1 Heterogeneity

Given that the variance of estimates is too large to be explained by the disturbance terms, we speak of between study heterogeneity. As Christensen (2003) points, there are two types of heterogeneity: factual and methodological. Factual heterogeneity concerns actual population differences, for example based on countries where the research was conducted. In our case of gasoline demand elasticities we examine if there is a difference between more saturated markets and developing countries, expecting the latter to be more sensitive. Methodological heterogeneity stems from different procedures used, be it econometric methods, data frequency, or models. As was described earlier in Subsection 2.5.3, this may be our case as well.

3.2 Publication bias

Apart from the obvious characteristics that influence the results, there is one factor that may bias the outcome – the researcher himself. If the result of his study is not in line with the theory or previous results, he may withhold his findings. Another practice is to keep modifying the specification or data until the results are consistent with the standard outcomes. Insignificant estimates may result in tampering with the model as well, this can be observed as a lot of estimates are just significant at the 5% level. All of these practices need to be measured and accounted for as they bias inference from the sample of collected observations.

3.3 Graphical approach

Before testing for publication bias using econometric methods, a simple visualization may benefit the analysis. While this approach is less objective and less informative in the sense of finding the underlying population value, it is helpful in order to get the overall picture of the results. The so-called funnel plot, extensively covered in Stanley & Doucouliagos (2010), visually describes individual observations on the horizontal axis along with their measure of precision, most commonly inverted standard error, on the vertical axis. The idea behind this graph is that the most precise estimates, those with the shortest confidence intervals, will be on top what should be an inverted funnel.

The problem is that often in practice the plot does not resemble a funnel, because there is a part of it missing. One of the reasons of this may be the lack of estimates that are inconsistent with theory, this is true especially if the population value is close to zero. In case of this funnel asymmetry, suspicion of publication bias may arise. The whole theory of funnel resemblance stems from the idea that there is one underlying population value. If this assumption is violated, funnel asymmetry no longer indicates publication bias. That is one of the reasons this graphical approach is merely informative, judgments should be made only after they have been checked using econometric methods.

3.4 Econometric models

Explaining an underlying population value using qualities of individual studies leads us to the following equation suggested by Stanley & Jarrell (1989). The estimate being a dependent variable, explained by various factors Z_k, the population value in question, β, and the iid disturbance e_j

$$b_j = \beta + \sum_{k=1}^{K} \alpha_k Z_{jk} + e_j \qquad j = 1, 2, \ldots, L. \tag{3.1}$$

The variables Z_{jk} may include information about model specification, publication outlet, number of observations, and so on.

As we will see later, this simple model has been extended and adjusted for various innovations in meta-analysis that have occurred since the late 1980s.

3.4.1 Funnel asymmetry

As we combine econometric modeling with publication bias investigation, we are essentially testing the asymmetry of our funnel plot. Building upon the asymmetry, we can model this dependence in the following way:

$$b_j = \beta + \alpha_0 se_j + \sum_{k=1}^{K} \alpha_k Z_{jk} + e_j, \tag{3.2}$$

where the estimate depends not only upon the characteristics from (3.1), but also on its standard error se_j. In this specification, α_0 measures the degree of publication bias.

The problem with variable se_j is that it is measured with an error, so the model suffers from error-in-variable bias. Stanley (2005) suggests using an IV, specifically the square root of the sample size or number of degrees of freedom. Such a variable is by definition highly correlated with the standard error and exogenous in the model. Thus it constitutes a valid instrument. Stanley (2005) further notes that it may not be so every time, especially if the correlation between $\sqrt{n}$ and se_j is low. It is also important to stress that IV regression yields consistent, not necessarily unbiased estimates.

Given the nature of the data, the disturbance e_j is unlikely to be homoskedastic, this was also confirmed in our sample.[1] While this does not affect our estimates, the standard error estimates will be biased. As a remedy, Stanley (2008) suggests to use WLS instead, using the standard errors as weights. The response variable becomes the t-statistic, together we obtain the following specification

$$t_j = \beta/se_j + \alpha_0 + \sum_{k=1}^{K} \alpha_k \frac{Z_{jk}}{se_j} + \varepsilon_j. \tag{3.3}$$

In this model the publication bias is treated as constant throughout the sample, but the constant may be decomposed using additional moderator variables

$$t_j = \beta/se_j + \alpha_0 + \sum_{l=1}^{L} \delta_l Z_{jl} + \sum_{k=1}^{K} \alpha_k \frac{Z_{jk}}{se_j} + \varepsilon_j. \tag{3.4}$$

This constitutes a quality model for modeling heterogeneity both in the estimate and the publication bias. For the best estimation of the effect beyond publication bias, Stanley & Doucouliagos (2007) suggest that since the effect of standard errors may be quadratic, we can model the asymmetry in the following way

$$t_j = \beta/se_j + \gamma_0 se_j + \varepsilon_j. \tag{3.5}$$

This specification helps us estimate the effect throughout the whole sample, an improved average after publication bias treatment.

3.4.2 Between and within study heterogeneity

Heteroskedasticity in the original model was dealt with, but only between the studies themselves. Estimates from the same study often share the same qualities in terms of estimation methods, data, ... This may result in correlation of

[1] Standard Breusch-Pagan test yielded p-values of 0.003 and 0.0231 for short-run and long-run estimates, respectively.

these estimates, a problem described early in Stanley & Jarrell (1989). This is even more pronounced as the number of estimates per study increases. In meta-analysis surveys, median of this value is reported to be three (Nelson & Kennedy 2009), in our case it is eight, both for short- and long-run estimates.

As a remedy, Nelson & Kennedy (2009) suggest a mixed-effects multilevel model. This setting allows us to have an error term for each study, obtaining a composite error together with the estimate-level disturbance. Extending (3.4) and (3.5), respectively, we obtain

$$t_{ij} = \beta/se_{ij} + \alpha_0 + \sum_{l=1}^{L} \delta_l Z_{ijl} + \sum_{k=1}^{K} \alpha_k \frac{Z_{ijk}}{se_{ij}} + u_i + \varepsilon_{ij}, \tag{3.6}$$

$$t_{ij} = \beta/se_{ij} + \gamma_0 se_{ij} + u_i + \varepsilon_{ij}. \tag{3.7}$$

In the final specifications, we can see the quadratic effect of publication bias through the estimate of γ_0, β reflects the underlying population value. Two sets of moderator variables α_k and δ_l denote the effect on the estimate itself and on publication bias, respectively. Our disturbance is split between a study-level error u_i and an estimate-level disturbance ε_{ij}. With $u_i|se_{ij} \sim N(0, \theta)$, $\varepsilon_{ij}|se_{ij} \sim N(0, \phi)$, and these errors being uncorrelated, variance of the composite error is a simple sum, $\text{var}\,(u_i + \varepsilon_{ij}) = \theta + \phi$. The closer the study-level variance θ is to zero, the weaker is the case to use the mixed-effects framework instead of OLS.

3.5 Criticism

As originally proposed by Glass (1976), meta-analysts often try to include all studies available in order to be as objective as possible. Such approach has been criticized for the lack of control of quality. As the final analysis is only as good as the estimates used, contaminated data may render the outcome questionable. To alleviate this, Slavin (1986; 1995) proposes to include only quality estimates through a careful selection based on a priori selected inclusion criteria. This approach, best-evidence synthesis, aims to combine the best of classical literature reviews and meta-analysis by not being too distant while still employing tools that are objective.

Chapter 4

MRA of gasoline demand income elasticity

There are several aspects of this thesis and its take on meta-analysis that differ from previous papers. These include publication bias treatment, in this research area performed earlier only in Havranek *et al.* (2012) on price elasticity, modern econometric framework taking into account between study heterogeneity, a very large dataset, and the parsimony of our models. The last point is of great importance, we will not try to unravel all the causes of heterogeneity of our sample, that has been thoroughly covered before (Dahl & Sterner 1991; Dahl 2012; Espey 1998), we will focus on the main influencing factors that are crucial and comment on their effect.

4.1 Previous analyses and surveys

With so much research interest in energy demand, various surveys and analyses emerged early on, non-econometric surveys include Dahl & Sterner (1991), Dahl (2012) or Graham & Glaister (2002). These papers stress the importance of model specification, grouping the studies by their choice of explanatory variables or lag structures. Having compared the studies within these clusters, not as a whole, Dahl & Sterner (1991) conclude *"[...] by a careful comparison we find that if properly stratified, compared and interpreted, different models and data types do tend to produce a reasonable degree of consistency."* Apart from arithmetic means, medians, and visualization of the results, these surveys do not use any statistical frameworks to estimate the population value.

To mitigate this problem, several meta-analyses on the topic have been performed, namely Espey (1998), Brons *et al.* (2008), and Havranek *et al.* (2012). These analyses differ in various factors, including the choice of data, econometric toolset, or treatment of publication bias. Only Espey (1998) deals with income elasticity, averaging 0.47 in the short-run and 0.88 in the long-run with medians close, namely 0.39 and 0.81, respectively. Basic features of the three meta-analyses are summarized in Table 4.1.

Table 4.1: Overview of previous meta-analyses.

	Espey	Brons *et al.*	Havranek *et al.*
# of studies	101	43	41
Range	1966-1997	1974-1999	1974-2011
# of estimates	LR price 277, LR income 245, SR price 363, SR income 345	SR price 191, LR price 79	SR price 110, LR price 92
Approach	OLS	Seemingly unrelated regressions	mixed-effects, clustered OLS

LR and SR stand for long-run and short-run, respectively

None of the previous surveys or analyses on income elasticity treated for publication bias, such practice can only be found in Havranek *et al.* (2012) who researched price elasticity. That study was the first one to do so in gasoline demand research, this thesis aims to fill in the gap with regards to income elasticity.

4.2 Data

The three meta-analyses mentioned in the previous section did use mostly peer-reviewed journal articles, together using 150 unique studies. This thesis uses a dataset previously used in Dahl (2012). It is a set containing several thousand observations, complied from 240 papers, books, working papers, mimeographs, . . . Together it constitutes arguably the widest sample used in gasoline demand research, certainly the largest used for a meta-analysis in in this field. Based on Nelson & Kennedy (2009), this dataset greatly exceeds the mean and median of the 125 meta-analyses surveyed in terms of the number of observations.[1]

[1]Nelson & Kennedy report the mean and median to be 191 and 92, respectively. Number of studies is 42 and 33 for mean and median, respectively.

Several aspects of this dataset help us in our inference.

- Pooling of published and unpublished studies will allow us to control if such stratification has some impact on either the estimates themselves or on the degree of publication bias.
- The size of the dataset will cause lower standard errors, hence more precise estimates. The number of degrees of freedom will not limit our selection of explanatory variables.
- Recorded t-statistics will allow us to treat for publication bias.

Apart from the elasticities themselves, the following information (among other) are recorded:

- t-statistic as a measure of precision,
- econometric method of estimation,
- lag dynamics used to estimate the long-run effect,
- goodness-of-fit, mostly R^2 or $\bar{R}^2$,
- estimate and the t-statistic of lagged dependent variable, if used,
- countries included in the study.

The dataset was further edited as entries without valid t-statistics were removed, model specifications were transformed into dummy variables, countries were pooled to groups of OECD and non-OECD countries, and information about journal publication was added. This augmented version is included on the attached DVD, the original version can be found online (Dahl 2010).

4.2.1 Car stock information

Even though data should not be further stratified in meta-analysis as usage of moderator variables is preferred, this case is an exception. As noted earlier in Section 1.2, survey authors the usage of car stock indicator to be a major influence on the long-run estimates of income elasticity. A quick look at Figure 4.1 suggests that this might be true in our dataset as well.

As will be shown later, the long-run elasticities without car stock information are almost twice as large as the rest, based both on weighted and

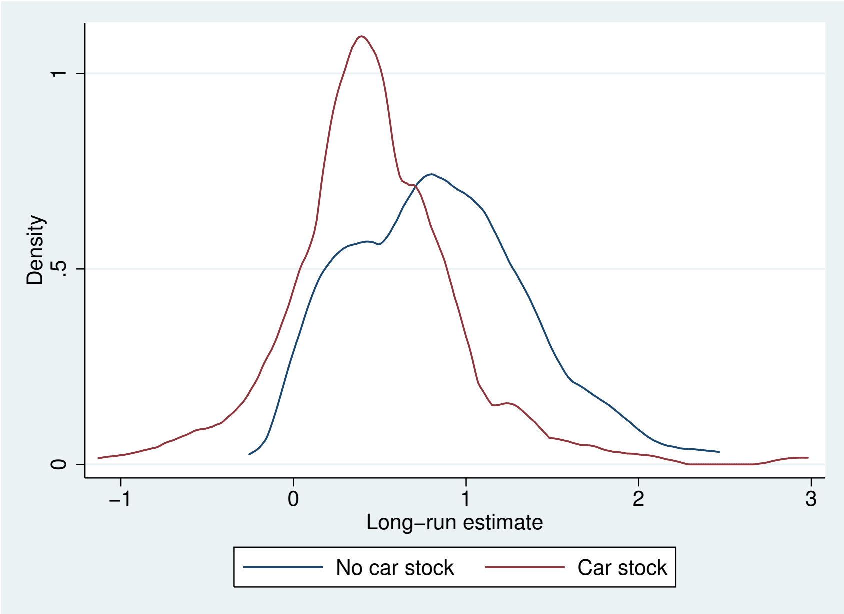


Figure 4.1: Densities of long-run estimates with and without car stock information.

unweighted averages. This strengthens our belief that these two subsamples represent two different phenomena. As the car stock adjustment forms a major part of the effect based on income change, the estimates with this information present report the long-run income elasticity beyond car stock adjustment. The other estimates reflect the total adjustment to income variation. From now on, these two subsamples will be treated separately.

Even though some surveys and analyses point to this discrepancy (Dahl & Sterner 1991; Dahl 2012; Espey 1998), primary studies rarely acknowledge this and publish their results as "long-run" regardless of inclusion of car stock information.

4.2.2 Lack of standard errors

Our analysis of publication bias is based on the knowledge of standard errors of individual estimates, but they are not always reported. As noted earlier, when models with a lagged dependent variable are employed, long-run effects are calculated as a non-linear combination of two or more regression estimates. As researchers in the vast majority of cases do not approximate the standard errors for these estimates, we are faced with a thousand long-run observations where precision information is unavailable.

As noted earlier, measurement error is sometimes alleviated by the use of instrumental variables for the inverted standard error variable (Stanley 2005), specifically the square root of number of degrees of freedom or observations. We proceeded with the number of observations that could be computed for the majority of our dataset. As the standard errors are missing, we cannot check the validity of our instrument directly, but comparison with inverted standard errors of our short-run estimates can be carried out. Stanley illustrates a poor instrument when correlation is 0.24. Our estimation yielded a rather low correlation as well, 0.31. As there is no reason for us to believe the correlation between this instrument and our missing standard errors would be any higher, we did not proceed any further with this IV estimation.

An alternative solution is as follows. If a researcher is presented only with standard errors of short-run estimates, they may influence the selection of long-run estimates. It would make sense to regress the long-run estimates on a measure of precision of the short-run estimate that comes from the same equation. We ran three separate tests, controlling for car stock in all of them and further regressing on various information of the short-run estimate. Measures were standard errors, t-statistics, and a dummy reporting significance at the 5% level.

Table 4.2: Indirect bias detection

	Standard errors		t-statistics		Significance	
auto	-0.281***	(-4.22)	-0.158**	(-2.35)	-0.103	(-1.52)
se	2.505***	(6.99)				
se^2	-2.601***	(-5.18)				
t-statistic			0.0712***	(7.13)		
t-statistic^2			-0.00228***	(-4.34)		
significance					0.339***	(8.80)
Constant	0.743***	(15.13)	0.685***	(12.28)	0.655***	(11.56)
N	741		741		741	

Response variable: long-run estimate

t statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

As regression results in (4.2) indicate, there is a very significant effect of short-run estimate significance on the magnitude of the long-run elasticity estimate. While these tests are not suitable for estimating the underlying value, they hint at possible misconduct. Given these information, simple arithmetic means or medians reported in surveys should be taken with a grain of salt.

4.3 Funnel asymmetry

Before the econometric analysis itself, we may inspect the dataset using several methods. The phenomenon of publication bias is expected to occur and there are multiple indicators of it just from the data visualization itself.

First, funnel plots in Figure 4.2 are heavily skewed. We can clearly see that the left part of the graph is almost completely missing in the funnel for short-run estimates, suggesting publication bias towards positive results that are more consistent with theory. The second funnel with long-run estimates shows two spikes, each for one of the two subsamples described in the previous section.

This asymmetry causes estimators such as arithmetic mean or median to report biased estimates, in our case these estimates will be positively biased as negative estimates of short-run effect and low positive estimates of long-run effect are not generally reported. This funnel asymmetry strengthens our case to use econometric methods that deal with publication selectivity.

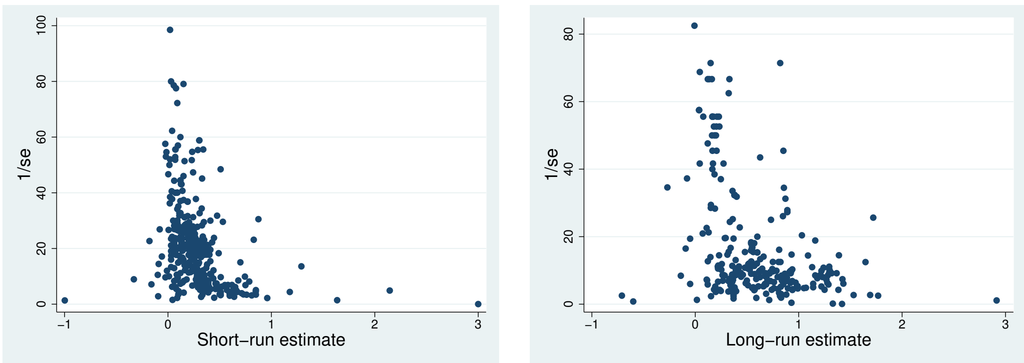


Figure 4.2: Funnel plots, only published estimates plotted.

Second, the densities of t-statistics of our estimates, depicted in Figure 4.3, exhibit a sharp increase around the value of 2. That roughly corresponds to a 5% significance level in a two-tail t-test for positive estimates – a measure that is a rule of thumb rather than a population value. High occurrence of t-statistics around plus or minus two therefore indicates possible intentions of researchers to alter their models in order to make their estimates more significant.

Third, as Stanley *et al.* (2010) suggest, a quick way to test for funnel asymmetry is to take estimates with the highest inverted standard errors, those at the top of the funnel, and compute their mean, usually 10% of the whole sample. These points should represent the most precise estimates from the whole

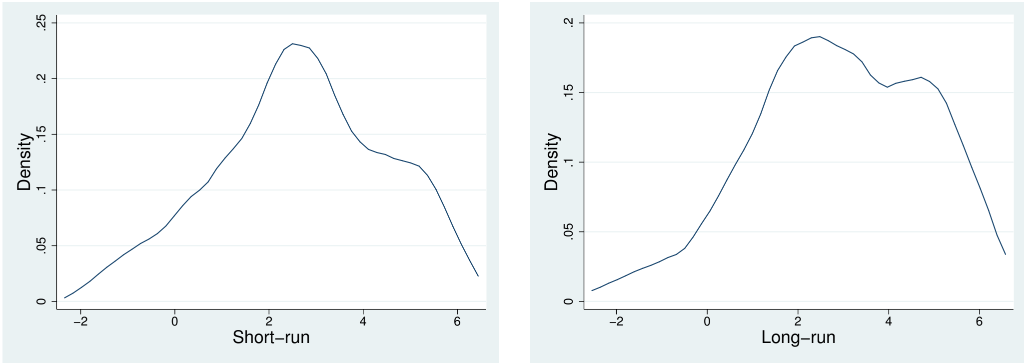


Figure 4.3: Kernel densities of t-statistics. Data truncated.

sample, thus their average should be close to the population value. Computing a weighted average for these subsamples, we arrived at 0.138, 0.329, and 0.636 in order for short-run, long-run with car stock, and long-run without car stock. These values are somewhat lower than the means and medians reported in the summary statistics in Table 4.3. This test is by no means rigorous as underestimated standard errors or outliers caused by typos may severely affect this estimate.

Table 4.3: Summary statistics

	Count	Avg.	Median	Std. Dev.	Min.	Max.
Short-run	831	0.284	0.250	0.326	-1.17	3
Long-run, car stock	346	0.465	0.395	0.509	-1.13	2.98
Long-run, no car stock	346	0.861	0.838	0.519	-.256	2.466

All these tests suggest the samples are skewed and reporting means or medians is not sufficient when looking for the underlying population value.

4.4 Model specification

First we will use the simplified funnel asymmetry test (3.7) and then the extended model (3.6) with moderator variables from our dataset. The FAT test only requires t-statistics, the estimates themselves, and stratification by studies as mixed-effects multilevel framework is used. As the extended model is concerned, even though there are dozens of potential explanatory variables, we focus only on the following.

OECD membership As noted earlier, demand in lower income countries may be affected by higher marginal rate of consumption, thus it is expected the long-run income elasticity will be higher as well.

Publication Publishing a paper in a peer-reviewed journal signals the study had to be approved by several people, suggesting certain expected quality. As the effect is not expected to be large, further specification (e.g. impact factor) was not considered.

Time dimension Studies employing annual data may yield different results than cross-section time series or pure cross sections that usually have significantly larger numbers of observations.

Car stock As noted earlier in the chapter, car stock information inclusion greatly affects the resulting estimate, so for the long-run elasticity regression, a dummy will be included to account for this information about the original model specification.

4.5 Results

All the results presented below come from regressions using the mixed-effects framework presented in the previous chapter.

After correcting for several outliers, the basic funnel asymmetry test yielded results depicted in Table 4.4. Publication bias extent represented by the constant term is significant at the 10% level for all models but the long-run model pooling all the data.[2] This is not unexpected as we saw the two spikes in the funnel plot in the previous section. This insignificance and subsequent significance when stratified strengthens our case about the division based on car stock information.

To judge the extent of publication bias, Doucouliagos & Stanley (2008) run Monte Carlo simulations and construct thresholds for the value of the constant in a funnel asymmetry test to distinguish the degrees of publication bias. By their terminology our short-run and long-run without car stock samples exhibit "severe" cases of publication bias, long-run with car stock contains "substantial" amount of bias.

To estimate the true underlying effect beyond publication bias, we employ the Heckman meta-regression with a quadratic effect, its results are summarized in Table 4.5. As expected, all estimates of the underlying value are highly

[2]Not included in the table.

Table 4.4: Funnel asymmetry test

	Short-run		Long-run			
	Whole sample		Car stock		No car stock	
1/se	0.0837***	(10.06)	0.209***	(7.71)	0.592***	(15.50)
Constant	2.997***	(7.97)	1.573**	(1.97)	3.032*	(1.77)
Observations	831		346		346	

Response variable: t-statistic
t statistics in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

significant, p-values need not be reported as the t-statistics exceed 9 in all cases.

Table 4.5: Overall results without moderator variables

	Short-run		Long-run			
	Whole sample		Car stock		No car stock	
1/se	0.0999***	(12.47)	0.234***	(9.76)	0.644***	(17.38)
se	-0.140	(-1.24)	-0.0501	(-0.11)	0.965***	(2.73)
Observations	831		346		346	

Response variable: t-statistic
t statistics in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

The difference in the estimates of underlying values of long-run elasticity allows us to approximate the decomposition of income elasticity. When adjusted for change of car stock effect, the estimated elasticity shrinks to about a third, this suggests that two thirds of the income effect affect the car fleet, the rest results in altered usage of the vehicles.

Table 4.6 offers a comparison of our regression results to widely used metrics. The weighted mean in our table is a result from a mixed-effects regression without publication bias treatment. Looking at the discrepancy between the values, we can see that classic tools do overestimate the effect due to asymmetry, due to the fact that the estimates from studies are not randomly distributed around the population value as their distribution is highly skewed. This affects inference based on these metrics, for example Dahl (2012) estimates the decomposition of long-run income elasticity based on the discrepancy in the two subsamples based on the car stock information. While we found the car stock sample to be about one third of the other estimate, Dahl found it to be one half, based on a sample average. The magnitudes in absolute values are also somewhat higher in her case.

Table 4.6: Comparison of regression results with sample means.

	Short-run	Long-run		
	Whole sample	Whole sample	Car stock	No car stock
MRA estimate	0.0999	0.457	0.234	0.644
Sample mean	0.284	0.663	0.465	0.861
Weighted mean	0.349	0.614	0.424	0.857

Comparing our MRA results with the only meta-analysis performed (Espey 1998), we find her results, 0.47 and 0.88 for short- and long-run, respectively, to be much closer to the sample averages rather than the estimates from our analysis. While the long-run estimate is comparable with our sample mean, the short-run estimate is not only higher than our average, it is four times larger than our estimate beyond publication bias. This overestimation may be partly due to the lack of publication bias treatment, but also due to the fact that Espey included estimates with unknown time structure to both short- and long-run samples, she also truncated her dataset by removing any negative estimates as inconsistent with theory, thus possibly biasing upwards both estimates.

The extended model using the short-run sample did yield only insignificant estimates of all possible moderator variables suggested in the previous section. Thus only the long-run estimation will follow.

Results from an augmented regression of the long-run effect are in Table 4.7. This analysis was not separated into samples by their car stock treatment as the influence of other factors is assumed to be constant throughout the whole sample.

Four variables are weighted by the standard errors so that we can observe their influence on the estimate itself. One variable is included unweighted so that we can estimate its effect on publication bias.

To control for the separation by car stock, only a dummy variable is added, attaining one if car stock information is included, zero otherwise. Other variables include OECD membership (=0 if studied country is a member), publication information (=0 if published in a journal), and time frame (=1 if estimating a pure time series, =0 in case of cross sections or cross-section time series).

As estimated before, the gap between the samples based on car stock in-

Table 4.7: Long-run estimation with moderator variables

	Estimate	
1/se	0.498***	(11.79)
unpublished	−0.0457	(−0.02)
unpublished (weighted)	0.206***	(4.19)
non-OECD (weighted)	0.149***	(2.64)
auto (weighted)	−0.460***	(−11.30)
time series (weighted)	0.150***	(3.32)
constant	1.103	(0.84)

Response variable: t-statistic
t statistics in parentheses
* $p < .1$, ** $p < .05$, *** $p < .01$

formation is substantial and this regression does not prove otherwise. Also our hypothesis about lower income countries cannot be rejected as the results indicate higher estimates in non-OECD countries. Models based on time series tend to yield slightly higher estimates. Due to their limited degrees of freedom, they are usually less significant[3] and form the majority of the right side of the funnel, they are more spread out. Given the selectivity and funnel asymmetry, this higher average elasticity can be explained.

Extreme insignificance (p-value of 0.98) of the unweighted `unpublished` variable shows that publication bias is common to both published and unpublished studies. However, the weighted variant does indicate that published studies tend to be more conservative, yielding lower estimates.

[3] Average t-statistic for time series models is half of the rest.

Chapter 5

Conclusion

The overview of econometric methods in the second chapter shows how complicated the subject has become. Every approach contains multiple drawbacks and it remains unclear what technique is the most suitable for gasoline demand modeling. Even though the major breakthroughs in econometrics were developed in the 1970s and 1980s, new estimators based on these techniques arise and offer new perspectives.

The meta-analysis itself differs in several aspects from previous research. First, I employed a much larger dataset that did let me employ asymptotic qualities of my estimators and I did not need to worry about degrees of freedom and other small sample issues. Second, I took publication bias into account. It is an issue present in a lot of areas of research, a practice that may skew the pool of results and disallow a survey author to employ traditional means of estimating the underlying effect. Third, I stratified the long-run estimates according to one model specification detail – car stock information. As it turns out, this little detail in model specification changes the interpretation of a long-run income effect. Without acknowledging car fleet information, researchers estimate the whole income effect, with that information included, only adjustment *beyond* vehicle stock change is estimated.

With that accounted for, regressions yielded elasticities of 0.1, 0.234, and 0.644 for short-run, long-run with car stock, and long-run without car stock, respectively. The importance of these results is two-fold. First, it shows how arithmetic and weighted means are unreliable in the presence of publication bias as our findings differ greatly from values previously reported. Second, the division by car stock shows great discrepancy between the estimates and makes a very strong case against pooling these together.

After estimating the underlying population values, I proceeded to inspecting some of the heterogeneity in the dataset. Data show higher sensitivity to income effects in countries with lower vehicle saturation, generally lower income countries. This corroborates the theory behind higher marginal rate of consumption in these regions. It was also shown that unpublished studies generally yield higher estimates, but the publication outlet does not affect publication bias.

This analysis reveals that pure income elasticity is generally low and increases considerably only when vehicle stock is not accounted for. This suggests a new category of estimates to be reported in gasoline demand research.

As our analysis was dependent on the knowledge of standard errors of individual estimates, we had to exclude long-run estimates resulting from ARDL models as they lack this information. To utilize this rather large part of our dataset, we devised a new approach of indirect publication bias testing, employing precision information of the short-run estimate, that revealed this subsample may be skewed as well. This suggests that even if the precision information is not known, bias may occur as the researcher acts upon the standard error and magnitude of the short-run elasticity.

Even though meta-analysis has the advantage of aggregating dozens or hundreds of studies and thus produces robust results, importance of the underlying research should not be underestimated. While surveys and analyses try to look for the population value and explain the heterogeneity, individual research papers are essential as they usually cover a specific geographical area and consider specific issues or events, e.g. changes in legislation. This level of detail and rigor is something that cannot be leveraged in a meta-analysis and it is exactly why individual research and analyses complement each other.

Bibliography

Akinboade, O. A., E. Ziramba, & W. L. Kumo (2008): "The demand for gasoline in South Africa: An empirical analysis using co-integration techniques." *Energy Economics* **30(6)**: pp. 3222–3229.

Al-faris, A.-r. F. (1997): "Demand for oil products in the GCC countries." *Energy Policy* **25(1)**: pp. 55–61.

Alves, D. C. O. & R. D. L. D. S. Bueno (2003): "Short-run, long-run and cross elasticities of gasoline demand in Brazil." *Energy Economics* **25(2)**: pp. 191–199.

Archibald, R. & R. Gillingham (1981): "A Decomposition of the Price and Income Elasticities of the Consumer Demand for Gasoline." *Southern Economic Journal (pre-1986)* **47(4)**: pp. 1021–1021.

Baltagi, B. H. & J. M. Griffin (1997): "Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline." *Journal of Econometrics* **77(2)**: pp. 303–327.

Baltagi, B. H. & C. Kao (2001): "Nonstationary panels, cointegration in panels and dynamic panels: A survey." .

Banerjee, A., J. J. Dolado, D. F. Hendry, & G. W. Smith (1986): "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence." *Oxford Bulletin of Economics and Statistics* **48(3)**: pp. 253–277.

Bentzen, J. & T. Engsted (2001): "A revival of the autoregressive distributed lag model in estimating energy demand relationships." *Energy* **26(1)**: pp. 45–55.

BP (2011): "Statistical Review of World Energy 2011." `http://www.bp.com/statisticalreview`. [Online; accessed 21-January-2012].

Brons, M., P. Nijkamp, E. Pels, & P. Rietveld (2008): "A meta-analysis of the price elasticity of gasoline demand. A SUR approach." *Energy Economics* **30(5)**: pp. 2105–2122.

Christensen, P. (2003): "Topics in meta-analysis: a literature survey." *Institute of Transport Economics, Oslo* .

Dahl, C. & T. Sterner (1991): "Analysing gasoline demand elasticities: a survey." *Energy Economics* **13(3)**: pp. 203–210.

Dahl, C. A. (1982): "Do Gasoline Demand Elasticities Vary?" *Land Economics* **58(3)**: pp. 373–382.

Dahl, C. A. (2010): ""DEDD-G2010.xls," in Dahl Energy Demand Database." `http://dahl.mines.edu/courses/dahl/dedd`. [Online; accessed 2-January-2012].

Dahl, C. A. (2012): "Measuring global gasoline and diesel price and income elasticities." *Energy Policy* **41(0)**: pp. 2–13.

Dickey, D. A. & W. A. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." *Journal of the American Statistical Association* **74(366)**: pp. 427–431.

Doucouliagos, H. & T. Stanley (2008): "Theory Competition and Selectivity: Are All Economic Facts Greatly Exaggerated?" *Economics series*, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

EIA (2012): "Gasoline and Diesel Fuel Update." `http://www.eia.gov/petroleum/gasdiesel/`. [Online; accessed 3-February-2012].

Engle, R. F. & C. W. J. Granger (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* **55(2)**: pp. 251–276.

Engle, R. F. & B. S. Yoo (1987): "Forecasting and testing in co-integrated systems." *Journal of Econometrics* **35(1)**: pp. 143–159.

Engsted, T. & J. Bentzen (1997): "Dynamic Modelling of Energy Demand: A Guided Tour Through the Jungle of Unit Roots and Cointegration." *OPEC Review* **21(4)**: pp. 261–293.

Espey, M. (1998): "Gasoline demand revisited: an international meta-analysis of elasticities." *Energy Economics* **20(3)**: pp. 273–295.

Glass, G. (1976): "Primary, Secondary, and Meta-Analysis of Research." *Educational Researcher* **5(10)**: pp. 3–8.

Graham, D. J. & S. Glaister (2002): "The Demand for Automobile Fuel: A Survey of Elasticities." *Journal of Transport Economics and Policy* **36(1)**: pp. 1–25.

Granger, C. & P. Newbold (1974): "Spurious regressions in econometrics." *Journal of Econometrics* **2(2)**: pp. 111–120.

Greening, L. A. (1995): "Use of Region, Life-Cycle and Role Variables in the Short-Run Estimation of the Demand for Gasoline and Miles Travelled." *Applied Economics* **27(7)**: pp. 643–656.

Havranek, T. & Z. Irsova (2010): "Which Foreigners Are Worth Wooing? A Meta-Analysis of Vertical Spillovers from FDI." *Working Papers 2010/03*, Czech National Bank, Research Department.

Havranek, T., Z. Irsova, & K. Janda (2012): "Demand for Gasoline Is More Price-Inelastic than Commonly Thought." *Energy Economics* **34(1)**: pp. 201–207.

Houthakker, H., P. K. Verleger Jr., & D. P. Sheehan (1974): "Dynamic Demand Analyses for Gasoline and Residential Electricity." *American Journal of Agricultural Economics* **56(2)**: p. 412.

Kayser, H. A. (2000): "Gasoline demand and car choice: estimating gasoline demand using household information." *Energy Economics* **22(3)**: pp. 331–348.

Kennedy, M. (1974): "An Economic Model of the World Oil Market." *The Bell Journal of Economics and Management Science* **5(2)**: pp. 540–577.

Kesicki, F. (2010): "The third oil price surge – What's different this time?" *Energy Policy* **38(3)**: pp. 1596–1606.

Kravis, I. B., A. W. Heston, & R. Summers (1978): "Real GDP Per Capita for More Than One Hundred Countries." *The Economic Journal* **88(350)**: pp. 215–242.

Nelson, J. & P. Kennedy (2009): "The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment." *Environmental and Resource Economics* **42**: pp. 345–377.

Nicol, C. J. (2003): "Elasticities of demand for gasoline in Canada and the United States." *Energy Economics* **25(2)**: pp. 201–214.

Ostro, B. D. & J. L. Naroff (1980): "Decentralization and the Demand for Gasoline." *Land Economics* **56(2)**: pp. 169–180.

Pesaran, M. & Y. Shin (1998): "An autoregressive distributed-lag modelling approach to cointegration analysis." *Econometric Society Monographs* **31**: pp. 371–413.

Pesaran, M. H., Y. Shin, & R. J. Smith (2001): "Bounds testing approaches to the analysis of level relationships." *Journal of Applied Econometrics* **16(3)**: pp. 289–326.

Sa'ad, S. (2009): "An empirical analysis of petroleum demand for Indonesia: An application of the cointegration approach." *Energy Policy* **37(11)**: pp. 4391–4396.

Slavin, R. E. (1986): "Best-Evidence Synthesis: An Alternative to Meta-Analytic and Traditional Reviews." *Educational Researcher* **15(9)**: pp. 5–11.

Slavin, R. E. (1995): "Best evidence synthesis: An intelligent alternative to meta-analysis." *Journal of Clinical Epidemiology* **48(1)**: pp. 9–18.

Stanley, T. & C. Doucouliagos (2007): "Identifying and correcting publication selection bias in the efficiency-wage literature: Heckman meta-regression." *Economics Series* **11**.

Stanley, T., C. Doucouliagous, & S. B. Jarrell (2006): "Meta-Regression Analysis as the Socio-Economics of Economic Research." *Economics series*, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

STANLEY, T. D. (2008): "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

STANLEY, T. D. & H. DOUCOULIAGOS (2010): "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economic Surveys* **24(1)**: pp. 170–191.

STANLEY, T. D. & S. B. JARRELL (1989): "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* **3(2)**: pp. 161–170.

STANLEY, T. D., S. B. JARRELL, & H. DOUCOULIAGOS (2010): "Could It Be Better to Discard 90% of the Data? A Statistical Paradox." *The American Statistician* **64(1)**: pp. 70–77.

STORCHMANN, K. (2005): "Long-Run Gasoline demand for passenger cars: the role of income distribution." *Energy Economics* **27(1)**: pp. 25–58.

U.S. CENSUS BUREAU (2009): "Commuting in the United States: 2009." `http://www.census.gov/prod/2011pubs/acs-15.pdf`. [Online; accessed 20-April-2012].

WHEATON, W. C. (1982): "The Long-Run Structure of Transportation and Gasoline Demand." *The Bell Journal of Economics* **13(2)**: pp. 439–454.

WILLIAMS, H. R. & R. I. MOUNT (1987): "OECD gasoline demand elasticities: An analysis of consumer behavior with implications for U.S. energy policy." *Journal of Behavioral Economics* **16(1)**: pp. 69–79.

Appendix A

Oil consumption by regions

Historical data were taken from BP (2011). The document contains annual data from 1965 to 2010, covering both production and consumption, prices, volumes, etc. The following calculations and tables are not based on the original data due to an error. World consumption up until 1991 did not include former Soviet Union, so the analysis was not valid. The corrected dataset is included on the DVD.

Summaries of all calculations are in Table A.1 and Table A.2. All the percentage representations, except for the shares, depict average annual growth.

	1960s	1970s	1980s	1990s	2000s	1965-2010	1965 share	2010 share
North America	5.34%	2.82%	0.32%	1.54%	-0.30%	1.33%	41.99%	26.80%
South and Central Am.	5.67%	5.18%	0.96%	3.53%	2.05%	3.01%	5.22%	6.99%
Europe	10.04%	2.50%	-0.74%	0.80%	-0.39%	1.37%	26.74%	17.35%
Former Soviet Union	7.20%	5.73%	-0.02%	-8.46%	1.16%	0.61%	10.77%	4.98%
Middle East	3.86%	6.71%	5.36%	3.39%	4.45%	4.79%	3.10%	8.95%
Africa	5.32%	6.66%	3.61%	2.57%	3.05%	4.15%	1.71%	3.77%
Asia Pacific	15.58%	5.82%	2.41%	4.46%	2.27%	4.86%	10.47%	31.17%
US	5.28%	2.54%	0.17%	1.55%	-0.54%	1.14%	37.43%	21.91%
China	16.76%	14.17%	3.67%	7.51%	6.22%	8.66%	0.70%	10.36%
Japan	17.81%	3.86%	0.15%	0.70%	-2.53%	2.16%	5.54%	5.09%
OECD	8.12%	2.82%	0.05%	1.59%	-0.51%	1.56%	75.20%	53.14%
Non-OECD	7.63%	6.51%	1.99%	1.22%	3.48%	3.80%	24.80%	46.86%
European Union	10.01%	2.37%	-0.72%	0.73%	-0.44%	1.29%	25.31%	15.90%
World	8.00%	3.83%	0.73%	1.45%	1.12%	2.35%	100.00%	100.00%

Table A.1: Historical oil consumption growth, all oil products pooled.

	1960s	1970s	1980s	1990s	2000s	1965-2010	1965 share	2010 share
North America	4.77%	2.29%	0.99%	1.43%	0.80%	1.51%	61.15%	38.58%
South and Central Am.	7.04%	5.16%	2.65%	3.83%	2.16%	3.44%	4.38%	6.47%
Europe	11.23%	4.15%	1.42%	0.20%	-2.45%	1.64%	17.91%	11.98%
Former Soviet Union	7.55%	6.69%	0.15%	-5.51%	3.04%	1.71%	6.57%	4.53%
Middle East	3.50%	6.49%	4.50%	5.17%	6.60%	5.48%	1.77%	6.28%
Africa	8.16%	6.26%	3.74%	2.08%	3.33%	4.36%	1.29%	2.83%
Asia Pacific	14.02%	6.45%	4.39%	6.77%	3.37%	5.90%	6.92%	29.33%
US	4.65%	1.91%	1.07%	1.46%	0.55%	1.34%	55.94%	32.78%
China	16.96%	15.50%	11.08%	8.24%	6.29%	10.03%	0.36%	8.64%
Japan	18.19%	5.15%	1.81%	3.77%	-0.18%	4.01%	3.25%	6.14%
OECD	6.82%	3.03%	1.28%	1.73%	0.08%	1.87%	83.88%	62.09%
Non-OECD	8.00%	6.71%	2.91%	2.88%	4.50%	4.52%	16.12%	37.91%
European Union	11.33%	4.06%	1.41%	0.14%	-2.47%	1.61%	17.00%	11.22%
World	7.01%	3.74%	1.68%	2.05%	1.50%	2.55%	100.00%	100.00%

Table A.2: Historical oil consumption growth, only light distillates included.

Appendix B

Contents of the attached DVD

`thesis.pdf`	This thesis in PDF format.
`data.xlsx`	Various data, mostly on gasoline consumption, used for graphs and computations in Chapter 1. Corrected data from BP (2011) are included as well.
`init.do`	Stata `do` file funnel and kernel density plots, generating variables, summary statistics, and running regressions and exporting regression outputs.
`mra-data.xlsx`, `mra-data.dta`	Excel and Stata files with all observations needed for the meta-analysis.