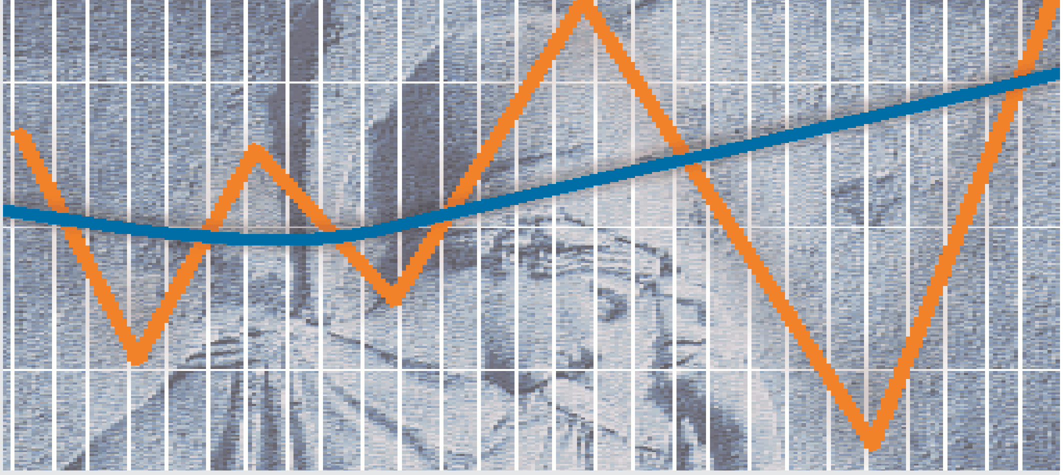


ifo Beiträge
zur Wirtschaftsforschung

68

Effectiveness of Climate Policies: Empirical Methods and Evidence

Julian Dieler

ifo Institut
Leibniz-Institut für Wirtschaftsforschung
an der Universität München e.V.

Herausgeber der Reihe: Clemens Fuest
Schriftleitung: Chang Woon Nam

68 ifo Beiträge zur Wirtschaftsforschung

Effectiveness of Climate Policies: Empirical Methods and Evidence

Julian Dieler

ifo Institut
Leibniz-Institut für Wirtschaftsforschung
an der Universität München e.V.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über
http://dnb.d-nb.de
abrufbar

ISBN-13: 978-3-95942-013-6

Druck: ifo Institut, München

ifo Institut im Internet:
http://www.cesifo-group.de

Preface

This volume was prepared by Julian Dieler while he was working with the Ifo Institute for Economic Research. It was completed in December 2015 and accepted as a doctoral thesis in May 2016.

At the latest since the adoption of the Kyoto Protocol in 1997 climate policies are permanently on the international policy agenda. The urgency to find effective and feasible strategies to curb greenhouse gas emissions increases as the carbon budget to reach the 2°C goal will be exhausted in 2045 according to current estimations by the Intergovernmental Panel on Climate Change (IPCC). Therefore decision-makers should be provided with thorough knowledge of climate policies and their effects by the scientific community. This thesis contributes to the scientific discourse by analyzing the effectiveness and the development of climate policies.

Chapter 1 analyzes the degree of effectiveness of gasoline and diesel taxes in Europe by estimating price and tax elasticities of fuel demand. The price or the tax elasticity is a typical measure to assess the effectiveness of policies which are designed as price mechanisms. Besides the insights into the European motor fuel market the analysis led to the more general finding that anticipation effects have to be taken into account while analyzing the impact of a tax introduction or increase. Chapter 2 makes a further methodological contribution in the area of fuel demand estimation. Especially in case of analyzing micro-data, an often encountered problem in demand estimation is the large number of zero-observations, which poses problems for standard regression methods. The study which is the basis for Chapter 2 provides alternative empirical methods which constitute a remedy to the problem of zero-observations. Chapter 3 introduces a new climate policy indicator which provides information about the stringency of climate policies in the OECD countries and can serve itself as an input in empirical analyses because of its empirical foundation.

Keywords: Anticipation effect, inventory effect, fuel demand, fuel price, fuel tax price elasticities, tax elasticities, log of zero, model comparison tests, two-part models, climate policies, carbon price, renewable subsidies, energy taxation, climate policy index

JEL-Codes: C01, C18, C21, C23, D12, H23, Q31, Q41, R41

Acknowledgements

I am sincerely grateful to all the people who have supported me on the way to my dissertation.

First and foremost, I would like to thank Karen Pittel for giving me the opportunity to write my dissertation under her supervision. Karen was always approachable for questions and concerns and she helped me throughout my dissertation with her invaluable ideas and comments which always base on an impressive analytical rigor. Furthermore I am very grateful for the positive atmosphere in our group at Ifo, which to a large extent goes back to her. I also owe to my second supervisor, Frank Goetzke, with whom I had many fruitful discussions, leading to a substantial extension of my skills in econometrics and to a nice joint paper with our colleague Colin Vance whom I also want to thank for valuable comments and advice. Also Gebhard Flaig deserves my thankfulness for being my third supervisor and for the pleasant collaboration on the empirical research seminar for students we organize together.

Special thanks to Hans-Werner Sinn and Meinhard Knoche, representative for the Ifo Institute, which was over all years an inspiring and motivating environment for conducting my research.

In addition I wish to thank present and former colleagues at the Ifo Institute, who also became friends during the years, Anna Ciesielski, Marc Gronwald, Niko Jaakkola, Jana Lippelt, Waldemar Marz, Ana Maria Montoya Gomez, Johannes Pfeiffer, Markus Reischmann, Marina Riem, Marie-Theres von Schickfus, Christoph Schinke, Alex Schmitt, Suphi Şen, Christoph Weissbart and Markus Zimmer. I was also fortunate enough to make some good friends at the University of Munich. Especially I want to thank Nadjeschda Arnold, Jakob Eberl, Wolfgang Habla, Martin Watzinger, Christopher Weber and Susanne Wildgruber for valuable conversations and experiences.

Many thanks go to Christian Beerman and Darko Jus for the exciting economic and non-economic discussions we had and all the memorable experiences we made together. For many unforgettable moments and conversations in the mountains and off the mountains I wish to thank Wolfgang Brummeisl and Steve Heinke.

Furthermore I would like to express my deep gratitude to Marion Handtke, who always supported and encouraged me and with whom I share unforgettable memories.

Lastly, I would like to thank my family for their love and support. I am highly indebted to my brother, as it was him, who gave me the idea of studying economics in first place. In addition, I am very grateful for all the joint adventures and great conversations. The biggest thank goes to my parents as they always supported all my endeavors and therefore without them, this dissertation would not have been possible. I dedicate this dissertation to my family. Thank you.

Julian Dieler
Munich, December 2015

Effectiveness of Climate Policies: Empirical Methods and Evidence

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2015

vorgelegt von
Julian Dieler

Referentin: Prof. Dr. Karen Pittel
Korreferent: Prof. Frank Goetzke, Ph.D.

Promotionsabschlussberatung: 11. Mai 2016

Contents

List of Figures

List of Tables

Introduction

In 1990 the Intergovernmental Panel on Climate Change (IPCC) stated in its First Assessment Report "*Emissions resulting from human activities are substantially increasing the atmospheric concentrations of the greenhouse gases [...]. These increases will enhance the greenhouse effect, resulting on average in an additional warming of the Earth's surface*".[1] This statement by a body of the United Nations Framework Convention for Climate Change (UNFCCC) cannot be praised highly enough: First, because it unambiguously stated this causality in such an early phase of the climate debate and second, because it implies a large degree on consensus within the IPCC. Nowadays the prevailing consent among the scientific community is that man-made climate change exists. The literature on anthropogenic global warming (AGW) has been extensively analyzed by Cook et al. (2013). Their study confirms that out of almost 12 000 scientific publications on climate in peer-reviewed journals, the ones which express a position on AGW, 97.1% "endorsed the consensus position that humans are causing global warming".[2] While occasionally climate sceptics still struggle to reason against global warming, there is no doubt that a global temperature rise will substantially harm all natural and human systems worldwide. From an economist's perspective the public good character of the global climate is the basic cause of inefficient high emissions of greenhouse gases (GHG), and thus not only justifies but requires political intervention to cope with climate change.

The adoption of concrete policy measures gained momentum in the 1990s with the implementation of the UNFCCC in 1994 and the adoption of the Kyoto Protocol in 1997 which established binding emission reduction targets for 37 industrialized countries. How to achieve the national reduction goals and which instruments to implement was left open and was, to prevent to further curtail national sovereignty, perceived as a matter of the participating countries. The national reduction goals led to increased effort in the adoption of national and international climate policies, which is illustrated in Figure 1.

[1] IPCC (1990), "Working Group I: Scientific Assessment of Climate Change (Policymakers Summary)", *IPCC First Assessment Report 1990.*

[2] Cook et al. (2013), "Quantifying the consensus on anthropogenic global warming in the scientific literature", *Environmental Research Letter* 8.

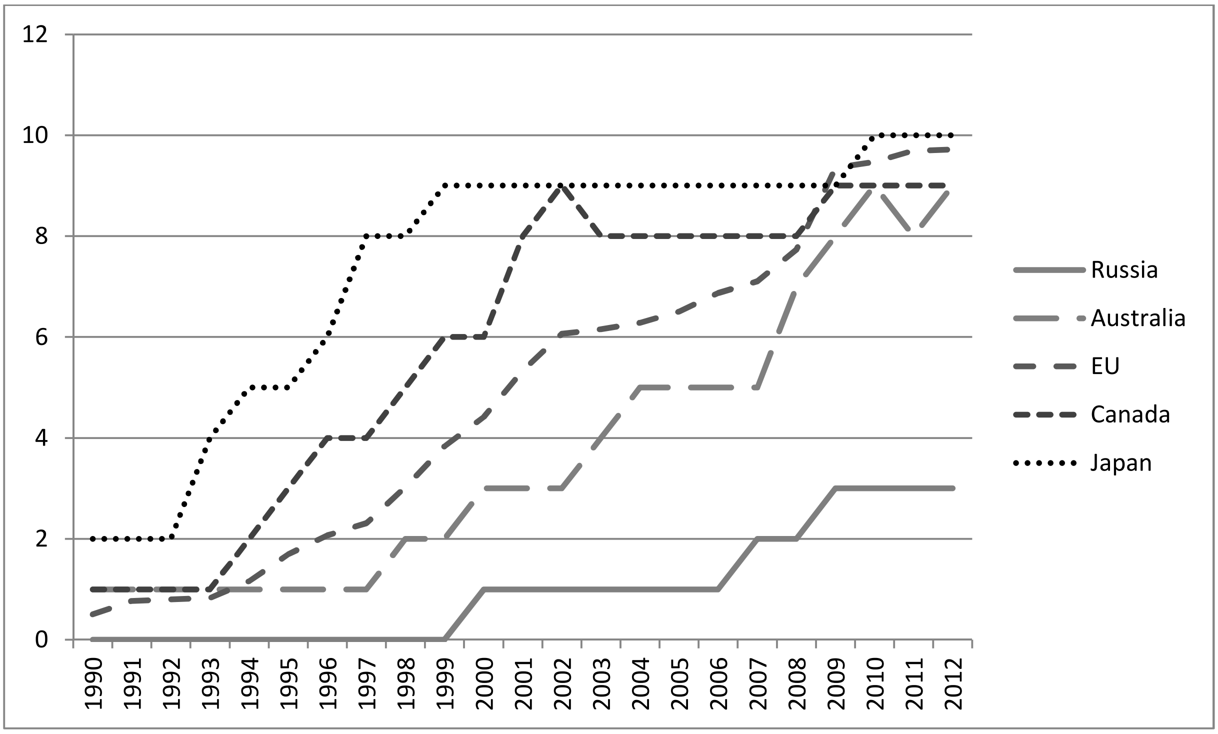


Figure 1: Number of policy categories in which the countries apply policy measures. The time series for the EU is an emission weighted average (own illustration).

Figure 1 depicts the number of different policy categories in which the respective governments have adopted policy measures. This representation serves as a simple policy diversity indicator for the 5 largest CO_2 emitters among the countries that committed themselves to binding emission targets for the first commitment period (2008-2012) of the Kyoto Protocol. Of course, the pure number of adopted climate policies is not very meaningful as long as their actual impact on the reduction of GHG emissions is neglected. The policy categories shown in Figure 1 represent only policy measures which firstly, are adopted in all countries considered in a comparable way and secondly, have the main goal to reduce GHG emission. Among other measures, efficiency standards and subsidies on renewable energy are included. In addition to the policies which primarily aim at climate protection, every national government implements dozens of policies which affect GHG emissions but were introduced for different reasons. The finest example is the consumption tax on petroleum products. Though having an impact on emissions its primary purpose is commonly of purely fiscal nature. Such policies are also referred to as climate related policy measures. Given the multitude of climate policies an assessment of the performance of each policy with respect to emission reduction is necessary.

An assessment of climate policy performance can be conducted in several dimensions like the aggregation and the objective. Considering the aggregation level it can be distinguished between analyzing countries with their aggregated policy set to reduce GHG emissions and analyzing specific policy measures. On the country-level an assessment of the policy performance equals in most of the cases an analysis of a policy mix as in most countries multiple policy measures affecting the climate are in force. In the analysis of policy mixes it is important to account for potential interactions of the policies. An extensively debated interaction in Germany for instance, is the impact of the Renewable Energy Act (EEG) on the EU Emission Trading System (ETS). The EEG is blamed to be – to a certain extent – accountable for the price erosion in the EU ETS. While these policy measures arguably act in opposite directions, other measures exhibit mutually beneficial tendencies. Given this interaction of policies, most empirically founded policy analysis are done on the policy level, analyzing the effects of single policy measures on a specific outcome variable. One of the finest examples is the analysis of fuel taxes with regard to their effect on consumption reduction.[3] Besides the entity of the analysis, a further dimension in which policy analysis differs is the objective used to evaluate climate policies. The evaluation measure could for example, be the pure number of certain policies like in Figure 1, the effectiveness of policy measures or even the policy efficiency. The first method is rather naïve with limited value for the assessment or comparison of policy measures. Effectiveness as a measure accounts for the impact of the analyzed instrument in respect of a specific policy target. A common objective for the evaluation of the effectiveness of climate policies is the quantity of avoided GHG-emissions. The commonly used method in the analysis of the policy effectiveness is the empirical analysis of historical data. Efficiency takes one step forward by considering the actual costs of the achieved emission reductions. An additional assessment of the costs of a policy measure requires the consideration of general equilibrium effects. This is typically done by using computable general equilibrium (CGE) models.

This dissertation focuses on the analysis and comparison of the effectiveness of climate related policies by applying empirical methods. Chapter 1 analyzes the degree of effectiveness of gasoline and diesel taxes in Europe by estimating price and tax elasticities of fuel demand. The price or the tax elasticity is a typical measure to assess the effectiveness of policies which are designed as price mechanisms. Besides the insights into the European motor fuel market the analysis led to the more general finding that anticipation effects have to be taken into account while analyzing the impact of a tax introduction or increase. Chapter 2 makes a further methodological contribution in the area of fuel demand estimation. Especially in case of analyzing micro-data an often encountered problem in demand estimation is the large number of zero-observations which poses problems for standard regression methods. The study which is the basis for Chapter 2 provides alternative empirical methods which constitute a remedy to the problem of zero-observations. Chapter 3 introduces a new climate policy indicator which provides information about the stringency

[3] Brons et al. (2008) provides an overview of the existing literature on the price elasticity of gasoline demand (see Chapter 1 for the reference).

of climate policies in the OECD countries and can serve itself as an input in empirical analyses because of its empirical foundation.

In the following sections I will give a brief overview of the dissertation. Each chapter is based on an individual research paper and thus is self-contained. To emphasize that Chapter 1 and 2 originate out of joint work, I have kept the pronoun 'we'.

Chapter 1: Anticipation and Inventory Effects on Fuel Demand: An Analysis of Prices and Taxes in Europe

In a panel study for 11 European countries covering the period between 1990 and 2012 we analyze the effectiveness of fuel taxes using monthly data. Because of its importance in Europe our data also includes data on diesel consumption, which typically is neglected in this kind of studies, since most of them are conducted for the United States, where Diesel is less relevant. Our main contribution to the existing literature however, is the disclosure and quantification of a tax anticipation effect. This anticipation effect lets consumers fuel up their cars shortly before the implementation of a tax introduction or increase. Thereby in the month of the tax increase the motor fuel purchases decrease for two reasons. First, the consumers have to fuel less because they filled up their tanks at the end of the previous month and possibly second because they face a higher price due to the tax increase. Taking anticipation effects into account therefore leads to decisively smaller tax effects than it has been postulated in the literature so far.

This overestimation of the tax effect can also occur using quarterly data as many changes in the tax regime are coming into force at the beginning of a year. If that is the case, the anticipation effect overlaps two different quarters and leads thereby to overestimated tax elasticities, similar to the case of monthly data.

By accounting for the anticipation effect the result of recent studies by Davis and Kilian (2011)[4] and Li et al. (2012)[5], stating that there is a difference between tax and tax exclusive price effect, cannot be confirmed, at least not in the short-run.

Chapter 2: The Problem of too many Zeros: Methods to Estimate Fuel Price Elasticities

Whereas the analysis in Chapter 1 was conducted using macro-data, data on country level, the focus of the second study in the area of fuel demand estimation is on the analysis of micro-data. Micro-data in this case means data on household level. A special feature of household driving, or to put it more general consumption data, is the large number of zero

---

[4] Davis, Lucas W., and Lutz Kilian (2011). "Estimating the effect of a gasoline tax on carbon emissions". Journal of Applied Econometrics 26(7).

[5] Li, Shanjun, Joshua Linn, and Erich Muehlegger (2012). "Gasoline taxes and consumer behaviour". NBER Working Paper 17891.

observations for consumption. In our dataset, a German household survey on driving behavior, the zeros account for a share of 15% of all observations. That means 15% of the households in the sample did not drive in the observation period. This share is representative for the German population and non-negligible from an econometric point of view, as the workhorse model in this field, the log-linear model estimated by ordinary least squares (OLS), is not able to deal with zero-observations. Thereby estimating the price effect on the demand of car driving with the log-linear model, leads to the omission of a substantial fraction from the sample and therewith to biased estimates. The log-linearized OLS model also suffers from an additional drawback which is of statistical origin. Log-linearizing leads in most cases to an endogeneity of the error term, which in turn leads to a violation of the exogeneity assumption of OLS.1 and thereby again to biased estimates. However our analysis shows that the zero-problem is the prevailing issue.

In our study we analyze both problems theoretically as well as empirically by conducting a case study using the above mentioned German household mobility survey. Besides illustrating the problems of the log-linear model, we also propose alternative econometric approaches to estimate the effect of prices or taxes on the demand for driving. The alternatives comprise count-data models like the Poisson or Negative Binomial model and Two-Part models like the Hurdle or Zero-Inflated models. To be able to establish a statement which model performs best with respect to the dataset we have at hand, we conduct model selection tests. The tests we apply have to be able to deal with models which base upon different distributional assumptions (linear vs. poisson). Therefore we cannot use model comparison tests which rely on the likelihood, like the Akaike Information Criterion (AIC) or the Vuong test, for example. Our preferred model selection test is the Model Confidence Set (MCS) developed by Hansen et al. (2011)[6]. The MCS even allows for a ranking of the compared models. We find that the Two-Part models which model the data generating process (dgp) as a two-step model fit the data best. The Two-Part models divide the dgp of how many kilometers the households drive into two decisions: The first decision households have to make is to buy a car or not and the second one is, once they bought a car how much to drive. As this theoretical structure of the Two-Part models represents the real decision process rather well, it is convincing that these models also empirically perform better than one-step linear models.

Chapter 3: Climate Policy Measure Index

Chapter 3 introduces a climate policy index measuring the stringency of the OECD countries with respect to climate related policies, taking into account not only the level of different policy categories, but also the effectiveness of the respective policy instrument with

---

[6] Hansen, Peter R. et al. (2011), "The Model Confidence Set", *Econometrica* 79 (2).

regard to the reduction of GHG emissions. Thereby countries which are very active in a policy category which at the same time is also very effective in reducing GHG emissions perform best in the index called Climate Policy Measure Index (CPMI). In contrast to most of the existing climate policy indices, the CPMI focuses on quantitative policy measures to enhance the comparability of the different policies. The policy categories which represent the sub-indicators of the CPMI are consumption taxes on fossil fuels, policies to promote the expansion of renewable energies and carbon pricing mechanisms. Another distinguishing feature of the CPMI is the focus on policy variables only, which makes it a purely input-variable-oriented index. Many of the existing indices, on the contrary, mix input variables like concrete policy measures with output variables like the progress in emission reduction or the degree of preserved biodiversity. Currently there are only two other indices, the Climate Laws and Institutions Measure Index (CLIMI)[7] and the Climate Performance Index (CPI)[8], which also solely include input variables in their indices. But in contrast to these two indices the weighting of the sub-indicators of the CPMI, which is based on empirics, varies over the policy categories. The CLIMI and the CPI assume equal weights for all sub-indicators. For the weighting of the single policy categories inside the CPMI, the impact of each policy category on the emissions is estimated. The resulting elasticities are used to calculate the weighting factors of the three policy categories.

The CPMI does not only allow a cross-country comparison between the OECD countries, but also a comparison of the climate policy performance over time for each country, as the CPMI provides data for the period between 1991 and 2012, so far.

The general findings of the CPMI are, that the policy stringency increases in all of the OECD countries over the observation period and that there are seven countries[9] out of the 33 OECD countries[10], which always perform above the average of all OECD countries.

[7] Steves, Franklin and Alexander Teytelboym (2013), “Political Economy of Climate Change Policy”, Smith School Working Paper Series, Working Paper 13-06.

[8] Künkel, Nana, Klaus Jacob and Per-Olof Busch (2006), “Climate policies – (the feasibility of) a statistical analysis of their determinants”, Manuscript.

[9] Denmark, Italy, Greece, Germany, Norway, Portugal and Spain

[10] Iceland is missing due to data availability.

Chapter 1

Anticipation and Inventory Effects on Fuel Demand: An Analysis of Prices and Taxes in Europe

Julian Dieler[*], Darko Jus[+], Markus Zimmer[#]

In recent years several studies improved the assessment of the effects of fuel prices and taxes on fuel demand. We further contribute to this literature by including anticipation and inventory effects in the analysis of motor fuel tax elasticities. Recent studies have focused on US data and gasoline consumption. By constructing a new data set we provide a complementary study for European countries and extend the analysis to diesel fuel, which is essential for the European market. Neglecting the reaction on anticipated tax increases leads to upward biased estimates for the tax elasticities. We also reveal a difference in the ability of gasoline and diesel consumers to react to announced tax changes. By including anticipation effects we are able to narrow down the band for the tax elasticity from 0 - 0.39. Lastly we show that using quarterly data in the presence of anticipation effects may also lead to upward biased results.

[*] Ifo Institute for Economic Research at the University of Munich, Germany, +49/89/9224-1346, dieler@ifo.de

[+] Center for Economic Studies, University of Munich

[#] Ifo Institute for Economic Research at the University of Munich

1.1 Introduction

Fuel taxes have received increased attention in today's policy agendas as they are a passionately debated topic in public. They are seen as an important instrument to raise tax revenues on the one hand but also to curb carbon emissions on the other hand. Consequently in recent years the economic literature has advanced from assessing price elasticities only to include the effects of taxes on fuel consumption.

We analyze the effect of fuel taxes in Europe in the short-run by estimating the response of gasoline consumption triggered by fuel price changes. More specifically, we derive the reactions to price and excise tax changes using a unique data set covering 11 European countries. The data includes diesel consumption and prices in addition to gasoline, extending previous literature which has mainly focused on gasoline[11]. While excluding diesel from the analysis for the United States seems to be valid, it is not the case for the European countries where major shares of private cars use diesel fuel. Consequently our study analyses and compares gasoline and diesel consumption in a panel that covers the period between 1990 and 2012 on a monthly basis.

We make five contributions to the existing literature. Our main contribution is the disclosure and the quantification of an anticipation effect in the demand for motor fuel that results from an increase in fuel taxes[12]. A potential mechanism might originate from consumers fuel up their cars shortly before the tax increase is implemented. Consequently they refuel less in the month when the tax increase comes into force. This interpretation needs some more clarification, since the sales volumes communicated to the statistical offices are rather wholesale figures than actual private fuel demand. Observed reactions in the data thus reflect the gas station manager's response to expected tax changes. Since taxes are determined at the date when the product is sold to the consumer, the tax change itself should not motivate increased wholesales, since the gas-station owner cannot save taxes through an early purchase. An increase in the reported fuel sales rather indicates the reaction of the gas station manager to an expected increase of fuel sales prior to the tax increase. While it is not clear if the gas station manager correctly anticipates the consumer reaction it is certainly valid to assume that she is well informed and aware of fuel tax changes. Most likely especially the gas stations which would have had to restock soon will pre-draw their order. Nevertheless the effect is potentially essential, since gas stations hold notable storage capacities. Thus there are two reasons for less consumption in the tax month. In addition to the increase in the retail price of the motor fuel we would expect a

[11] See e.g. Hughes et al. (2008), Davis and Kilian (2011) and Li et al. (2012).

[12] Ours is the first paper to present this idea and we apply it to the European context. Anticipation effects should be included in any empirical analysis of the effects of announced tax changes. This point is considerably strengthened by the fact a corresponding analysis from the American context by Coglianese, J. et al. (2015) is forthcoming in the Journal of Applied Econometrics.

behavior that leads to fuel tanks being fuller before the tax change than in months without a tax increase (be it in the cars itself or in the reservoirs of the gas-stations). In terms of the effect of taxes on fuel consumption, only the reaction to the increased retail price represents an actual decline in fuel consumption. The fuller tank in the beginning of a tax month represents an *anticipation effect,* or an intertemporal shifting of fuel purchases with no change in overall fuel consumption. Beneficial for the analysis is also the fact, that the vast majority of tax changes come into effect on the first of a month and therefore the reported monthly sales volumes nicely separate the pre- and post-tax effect. Disregarding the anticipation effect leads to an overestimation of the tax impact. Controlling for the anticipation effect we find tax elasticities of –0.30 for gasoline and –0.36 for diesel. This is decisively smaller than the elasticities Davis and Kilian (2011) and Li et al. (2012) find in their studies for the USA without controlling for the anticipation effect. Their estimates for the tax elasticities lie between –0.46 and –0.765. When we don't control for the anticipation effect we find similar results for the European gasoline market (–0.64).

In the literature of modeling crude oil markets, anticipation effects are already acknowledged to have an impact on demand and thereby on the crude oil price. Studies like Kilian and Lee (2014) and Kilian and Murphy (2014) argue for the necessity of including the stock of above-ground oil inventories in modeling the oil price, as the stock of inventories is an indicator for the anticipation of future oil price movements.

A second contribution of our study is the construction of a new European dataset on a monthly basis. Previous literature has either focused on US data (e.g. Hughes et al. 2008) or on yearly data for the European countries (e.g. Pock 2010 and Liu 2004) as monthly data was not readily available for Europe before. We collected data from many different national institutions and compiled them into a single data set of comparable fuel tax measures in 11 European countries.[13]

As a third contribution we analyze – in addition to the commonly studied gasoline consumption – the effects of fuel taxes on diesel consumption. Including diesel use in the analysis is crucial for Europe as the share of diesel cars in the EU fleet amounts to 36.8% and is increasing further.[14] In the United States, diesel-fueled cars are also becoming more and more important, but currently they make up only 3% of the entire market. In Europe the mentioned 36.8% of diesel cars are particular popular with frequent drivers. For frequent drivers it pays off to incur the higher initial costs of acquiring a diesel car as the fuel costs are lower due to higher mileage and lower fuel prices. Additionally, most light and heavy duty trucks, typically used for commercial purposes, use diesel engines. Fuel

---

[13] See table 11 in the appendix for detailed information on the data sources. The dataset is available on request.

[14] Newest figures from 2011 (ACEA 2014).

costs account for a major share of total expenses of these commercial users. Assuming they react differently from people who only use their car to drive to the supermarket, we expect to find a difference between the consumption behavior of diesel and the consumption behavior of gasoline. In addition to being more sensitive to fuel costs, commercial fuel consumers also have better opportunities for intertemporal shifting due to bigger fuel tanks or even large central fuel reservoirs (e.g. in trucking companies or farms).

Fourthly, taking the anticipation effect into account we find no significant difference between the effect of a tax change on fuel consumption and the effect of a net of tax price change on fuel consumption. This is in line with a finding by Anderson et al. (2013) who analyze consumer beliefs about future gasoline prices in Michigan. They find that consumers' price forecasts do not respond differently to net of tax price changes and to tax changes. This means consumers do not view tax increases as a more important reason than pre-tax price changes to invest in more fuel efficient vehicles. Consequently a tax increase should lead to similar gasoline consumption effects as a net of tax price increase does. Using a similar approach to the one used in this paper, but without controlling for the anticipation effect, Davis and Kilian (2011) and Li et al. (2012), however, find indication for differential impacts of tax changes and pre-tax price changes. Analyzing the impact of the introduction of a carbon tax in British Columbia, Rivers and Schaufele (2013) also find evidence for differences between tax exclusive price effects and tax effects. The authors of these studies advance two arguments as possible explanations for their results. First, they argue that tax changes attract more media attention than pre-tax price changes, and that thereby tax changes are more salient and lead to larger responses of consumers. The second possible explanation is that of Anderson et al. (2013): namely, that tax changes are perceived to have a more persistent impact on the retail price than net of tax price changes and, thus, a stronger influence on the vehicle choice. In their survey of consumer beliefs, however, Anderson et al. do not find statistical evidence in support of this hypothesis.

Lastly, we make a further methodological contribution and show that using quarterly data, in a calendar quarterly sense and in the presence of anticipation effects, may also lead to upward biased results. When the anticipated event takes place in the first month of the calendar quarter, which is rather often the case for tax changes, it might impact the behavior in the quarter before through the anticipation effect, and that this effect should thus be controlled for.

The structure of our paper is as follows. We proceed in the next section by presenting a summary of the quantitative results and relating them to the findings of the previous literature. In section we 3 explain our data in detail and present some summary statistics before we address the different empirical methods applied and their respective results in section 4. In the remaining section we draw some conclusions.

1.2 Data

The empirical analysis of the effect of vehicle fuel taxes on gasoline and diesel consumption is based on an original panel data set which has been compiled from a wide range of sources. These sources include national governments and ministries, national statistical offices, research institutes, private companies, and supranational organizations (cf. Table 8 in the Appendix). In total, our data set with monthly observations covers 11 countries (Austria, Belgium, Denmark, France, Germany, Ireland, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom). The observation period is from 1990 until 2012. Unfortunately, the data are not available for all countries over the entire period. The key variables in our econometric analysis are the consumption of the two main vehicle fuel types ("Diesel" and "Eurosuper", a 95 octane gasoline, which we simply refer to as gasoline), the respective consumer prices, and the taxes levied on these two products. The taxes consist of excise taxes, the value-added tax (VAT), and additional fees that act like excise taxes. In the remainder of the article tax stands for the sum of excise taxes, the VAT levied on excise taxes and additional fees that act like excise taxes. Consequently taxes according to our definition comprise only price exogenous elements and no elements which change automatically with pre-tax price like the VAT, for example. **Figure 2** illustrates the evolution of taxes and retail price in the analyzed countries.

Figure 7 and 8 in the Appendix show exemplarily the composition of the gasoline and diesel retail price. **Figure 3** shows the development of consumption of gasoline and diesel. The sales of diesel grew due to the increasing spread of turbo-charged engines in passenger vehicles since 1988. Comparing the evolution of the consumption of diesel to that of gasoline, we observe in all countries – but especially in France, Sweden and the United Kingdom – a trend away from gasoline fueled cars to more fuel efficient diesel fueled cars. In the case of gasoline, the increase stems from the fact that we use 95 octane gasoline, which was not used widely before the end of the 1980s. In the subsequent years, it steadily replaced 91 octane gasoline, previously the main type of gasoline. Therefore, the increase in gasoline consumption in **Figure 3** does not imply an overall increase in consumption, but is the result of the substitution of 95 octane for 91 octane gasoline.

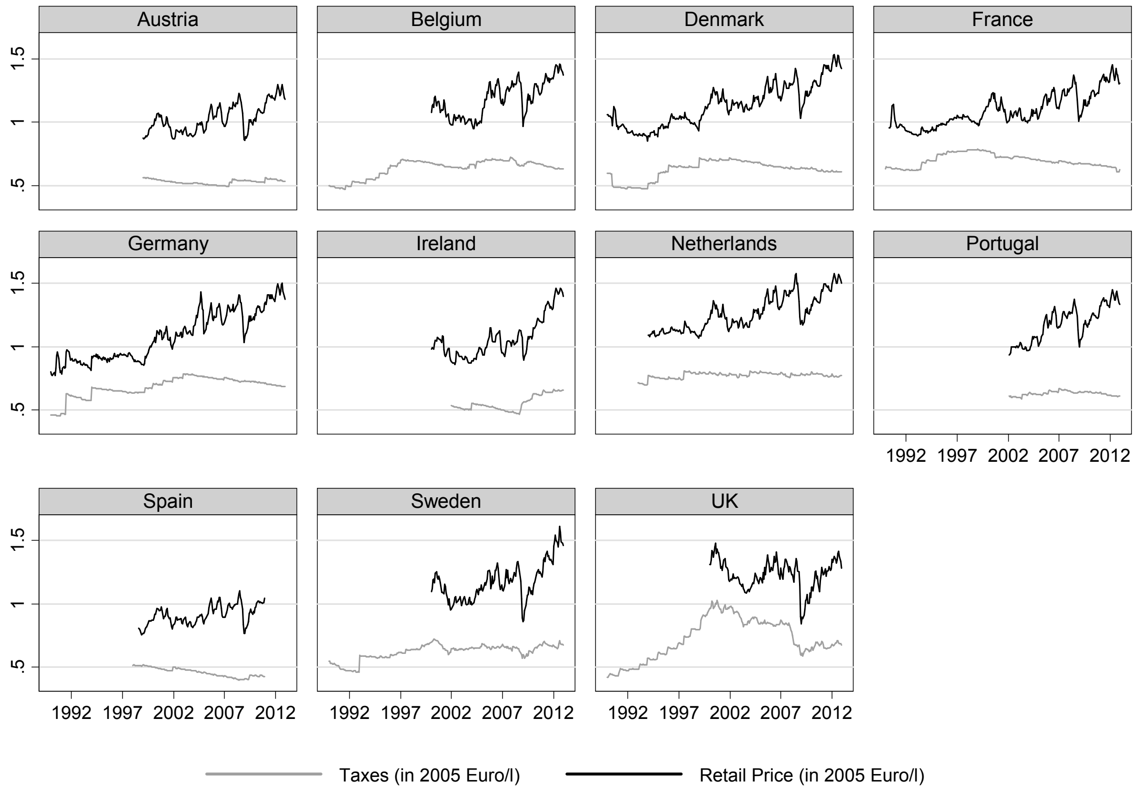


Figure 2: Taxes and prices per liter in 2005 Euros for gasoline

The consumption curves of those countries that were severely affected by the financial crisis (Ireland, Portugal, and Spain) exhibit a discontinuity in 2008. German consumption of gasoline has two discontinuities: one at the end of 2007 when 91 octane gasoline was taken "off the shelf" by most gas stations and a second one at the beginning of 2011 when a new type of gasoline with a higher degree of bioethanol was introduced. Another observation from **Figure 3** is the high seasonality of the consumption data which implies the necessity to control for the month of the year.

Data on the other control variables, including the number of working days per month, the monthly oil price, and the monthly national currency to dollar exchange rate are taken from OECD and World Bank databases, respectively.

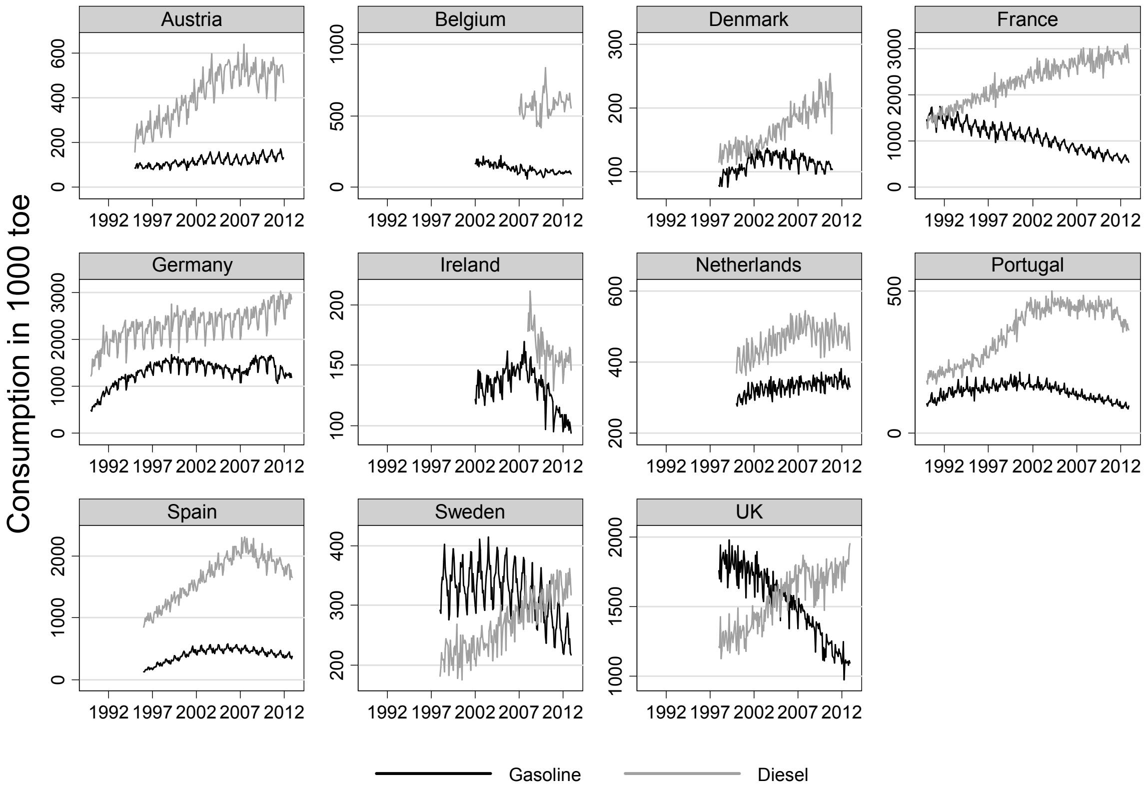


Figure 3: Gasoline and diesel consumption

Gasoline

Variable	Mean	Std.Dev.	Min	Max
Consumption (in 1000 toe)	609.8	546.1	55	1935
Tax exclusive price (in current Euros)	0.409	0.152	0.143	0.821
Tax (in current Euros)	0.663	0.123	0.336	0.974
Working days (monthly differences)	-0.00430	1.908	-6	5
Absolute change of unemployment rate	0.0195	0.374	-2.400	2.200
Oil price (in current Dollar/barrel)	54.58	32.63	9.820	132.7

N=1861

Diesel

Variable	Mean	Std.Dev.	Min	Max
Consumption (in 1000 toe)	1187	922.6	111.8	3086
Tax exclusive price (in current Euros)	0.410	0.186	0.136	0.854
Tax (in current Euros)	0.467	0.155	0.266	1.034
Working days (monthly differences)	-0.00505	1.910	-6	5
Absolute change of unemployment rate	0.0182	0.376	-2.400	2.200
Oil price (in current Dollar/barrel)	54.24	33.43	9.820	132.7

N=1829

Table 1: Summary statistics

Table 1displays summary statistics for the variables used in the regression analysis. The "working days" variable is defined as the absolute difference in the working days from one month to the next. All variables show fairly high variation. Especially with regard to taxes one could be concerned about sufficient variation. Surprisingly tax changes are rather common. In fact we observe about 1.8 nominal tax changes per country and fuel type per year or a total of 262 changes for gasoline and of 230 changes for diesel over the entire panel. For gasoline as well as for diesel the bulk of the changes are increases: 75% for gasoline and 79% for diesel.

1.3 Models of Fuel Consumption

We extend the analytic structure used by Davis and Kilian (2011) by controlling for forward looking consumption behavior, in order to study changes in purchasing behavior *before* the tax is implemented. Estimating this tax anticipation effect leads us to the conclusion that taxes are less effective in the short run than previously stated.

1.3.1 Least Squares Estimates of Fuel Consumption

To establish a common benchmark with similar studies, we start with the analysis of the impact of price changes on monthly fuel consumption (cf. e.g., Hughes et al. 2008; Davis and Kilian 2011).

A frequently employed specification in the literature links y_t, the logarithm of fuel consumption in month t, linearly to p_t, the logarithm of the average price of the respective fuel in the same month, along with a vector Z of control variables and an unobserved idiosyncratic time-varying factor ε_t.

$$y_t = \alpha_0 + \alpha_1 p_t + Z_1 + \varepsilon_t$$

In the basic specification, Z_1 contains month-of-the-year fixed effects that are intended to control for the strong seasonality of the consumption data.

However, since y_t and p_t are likely to follow a long-run trend, we estimate the equation in differences of logs, i.e., in growth rates ($\Delta x_t = \ln \frac{x_t}{x_{t-1}} = \ln x_t - \ln x_{t-1}$):

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta p_t + Z_1 + \varepsilon_t \tag{1}$$

This equation uses the pooled data set, meaning that country-specific effects are ignored. In the second step, we make use of the panel structure for the 11 countries in our data set:

$$\Delta y_{it} = \alpha_0 + \alpha_1 \Delta p_{it} + Z_2 + \omega_{it} \tag{2}$$

In this specification, Δy_{it} is the monthly growth rate of fuel consumption in month t in country i. Accordingly, Δp_{it} is the monthly growth rate of the price of the respective fuel in the same month t^{15} and in the same country i. In addition to the month-of-the-year fixed effects included in Z_1, Z_2 contains country and year fixed effects and other controls, such as the change in the unemployment rate and the month-over-month change in working days. Their relevance will be explained below. Thereby, the unobserved idiosyncratic error term varies across countries and over time on a monthly basis.[16]

	Gasoline			Diesel		
% change in consumption	(1)	(2)	(3)	(1)	(2)	(3)
% change in retail price	-0.008	-0.224***	-0.211***	-0.101	-0.198**	-0.181***
	(0.083)	(0.076)	(0.071)	(0.084)	(0.079)	(0.065)
change in working days			0.014***			0.027***
			(0.001)			(0.001)
change in unemployment			-0.015***			0.000
			(0.005)			(0.007)
Month of the year indicators	No	Yes	Yes	No	Yes	Yes
Country indicators	No	Yes	Yes	No	Yes	Yes
Year indicators	No	Yes	Yes	No	Yes	Yes
R^2	0.00	0.35	0.42	0.00	0.23	0.46
N	1861	1861	1861	1829	1829	1829

* p<0.1; ** p<0.05; *** p<0.01

Table 2: OLS estimates

Table 2 shows the results for the above mentioned estimation equations for both fuel types. Columns (1) give the results of the basic model, i.e., the relationship between the change in fuel consumption and change in its price using a pooled OLS regression (cf. Equation (1)). The explanatory power of this model is rather poor, mainly due to the strong seasonality of consumption (cf. **Figure 3**) which explains much of the variation. We address this issue by using the panel structure of the data in column (2). Columns (2) depict the results of Equation (2) without controlling for changes in working days and changes in the unemployment rate. The model's explanatory power increases and the price coefficient becomes highly significant. In the Columns (3) we add changes in working days and changes in the unemployment rate. The number of working days in a month has a decisive impact on gasoline and diesel consumption. As described in Section 1.2, the number of working days sometimes varies quite substantially (up to a 6-day difference from the previous month). The positive impact of working days on fuel consumption may

[15] Whereas t counts the number of the month from the beginning of the observation period (January 1990) till the end (December 2012).

[16] When we discuss a variable in the remainder of this article we refer to its growth rate, month-to-month growth concerning working days, or absolute change concerning the unemployment rate, unless otherwise stated.

be attributed to commercial activity that requires fuel or simply because many individuals are commuting to work by car. Similar reasoning drives the observed effect of including changes in the unemployment rate. When the unemployment rate increases, fewer individuals commute to work. Unemployment can also be seen as a more general indicator for the economic performance of a country. In an economic downturn, which is mostly accompanied by increasing unemployment there is less economic activity in general which leads to decreasing fuel demand. Additionally a shrinking economy leads to tighter budgets of the individuals in the economy. In these phases it is plausible that the individual substitute away from fuel purchases in their spending decisions. And, indeed, for gasoline consumption, we find empirical evidence in support of this argument. Unemployment can also be seen as a more general indicator for the economic performance of a country. In an economic downturn, which is mostly accompanied by increasing unemployment there is less economic activity in general which leads to decreasing fuel demand. Additionally a shrinking economy leads to tighter budgets of the individuals in the economy. In these phases it is plausible that the individual substitute away from fuel purchases in their spending decisions. And, indeed, for gasoline consumption, we find empirical evidence in support of this argument.

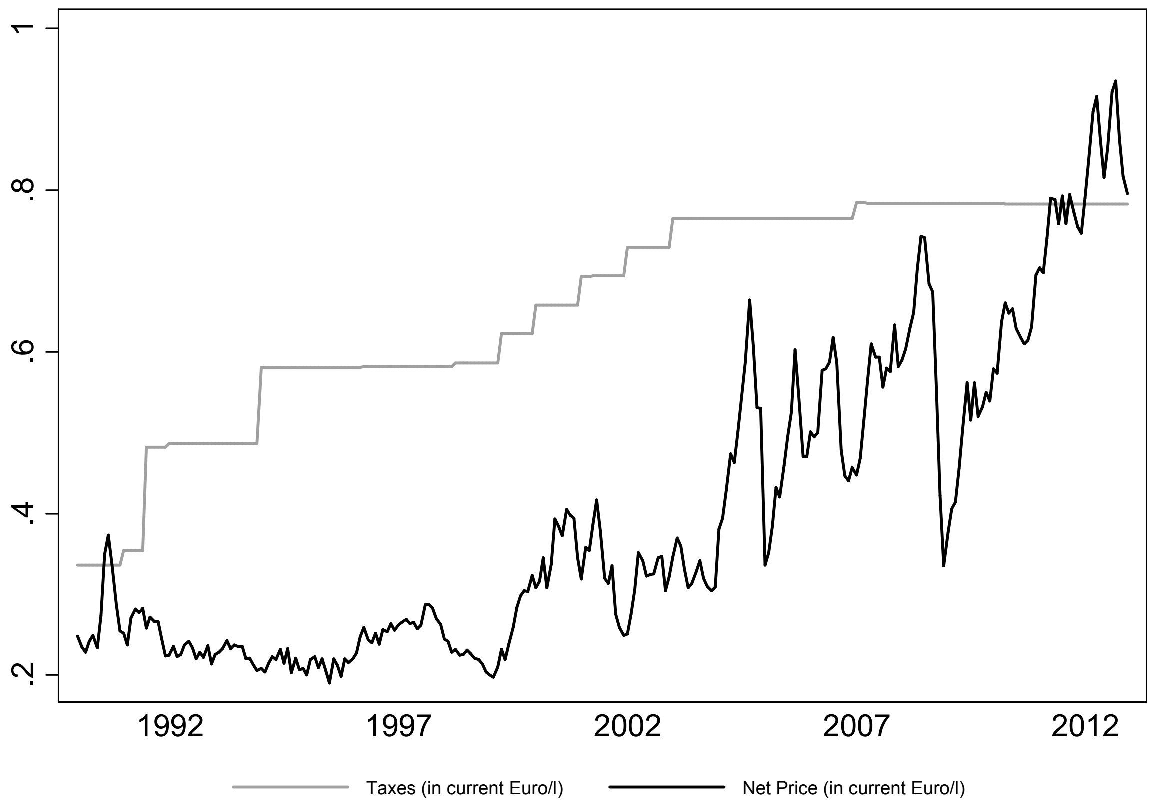


Figure 4: Gasoline taxes and net price in current values (Germany)

Our short-run price elasticity estimates of around –0.2 (for both gasoline and diesel) are very similar to those found in the literature (cf. the review study by Dahl and Sterner 1991).[17] While the analysis of price elasticities may serve as a proxy on how consumption is affected by tax changes (which naturally translate into price changes), Davis and Kilian (2011) and Li et al. (2012) propose a different approach. They point out that decomposing the effects of taxes and tax exclusive prices, instead of looking at the overall effect of a changing gross price, provides further insight into consumer behavior. Such disentanglement allows separating the reaction to difficult-to-predict and highly volatile net price changes and the reaction to more easily foreseeable and persistent tax changes (cf. Figure 4).

Li et al. (2012) find evidence for a stronger reaction to salient retail price changes due to tax changes (elasticity of –0.77) than to net price changes (–0.37). Their explanation for this difference is motivated firstly, by tax changes being announced before they become effective and secondly, by consumers perceiving them to be more persistent than tax exclusive price changes, as these market driven price changes, are always difficult to predict in advance and fluctuate strongly. Therefore, at least in the short run, changes of consumption behavior should be larger with respect to tax changes. Differentiating between individuals' reactions to tax changes and to changes in the net price also raises an additional question: do individuals form expectations about the retail price and, if so, what does this imply for their purchasing decisions? The variable y_{it} is commonly interpreted as *consumption* of gasoline and diesel, while it is actually stating the *purchase* of fuel and not its consumption. As a result the demand for motor fuel possesses features of stock demand. Such features are not as pronounced as for crude oil (Kilian and Murphy 2014, Kilian and Lee 2014), but up to a certain degree, car drivers can use their fuel tank and gas stations and large commercial consumers can use their reservoirs for storage. Therefore the observed reduction of purchases in the month the tax increases can be either the consequence of a reduction in consumption or of a depletion of fuel stocks. Figure 5 and Figure 6 – which will be discussed in depth later – shed light on this issue. They show a significant increase in purchases in the month before and after the tax increase. That indicates the accumulation of fuel stocks, which are subsequently depleted in the month of the tax increase. Thus it seems unlikely that individuals really consume less fuel by e.g. not using their cars to go to work or to take their children to school. Especially towards the end of the month, before an announced tax increase is implemented, individuals could choose to increase their purchase volume in anticipation of the tax increase. We control for this by including the first lead of tax in our estimation equation. To analyze consumption or, rather, purchase behavior after the month of the tax increase, we also include the

[17] Dahl and Sterner (1991) reveal a short-run price elasticity of 0.26.

lag of tax in our regression. These considerations lead to the following estimation equation:

$$\Delta y_{it} = \beta_0 + \beta_1 \Delta\overline{p_{it}} + \beta_2 \Delta tax_{it} + \beta_3 \Delta tax_{it\ \ 1} + \beta_4 \Delta tax_{it+1} + Z_2 + \omega_{i\tau} \tag{3}$$

As in Equations (1) and (2), Δ stands for the first differences of the logs. The tax exclusive fuel price is denoted by $\overline{p_{\iota t}}$ and the tax of the current, previous, and next month are denoted by Δtax_{it}, Δtax_{it-1}, and Δtax_{it+1}, respectively. Z_2 again contains the growth of the unemployment rate and the difference in working days of the current month compared to the previous one.

Table 3 summarizes the results of our disentangled analysis of the tax exclusive price and the tax. Columns (1) show that we observe even larger differences between consumption reactions to net price changes and consumption reactions to tax changes in the European data than Li et al. (2012) found in the U.S. data. The short-run tax elasticity of –0.64 is decisively larger than the price elasticity of –0.07.

	Gasoline		Diesel	
% change in consumption	(1)	(2)	(1)	(2)
% change in net price	-0.068**	-0.063**	-0.074**	-0.071**
	(0.028)	(0.028)	(0.029)	(0.029)
% change in unit tax (lead1)		0.373**		0.323**
		(0.145)		(0.152)
% change in unit tax	-0.635***	-0.837***	-1.029***	-1.044***
	(0.182)	(0.171)	(0.211)	(0.211)
% change in unit tax (lag1)		0.250**		0.371***
		(0.121)		(0.131)
change in working days	0.014***	0.014***	0.026***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)
change in unemployment	-0.014***	-0.014***	0.002	0.002
	(0.005)	(0.005)	(0.007)	(0.007)
Month of the year indicators	Yes	Yes	Yes	Yes
Country indicators	Yes	Yes	Yes	Yes
Year indicators	Yes	Yes	Yes	Yes
R^2	0.42	0.43	0.47	0.47
N	1861	1849	1781	1769

* p<0.1; ** p<0.05; *** p<0.01

Table 3: OLS estimates for decomposed retail price

A tax elasticity of –0.64 translates into a 5% consumption reduction in the month of a 5 cent tax increase, which is a large effect. These large values confirm the results Davis and Kilian (2011) and Li et al. (2012) found for the United States. Next we check for intertemporal shifting of fuel purchases around the month of a tax change. The tax change next month and the tax change in the previous month are statistically significant and non-negligible in magnitude (cf. Columns (2)). The coefficient of the tax lead depicts the anticipation effect (cf. β_4 in Equation (3)) and the lag of tax stands for the catch-up effect after a tax change (cf. β_3 in Equation (3)).

It seems plausible that at least some consumers refuel their cars during the last few days before a tax increase comes into force resulting in an adequate purchasing behavior of the intermediate suppliers. Consequently, these consumers need to buy less fuel in the month of the tax increase, independent of any changes in their consumption behavior. Thus the effect, which we measure in the tax-change month is a combination of less fuel purchased due to having refueled at the end of the previous month and less fuel purchased due to a change in consumption behavior. According to these results, a 1% gasoline tax increase in the next month leads to a 0.38% consumption increase in the current month. In the month of the tax change, consumption decreases by 0.84% before increasing again in the following month by 0.25%.

Table 3 shows similar results for diesel. The tax and net price elasticities tend to be slightly higher for diesel than for gasoline. A possible explanation for this finding is that diesel is to a large extent used commercially and commercial fuel customers, such as freight carriers, might be better positioned to shift fuel purchases across time (e.g. due to having more storage capacity). However, the differences in the elasticities are not statistically significant. Based on the arguments above, we should observe smaller tax effects after aggregating the monthly data to calendar quarterly data (Jan-Mar, Apr-Jun, etc.). That is, some of the intertemporal shifts should vanish over the course of a quarter. We use the same methodology as we did for the monthly data analysis above (see Equation (3)), but omit the lag and the lead of the tax variable and replace month of the year indicators with quarter of the year indicators. As expected, the tax elasticity of consumption decreases from –0.84 (cf. Table 3) to –0.31 for gasoline and from –1.04 to –0.50 for diesel (see column 1 and 2 in Table 4). Although aggregating to calendar quarters diminishes the anticipation effect, the tax effect is still overestimated, as more than 60% of the tax changes occur at the beginning of a quarter. Thereby the anticipation and the tax effect act in two different quarters and using quarterly data still exhibits a large tax effect driven by anticipation. Moving the calendar quarter one month forward so that the first quarter starts in December instead of January leads to a decisive reduction of the tax effect (see column 3 and 4 in Table 4).

Now the estimated tax effect is statistically not differentiable from zero any more. An even more precise way of identifying the tax effect net of the intertemporal shifting is to pretend that the pre-fueling in the month before the tax change, the reduced purchases in the tax month, and the return to the pre-tax behavior in the month after the tax change all happen in one month. We implement this test aggregating over these three months[18] by averaging the consumption, tax, and net price data. We call this period the tax period. [19] By leaving the estimation methodology unchanged compared to Equation (3), except for excluding the lead and the lag of tax, we again obtain a smaller coefficient for the tax elasticity than without aggregating (cf. Table 3 and Table 4). Thus, we find that the tax elasticity is overestimated without controlling for the anticipation effect.

	Quarterly		Quarterly (shifted)		Tax period	
% change in consumption	Gasoline	Diesel	Gasoline	Diesel	Gasoline	Diesel
% change in net price	-0.015	0.009	-0.090*	-0.043	-0.114***	-0.044
	(0.035)	(0.028)	(0.049)	(0.035)	(0.029)	(0.033)
% change in unit tax	-0.308**	-0.503***	-0.195	-0.200	-0.296*	-0.356**
	(0.121)	(0.133)	(0.149)	(0.166)	(0.162)	(0.170)
R^2	0.52	0.45	0.71	0.74	0.37	0.23
N	614	587	623	596	1537	1476

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 4: OLS estimates of aggregated data.

To check whether the anticipation effect is driven by consumer's behavior towards tax increases or by their reaction towards tax decreases we run the regression separately for tax increases and decreases. Table 9 in the appendix shows that the fuel pattern of pre-drawing fuel purchases which we discussed above is dominantly driven by tax increases. This is not very surprising since tax increases dominate the tax changes (in total for gasoline and diesel: 507 increases against 134 decreases). As a further robustness check we compare the OLS results with those of a feasible generalized least squares (FGLS) approach (see Table 5).

Allowing for autocorrelation of the AR(1) type and heteroskedasticity across panels, we obtain similar results. The net price elasticity decreases for gasoline and stays the same for diesel. For both fuel types, the net effect of the tax over the three months decreases as the purchase-reducing effect of the tax in the tax month shrinks and the anticipation and catch-up effects increase. As the results of the two methods do not differ substantially, we are confident that our results are robust.

Estimating demand functions using the price as explanatory variable is routinely met with the concern of biased coefficients due to endogeneity. Endogeneity could be caused by

[18] In reality, this is often a period of four months since in countries like Austria, for example, an increase in the mineral oil tax occurs in two steps over two consecutive months.

[19] The "tax period" includes the month before the tax change, the month of the tax change, and the month after the tax change.

reverse causality of price and demand. If the amount of fuel consumed locally has an influence on its price, the price elasticity would be underestimated by a simple least squares approach. To account for the possible endogeneity of the price we apply a two-stage least squares (2SLS) approach in the next section. Skeptics might argue that taxes are endogenous as well as there might be the tendency of politicians to decrease taxes during periods of economic downturn and to increase taxes during periods of economic boom. While this might be true in the long run it is more than unlikely for monthly data. The subsequent section will reflect on these concerns and evolve the analysis with a focus on the possible endogeneity of the tax exclusive price.

	Gasoline		Diesel	
% change in consumption	OLS	FGLS	OLS	FGLS
% change in net price	-0.063**	-0.036**	-0.071**	-0.071***
	(0.028)	(0.018)	(0.029)	(0.020)
% change in unit tax (lead1)	0.373**	0.428***	0.323**	0.402***
	(0.145)	(0.106)	(0.152)	(0.101)
% change in unit tax	-0.837***	-0.699***	-1.044***	-0.952***
	(0.171)	(0.111)	(0.211)	(0.107)
% change in unit tax (lag1)	0.250**	0.277***	0.371***	0.379***
	(0.121)	(0.091)	(0.131)	(0.099)

* p<0.1; ** p<0.05; *** p<0.01

Table 5: Favored OLS estimates compared to heteroskedastic- and autocorrelation-robust FGLS estimates

1.3.2 IV Estimates of Fuel Consumption

In the spirit of the recent approaches to separate the effects of tax exclusive price changes and tax changes on consumption we instrument the net of tax gasoline price with the crude oil price and with the Dollar-Euro exchange rate. For the crude oil price we use the spot price of Brent, but this choice is arbitrary and of no special importance as the spot prices of all reference oil grades (WTI, Brent, and Dubai crude) are highly correlated with a correlation coefficient larger than 0.95 (World Bank 2013). For the oil price to be a valid instrument for the fuel price in the estimation of the consumption equation, consumption of gasoline and diesel should have no impact on the crude oil price, which is an assumption our data appear to support. On average, each of the observed countries account for 0.9% of the world crude oil consumption. Out of our sample countries, Ireland consumes the smallest fraction (0.07% of world consumption) and Germany the largest (2.3%) (IEA 2013).[20] Moreover, the amount of crude oil used to produce car fuels is even less important. Considering only the use of crude oil for light- and heavy-duty vehicle fuels,

[20] The figures are for 2012 but appear representative of the observation period. Germany consumed the biggest fraction of world consumption in the observation period in 1994 with 3.4%.

Germany's fraction of 2.3% of world consumption shrinks to 1.7%.[21] Thus, from the perspective of a single country, the oil price should be assumed to be exogenous.

As second instrument serves the dollar exchange rate as it is an important determinant of gasoline and diesel prices in European countries since oil, as the main input for producing gasoline and diesel, is traded in U.S. dollars, but the final products are sold in national currencies. To account for potential timing issues related to the price setting of mineral oil firms we include the first lags of the oil price and the exchange rate. Hence, the first-stage equation is:

$$\Delta\overline{p_{it}} = \delta_0 + \delta_1 \Delta oilpr_t + \delta_2 \Delta oilpr_{t\ 1} + \delta_3 \Delta oilpr_{t\ 2} + \delta_4 \Delta exc\ \_r_{it} + \delta_5 \Delta exc\ \_r_{it\ 1} + Z_2 + u_{i\tau} \tag{4}$$

Table 6 shows the estimates of the first stage for both fuel types. For our instruments to be valid two conditions must hold: Firstly, the oil price as well as the dollar exchange rate need to contribute to the explanation of the tax exclusive price of gasoline and diesel. This condition holds as the F statistic of the first stage amounts to 259 and 131 for gasoline and diesel respectively. Secondly, the instruments need to be exogenous. Since the number of instruments exceeds the number of endogenous variables in the presented model it is advisable to apply an Overidentifying Restrictions Test for exogeneity. The result of the test indicates the instruments to be exogenous in case of our IV approach.

% change in net price	Gasoline	Diesel
% change in oil price	0.372	0.240
	(0.014)***	(0.014)***
% change in oil price (lag1)	0.250	0.181
	(0.014)***	(0.014)***
% change in oil price (lag2)	0.033	0.095
	(0.014)**	(0.014)***
% change in exchange rate	0.291	0.159
	(0.052)***	(0.052)***
% change in exchange rate(lag1)	0.246	0.253
	(0.052)***	(0.053)***
R^2	0.38	0.24
N	2084	2136

* p<0.1; ** p<0.05; *** p<0.01

Table 6: First stage estimates

The second stage estimation is identical to Equation (2), except for p_{it} being instrumented as explained above.

[21] These figures are from December 2012.

$$\Delta y_{it} = \mu_0 + \mu_1 \Delta \widehat{p_{it}} + \mu_2 \Delta tax_{it} + \mu_3 \Delta tax_{it\;\;1} + \mu_4 \Delta tax_{it+1} + Z_2 + \omega_{i\tau} \tag{5}$$

The $\widehat{p_{it}}$ stands for the instrumented pre-tax price based on Equation (4). Table 7 shows the results of the IV regression for both fuel types.

Our findings for this sample of European countries differ only slightly from previous studies, e.g. those of Davis and Kilian (2011) and Li et al. (2012) for the United States. Our tax elasticity for gasoline consumption, at –0.82, lies in the range of the results of Davis and Kilian (–0.46) and those of Li et al. (–0.77). What distinguishes our results from previous works is that we also quantify the anticipation and the catch-up effect. Li et al. (2012) and Davis and Kilian (2011) only hint at intertemporal effects.[22] By not considering the effects in the previous and the following month of a tax change, the consumption reaction to a tax change is overestimated. Therefore quantification of the anticipation effect is necessary.

% change in consumption	Gasoline	Diesel
% change in net price	-0.040	-0.012
	(0.041)	(0.048)
% change in unit tax	-0.821***	-1.004***
	(0.173)	(0.223)
% change in unit tax (lag1)	0.371**	0.370***
	(0.144)	(0.126)
% change in unit tax (lead1)	0.380***	0.332**
	(0.145)	(0.147)
change in working days	0.014***	0.026***
	(0.001)	(0.001)
change in unemployment	-0.014***	0.002
	(0.005)	(0.006)
month of the year indicators	Yes	Yes
country indicators	Yes	Yes
year indicators	Yes	Yes
R^2	0.43	0.47
N	1847	1767

* p<0.1; ** p<0.05; *** p<0.01

Table 7: IV estimates

The results from the IV approach do not substantially differ from those of the OLS estimation (see columns 2 in Table 3). If the net price was endogenous due to a reverse causality with the consumption we would expect our price coefficient be biased towards zero in the OLS estimation. But as the price coefficient becomes even smaller in the IV estimation it

22 Li et al. (2012, p. 11).

is questionable whether reverse causality is an issue in our consumption and price data. As OLS is the most efficient estimator given no endogeneity problems we regard the OLS approach as preferred to the IV approach.

1.3.3 Anticipation Effect and Temporary Persistency of a Motor Fuel Tax

Figure 5 and Figure 6 illustrate the change in consumption behavior prior to and after a tax change for gasoline and diesel, respectively. This graph was generated by including additional lags and leads in the OLS model (Equation (3)) and then plotting the point estimates (solid line) and the 95% confidence interval around it (dashed lines). The time of the tax change is indicated by month 0. In the observed period around a tax change, we find significant tax effects only in the month before a tax change, in the month after a tax change, and in the month of a tax change. The grey shaded area marks the period for which we do not find statistical significant coefficients for the percentage change in tax. Making more long-run-oriented statements about tax elasticities is difficult, especially with monthly data.

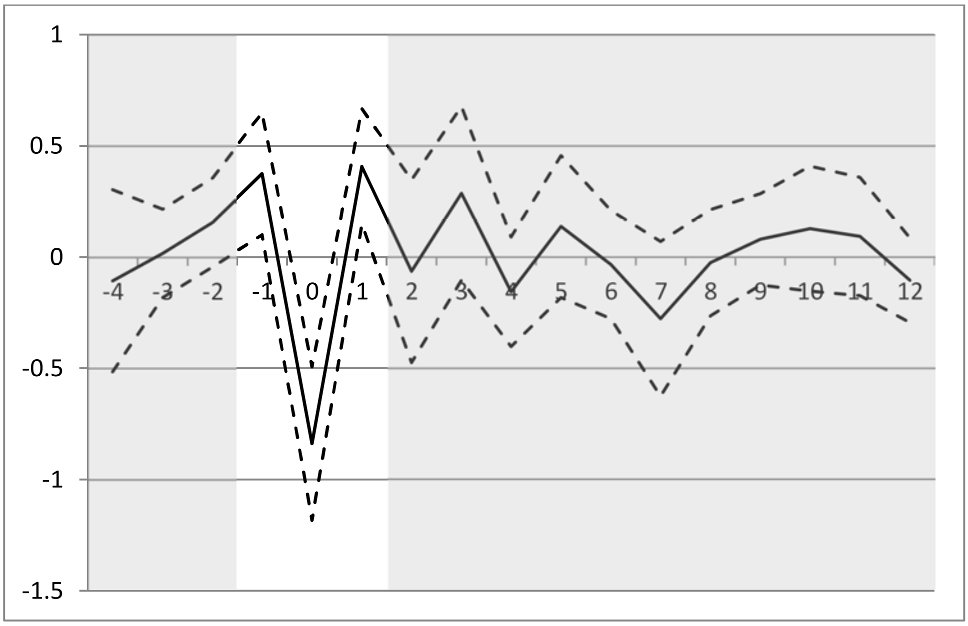


Figure 5: Development of gasoline purchases four months before and twelve months after a tax change

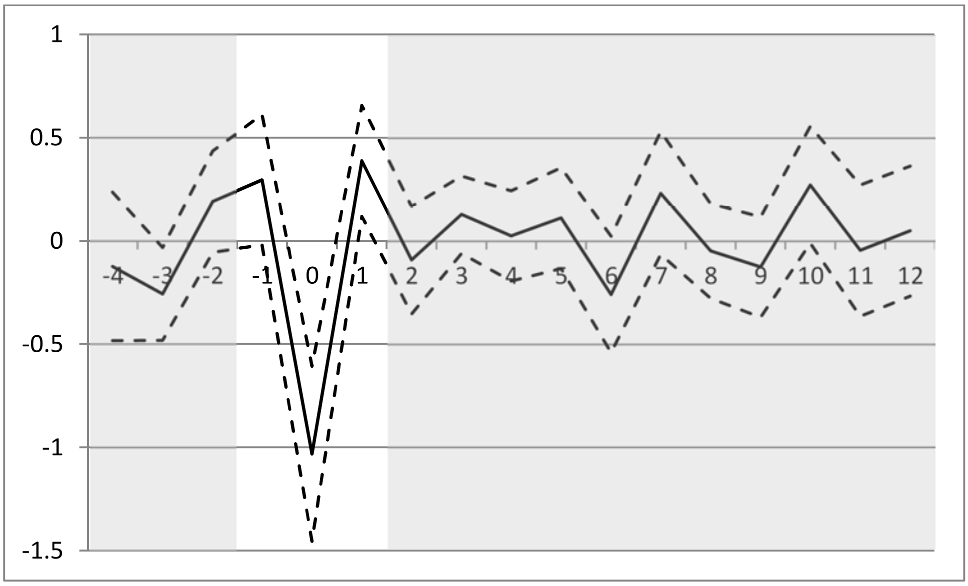


Figure 6: Development of diesel purchases four months before and twelve months after a tax change

1.3.4 Policy Implications

Comparing our results e.g. with those of the studies for the USA by Davis and Kilian (2011) and Li et al. (2012) we find similar elasticities. The Least Squares point estimate of the gasoline demand's tax elasticity lies with 0.84 in the ballpark of the results of Davis and Kilian (0.46) and Li et al. (0.77). But considering a concrete absolute tax change of 10 US-cents per gallon, however, discloses the fundamental differences between the EU and the USA. A 10 US-cents increase of the gasoline tax per gallon leads according to Davis and Kilian's results for the USA to a consumption reduction of 1.4%. Imposing a 10 US-cents tax increase per gallon which equates a 1.9 Euro cents tax increase per liter in the EU would lead to 2.5% decrease of gasoline consumption and to a 4.5% decrease of diesel consumption. Although the tax elasticities look rather similar at first glance, the impact of the same absolute tax change differs a lot: on the one hand between the USA and the EU, and on the other hand between gasoline and diesel consumption in the EU. These differences are due to discrepancies in the tax levels. Whereas the average gasoline tax in the USA amounts to 4.5 Euro-cents per liter the same tax in Europe averaged over the countries in our sample is more than ten times larger. The gasoline tax in Europe totals 52.5 Euro-cents and the diesel tax 37.3 Euro-cents.[23]

1.4 Conclusions

From our analysis of 11 European countries we draw three main conclusions. Firstly, in analyzing short run tax elasticities of fuel demand, it is important to control for anticipation and inventory effects since neglecting these two effects leads to overestimation of the

[23] These figures refer to the excise taxes only without VAT.

tax effect. Secondly, the intertemporal shifting of fuel consumption only occurs in connection with tax changes and not with net of tax price changes. Thirdly, there is an indication for a difference between gasoline and diesel consumption regarding the ability to shift fuel purchases over time.

The main contribution of our analysis to the existing literature in the area of fuel demand estimation is the identification and quantification of an anticipation effect. We show that consumers and intermediate suppliers who expect a fuel price increase induced by a tax increase in the following month bring forward part of their fuel purchases of the next month to the extent of their storage capacities. This behavior is likely motivated by consumers fueling up shortly before a tax increase becomes effective, even if their tank is not completely empty yet. They simply try to exploit the lower retail price once more. Accordingly we identify an increase in the consumption or more precisely in the amount of fuel purchased by 0.37% (0.32%)[24] in the month before the implementation of a 1% tax increase. In the month of the tax increase these consumers naturally have to fuel less as they have brought forward part of their usual monthly fuel purchase, though this reaction has only little to do with a change in the consumption behavior. As a consequence of the tax mainly causing an intertemporal shift in the purchasing pattern, the amount of fuel purchased in the month after the tax increase rises again. Or stated differently: Drivers return to their pre-tax-increase purchase behavior. The after-tax-increase effect of a 1% tax increase amounts to a 0.25% (0.37%) consumption increase. Aggregating over the entire "tax period" we do not find a statistically significant effect of the tax, leading us to conclude that considering only the consumption effect of the tax in the month of the tax increase results in an overestimation of the tax impact. Regarding net of tax price changes we do not find an anticipation effect which is plausible since pre-tax price changes are difficult to predict.

These results provide the following insight: In the month in which a tax increase becomes effective we observe a large reduction in demand which is mainly triggered by consumers shifting fuel purchases to the prior month. As we do not observe such shifting prior to net of tax price changes, we find different demand reactions to retail price changes dependent on whether they are caused by tax changes or by pre-tax price changes. This is in line with the findings of Davis and Kilian (2011), Li et al. (2012) and Rivers and Schaufele (2013). According to these studies, larger tax elasticities are rationale since tax changes are announced and predictable and they are more persistent, especially compared to net of tax price changes. The price elasticities in our preferred specification are –0.06 for gasoline and to –0.07 for diesel, whereas the estimated tax elasticities are –0.84 for gasoline and –1.04 for diesel. Our interpretation of these results, however is somewhat different. We

[24] Figures in brackets refer to diesel consumption.

find that the announcement plays the biggest role in the difference between price and tax elasticities, and that controlling for the anticipation effect results in the estimates of tax changes and tax exclusive price changes being statistically undistinguishable from each other.

As a third major result we find diesel purchases to react more strongly to a tax change than gasoline purchases. We explain this finding by the large fraction of commercial users among the diesel consumers. Firms like freight carriers might have more possibilities to shift fuel purchases over time, for example, due to onsite fuel reservoirs, whereas private users usually only use their car's gas tank as storage capacity.

As an additional result of this study we show price and tax elasticities of fuel demand to be rather constant as the results for our European dataset are in the same ballpark as those from Davis and Kilian (2011) for an American dataset. This should not be expected upfront as the price and tax levels differ tremendously between the European countries and the United States. Still nonlinearities in the reaction to price and tax changes are not dominant as the elasticities in comparable estimations do not vary significantly.

This study confirms the importance of accounting for intertemporal purchase shifting in estimating tax elasticities of fuel demand. Neglecting the forward-looking purchase decisions of fuel consumers results in an overestimation of the tax impact on consumption behavior.

1.5 References

Anderson, Soren T., Ryan Kellogg and James M. Sallee (2013), "What do consumers believe about future gasoline prices?", *Journal of Environmental Economics and Management* 66: 383–403.

European Automobile Manufacturers Association (ACEA) (2014), *The Automobile Industry Pocket Guide*, p. 68.

Brons, Martijn, Peter Nijkamp, Eric Pels, and Piet Rietveld (2008), "A meta-analysis of the price elasticity of gasoline demand. A SUR approach", *Energy Economics* 30: 2105–2122.

Coglianese, John et al. (2015), "Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand", *Journal of Applied Econometrics* (forthcoming).

Dahl, Carol, and Thomas Sterner (1991), "Analysing gasoline demand elasticities: A survey", *Energy Economics* 13: 203–210.

Davis, Lucas W., and Lutz Kilian (2011), "Estimating the effect of a gasoline tax on carbon emissions", *Journal of Applied Econometrics* 26(7): 1187–1214.

Hughes, Jonathan E., Chritopher R. Knittel, and Daniel Sperling (2008), "Evidence of a shift in the short-run price elasticity of gasoline demand", *Energy Journal* 29(1): 113–134.

IEA (2013), "OECD crude supply", *IEA Oil Information Statistics* (database), doi: 10.1787/data-00475-en (accessed November 21, 2013).

Kilian, Lutz and Thomas K. Lee (2014), "Quantifying the speculative component in the real price of oil: The role of global oil inventories", *Journal of International Money and Finance* 42: 71-87.

Kilian, Lutz and Daniel P. Murphy (2014), "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil", *Journal of Applied Econometrics* 29 (3): 454-478.

Li, Shanjun, Joshua Linn, and Erich Muehlegger (2012), "Gasoline taxes and consumer behaviour", *NBER Working Paper* 17891.

Liu, Gang (2004), "Estimating Energy Demand Elasticities for OECD Countries: A Dynamic Panel Data Approach", *Statistics Norway Discussion Papers* 373.

Pock, Markus (2010), "Gasoline demand in Europe: New insights", *Energy Economics* 32: 54–62.

Rivers, Nicholas and Brandon Schaufele (2013), "Salience of Carbon Taxes in the Gasoline Market", Available at SSRN: http://ssrn.com/abstract=2131468

World Bank (2013), "Crude oil, Brent", "Crude oil, Dubai", and "Crude oil, WTI", *Global Economic Monitor (GEM) Commodities* (database) (accessed November 21, 2013).

1.6 Appendix

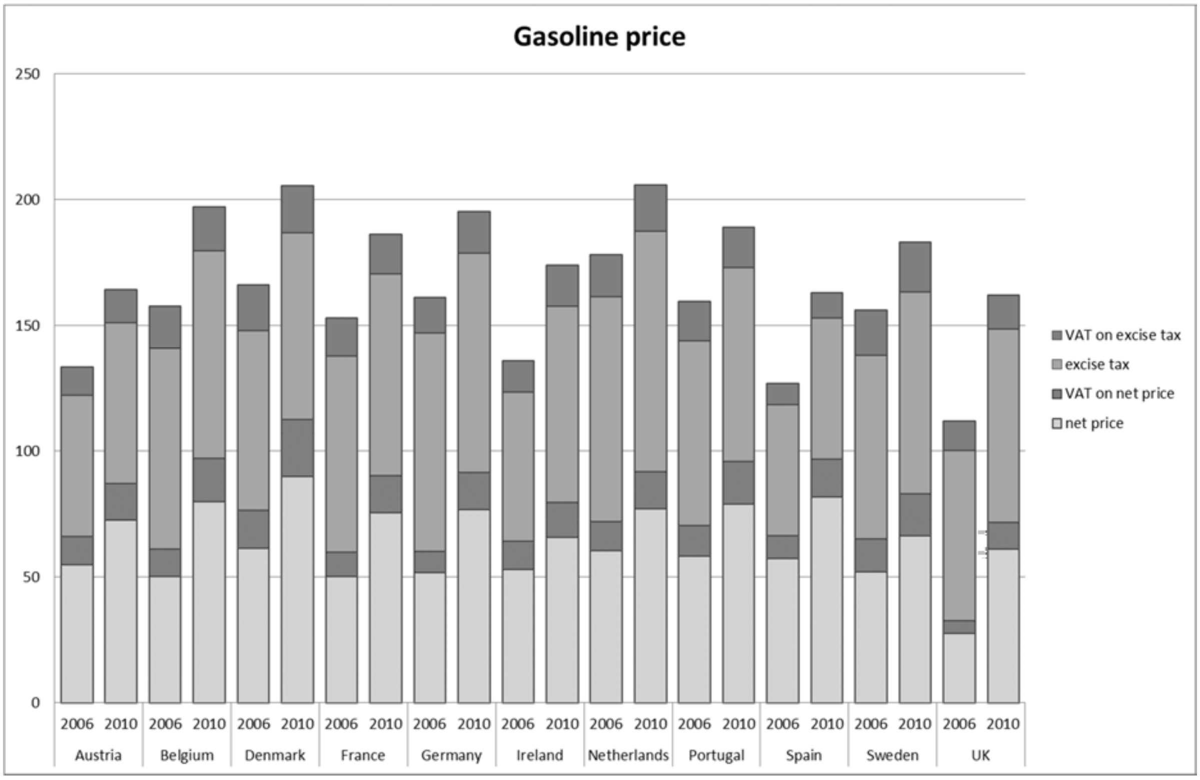


Figure 7: Gasoline retail price composition in December 2006 and 2010 (nominal values in € cents)

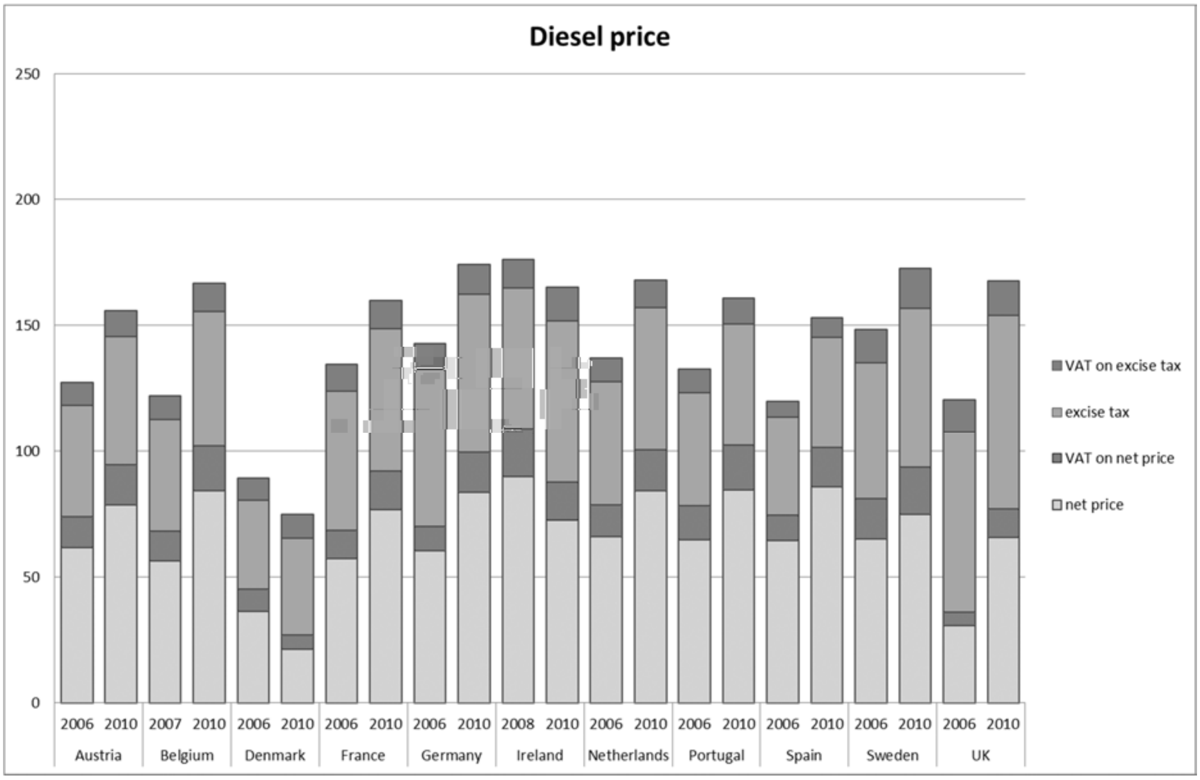


Figure 8: Diesel retail price composition in December 2006[25] and 2010 (nominal values in € cents)

[25] As the time series in Belgium and Ireland starts later, the first observation is used, which is February 2007 and February 2008, respectively.

	Time period	Consumption	Taxes	Prices
Austria	1999–2012	Federal Ministry of Economics		
Belgium	2002–2012	Federal Public Service Economy	Belgian Federation of Petrol	Europe's Energy Portal
Denmark	1990-2012	Danish Oil Industry Association		
France	1990–2012	Federal Ministry of Ecology, Durable Development and Energy	French Association of the Petrol Industry	Federal Ministry of Ecology, Durable Development and Energy
Germany	1990–2012	Federal Office of Economics and Export Control	Association of the German Petroleum Industry	ARAL, ADAC
Ireland	2002–2012	National Oil Reserves Agency	Irish Tax and Customs	Europe's Energy Portal
Netherlands	2000–2012	Statistics Netherlands		Thomson-Reuters
Portugal	1998–2012	Thomson-Reuters	Ministry of Economy	Thomson-Reuters
Spain	1990-2010	National Oil Reserves Agency	Spanish Tax Agency	Spanish Ministry of Industry, Agriculture and Tourism
Sweden	2000–2012	Swedish Petroleum and Biofuel Institute	Federal Tax Agency	Europe's Energy Portal
UK	2000–2012	Government Department of Energy & Climate Change	Government Department of Revenue & Customs	Europe's Energy Portal

Table 8: Data sources

	Gasoline					Diesel				
% change in consumption	all	increase	increase > 1 cent	decrease	decrease > 1 cent	all	increase	increase > 1 cent	decrease	decrease > 1 cent
% change in net price	-0.063**	-0.072***	-0.072***	-0.068***	-0.068**	-0.071**	-0.079***	-0.085***	-0.073***	-0.082***
	(0.028)	(0.028)	(0.027)	(0.026)	(0.026)	(0.029)	(0.030)	(0.030)	(0.028)	(0.030)
% change in unit tax	-0.837***	-0.869***	-0.840***	0.234	0.106	-1.044***	-1.297***	-1.286***	0.060	0.109
	(0.171)	(0.160)	(0.158)	(0.380)	(0.399)	(0.211)	(0.226)	(0.287)	(0.246)	(0.329)
% change in unit tax (lag1)	0.250**	0.281**	0.329**	-0.136	-0.126	0.371***	0.480***	0.493***	0.002	-0.027
	(0.121)	(0.130)	(0.138)	(0.311)	(0.289)	(0.131)	(0.153)	(0.180)	(0.173)	(0.202)
% change in unit tax (lead1)	0.373**	0.328**	0.247**	1.504***	1.335***	0.323**	0.396**	0.375*	-0.045	-0.244
	(0.145)	(0.143)	(0.124)	(0.357)	(0.327)	(0.152)	(0.198)	(0.202)	(0.212)	(0.267)
change in working days	0.014***	0.014***	0.015***	0.015***	0.015***	0.026***	0.027***	0.027***	0.027***	0.026***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
change in unemployment	-0.014***	-0.012***	-0.012***	-0.007	-0.009**	0.002	0.005	0.008	0.012*	0.016**
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Month of the year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.43	0.46	0.49	0.51	0.51	0.47	0.49	0.49	0.47	0.48
N	1849	1717	1481	1401	1323	1769	1638	1395	1348	1270

* p<0.1; ** p<0.05; *** p<0.01

Table 9: All tax changes in comparison to tax increases and decreases only

Chapter 2

The Problem of too many Zeros: Methods to Estimate Fuel Price Elasticities

Julian Dieler[*], Frank Goetzke[+] and Colin Vance[#]

The workhorse model to estimate price and income elasticities in the fuel demand literature is the log-linear model estimated using least squares. This method however, has to deal with two flaws. The first one is general in nature and applies to all log-linearized regression equations. It is the fact that the log-linearization leads to violation of the OLS.1 assumption and thereby to biased estimates. The second problem of the log-linear model is not specific to the estimation of fuel demand but it applies to many types of demand estimation on micro-level. It is the problem of a substantial share of zero observations for the dependent variable, which also causes biased demand elasticity estimates.

This study presents a set of models, which constitute remedies to the drawbacks of the log-linear model estimated by OLS and model selection tests to compare the goodness of fit of non-nested models, which base upon different distributional assumptions. Applying the tests to a German household travel survey, we show that the Hurdle Poisson model performs best.

[*] Center for Energy, Climate and Exhaustible Resources, Ifo Institute for Economic Research at the University of Munich, Germany, +49/89/9224-1346, dieler@ifo.de

[+] Department of Urban and Public Affairs, University of Louisville, Louisville, KY, USA

[#] Department of Environment and Resources, RWI Essen, Germany and Jacobs University Bremen, Germany

2.1 Introduction

Accurate estimates of income and gasoline price elasticity of demand for vehicle miles traveled (VMT) are important for designing efficient and successful transportation policies. Many, if not most, empirical gasoline price elasticity values are derived by constant-elasticity regression models with the convenient log-log (sometimes also called log-linear) functional form (Basso and Oum 2007), or some other kind of related equation type, such as the translog functional form (Wadud et al. 2010). This study provides an analysis of the methods used most frequently to estimate gasoline price elasticities. Additionally we present model selection procedures to compare the different econometric approaches and illustrate the results using the German Mobility Panel (MOP 2013).

Review and meta studies like Espey (1998) and Hanly et al. (2002) show that the estimation method chosen makes a difference – although not a huge one but with regard to the use of price elasticities for forecasts and policy advice also minor differences matter. For the decision between fuel taxes or carbon prices and fuel efficiency standards it is important to be able to rely on as accurate elasticity estimates as possible. For example, Frondel et al. (2008) claim that the rebound effect is analytically equivalent to the gasoline price elasticity. Therefore they argue that price mechanisms like taxes are superior to efficiency standards. Gillingham et al. (2013) suggest that gas mileage regulation policies are more effective than they are said to be. They argue that the rebound effect, which counteracts efficiency standards, is overestimated relative to the price elasticity. Nevertheless the discussion about the appropriateness of the applied econometric methodology is mostly not accorded the weight it deserves. With this paper we would like to contribute to the methodological debate in the literature of estimating fuel price elasticities. Therefore we analyze and compare several estimation methods which have already been used for the estimation of price elasticites in vehicle travel or fuel demand in the past. Next to those established methods we also include methods such as Poisson, Negative-Binomial regression and Two-Part models, which have been also applied for elasticity estimations in other fields of economics like in the migration or trade literature (e.g. gravity models).

One of the most widely applied methods is the estimation of log-linearized constant elasticity demand functions. However since Haworth and Vincent (1979), it has become clear that log-linearizing estimation regression models using a log transformation may lead to biased estimates, and thus faulty elasticity values. Flowerdew and Aitkin (1982) took up on Haworth and Vincent's findings and showed that they also hold true for gravity models. As an example Flowerdew and Aitkin used a gravity model explaining migration in Great Britain. More recently Silva and Tenreyro (2006) argued that the concerns of biased and also inconsistent estimates due to log-linearization also play a role in log-normal models to estimate bilateral trade. The possible inconsistency of elasticity estimates of log-linear estimations has two reasons: Firstly, the log-linearization of originally multiplicative relations leads most probably to heteroskedasticity of the error term and secondly, the existence of a considerable number of zero observations for the dependent variable. The second problem arises only analyzing micro data as the occurrence of zeros in the dependent variable is rather unlikely in case of aggregated car driving data.

To select the best performing model or set of best performing models we use three different statistical tests. All three tests use the predictive ability of the models as selection criteria to evaluate their performance. Since we are comparing models which do not all rely on the same distributional assumptions we cannot resort to criteria based on likelihood, e.g. the Akaike Information Criterion (AIC) or the Vuong test (Vuong 1989). The first model selection procedure we apply is the Model Confidence Set (MCS) developed by Hansen et al. (2011). The MCS is a model selection procedure which chooses a superior set of models out of all models which are to be compared by repeated testing of equal predictive ability (EPA) of the analyzed models. Crucial for the MCS is the possibility to choose freely a loss function serving as model selection criterion. To check the robustness of these results we also apply the Leave-One-Out Cross-Validation and the comparison of the count probability of the observations against the count probability of the predictions.

For the German Mobility Panel we find that Two-Part models fit best as it accounts well for the two decisions which are behind the data generating process (dgp) of the German driving data. More precisely the Hurdle and the Zero-Inflated model using the Poisson regression model on the second stage perform best, which is surprising since both models are rarely applied in the literature. An exception provides a study by Frondel and Vance (2011) which analyzes the determinants of ridership in public transport using a count-data model. Another important finding of this study is, that it obviously matters which model is used. Looking at all analyzed models, we find a quite substantial divergence between the estimates of the price elasticity. It ranges from –0.115 until –0.255.

The structure of the paper is as follows. The following section discusses the different estimation methods and their assumptions from a theoretical point of view. Section 3 shortly presents the data we use for our empirical comparison of the different estimation models followed by a description of the model selection procedures in section 4. In section 5 we will present the estimation results of the elasticities as well as the results of the model comparison tests. Finally we discuss these results and conclude in section 6.

2.2 Theoretical background of fuel price elasticity estimation

The theoretical foundation for fuel price elasticity regression models is a constant elasticity demand function like $q_i = \alpha p_i^{\beta_1} m_i^{\beta_2}$. With q as amount of fuel purchased or distance driven, p standing for the fuel price and m for income. As this relation is not deterministic and the respective elasticities β_1 and β_2 cannot simply be measured a stochastic term u accrues. Most commonly one thinks about additive error terms which would lead to the following stochastic model:

$$q_i = \alpha p_i^{\beta_1} m_i^{\beta_2} + \varepsilon_i \tag{1}$$

The alternative would be a multiplicative connection of the error term.

$$q_i = \alpha p_i^{\beta_1} m_i^{\beta_2}\, \varepsilon_i \tag{2}$$

From Espey's (1998) meta-analysis it becomes clear that most of the studies, she included in her comprehensive fuel price elasticity analysis, use **Ordinary Least Squares (OLS)** regression methods to estimate price elasticities. OLS regression implies log-linearization of the Cobb-Douglas type estimation equation. Independent from the assumption made about the connection of the error term log-linearization of the demand function leads to violation of the OLS.1 assumption. OLS.1 requires strict exogeneity of the errors for the OLS estimates being unbiased. That means $E(X'\,\varepsilon) = 0$ with X as the matrix of explanatory variables. After log-linearization the OLS.1 assumption for unbiasedness turns into the following relation: $E(\ln X\,'\,\ln\varepsilon\,) = 0$ The latter relation is rather unlikely to hold as the expected value of the logarithm of ε depends both on its mean and on the higher order moments of distribution (Silva and Tenreyro 2006, p. 642), which includes the variance. The variance of ε again depends most likely on X since most datasets have the property of increasing variance of the errors with increasing value of the explanatory variables.

The second problem of log-linearization is the problem of the occurrence of zeros in the dependent variable. In the estimation of fuel price elasticities this problem mainly arises if the analysis bases on micro data, since data of driving or fueling behavior from individuals or households is mostly derived from surveys which only cover a limited observation period, like a certain day or a certain week. That makes it rather likely to observe individuals who did not drive or fuel their car in that specific period. For more aggregated data on the contrary this is a rather unlikely observation. Therefore the zero-problem primarily arises analyzing micro-level data. The problem of zero observations in the dependent variable lies in the simple fact that the logarithm of zero is undefined and thereby the log-linearization leads to omission of all individuals who did not drive in the observation period no matter whether they do not drive at all or whether it was simply a coincidence that they were not driving in the observation period. Leaving out this information leads to an underestimation of the fuel price elasticity since the decision of not having a car or not driving might also be a consequence of the fuel price. To circumvent this problem empiricists typically add an arbitrarily small constant to the dependent variable. But adding a constant to a non-linearly transformed variable in turn leads to biased coefficients again as we demonstrate in the Appendix.

As a solution to the problems of the log-linearization of the multiplicative demand function Silva and Tenreyro (2006) suggest the application of a **pseudo-maximum-likelihood (PML)** estimator which estimates the fuel or travel demand directly in multiplicative form. Another possibility would be to use the **nonlinear least squares (NLS)** estimator but this method weights large observations more which is not very efficient since they exhibit also a larger variance and thereby calculating the coefficient one would rely more on noisier observations. This flaw could be remedied by estimating a weighted NLS estimator but therefore a nonparametrically estimated function of the conditional variance would be needed. The PML estimator Silva and Tenreyro (2006) propose gives equal weight to all observations and does not require a nonparametric estimation of the conditional variance of the dependent variable. It is numerically equal to the **Poisson pseudo-maximum likelihood (PPML)** estimator and assumes the conditional variance of the dependent variable y_i to be proportional to the conditional mean of y_i.

$$Var(y_i|x_i) = \sigma^2 E(y_i|x_i) \text{ with } \sigma^2 = 1 \tag{4}$$

Therefore the PPML estimator is completely identified by its conditional expectation $E(y_i|x_i) = \exp(x_i\beta)$ (with $y_i = q_i$ and $x_i = (p_i, m_i)$ in our application).

For $\sigma^2 > 1$ Cameron and Trivedi (1986) propose a full maximum likelihood estimator – the **Negative Binomial (NB)** model. $\sigma^2 > 1$ means that the data exhibits overdispersion. Under this special assumption, which can be tested, the NB model constitutes a more efficient estimator than the PPML estimator. Since variance and conditional expectation are not equal for overdispersed data the NB model additionally to the conditional expectation needs its variance ($Var(y_i|x_i) = (1+\alpha)\exp(x_i\beta)$) to be identified unambiguously. If, however, the variance condition of the NB model ($\alpha > 0$) is not met, its coefficient estimates become inconsistent (Wooldridge 2002, p. 657).

An alternative approach to model automobile travel demand is to think of the dgp as a two-step process. The first decision consumers have to make is to own a car or not. Then, given the consumers possess a car, the second decision is how much to use it. There are two models from the family of **Two-Part** models which are predestined to describe this dgp: **Zero-Inflated** models and **Hurdle** models. By allowing two different dgp's Zero-Inflated and Hurdle models account for excess zero kilometers driven caused by people who do not own a car.

The first step of the Zero-Inflated model consists of a binary count model like the Probit model

$$\Pr(y_i = 0|x_i) = \int_{-\infty}^{x_i} \frac{1}{2\pi} \exp\left(\quad \frac{1}{2}(x_i\beta)^2\right) dx = \Phi(x_i\beta) = \pi_i.$$

Where π_i stands for the probability of the household having no car and π simply depicts the number π. Step two is composed of a count model $g(j)$. Typically Poisson or Negative Binomial models are applied as second step. The specialty of Zero-Inflated models compared to Hurdle models is that zeros can materialize on both steps. That means consumers either have no car and therefore drive zero kilometers, or they have a car but they decide not to drive in the observation period. Equation 5 shows the count distribution of the Zero-Inflated model.

$$\Pr(y_i = j) = \begin{cases} \pi_i + (1 \quad \pi_i)g_i(0) & j = 0 \\ (1 \quad \pi_i)g_i(j) & j = 1, 2, \dots \end{cases} \tag{5}$$

Following equation 5 we can calculate the expected value as

$$\mathrm{E}(y_i|x_i) = (1 \quad \pi_i)\mu_i \text{ with } \mu_i = E\big(g_i(j)\big) \text{ for } j = 1, 2, \dots . \tag{6}$$

The variance which characterizes Zero-Inflated models is depicted in equation 7.

$$Var(y_i|x_i) = \pi_i\mu_i + \pi_i\mu_i^2(1 \quad \pi_i) \tag{7}$$

From equation 7 can be seen that the variance is larger than the expected value μ_i and thereby zero inflated models are similar to the NB model able to account for overdispersion. The second Two-Part model, the Hurdle model accounts for over- as well as for underdispersion. In return the Hurdle model does not allow zeros to occur on the second step. In our case that would mean all people who have a car also drive in the observation period. This restriction also gets clear from the count distribution:

$$\Pr(y_i = j) = \begin{cases} \pi_i & j = 0 \\ k_i g_i(j) & j = 1, 2, \dots \end{cases} \tag{8}$$

With $k_i = (1 \quad \pi_i)/(1 \quad g_i(0))$. Which leads to

$$\mathrm{E}(y_i|x_i) = k_i\mu_i \text{ with } \mu_i = E\big(g_i(j)\big) \text{ for } j = 1, 2, \dots \tag{9}$$

and

$$Var(y_i|x_i) = k_i\mu_i + k_i\mu_i^2(1 \quad k_i). \tag{10}$$

Since k_i can be both, smaller or larger one, the Hurdle model can additionally to overdispersed data deal with underdispersed data. However, the latter case might be rather rare in car use data (Frees 2010).

Similar to the Zero-Inflated and the Hurdle model the so called **Heckit** model, which goes back to the seminal work by Heckman (1979), as well assumes a two-stage dgp and is regularly applied in the analysis of datasets with a large share of zero observations. But there is a decisive difference between the Two-Part models and the Heckit procedure: Dow and Norton (2003) make this distinction clear by differentiating between *actual* and *potential* outcomes. *Actual* outcomes describe a dependent variable which can be fully observed. For *actual* outcomes zero observations are a consequence of the true dgp. For car use data that means that people can drive a positive amount of kilometers or they cannot drive which results in a zero observation. The zero is therefore the true observation in this dgp. The *potential* outcome in contrast, cannot be fully observed. It is a latent variable which is only observed for a fraction of the sample. For the rest of the sample it is often set to zero and thereby creates excess zeros. The prime example for a *potential* outcome is the wage offer (Gronau 1974) which can only be observed from the working part of the population. Estimates of certain characteristics' effects would be biased if the working part of the population would differ systematically from the not working part of the population after controlling for the covariates. The original purpose of the Heckit model is to correct for this selection bias. Its first stage consists similarly to the Two-Part models of a Probit model:

$$\Pr\left(y_{1,i} = 1 \middle| x_i\right) = \Phi(x_i\beta_1, \varepsilon_1) \tag{11}$$

And the second stage is an OLS regression among the subsample for which $y_{1,i} = 1$ holds. The overall marginal effect of x_i can be deducted from the following expected value.

$$\begin{aligned} \mathrm{E}\left(y_{2,i} \middle| y_{1,i} = 1, x_i\right) &= x_i\beta_2 + E(\varepsilon_2 | y_i > 0, x_i) \\ &= x_i\beta_2 + \rho\sigma\lambda(x_i\beta_1) \end{aligned} \tag{12}$$

With $\lambda(\,) \equiv \phi(\,)/\Phi(\,)$ being the inverse Mills ratio, which controls for potential correlation between the error terms from both stages. It is the estimation of $E(\varepsilon_2 | y_i > 0, x_i) = x_i\beta_2 + \rho\sigma\lambda(x_i\beta_1)$ where $(\varepsilon_1, \varepsilon_2) \sim N(0, \Sigma)$ and

$$\Sigma = \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}. \tag{13}$$

From Σ it can easily be seen that the Heckit model reduces to an OLS regression as the inverse Mills coefficient $\rho\sigma$ gets zero when the errors of both stages are uncorrelated. Although the Heckit model originally was developed for analyzing potential outcome variables it can also be used to estimate regression equations with actual outcome variables as it is the case in the fuel or travel demand literature. The marginal effects calculation of the explanatory variables on the actual outcome variable, however, must be based on a different expected value which combines both stages (Dow and Norton 2003):

$$\mathrm{E}(y_i | x_i) = \Phi(x_i\beta_1)[x_i\beta_2 + \rho\sigma\lambda(x_i\beta_1)] \tag{14}$$

Of course, this selection of models does not cover the whole range of models used in the fuel demand or automobile travel demand literature but it comprises the workhorse model in this literature – the loglinear regression model – and a few models which were brought up to alleviate the caveats of the loglinear model. Another model being applied in this strand of literature is the **Tobit** model. Due to theoretical arguments the applicability of the Tobit model must be doubted in the analysis of automobile travel demand. Maddala (1992, p.341) argues that the Tobit model should only be applied to censored data. That means data whose observations are stacked into one category in case they are under or above a certain threshold. This is, for example, often the case for pollutant data when the measuring instrument can only detect pollutant concentrations above a certain level. However this is not appropriate for the type of data used for travel or gasoline demand analysis. The zeros are not a consequence of censoring but of the decision of people not to drive. Maddala argues that instead of mechanically falling back to the Tobit model the appropriate procedure would be to model the decision that produces the zeros. That is exactly what those models like the already described Hurdle and Zero-Inflated models or the Heckit model do.[26]

From a theoretical perspective the log-linearized demand function estimated with OLS will presumably not perform well applied to micro driving data due to the endogeneity of the logged error term and the occurrence of zero observations for kilometers driven. If the driving data can be described as one decision by households, meaning that the analyzed explanatory variables have equal impact on the decision of owning a car and driving, the Poisson and the Negative Binomial model will be the models of choice. To decide between Poisson and Negative Binomial the crux of the matter is the conditional variance of the kilometers driven. If it is larger than the conditional expectation, meaning the data is overdispersed, the Negative Binomial model outperforms the Poisson regression. For the case the dgp is assumed to be a stepwise process however, a switch to the Two-Part models is indicated. They model the dgp of the driving data as a two-step process. On the first step consumers decide upon buying a car or not and on the second step, given they have

[26] Nevertheless, we included the Tobit model in our analysis but we will not go into details about this model in the main part.

bought a car, how much to drive with that car. Again there are two options with Two-Part models. First the Hurdle models which do not allow zeros to occur on the second step which means that Hurdle models assume every household which has a car to drive. Second there are the Zero-inflated models which relax this assumption. Finally we include the Heckit model in our analysis. The Heckit model would be a good choice if there is a substantial share of households in the data who do not own a car and who would react differently to price or income changes if they had a car compared to those households in the dataset who own a car. Before coming to the characteristics of our dataset we present the model comparison tests in the next section.

2.3 Model Comparison Tests

One of the main contributions of this study is the proposition of tests for comparing the suitability of different models for vehicle travel or fuel demand. For being able to compare not only *non-nested* models but also models which base on different distributional assumptions. This is not possible with the standard model selection criteria like the Akaike Information Criterion (AIC) or the Vuong test (Vuong 1989) as they hinge on the maximum likelihood. Comparing likelihoods from two models with different distributional assumptions however is not very meaningful without some standardization. Therefor we resort to model selection tests which rely on the model's predictive ability.

Hansen et al. (2011) propose a convincing approach to determine the best performing models given a certain dataset. They developed the **Model Confidence Set** (MCS) which sorts out "bad performing" models and thereby leads to a selection of "good performing" models – the MCS – that contains the best model(s) on a chosen significance level. The structure of the procedure is as follows:

1. Definition of a set of models to be compared.
2. Test of the null hypothesis of **equal predictive ability** (EPA) of the models.
3. If the null hypothesis is rejected, the worst performing models are eliminated from the comparison set and the null hypothesis is tested again. This process is repeated until the null hypothesis is accepted und thereby the MCS is determined.

The basis for the elimination rule is provided by a loss function, which in our case is defined as a measure of the prediction's deviance from the observed values. Next to revealing the best models the MCS also provides information about the data's informative quality. An informative dataset leads to a MCS containing few, or at best only one model whereas a less informative dataset leads to a MCS that contains more models. Besides the differentiation between models being inside and outside the MCS, it allows a continuous ranking of all models, since the MCS is a multilateral comparison method and no bilateral method like the Vuong test, for example.

To check whether this ranking is robust, we firstly use different types of loss functions and secondly we apply further measures to compare the model's goodness of fit. As first robustness check we use a resampling method – the **Leave-One-Out Cross-Validation** (LOOCV). The principal idea of LOOCV is to take a subsample of the analyzed dataset which consists of all observations but one, fit the model to the subsample and use the fitted model to predict the response for the left-out observation. This procedure is repeated for all observations what leaves you with n predictions for the n observations. From these n pairs a mean squared error (MSE) is calculated. In this way we derive a MSE for each model and use it as indicator for the comparison of the goodness of fit of the different models for the given dataset.

Another comparative indicator is the mean of the differences between the observed count probabilities of kilometers driven and the predicted count probabilities of kilometers driven (Liu and Cela 2008). For this purpose we first calculate the categorical distribution for the kilometers driven, once for the observed parameter values and once for the predictions of each of the models being compared. Then we take the differences between the observed count probabilities and the predicted count probabilities at each value of kilometers driven. Taking the mean of these differences provides us with another goodness of fit measure for the different models focusing on the probability distribution. For the remainder of this paper this test is referred to as the Difference of Count Probabilities (DOCP) test.

2.4 Data

As case study for our model comparison we use a survey-based micro dataset on the mobility behavior of German households – the German Mobility Panel (MOP 2013). The mobility panel results from two surveys. The first one asks facts about everyday mobility and the second one is focused on automobile mobility only. The respondents of the everyday mobility survey are asked to fill in a travel diary over the time span of a week containing questions on the purpose of the travel, starting- and end time, distance and so on. Here all travels are questioned independent of the means of transport. In the second survey however, only a subsample of car owning households is interviewed and asked to fill in a refueling diary over a period of two months with information about kilometers driven between each refueling stop, the amount of petrol purchased and the price paid for the petrol. In our observation period from 2000 until 2013 there are 2752 households in the sample. Since the survey is organized as a panel, some households emerge several times. At most one household stays in the panel for three years, which is rather short for making use of panel methods. Additionally only 703 households stayed in the panel for all three waves.[27] Therefore we make a pooled cross-section analysis with household clustered errors.

Variable	Variable definition	Mean	Std.Dev.	Min	Max
Daily kilometers	Daily kilometers driven in km on average in the observation period	28.43	24.57	0	202.6
Fuel price	Real fuel price in € per liter	1.389	0.134	1.085	1.687
Income (real)	Real net monthly household income in €	2052	887.2	236.5	4376
Household size	Number of people living in the household	1.863	0.967	1	7
Employed	1 if person is employed in full- or half-time job, 0 otherwise	0.487	0.500	Binary variable	
N = 4891					

Table 10: Descriptive Statistics

From the 2752 households there are 678 households which do not own a car. And these 678 households without a car are responsible for all the zero observations of kilometers driven in our dataset. That means there is no household possessing a car which does not drive in the observation period. The carless households stem from a randomly selection from the general MOP survey such that they comprise 15% of the observations in a given year. This share roughly corresponds to the actual share of carless households in Germany.

The participants of the survey are also asked to give some background information on their household and personal characteristics including age, gender, amount of household members and employment status among others. But since we are not after measuring the size of a certain variable's effect on the demand for driving but more after comparing different models with respect to their predictive ability, we keep the demand equation rather basic. We concentrate on four main drivers of the demand for driving: fuel price, income, household size and the employment status (see Table 10Table 1 for summary statistics of these variables). Because in the original dataset the fuel price is only available for households which took part in the automobile mobility survey we use spatially interpolated price data covering all of Germany by Frondel and Vance (2011). By spatial interpolation Frondel and Vance succeeded in assigning fuel prices also to households which were not asked for the fuel price in the survey. Having this data is very important for such kind of analysis, as the fuel price might also be an important driver in the decision of buying a car or not.

[27] 733 households responded twice and the majority of 1316 households only took part once. That makes a total number of observations of 4891.

2.5 Results

Most of the literature in fuel or travel demand estimation in the past used the log-linear regression without testing the suitability of the model to the data at hand. In this section we compare the goodness of fit of the models presented in section 2.2 on the basis of our case study. We also present the estimation results of the different models to illustrate the size of the potential deviations between the estimates of the different models.

Table 11 shows the rankings which result from the different model comparison tests. For each of the test we specified two different loss functions – the squared error and the absolute value of the error – to check the robustness of the results towards changes of the loss function. And it turns out that changing the loss function inside the testing procedures leads to minor changes in the rankings. Comparing the ranks of the three tests with each other it becomes clear that especially the MCS and the LOOCV test result in rather similar rankings for the twelve different models analyzed. The reason for this finding is that these two tests focus on a similar selection criterion which is the predictive ability whereas the DOCP test puts more emphasis on the correspondence of the distributions of the observed and predicted outcome variable.

	Model Confidence Set		**LOOCV**		**DOCP**	
Model	squared error	absolute error	squared error	absolute error	squared error	absolute error
OLS	12	10	12	10	10	4*
Zero-truncated OLS	8	8	7	7	8	10*
PPML	5	5	5	5	2*	2*
Zero-truncated PPML	10	12	10	12	11	12*
NB	6	6	6	6	1*	1*
Zero-truncated NB	9	11	9	11	9	11*
Hurdle OLS	11	9	11	9	12	8*
Hurdle PPML	2*	1*	1*	1*	5*	6*
Hurdle NB	4*	4*	4*	4*	3*	3*
ZIP	1*	2*	2*	2*	6*	7*
ZINB	3*	3*	3*	3*	4*	5*
Heckman	7	7	8	8	7	9*

Table 11: Rankings of the model comparison tests (models whose ranks are marked with a * are similar on the 95% significance level)[28]

[28] See Figure 9 in the Appendix for an explanation of the significance level of the LOOCV and the DOCP test. The MCS test produces the significance levels automatically as it is a joined test (see Section 3).

The models can be divided roughly into two groups – the one-step and the two-step models. The one-step models consist of OLS, PPML and NB. For each of them we estimated one version including all observations and one version excluding the zero observations, since both approaches are used for estimating fuel price elasticities from a log-linearized demand function by OLS. For being able to do the analysis for all observations using OLS, we applied the Inverse Hyperbolic Sine (IHS) transformation which adds a constant to the outcome variable relative to its size. Thereby the zero observations are not dropped because of taking logs. The IHS transformation reduces the induced bias by the added constant and was first introduced by Johnson (1949). The two-step models include the Hurdle model with OLS, Poisson and NB as second-step model, the Zero-inflated model with Poisson and NB as second step model and the Heckit model. Since Zero-inflated models allow for zero-observations on the second step as well, it does not make sense to combine it with OLS like the Hurdle model.

Taking the MCS and the LOOCV test into account, which means looking at the predictive ability of the models, the result which can be derived from theory gets confirmed. All models which base on OLS do not perform well due to the zero-problem but also due to the endogeneity of the error. The endogeneity issue of the errors is illustrated by the ranks of the zero-truncated versions of the one-step models as they abstract from the zero-problem by looking at non-zero observations only. Also here PPML and NB outperform OLS. The same holds true for the two-step models. Again the OLS version of the Hurdle model ends up behind the other Hurdle models and the Zero-Inflated models. So does also the Heckit model although, we included an additional identifying variable to estimate the first step of the Heckit model. We used the costs of the household's car insurance as identifying variable since it can be seen as a fixed cost which might influence the decision of buying a car but not the decision of how much to drive. Both the MCS and the LOOCV test give rise to the assumption that the Hurdle model performs best for our case study. This is completely coherent with theory. As has been stressed before the Hurdle model assumes the dgp to be a result of two decisions: First, the households decide to buy a car or not and once they bought a car they decide how much to drive and in contrast to the Zero-Inflated model it does not allow for zero observations on the second step which is true for our dataset since all the zero-observations stem from households without a car. Furthermore overdispersion does not seem to play a big role in the data with respect to the predictive ability. Although the likelihood ratio test finds the overdispersion parameter α being larger than zero and in general the NB model seems to make better distributive assumptions as the Poisson model as can be deducted from the DOCP test. Here the NB models always outperform the Poisson versions of the four models (PPML, zero-truncated PPML, Hurdle Poisson and ZIP). But as mentioned above, overdispersion leads merely to less efficiency of the Poisson estimator and not to biased estimates. Additionally the rankings of both Hurdle models and both Zero-Inflated models are not statistically significant on the 95% level which is indicated by the asterisks in Table 11. For the DOCP test the amount of models between which no statistical significant differences exist is even larger with 6, respectively all 12 models. That means there is no distinction possible according to the second DOCP test. But apart from that test, all tests show that models relying on OLS estimation perform significantly worse than the models which estimate the demand function directly in its multiplicative form like the Poisson and Negative Binomial regression models. The good performance of the OLS model in the DOCP test with absolute value of the error as loss function is relativized by the large variance of the mean compared to the other models.

To assess if it is worth the effort to compare different models one should look at the variance of the coefficients being estimated since they are the values of interest. When there are no notable differences in the size of the coefficients one could argue that the choice of the correct model is not too important. Table 12 depicts the marginal effects defined as elasticities calculated from the estimated coefficients.[29]

The first thing to be noticed is that there is a quite substantial variance in the estimated elasticities between the different models. The price elasticity ranges from –0.261 (zero-truncated PPML) until very high –0.741 (OLS).

One-step models:

Variables	OLS	Zero-truncated OLS	PPML	Zero-truncated PPML	NB	Zero-truncated NB
Fuel price	-0.741**	-0.345*	-0.473**	-0.261*	-0.483**	-0.324*
Income (real)	1.186**	0.171**	0.533**	0.193**	0.543**	0.181**

Two-step models:

Variables	Hurdle OLS	Hurdle PPML	Hurdle NB	Heckman	ZIP	ZINB
Fuel price	-0.603*	-0.351*	-0.418**	-0.348**	-0.351*	-0.414**
Income (real)	1.085**	0.467**	0.454**	0.178**	0.467**	0.455**

* $p<0.05$; ** $p<0.01$

Table 12: Marginal effects

For the income elasticity the range is even larger: 0.171 (zero-truncated OLS) until – 1.186 (OLS). So it is worth to compare different models to estimate fuel demand elasticities. Part of the variance however can be explained by the differences regarding the specification. Considering the one-step models, the zero-truncated versions always lead to smaller elasticities in absolute terms. This is due to the fact, that in these models only

[29] Table 13 in the Appendix shows the complete regression results of the analyzed models.

households are taken into account, which have a car. The one-step models which include all households combine the elasticities of the decisions of buying a car and how much to drive. Especially the income effect is by far bigger for the investment decision to buy a car than for the marginal decision of how much to drive once the car is in the garage. This finding already is a strong argument for the application of two-step models since most of the time researchers and politicians are interested in the overall price effect on driving demand and thereby on fuel demand which in turn causes GHG emissions.

The second noticeable finding is that the price elasticities estimated by OLS and Hurdle OLS which include all observations are beyond the range of elasticities found in the existing literature. Hanley et al. (2002) specify the range in their literature review beginning with 0.25 for short-run and ending with –0.6 for long-run elasticities. Since we are estimating a static model we cannot say something about the time dimension, but nevertheless the non-zero-truncated OLS model yields marginal effects beyond the scope typically found in the literature.

The third remarkable result is that the price elasticity estimate of the zero-truncated OLS differs only slightly from those of our most preferred model, the Hurdle PPML. This shows that the zero-problem seems to be the largest problem with respect to the extent of the bias as this is excluded by the zero-truncated OLS regression model.

Another noteworthy result is that the one-step zero-truncated count models (PPML and NB) have diverse results, while their non-zero-truncated versions lead to very similar results concerning the price elasticities. That hints at the fact that the zeros cover the overdispersion which exists in the non-zero part of the sample for which the NB model controls and the PPML model does not. This explanation is confirmed by testing the statistical significance of the overdispersion parameter α, which is positive and statistical different from zero. This is also the case for the sample including the zeros but here the impact of the overdispersion is most probably outweighed by the effect of the excess zeros. This explanation would also explain the fact that the elasticity estimates of the two-step models using the Poisson model as second step differ from those using the NB model as second step.

Another peculiarity attracts attention while comparing the marginal effects of the Hurdle PPML with the ZIP model and the marginal effects of the Hurdle NB with the ZINB model: The models using a Poisson regression as second step (Hurdle PPML and ZIP) lead to exactly the same results while this is not the case for the Hurdle NB in comparison with the ZINB model. For understanding this result one should recall some of the features of zero-inflation and the NB model in comparison to the Poisson regression. The two count models differ in the capability of incorporating excess zeros. The NB model is able to explain more zeros than the Poisson model as it controls for overdispersion. The first step in the Zero-Inflated models, in contrast to the first step of the Hurdle model, only explains the zeros which are not explained by the model of the second step. In case of the Zero-Inflated PPML all zero observations are attributed to the first step. This makes it equal to the Hurdle PPML model, since the Hurdle model explains all zero observations per definition. In case of the ZINB model not all zeros are attributed to the Probit model on the first step, as the NB model on the second step explains the zero observations to a certain extent itself. This can be seen by looking at the coefficients estimated by the Probit regression in the first step in each of the four models[30].

Last, not only the statistical tests prove that there is little to no difference between choosing Hurdle or Zero-Inflated models with count data models as second-step but also the estimation results of the Hurdle PPML model are rather similar if not equal to the ZIP model as well as the results of the Hurdle NB and the ZINB model.

2.6 Conclusion

This study presents empirical models which are used, or should be used for the estimation of fuel demand elasticities. At the same time we introduce methods to compare the empirical performance of the different models. Applying the different models and tests to a German household dataset we show the importance of assessing the goodness of fit of the respective model to the data at hand as estimates can differ quite substantially between different estimation models.

A common phenomenon in the applied econometrics literature is the existence of a workhorse model in each branch which is adopted again and again without scrutinizing. In the fuel or vehicle travel demand literature this is, for example, the log-linear model. Silva and Tenreyro (2006) showed however, that the log-linear model has some flaws in the application to right skewed data with a large fraction of zero observations in the dependent variable. This data pattern also applies to driving or fuel demand micro data as there are always people who do not drive in the observation period especially in Europe where a decisive fraction of households does not own a car. Therefore this paper provides an overview of empirical models which are applied in the estimation of price elasticities of travel demand and those models which possibly should be applied. We compare the different models with regard to their suitability in the estimation of price elasticities on theoretical grounds. The models discussed make different assumptions concerning the distribution of the analyzed data as well as concerning the data generating process which is behind the data. For our specific dataset a two-step approach seemed to be most promising as it models the decision to drive and the decision how much to drive separately which fits very well the large fraction of zero observations for kilometers driven.

---

[30] See Table 14 in the Appendix for the regression results.

Besides the theoretical assessment of the models we present three different statistical methods to compare the model's empirical goodness of fit. The Model Confidence Set (MCS), the Leave-One-Out Cross-Validation and an approach which compares the count probability of the predictions with the count probability of the true observations. Since the selected models partly base upon different distributional assumptions we cannot resort to typical model comparison tests like the Vuong test or measures like the Akaike or Bayesian Information Criterion as they rely on the likelihood as comparison measure which does not allow the comparison of models with different distributional assumptions.

Applied to our case study – the German Mobility Panel – the model comparison tests identify the Two-Part models with a Probit- as first and Poisson regression as second step as the best performing model. Further well performing models are the Two-Part models with a Negative Binomial regression model as second step. The workhorse model – the log-linear model – however is not among the best models according to our comparison. One has to emphasize at this point that this is no general result as it can differ from sample to sample. But this result shall encourage researchers to test the goodness of fit of various models to their analyzed data. This is stressed even more looking at the variability of the results depending on the different estimation methods. In our case the estimations of the price elasticity vary between –0.261 and –0.741. The ratio of these estimates is larger than the factor 2 which is crucial when it comes to the derivation of policy advice.

An interesting expansion of this study would be the application to another case study and the question if the patterns shown for the German data hold true for other countries or regions as well.

In any case, this work shows that it is advisable to scrutinize the empirical model with respect to the data and the hypothesis to be tested.

2.7 References

Basso, Leonardo J. and Tae Hoon Oum (2007), "Automobile Fuel Demand: A Critical Assessment of Empirical Methodologies." *Transport Reviews* 27 (4).

Cameron, Colin A. and Pravin K. Trivedi (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics* 1 (1).

Dow, William H. and Edward C. Norton (2003), "Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions", *Health Services & Outcomes Research Methodology* 4.

Espey, Molly (1998), "Gasoline Demand Revisited: An International Meta-Analysis of Elasticities", *Energy Economics* 20 (3).

Flowerdew, Robin and Murray Aitkin (1982), "A Method of Fitting the Gravity Model Based on the Poisson Distribution", *Journal of Regional Science* 22 (2).

Frees, Edward W. (2010), "Regression Modeling with Actuarial and Financial Applications", *Cambridge University Press*, New York, NY.

Frondel, Manuel, Jörg Peters and Colin Vance (2008), "Identifying the Rebound: Evidence from a German Household Panel", *The Energy Journal* 29 (4).

Frondel, Manuel and Colin Vance (2011), "Rarely enjoyed? A count data analysis of ridership in Germany's public transport", *Transport Policy* 18.

Gillingham, Kenneth, Matthew J. Kotchen, David S. Rapson and Gernot Wagner (2013), "Energy Policy: The Rebound Effect is Overplayed", *Nature* 493.

Gronau, Reuben (1974), "Wage Comparisons-A Selectivity Bias", *Journal of Political Economy* 82 (6).

Hanly, Mark, Joyce Dargay and Phil Goodwin (2002), "Review of Income and Price Elasticities in the Demand for Road Traffic", *Economic & Social Research Council Transport Studies Unit publication* 2002/13

Hansen, Peter R. et al. (2011), "The Model Confidence Set", *Econometrica* 79 (2).

Haworth, J. M. and P. J. Vincent (1979), "The stochastic disturbance specification and its implications for log-linear regression", *Environment and Planning A* 11 (7).

Heckman, James J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica 47* (1).

Johnson, N. L. (1949), "Systems of Frequency Curves Generated by Methods of Translation", *Biometrika* 36 (1/2).

Liu, WenSui and Jimmy Cela (2008), "Count Data Model in SAS", *SAS Global Forum 2008, Paper* 371-2008.

Maddala, G.S. (1992), "Introduction to Econometrics", *Macmillan Publishing Company*, New York, New York.

MOP (2013), "Deutsches Mobilitätspanel", Federal Ministry of Transport.

Silva, J. M. C. Santos and Silvana Tenreyro (2006), "The Log of Gravity", *The Review of Economics and Statistics* 88(4).

Vuong, Quang H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econmetrica* 57 (3).

Wadud, Zia, Daniel J. Graham and Robert B. Noland (2010), "Gasoline Demand with Heterogeneity in Household Responses." *The Energy Journal* 31 (1).

Wooldridge, Jeffrey M. (2002), "Econometric Analysis of Cross Section and Panel Data", *MIT Press*, Cambridge, Massachusetts.

2.8 Appendix

A1. Bias after adding a constant

In the following we show how adding a constant to the dependent variable leads to biased coefficients:

$$ln(q_i) = ln(k) + \alpha\, ln(p_i) + ln\, u \;\; or\, y = \beta_0 + \beta_1 x_i + \mu \qquad \text{(A.1)}$$

$$\widehat{\beta_1} = (x'x)^{-1}x'y = \frac{Cov(x,y)}{Var(x)} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

insert $\ln(q)$ for y:

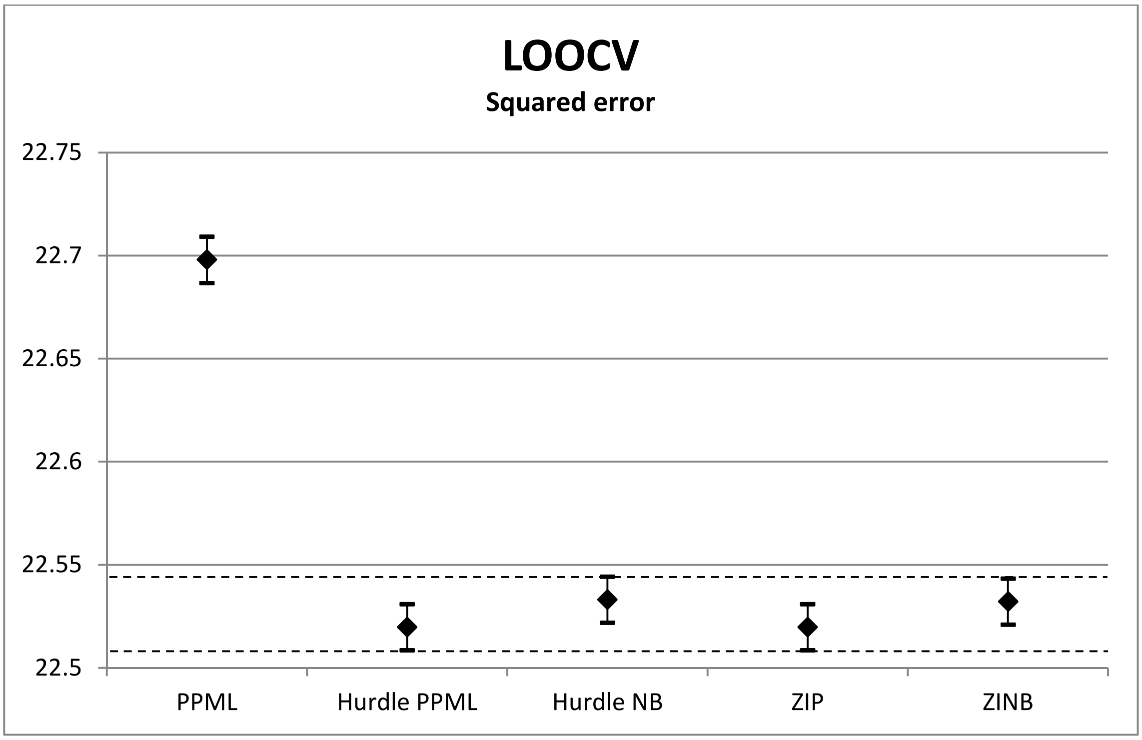
$$\widehat{\beta_1} = \frac{\sum(x_i - \bar{x})(\ln(q_i) - \overline{\ln(q)})}{\sum(x_i - \bar{x})^2} \qquad \text{(A.2)}$$

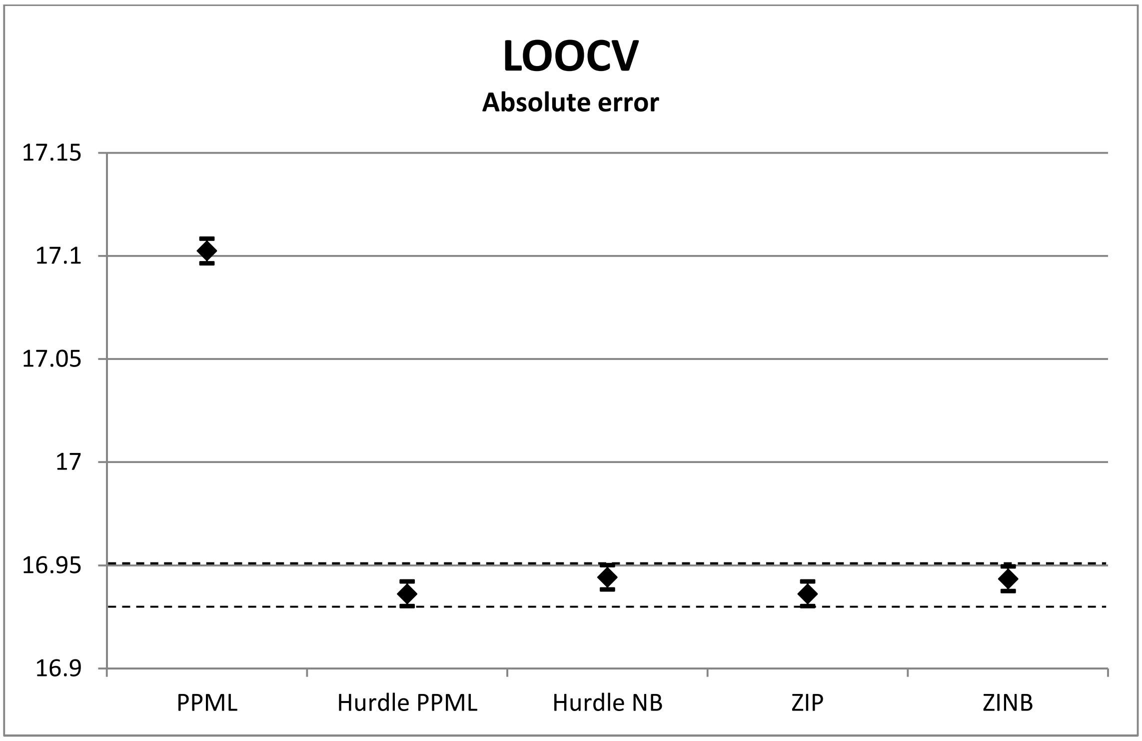
Adding an infinitesimal constant c to the dependent variable q to circumvent the problem of zeros leads to a biased estimator of β_1 after taking the logs.

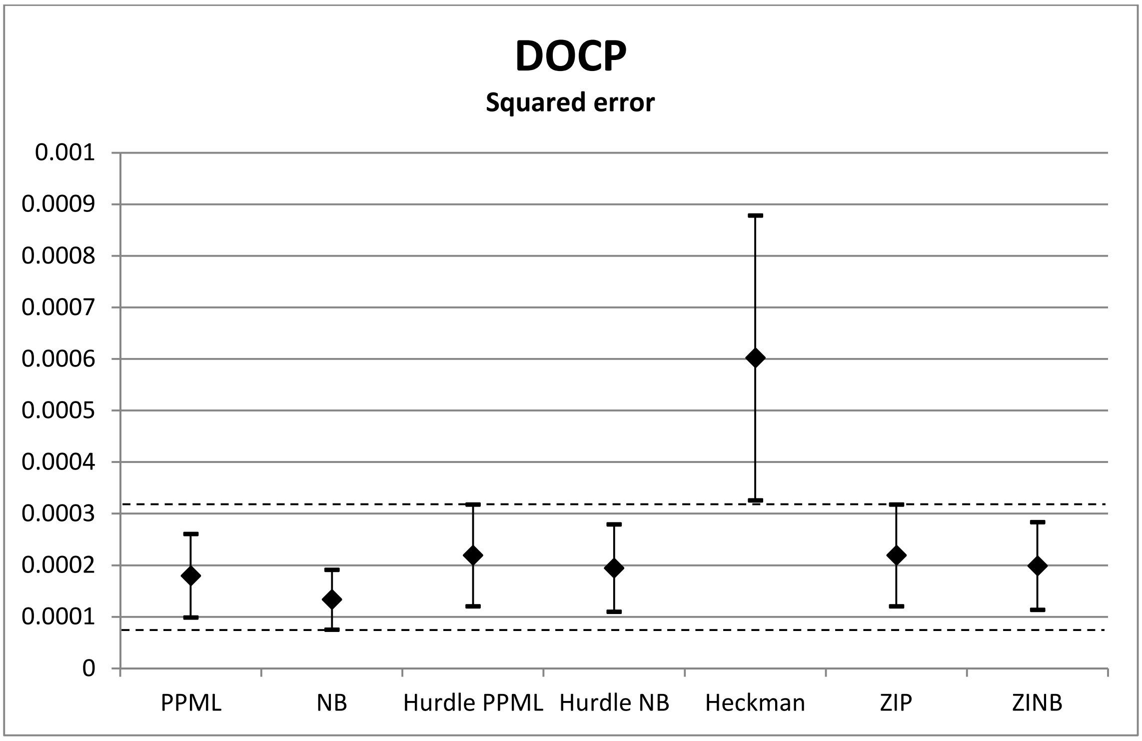
$$\sum(\ln(q_i) - \overline{\ln(q)}) = \sum\left(\ln(q_i + 1) - \overline{\ln(q+1)}\right) \qquad \text{(A.3)}$$

The equality in equation A.3 only holds in special cases. Generally adding a constant leads to a biased estimator of β_1.

A2. Confidence Intervals of the mean errors for LOOCV and the differences in count probabilities




LOOCV
Absolute error
17.15
17.1
17.05
17
16.95
16.9
PPML
Hurdle PPML
Hurdle NB
ZIP
ZINB


DOCP
Squared error
0.001
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
PPML
NB
Hurdle PPML
Hurdle NB
Heckman
ZIP
ZINB

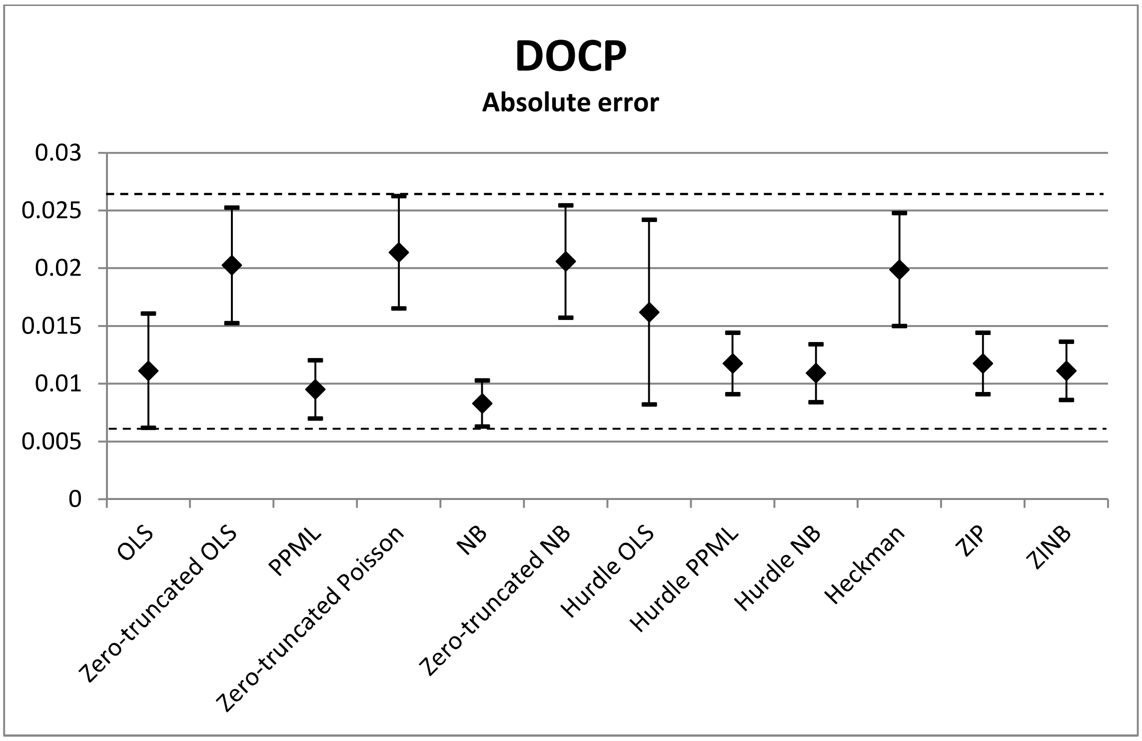


Figure 9: 95% confidence intervals of the mean errors for the test statistics of the LOOCV method and the DOCP method. The rationale behind the assessment of the statistically significant difference between the models is as follows: For each test models are included according to their rank until the confidence intervals overlap. Therefore the diagrams include the best models with overlapping confidence intervals plus the next model in the ranking whose confidence interval does not overlap with the confidence intervals of the preceding models. For the DOCP test with the absolute value of the error loss function this method does not allow a distinction between the models.

A3. Regression results of all models in the log-log specification

One-step models:

Variables	OLS	Zero-truncated OLS	PPML	Zero-truncated PPML	NB	Zero-truncated NB
Fuel price	-0.741**	-0.345*	-0.473**	-0.261*	-0.483**	-0.324*
	(0.267)	(0.134)	(0.146)	(0.126)	(0.163)	(0.127)
Income (real)	1.186**	0.171**	0.533**	0.193**	0.543**	0.181**
	(0.073)	(0.035)	(0.041)	(0.035)	(0.052)	(0.034)
Household size	0.627**	0.158**	0.186**	0.064*	0.269**	0.087**
	(0.065)	(0.032)	(0.037)	(0.032)	(0.038)	(0.031)
Employed	0.172**	0.327**	0.257**	0.287**	0.274**	0.289**
	(0.059)	(0.029)	(0.033)	(0.028)	(0.033)	(0.027)
Constant	-5.810**	1.893**	-0.792**	1.971**	-0.922*	2.069**
	(0.538)	(0.257)	(0.300)	(0.258)	(0.379)	(0.244)
N	4891	3968	4891	3968	4891	3968
AIC	17950.4	7990.2	114807.3	68383.3	41887.8	34327.4
BIC	17982.9	8021.6	114839.7	68414.7	41926.7	34365.2

Two-step models:

Variables	Hurdle OLS	Hurdle PPML	Hurdle NB	ZIP	ZINB	Heckit
Fuel price	-0.344**	-0.261*	-0.327*	-0.261*	-0.325*	-0.348**
	(0.133)	(0.126)	(0.128)	(0.126)	(0.128)	(0.134)
Income (real)	0.171**	0.193**	0.180**	0.193**	0.181**	0.178**
	(0.035)	(0.035)	(0.034)	(0.035)	(0.034)	(0.037)
Household size	0.157**	0.064*	0.089**	0.064*	0.087**	0.161**
	(0.032)	(0.032)	(0.031)	(0.032)	(0.031)	(0.033)
Employed	0.326**	0.287**	0.289**	0.287**	0.289**	0.326**
	(0.029)	(0.028)	(0.027)	(0.028)	(0.027)	(0.029)
Constant	2.588**	1.971**	2.077**	1.971**	2.067**	1.837**
	(0.256)	(0.258)	(0.243)	(0.258)	(0.245)	(0.279)
N	4891	4891	4891	4891	4891	4891
AIC	11867.0	72291.3	40038.9	72291.3	38230.8	11902.2
BIC	11932.0	72356.3	40103.8	72356.3	38302.3	11980.1

* $p<0.05$; ** $p<0.01$

Table 13: Complete regression results.

A4. First-step regression results of the Hurdle and the Zero-Inflated models

Variables	PPML	ZIP	NB	ZINB
Fuel price	-0.339	0.339	-0.339	0.336
	(0.272)	(0.272)	(0.272)	(0.273)
Income (real)	1.029**	-1.029**	1.029**	-1.031**
	(0.078)	(0.078)	(0.078)	(0.078)
Household size	0.627**	-0.627**	0.627**	-0.629**
	(0.072)	(0.072)	(0.072)	(0.072)
Employed	-0.085	0.085	-0.085	0.086
	(0.058)	(0.058)	(0.058)	(0.058)
Constant	-6.851**	6.851**	-6.851**	6.865**
	(0.566)	(0.566)	(0.566)	(0.569)

* $p<0.05$; ** $p<0.01$

Table 14: Probit regression results of the first step for the Hurdle and the Zero-Inflated models. Note that the Probit model of the Zero-Inflated models estimates the probability for the observation being zero and not 1 as it is the case for the Hurdle model. Therefore the signs of the coefficients shift between the Zero-Inflated and the Hurdle models.

Chapter 3

Climate Policy Measure Index

Julian Dieler[*]

Climate policy has been a prominent topic in the political arena as well as in several fields of research for many years already. Although protecting the climate is originally an international issue, the up to now limited success of international climate negotiations has resulted in the fact that most action against climate change has been taken on country level so far. To enable the analysis of countries' policies to reduce GHG emissions it is useful to have a composite measure which integrates the multitude of individual policy measures into one figure. This is what the Climate Policy Measure Index (CPMI) is designed for. It facilitates the analysis of the stringency and effectiveness of climate policy in the countries. The CPMI provides cross-country comparisons of the stringency of climate policy as well as a comparison of the development over time (1991-2012). The major contribution to the existing indices in this field is the entirely data-driven aggregation of the several sub-indicators to one composite indicator, which also considers the effectiveness of the respective policy category with respect to the reduction of the emission intensity of production. Thereby the CPMI provides a high degree of objectivity and transparency, which is of high importance for scientific applicability as well as for the informative value for the public.

[*] ifo Institute for Economic Research at the University of Munich, Germany, +49/89/9224-1346, dieler@ifo.de

3.1 Introduction

The main purpose of indices is to integrate large amounts of information into a format which provides an easily accessible overview of the underlying information. Well-known examples for indices are stock market indices like the Dow Jones or DAX. They comprise price information about all the different stocks which are part of the particular stock market index and in practice they are interpreted even broader, namely as indicator for the entire stock market's performance. In the field of climate related policies the need for information condensation is at least as high if not even higher than on the stock market. For climate policies, it is hardly possible to get a quick overview of the current effort of countries to curb greenhouse gas (GHG) emissions, not to mention of their success in doing so. This is even the case looking at one country only without cross-country comparisons. This has several reasons, first of all the quantity of different policies. The United States, for example, currently have more than 130 climate change related policies in place on national level.[31] Second, it is difficult to compare the different policy measures as they differ with regard to their type (e.g. economic, command and control or information and education instruments) and with regard to their target (e.g. electricity, heat or transport). And third it is difficult to assess the relative importance of a policy, meaning the potential GHG emission reduction by a certain instrument. Thus, it is not surprising that there is a long list of indices dealing with climate issues.[32] Surminski and Williams (2012) describe the history of climate related indices as an evolution which started with indices measuring outcome variables like emissions or degrees of environmental pollution. Nowadays the focus shifts more and more towards looking for indicators which give information about the input variables. Input variables can either be actual input factors like resources, or policy measures which can also be seen as input factors in the various processes which lead to GHG emissions. Among the 33 commonly referenced environmental indices (Surminski and Williams 2012) there are only 6 indices which include climate or environmental policies in their assessment. The way of incorporating policy measures varies considerably between the different indices. The **Climate Change Performance Index** (CCPI) by Germanwatch (2013), for example, compile the policy oriented sub-indicator based on expert opinions. By contrast the **Climate Laws, Institutions and Measures Index** (CLIMI) introduced by Steves and Teytelboym (2013) and the **Index of Sustainable Energy** (ESI) from the European Bank for Reconstruction and Development (EBRD, 2008) rely on scores they award to each considered policy in comparison to other countries' effort in

[31] This number is taken from the IEA Policies and Measures Database (IEA, 2015a). It depicts the number of policy measures against climate change, which are currently in place and were introduced after 1973. The figure stated must even be seen as a lower bound, as the list of the IEA is not completely exhaustive.

[32] See Annex 1 in Surmsinki and Williamson (2012) for a comprehensive list of environmentally related indices.

this measure. These three examples already clearly show the difficulties implied by the construction of a policy index: The first difficulty is finding a measurement method for the different policies which cannot be measured metrically. This includes all regulatory policies as well as information and education instruments. The second difficulty, once a measurement method is found, is to make the indicators comparable between countries (see Holzinger (2006) for an extensive discussion of comparability problems of indices). Another challenge in the construction of policy indices is, of course, the data availability problem which increases in the targeted number of countries and time periods the policy index shall cover. The less countries, or the shorter the time frame covered by the index, the better is the data quality, obviously. But at the same time the information value is very limited for an index consisting of two or three countries only, as it is prone to problems of sample selection. Another difficulty is finding meaningful weighting factors for the aggregation of the sub-indicators. This problem is in most cases solved by simply using equal weights. The first two problems can at least to some extent be circumvented by focusing on metrically measurable policy instruments like taxes, subsidies or tradable permits. This approach is already applied by Künkel and her coauthors (Künkel et al. 2006) in constructing their **Climate Performance Index** (CPI). In contrast to the indices mentioned earlier, the CPI concentrates on climate policy and does not assess the development of several other sub-indicators like the development of GHG emissions, or the ecosystem vitality which are outcome indicators of climate policies.

The Climate Policy Measure Index (CPMI) introduced in this study follows the basic approach of the CPI by focusing on metrically measurable climate policy measures. At the same time it makes two significant contributions: Firstly, the CPMI covers 33 countries[33] instead of 24 OECD countries covered by the CPI. Furthermore, the CPMI is constructed for the continuous period between 1991 and 2013 instead of three single years (1992, 1997 and 2005) as it is the case for the CPI. The second contribution of the CPMI is methodology wise. Instead of using educated guesses for the weights or equal weights the CPMI uses data based weights for the sub-indicators of the different policies considered. To be more precise, the weighting inside the sub-indicators is done by using shares reflecting the scope of the respective policy instrument. For the weighting of the sub-indicators, estimated elasticities are used, which describe the impact of the sub-indicator on the emission-intensity of production. Thereby the stringency measure of the CPMI consists of two parts: First, it consists of the level of the policy and second it consists of a measure for the effectiveness in reducing carbon intensity. This implies that a high tax on e.g. fossil fuel consumption alone is not enough to rank high in the CPMI. The tax also has to be effective with respect to reduction of the carbon intensity. Both aspects will be

[33] These are all OECD countries except for Iceland. The absence of Iceland is due to data availability.

referred to as stringency in the remainder of this study. The goal of the resulting index is not only to provide information about the development of the stringency of a country's climate policies over time and relative to other countries but also to serve as a sound basis for empirical analysis.

The overarching results of the index are that the stringency of climate policies seems to have increased in recent years and that despite the variation in the rankings over time there are few countries which always perform well with regard to the stringency of climate policies. The CPMI provides the possibility to track the strictness in climate policy for each country included over time as well as a cross-country comparison at specific points in time. A by-product of the empirical analysis for the calculation of weighting factors is the finding that the impact of climate policies is rather small expressed in elasticities and that subsidies on renewable energy are more effective than taxes on fossil fuel products and carbon prices in decreasing the emission intensity of GDP.

The paper is structured as follows: The next section briefly summarizes the existing climate policy indices with regard to the applied methodologies and results. In the subsequent section, the data which is used to construct the CPMI is presented. Section 3.4 looks at the theoretical framework behind the CPMI and the methodology applied to compile the index. The following section 3.5 summarizes the results. Furthermore section 3.5 will present some sensitivity analyses to check for the robustness of different standardization methods. Finally, section 3.6 discusses the results and concludes.

3.2 Short overview of existing climate policy indices

As already mentioned in the introduction there are a lot of indicators related to environmental and/or climate issues. Surminski and Williamson (2012) singled out 33 environment or climate related indices from a very comprehensive policy index survey conducted by Bandura (2008), who identified 178 different indices over all fields (economic, political, social or environmental topics). Most of the 33 environmentally related indices assess the environmental status, for example by measuring emissions of pollutants or the degree of water pollution. The motivation for the construction of these so called outcome or impact indicators was to measure the success or failure of governments in their efforts to improve the environmental status. When it comes to the measurement of countries' strictness in environmental policies, however, outcome indicators are only helpful to a certain degree, because the emissions of GHG, for example, have many other drivers apart from the government's effort to reduce emissions. A good example for the multiple dimensions of GHG emissions is the recent shale gas revolution in the United States. The development of shale gas and its partial substitution of coal in the electricity production was definitely one important factor in bringing down total CO_2 emissions in the United States.

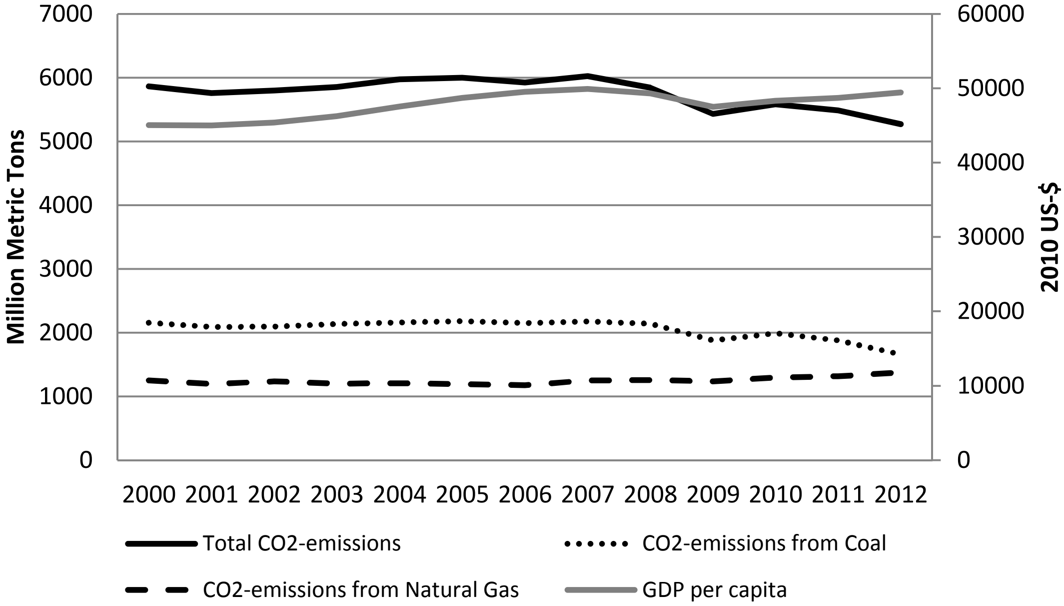


Figure 10: Emissions and GDP in the United States (source EIA (2015a) and OECD (2015a), own diagram)

Figure 10 depicts total CO_2 emissions as well as CO_2 emissions from Coal and Natural Gas which also includes shale gas. It can be seen that natural gas consumption is picking up from 2005 onwards, whereas coal consumption starts shrinking at still increasing economic activity measured by GDP per capita. At the same time total emissions start decreasing which could be due to a partial substitution of coal by gas which has lower carbon content. Another important driver of CO_2-emissions is definitely economic activity as can be seen in Figure 10 as well. Therefore the focus of indices which are supposed to measure the strictness of environmental policies shifted in recent years more and more to the attempt of measuring the policies impact directly as an input in the emission generating process. That means the strictness of climate policy is no longer measured by means of outcome, like the amount of emissions, but it is measured by the design of single policy measures which intend to curb emissions. The latter indicators will be called direct policy indicators for the remainder of this study. In the following six indices will be presented, which either consist of direct policy indicators mixed with other indicators, or solely focus on direct policy indicators.

Table 15 gives an overview of currently available indices including direct policy indicators. They are compared with respect to their geographic and temporal coverage, the actual policy measures underlying the direct policy indicator and the applied weighting and aggregation method. It is also depicted which indicator types are applied. Pure indicators solely consist of direct indicators measuring the policy variables and mixed indicators apply direct as well as outcome variables as indicators. Apart from the ISE (EBRD 2008) all indices cover a large share of the OECD countries if not all of them. The ISE has a special focus on the development of environmental policies in the transition countries and they include merely Germany, the Netherlands, Spain and the UK as benchmark countries.

Especially the **Environmental Performance Index** (EPI) (Hsu et al., 2013), the CLIMI (Steves and Teytelboym 2013) and the CCPI (Germanwatch 2013) cover a huge number of countries up to 178 including all OECD and the five BRICS countries. The indices can further be differentiated regarding the time horizons they cover: There are three indices (CCPI, CDI (Birdsall and Roodman 2003) and EPI) which are compiled on a regularly basis, meaning yearly or even twice a year (EPI). The other three indices were provided only once or three times in case of the CPI (Künkel et al. 2006). The policy types covered in the different indices vary a lot, as does the amount of policies included in the indices. All but one index contain metric as well as non-metric policies. Metric policies are, for instance, taxes or subsidies which put tangible numbers on the consumption or production of a special good which can be made comparable by means of normalization. This is not the case for non-metric policies like regulation on energy efficiency or information and education measures, for example. Therefore those indices which contain non-metric measures use scores, which are attributed to the particular policies. These scores either stem from surveys among experts (CCPI) or they are determined relatively in comparison to national goals or in comparison to the design of equivalent policies in other countries which are part of the sample. The only index which focuses on metrical measurable policies is the CPI.

Index	Coverage	Direct policy indicators	Methodology	Type
Climate Change Performance Index (CCPI)	58 countries (incl. OECD & BRICS), annually since 2004	**Policies:** promotion of renewable energies, increase of efficiency and measures to reduce GHG emissions **Measurement**: Expert ratings	**Structure:** Policy is one of three sub-indicators. **Weighting:** Fixed weighting factors based on educated guesses.	mixed
Commitment to Development index (CDI)	27 OECD countries, annually since 2003	**Policies:** fishing subsidies and regulation of illegal timber exports **Measurement:** relative scores	**Structure:** Environment is one of six sub-indicators **Weighting:** Fixed weighting factors based on educated guesses.	
Climate Laws, Institutions and Measures Index (CLIMI)	95 countries (incl. OECD & BRICS), 2011	**Policies:** Broad scope of policies divided into 4 policy areas (International cooperation, Domestic climate framework, Sectoral fiscal or regulatory measures or targets and Cross-sectoral fiscal or regulatory measures) **Measurement:** relative scores	**Structure:** Structured according the policy areas, no non-policy sub-indicators included **Weighting:** Fixed weighting factors based on educated guesses for Policy indicators and for most of the single policies except for the sectoral policies. They are weighted according to emissions affected.	pure
Climate Performance Index (CPI)	24 OECD countries, 1992, 1997 and 2005	**Policies:** 6 areas of policies (e.g. Industry, Energy Efficiency). As indicator for the climate policy in each of these areas there is at least one policy (e.g. taxes, subsidies). **Measurement:** Only metric indicators, can be measured directly	**Structure:** Structured according the policy areas, no non-policy sub-indicators included **Weighting:** The sub-indicators for the areas as well as the policy measures in the different areas are fix and equally weighted.	
Environmental Performance Index (EPI)	178 countries (incl. OECD & BRICS), biannually since 2002	**Policies:** agricultural subsidies and pesticide regulation **Measurement:** relative scores	**Structure:** Environmental policy indicators are included in the sub-sub-indicator agriculture which in turn is part of the sub-indicator ecosystem vitality. **Weighting:** Weights are set according to the assessment of the data quality of the single variables and the importance of the indicators and categories with respect to policymaking.	mixed
Index of Sustainable Energy (ISE)	33 countries (mostly former USSR and 4 European comparative countries), 2008	**Policies:** subsidies, taxes **Measurement:** relative scores	**Structure:** Index consists of three components: Energy Efficiency, Renewable Energy and Climate Change which are evaluated on basis of the three pillars institutions, market incentives and outcomes. Policy measures are included in the first two pillars. **Weighting:** The components are equally weighted as well as the policy measures inside the three pillars.	

Table 15: Methodological overview of existing climate policy indices (source: methodology reports of the indices, own illustration)

Country	CCPI	CDI	CLIMI	CPI	EPI	ISE
Australia	21	21	22	22	2	
Austria	13	11	13	11	6	
Belgium	8	8	11	16	22	
Canada	22	22	21	10	17	
Czech Republic	16	5	12	21	3	4
Denmark	1	6	6	3	10	
Finland	15	3	1	1	13	
France	5	10	2	7	19	
Germany	10	9	10	4	4	1
Greece	19	15	16	19	16	
Hungary	6	2	18	17	20	5
Ireland	7	17	9	12	14	
Italy	9	13	14	9	15	
Japan	20	20	15	13	18	
Netherlands	14	12	8	5	9	3
New Zealand	17	16	17	14	11	
Norway	12	19	5	6	8	
Portugal	2	4	19	20	12	
Spain	11	7	4	18	5	2
Sweden	3	1	7	2	7	
Switzerland	4	14	3	8	1	
United States	18	18	20	15	21	

Table 16: Rankings of the different indices applied to a subsample of countries which are covered by all indices except for ISE. The rankings stem from the most recent versions of the indices (source: own illustration).

The most prominent weighting method to aggregate all the single indicators is the use of fixed weights. The CPI and the ISE use fixed and equal weights for aggregating the single components of each indicator. The CLIMI uses fixed and equal weights as well, except for the sectoral policy indicators, which are calculated with the help of each sector's share in emissions. Another data-driven weighting method is to make use of the data quality of each policy indicator. This means to evaluate the indicators according to completeness, number of missing observations or representativeness, for example (Hsu et al. 2013). The Yale Center for Environmental Law & Policy uses data quality and expert opinion for their EPI. Germanwatch (2013) and Birdsall and Roodman (2003) rely on expert opinion only for determining the weighting factors of the CCPI and CDI.

In anticipation of section 3.4, which sets out the methodology of the CPMI, the contribution of the CPMI becomes very clear at this point. Similar to the CPI, the CPMI covers metric policy measures only, in order to be able to rely merely on figures. What distinguishes the CPMI, however, from all existing climate policy indices is the fully data-driven determination of the weighting factors and the inclusion of an effectiveness measure which is estimated empirically.

Regarding the multitude of different approaches and methods in compiling indices it is not surprising that the results of different indices vary quite substantially. Table 16 shows the country rankings for the reviewed indices applied to a subsample of 22 countries. This subsample was selected as it represents the countries which are included in all indices except in the ISE. The intersecting set of the ISE with all other indices is rather small due to ISE's focus on transition countries. Table 16 emphasizes the importance of sensitivity analysis of the outcome of an index with respect to the selection, weighting, normalization and aggregation of variables (Freudenberg 2003).

3.3 Data

The CPMI covers all OECD-countries except for Iceland, which is due to data availability. Thereby it covers nearly 40%[34] of current worldwide CO_2 emissions. The time horizon includes the years from 1991 until 2013. The year 1991 is chosen because hardly any climate policies were passed before that date, at least not with the intention to reduce GHG emissions. The *United Nations Framework Convention on Climate Change* (UNFCCC) was founded in 1992 and can be seen as the starting point of international as well as national efforts to curb anthropogenic climate change (UNFCCC 1992). The latest year for which it is possible at the moment to construct such an index due to data availability is the year 2012.

As pointed out several times already, the CPMI consists exclusively of metric measurable policy measures. Specifically, it is composed of three policy categories: Taxes on energy products (oil and coal products, natural gas and electricity), subsidies on renewable energy production and carbon taxes or emission trading systems (ETS), respectively.

The taxes on energy products stem from the *IEA Energy and Prices Statistics* database (IEA 2015b). The International Energy Agency (IEA) provides taxes[35] as well as prices on a very disaggregated basis regarding the different energy products. In total the *IEA Energy and Prices Statistics* comprise 14 different products, including multiple types of fuel oils, diesel, different grades of gasoline, LPG, natural gas, two types of coal products (steam and coking coal) and electricity. If there is more than one price or tax available per product and country, the IEA creates a representative price or tax series for that product (IEA 2012). Next to the differentiation according to the different product types, the taxes are also distinguished by the sectors in which they incur. These sectors are industry, households and energy generation. Considering the unit of the prices and taxes, the user can

[34] Calculated on the basis of emission levels from 2012 (EIA, 2015a).

[35] Subsidies on fossil fuel consumption are also included and embodied as negative tax rates.

choose between national currency (nat. cur.) per unit and nat. cur. per ton of oil equivalent (toe) which is a measure of the energy content.[36] For constructing an index, the latter unit is of course preferable as it is comparable over all energy products. The conversion into nat. cur. per toe is done by the use of country specific conversion factors (IEA 2012, pp. 11).

Unfortunately, there is not such a rich database for policies which foster the energy production from renewable sources as there is for the taxes on energy. Considering one of the most common policy measures to subsidize renewable energy production – feed-in tariffs – one would need differentiated tariffs according to the different technologies and the energy production of the respective technologies to calculate a weighted average feed-in tariff per country. So far there is no collection of such data for the OECD countries over time. Therefore policies which promote renewable energy production are measured by a proxy in the CPMI. The chosen proxy is the installed capacity of renewable energy production, or more precisely the growth of the installed capacity, which is the key figure when it comes to the measurement of investment or production subsidies. The growth of the installed capacities is a more direct measure than for example production shares of renewables, since the energy production depends on even more factors than the installed capacity, such as weather conditions or the structure of energy demand. In contrast to the energy tax indicator, which is an input measure which directly measurs the policy activity, the indicator for subsidies on renewables is a more indirect measure as it measures a first policy output. Surminski and Williamson (2012) call this type of indicator an intermediate indicator because although it measures a policy output it is not the intended final output of the policy, which would in this case be the reduction of GHG emissions. The indicator for the fostering of renewable energy sources includes all forms of renewable energy sources.[37] The data on installed capacities is taken from the IEA statistic *OECD – net Capacity of Renewables* (IEA 2015d). The capacity is measured in MW electricity equivalents (MWe) as the data includes plants which produce electricity as well as heat.

The third policy category consists of carbon pricing. This includes both carbon taxes and ETSs. In the OECD countries there are currently 7 different trading systems in force (see Table 20 in the Appendix for an overview of all carbon pricing mechanisms in OECD countries). An eighth ETS, the Carbon Pricing Mechanism (EPM) of Australia, is also covered by the CPMI. But this one has already been abolished after being in force for only two years (2012/2013). In 28 of the 33 countries considered, there are (or were) trading schemes in place. The CO_2 emissions covered by the 8 ETSs amount to 6.12 billion tons

---

[36] 1 toe corresponds to an energy content of 10^7 kcal.

[37] The renewable energy sources included are: Hydro, Geothermal, Solar, Tide/Wave/Ocean, Wind, Waste and Biofuels.

of CO_2. That makes up approximately 17% of global CO_2 emissions.[38] A carbon tax is applied in 12 of the countries included in the CPMI. Apart from Mexico, all carbon tax countries also make use of an ETS, with Canada's ETS and carbon tax affecting two different regions (Québec and British Columbia). Further ETSs which apply to parts of the country only are the Japanese (Tokyo and Saitama) and the US (California and Northeast) systems.

3.4 Theoretical Framework and Methodology

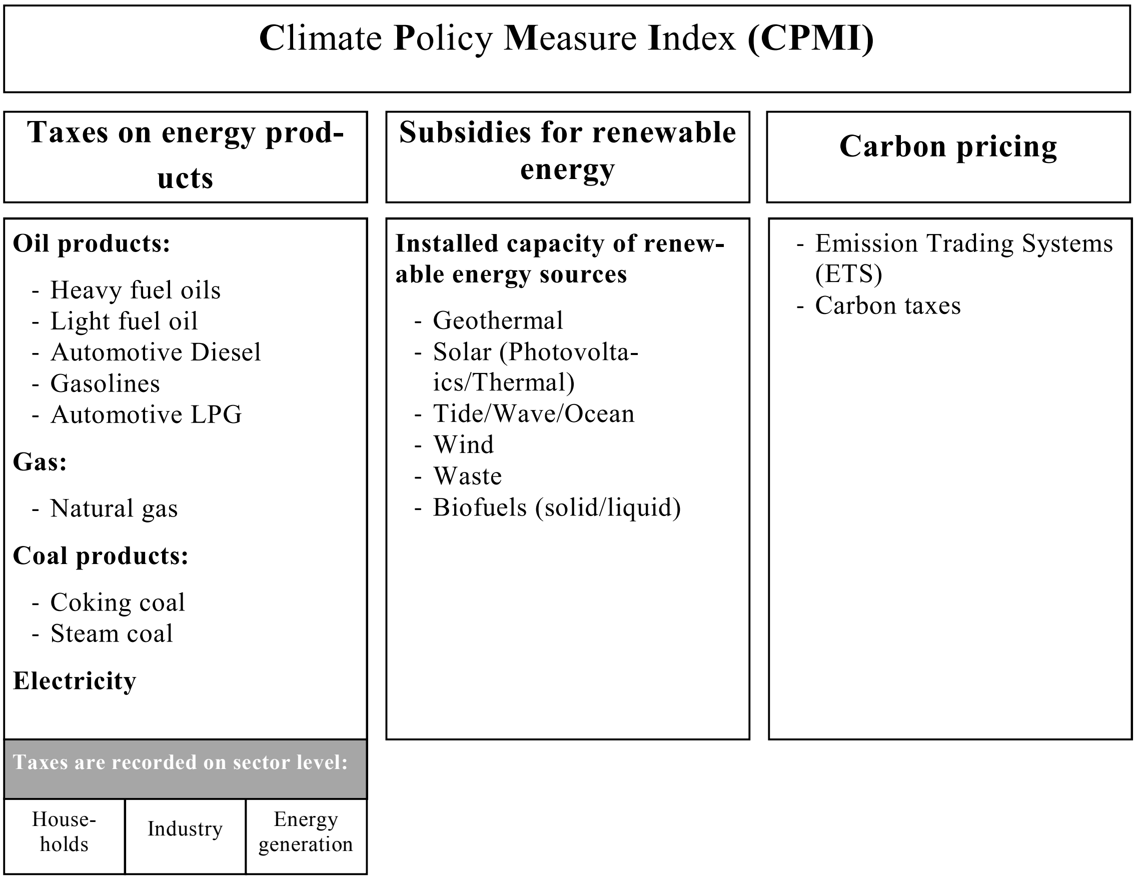


Figure 11: Structural overview of the CPMI (source: own illustration)

Following the classification by Freudenberg, the CPMI is a *Composite indicator* consisting of three *thematic indicators*, the three policy categories: taxes on energy products, subsidies on renewable energy sources and carbon pricing. They in turn consist of different *individual indicator sets* (Freudenberg 2003, p. 7). Figure 11 gives an overview of the structure of the CPMI. This section explains the structure in a bottom up style, meaning that the composition of the three policy categories with their individual policy indicators are explained first, and in a second step the aggregation to one single indicator is addressed.

[38] Calculation is made for the year 2013.

3.4.1 Taxes on Energy Products

The energy taxes category covers taxes on all fossil fuels (oil products, natural gas and coal products) and on electricity. In 2010 the combustion of fossil fuels accounted for 93% of all CO_2 emissions in the OECD countries.[39] As already pointed out in the data section, the tax data (IEA 2015b) is available on a highly disaggregated product level. Some of the products listed in Figure 11 consist themselves of several sub categories of products like Gasolines, for example. Figure 11 also illustrates the differentiation of the taxes with respect to the three sectors: households, industry and the energy generating sector. In total there are 23 different tax combinations between products and sectors, as not every sector uses all products. One drawback of the *Energy Prices and Taxes* statistic is the not negligible amount of missing values. Especially problematic for the construction of an index over time are missing values within a time series since this would be equivalent to the abolishment of a tax which is a rather rare event (cf. Bauer 2006). Furthermore the database indicates the reason for missing values. It states if a missing value is not observed due to the fact that it simply was not collected or due to the fact that the data point cannot exist. In the first case, which is the case of a missing value in its original sense, a data imputation method is applied in this study which perpetuates the tax in a constant way. This extrapolation is used until new information is available or until the end of the respective time series. This prevents the energy tax indicator from unrealistic jumps and by using constant extrapolation there is no artificial variation introduced to the data, which could harm the empirical analysis of the tax indicator. On average there were 2.5 data imputations necessary per country and tax time series which comprise of 36 years each, since the tax data is available for a longer time period[40] than the renewable energy capacity data, for example. In the second case, the case in which there is no tax, the observation is set to zero.

For transforming the taxes of all 23 different product-sector combinations into one single indicator it is necessary to weigh and aggregate the single taxes. The weighting inside this sub-indicator is done by using the respective year-specific consumption. For this purpose again, data from the International Energy Agency (IEA) is used[41]. This guarantees consistency with regard to the data collection and classification method. One could suspect that, using consumption shares as weighting factors leads to a certain cannibalization especially of effective energy taxes in the CPMI, as an increase of those would lead to a decrease in its share. But this is only the case if there is a high degree of between product

[39] The remaining 7% stem from fugitive emissions and industrial processes. Source: own calculation on the basis of IEA (2015e and 2015f).

[40] The *Energy Prices and Taxes* statistic (IEA 2015b) provides data for the period from 1978 until 2013.

[41] Consumption data sources for the different product types: Oil products (IEA 2015g), coal products (IEA 2015h), natural gas (IEA 2015i) and electricity (IEA 2015j).

leakage, meaning that the consumption of the highly taxed product is substituted to a large extent by another product; only then the denominator (i.e. total consumption) would stay constant. For the case in which the increased tax leads to a true reduction in the consumption of fossil fuels, the share more or less stays the same. For the leakage case, on the other hand, it is sensible under policy effectiveness considerations that the share decreases.

Since the consumption data is provided in product specific units like liters in case of gasolines or tons in case of heavy fuel oils, it has to be standardized before calculating the sector and product specific weights. This is done by using product and country specific conversion factors[42] which allow converting the product consumption units into toe. The weighting based on the consumption allows the energy tax indicator to be interpreted as the average tax paid on energy products per toe in the three observed sectors.

3.4.2 Subsidies for Renewable Energy

As explained in the data section, subsidies on renewable energy sources are represented by the sum of the installed capacity of renewable energy technologies producing either electricity or heat. To make both types of energy comparable the capacity is measured in MWe. Since the main interest of the CPMI lies in the measurement of the stringency of countries' climate policies, the annual growth of the installed capacity is used to represent subsidies on renewable energy technologies. If alternatively percentage changes were used, countries which start from very low levels would have huge percentage changes in the first years. Therefore absolute changes are applied. Since a 10 MWe expansion in the United States, however, is to evaluate differently from a 10 MWe expansion in Luxembourg the absolute changes are corrected by country size by using the absolute growth in installed capacity per capita.

3.4.3 Carbon Pricing

The carbon pricing indicator is composed of carbon prices achieved by ETSs and of carbon prices set by a carbon tax. For the cross-country comparison with regard to carbon prices two things are important next to the carbon price, provided that there is a carbon pricing mechanism. First, the coverage of the pricing mechanism, meaning how many CO_2 emissions are governed by the respective mechanism, and second, the number of pricing mechanisms as in some countries there is an ETS as well as a carbon tax in force (see Table 20 in Appendix). To generate an average carbon price for each country, the single carbon price indicators, similarly to the energy tax indicator, have to be weighted. First of all this is done by the share of emissions covered by the respective mechanism relative to the total GHG emissions of each country. Data on total emissions is taken from the IEA

[42] The conversion factors are taken from the IEA report "Energy Prices and Taxes: Methodology Notes" (IEA, 2012).

(2015e). Both are measured in CO_2 equivalents. In those 11 countries in which both instruments are applied simultaneously the average of the weighted carbon prices is taken. This excludes double charging of companies, meaning it is assumed that taxpayers either pay carbon taxes or buy emission certificates but not both. This simplifying assumption has to be made due to missing information about the potential overlap of the systems. Anecdotal evidence shows that in many countries considered, taxpayers are governed by one or the other system and not by both. The emissions covered by the ETSs are provided by the status reports of the *International Carbon Action Partnership* (ICAP 2015). For the European ETS (EU ETS) country specific data is available at the European Environment Agency (EEA 2015). The information about the carbon tax coverage stems from the World Bank (2014).

3.4.4 Aggregation of the Sub-indicators

To aggregate the three sub-indicators to one single index, the sub-indicators have to be standardized and weighted. Standardizing is necessary because of the different units and scales shown by the sub-indicators. Aggregating the taxes on energy products and carbon pricing mechanisms could easily be done since both indicators represent a cost per unit consumed, which could be converted in a way that they have equal basis. Transforming the subsidy indicator into a cost indicator, however, is not so trivial. Therefore the indicators are standardized before they are weighted. The standardizing method chosen is the most commonly applied method, namely the standard deviation from the mean as it has several advantages compared to other methods. It puts all sub-indicators on a common scale which is normally distributed with zero mean and standard deviation equaling one. The zero mean prevents the CPMI from being dependent on the different scales of the sub-indicators (Freudenberg 2003) and makes variables with completely different units comparable. The scale of the subsidy indicator, for example, is by far larger than those of the other two sub-indicators and also its unit is MWe compared to dollars per unit consumed. As the CPMI allows comparison of the policy performance between countries at a certain point in time and comparison over time for individual countries, the standardization has to be done in two dimensions. For the cross-country comparison, each sub-indicator is standardized per year over all countries. This means that each country's performance is assessed in each of the three policy categories in relation to those of the other countries in a specific year. To also account for differences in the purchasing power in the different countries, the originally national currencies are converted to U.S. dollars using Power Purchasing Parities (PPP). Thereby a 10$ tax in Chile can be compared with a 10$ tax in Denmark. For the within-country comparison over time, the three policy categories are standardized using each sub-indicator's time series of the respective country to depict the performance of a certain year relative to the performance of other years.

Not only the weighting inside the sub-indicators, but also the weighting between the sub-indicators of the CPMI is data driven, which is one of the major contributions to the climate policy literature. Many composite indicators use equal weights or educated guesses about the importance of the single components. For the CPMI the impact assessment for the three sub-indicators is done empirically by estimating elasticities. The theoretical background for the estimation equation is the Kaya identity which is oftentimes the foundation for the emission scenarios in Integrated Assessment Models (IAMs) (IPCC 2000). The Kaya identity decomposes the anthropogenic CO_2 emissions into four factors (Kaya 1990):

$$\underset{emissions}{F} = \underset{population}{P} \quad \underset{GDP\ per\ cap.}{\frac{Y}{P}} \quad \underset{\substack{energy\ intensity\\ of\ GDP}}{\frac{E}{Y}} \quad \underset{\substack{emission\ intensity\\ of\ energy}}{\frac{F}{E}} \tag{1}$$

These four factors constitute at the same time the leverages to reduce GHG emissions whereby the first two factors, population size and GDP per capita, are usually not part of the choice set of politicians while trying to reduce emissions. The factors left, which can be addressed by climate policy, are the last two: Energy intensity per GDP and emission intensity of energy production. Since climate policies impact both intensity components simultaneously, energy intensity of GDP and emission intensity of energy production are combined to emission intensity of GDP. This modified version of the Kaya identity constitutes the basis for the estimation equation which shall produce the weights of the three policy indicators. The empirical strategy follows a two-step approach in which the impact of the policies on the emission intensity of GDP is estimated in a first step, and in a second step the modified Kaya identity is estimated with an instrumented emission intensity of GDP.

First stage:

$$\frac{F}{Y}_{it} = \beta_0 + \beta_1 Energy\ tax_{it} + \beta_2 Renewable\ subsidy_{it} + \beta_3 Carbon\ pricing_{it}$$

Second stage:

$$F_{it} = \gamma_0 + \gamma_1 P_{it} + \gamma_2 \frac{Y}{P}_{it} + \gamma_3 \widehat{\left(\frac{F}{Y}\right)}_{it} \tag{2}$$

The regression model described in equation 2 is estimated by a two-stage least squares panel-data model using country-fixed effects in a log-log specification. The panel structure of the data is clarified by the sub-indices i and t, standing for the country and the year respectively. The log-log specification provides the nice property that the coefficients can directly be interpreted as elasticities. In the first stage, the individual impact of the three policy categories on emission intensity is estimated. The emission intensity is taken from an EIA database (EIA 2015b). In the second stage this estimated emission intensity[43] is regressed on the CO_2 emissions (IEA 2015e). It should be noted that this two-step regression is not used as an instrumental variable approach but the two-step structure should account for the channel through which policy measures can impact emissions in the framework of the Kaya identity, namely through the channel of emission intensity of GDP. The main goal of this exercise, however, is not to explain the determinants and their impact on CO_2 emissions, but to obtain weights for the three policy categories of the CPMI. Therefore the focus lies on the regression results of the first stage. For completeness Table 17, however shows the regression results of both stages.

	Emission intensity of GDP (1st stage)	Emissions (2nd stage)
Emission intensity of GDP		0.829**
		(0.096)
Emissions (t-1)	0.659**	0.163*
	(0.031)	(0.071)
Population	-0.519**	0.654**
	(0.081)	(0.077)
GDP per cap.	-0.795**	0.763**
	(0.031)	(0.089)
Taxes on Energy	-0.015**	
	(0.006)	
Subsidies on Renewables	-0.022**	
	(0.008)	
Carbon Pricing	-0.004**	
	(0.001)	
R^2	0.84	0.88
N	627	627

* $p<0.05$; ** $p<0.01$

Table 17: Regression results for the weighting factors of the three policy categories

In addition to the regression model depicted in equation 2, the final regression includes a lagged term of emissions to account for the dynamics in the development of CO_2 emissions. Although the results of the second stage are not of major significance in this study, they can serve as a plausibility check of the regression model. All factors of the second stage are statistically significant and show the expected signs. It is not surprising that the estimated emission intensity of GDP has a clearly positive impact on the emission level. The same is true for the emissions from the year before which makes clear the trending behavior. It is also not surprising that larger and richer countries emit more. Similarly to the emission level, also the emission intensity of GDP depends positively on the level on

[43] The estimated variable is labeled with a hat in equation 2.

the emissions from the past period. The negative effect of country size in terms of population size cannot be interpreted causally, as it is highly correlated with the lagged emissions (correlation coefficient: 0.93). Thereby the negative population coefficient represents a smaller overall effect of past emissions rather than an impact of population on emission intensity. This can be shown by leaving population out of the regression, which leads to a smaller impact of the past emissions on the emission intensity. The coefficient for GDP per capita, however, does measure the impact of GDP per capita as its correlation with the past emissions is relatively small (correlation coefficient: 0.018). The elasticity of GDP per capita shows the negative relation between income and emission intensity. The most important result of the regression analysis, however, is the measurement of the climate policies' impact on the emission intensity of GDP.

The first thing to be noticed is that the magnitude of the elasticities is relatively small. A tax increase on energy products by 10%, for example, leads on average only to a reduction of the emission intensity by 0.15%. This result is quite sensible as the demand for energy products is rather inelastic. Typically taxes are more effective in the long run, but since the regression model is a static model only year over year elasticities are estimated. Subsidies on renewable energy, here represented by growth of installed capacities of renewable energy plants, have a slightly larger elasticity in magnitude than taxes on energy consumption have. Nevertheless it is still small. This is surprising at first sight, as at least the emissions from the electricity and heat production are impacted rather immediately by more power plants using renewable energy sources. But two facts have to be kept in mind. First, the so called rebound effect, which states that e.g. higher fuel efficiency reduces consumption, which in turn leads to lower fuel prices, and as a result demand is increased again. This goes back to Jevons' Paradox (Jevons 1865). A second explanation might be that capacity is not automatically equal to power submitted to the network. The capacity installed has to be multiplied by the region- and technology-specific capacity factor, which for most renewable energy technologies is far from being equal to one. Even smaller than the effect of renewable energy capacity is the effect of carbon pricing mechanisms. This again is not so surprising looking at the design and the current prices of the ETSs. Applying the estimated elasticity of the carbon pricing mechanism to the EU ETS, for example, illustrates why the elasticity is so small. The current EU ETS price is around 8 US-Dollars per ton of CO_2. Assuming now an increase by 375%, which would bring the EU ETS price to 30 US-Dollars per ton of CO_2, the resulting emission intensity reduction would be 1.5%.

The weighting factors for each of the policy categories are simply calculated by the share of the respective elasticity of the sum of the three elasticities. As the elasticities are estimated using a panel approach, it is implicitly assumed that all policies are equally effective in all countries of the sample and also equal effective over time. This leads to weighting factors of around 38% for the taxes on energy products, 53% for the subsidies on renewable energies and 9% for the carbon pricing mechanisms.

In case of missing observations for one of the three sub-indicators, may it be due to missing data or due to the fact that a country does not apply policies of a certain policy category in a certain year, the respective sub-indicator is treated as it would have taken the value zero, which is equal to the mean of the distribution of the respective indicator. Given the fact that there are quite a few countries, especially in the first years of the observation period, which do not always apply policies from all three policy categories this is the most consistent way to treat missing values. Also with respect of the fairness of the index it seems sensible, as countries should have the possibility to choose which policies to apply and they can always make up the absence of one policy type by acting more stringent in the others.

The theoretical framework of the CPMI can be summarized by categorizing its components. As policy measures, the CPMI uses input variables in case of the taxes on energy products and in case of the carbon pricing mechanisms. The policies to promote renewable energy production are measured by an intermediate indicator, the growth in renewable energy capacity. For the weighting inside the sub-indicators, for energy taxes and carbon prices the coverage of the respective policy is used, which is also an intermediate indicator. The weighting of the sub-indicators is done using an impact indicator.

3.5 Results

For the interpretation of the results it is worth recalling the three main ingredients of the index: taxes on the consumption of fossil energy products, the annual growth of the installed capacities of renewable power plants as proxy for the subsidies on renewable energy and carbon price mechanisms. Considering these three factors, it becomes clear that the CPMI exclusively measures the political performance of governments to curb the emissions of GHGs and explicitly not the country's performance with regard to environmental friendliness in general. The latter issue is covered already by many environmental related indices. Prominent examples are the EPI (Hsu et al., 2013) and the CCPI (Germanwatch, 2013). In these more general indices, policy is only one of many sub-indicators. The other sub-indicators usually describe the current status of the environment and the consequences of anthropogenic intervention, like fossil fuel or land use. That means these indicators mix input and output variables to measure environmental performance. The CPMI, in contrast, focuses on the assessment of current policy's stringency, which is the input for the future development of country's environmental performance. This focus is important for policy evaluation since it abstracts from historical or physiographic reasons for countries being in a certain state of environmental friendliness. That implies that countries which already run their economies at very low levels of emission intensity and do not have to apply very strict policies to stay on this level will perform worse than countries which currently consume large amounts of fossil fuels per GDP but at the same time adopt strict policies. This is especially the case for the indicator of subsidies on renewable energy production. A country like Norway, for example, has very limited scope to further develop the production of electricity and heat from renewables as the Norwegians already cover 99%[44] of their heat and electricity generation by hydro energy. Norway is a good example to visualize the policy focus of the CPMI because the large share of energy production from renewables is clearly no consequence of a super strict environmental related policy. It is rather a consequence of the physiographic preconditions, which should not be included when it comes to policy assessment.

In the remainder of this section, the rankings produced by the CPMI are presented: First the cross-country rankings and their development over time and then the development of climate policy stringency within the 33 OECD countries. Subsequently the changes due to robustness checks are presented.

3.5.1 Comparison of Climate Policy Stringency between the OECD Countries

As explained in section 3.4.4, the composite index is calculated from the weighted sum of the standardized sub-indicators. The applied standardization method leads to zero mean and a standard deviation equal to one. With regard to this distributional property the composite indicator is to be interpreted as follows: The values of the composite indicator can, thanks to the standardization, be interpreted cardinally. Negative values indicate a below average stringency of climate policy and positive values an above-average stringency.

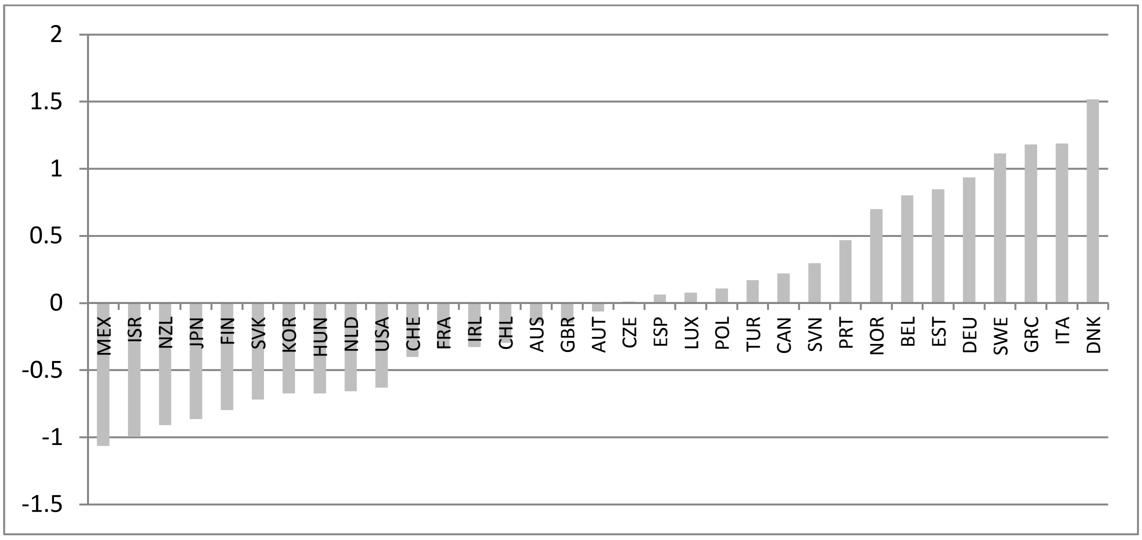


Figure 12: Cross-country comparison of the CPMI for the year 2012. (source: own illustration)

---

[44] IEA (2015c).

Figure 12 shows the result of the cross-country comparison of the CPMI. To get a feeling for the size of the index values it should be noted that using a standard normal distribution, the probability of observing larger values than the standard deviation in absolute terms amounts to around 16%, meaning that 68% of the observations lie between -1 and 1. The probability of values larger than 1.5, like it is the case for Denmark, is 6.5%. Denmark is leading the 2012 CPMI which is mainly due to its high index value for the sub-indicator on renewable energy subsidies as Table 18 shows.

Table 18 depicts the index values of all sub-indicators as well as the values of the composite indicator for the 5 leading and for the 5 lowest performing countries of the 2012 index. It illustrates nicely which sub-indicators are driving the composite indicator. In the first line of the table the percentage numbers in the brackets represent the weights of the respective policy category. Although renewable energy subsidies have the largest weighting factor, they are not dominating the overall result in all countries. In the case of Greece, for example, the taxes on energy consumption are outweighing a below average carbon price and a moderate index value for the promotion policies of renewables and let Greece end up on rank No. 3 in 2012. The last three ranks are occupied by countries which neither had a carbon tax nor an ETS in 2012. But this alone is not the reason for them being at the end of the list. A non-existent policy category is, as already mentioned in section 3.4.4, treated as if the performance in the missing policy category were equal to the mean performance of all countries.

#	Country	Composite Index	Taxes on energy products (38%)	Subsidies for renewable energy (53%)	Carbon pricing (9%)
1	**DNK**	1.516503	1.319203	1.947698	-0.179337
2	**ITA**	1.187681	1.922957	0.9507561	-0.4820045
3	**GRC**	1.181562	1.903111	0.8566154	0.0821725
4	**SWE**	1.115045	0.1823655	2.099523	-0.7556109
5	**DEU**	0.9351041	0.4165622	1.513126	-0.2838535
…	**…**	…	…	…	…
29	**FIN**	-0.7963237	-0.5202445	-1.012222	-0.6829728
30	**JPN**	-0.8630255	-1.614383	-0.7381766	1.526999
31	**NZL**	-0.9086341	-0.7575453	-1.171958	
32	**ISR**	-0.98918	-1.212378	-0.9999679	
33	**MEX**	-1.064404	-1.394248	-1.012201	

Table 18: Decomposition of the CPMI for selected countries in 2012

Another interesting aspect is the development of the cross-country rankings over time. Although the index values, which are standardized on the cross-country dimension, are not comparable over time, it is still informative to compare how the rankings change over time. Figure 15 and Figure 16 in the appendix depict the results from the cross-country comparison of the CPMI for the years 1995 and 2000. The order of the countries follows the order according to the 2012 ranking to allow for an easier comparison. The country which is depicted on the far right in both figures is the country which performs best in 2012. Comparing the three cross sections, one recognizes that there are seven countries which in all three years shown applied above average stringent climate policies. These countries are Denmark, Italy, Greece, Germany, Norway, Portugal and Spain. On the other end of the list there are nine countries[45] which in the three cross-country comparisons for 1995, 2000 and 2012 always lie below the average.

3.5.2 Climate Policy Stringency over Time

This sub-section focusses on the development of the stringency for each country over time. It should be noticed that these index figures are calculated differently from those in the cross-country comparison in the previous sub-section. They differ with respect to the standardization. The standardization for the index describing the development in a specific country over time is done using the country-specific time series of each sub-indicator: The mean and the standard deviation are calculated e.g. for the weighted energy tax of Australia over the years from 1991 until 2012. Using this approach, the index values are not comparable between countries, only within countries over time. Here, values below zero indicate that the country was less ambitious regarding climate policy in that year than it has been on average over the entire observation period. As an example Figure 13 shows the development of the climate policy stringency of a country which compared to other countries always performed above average (Germany), and the policy performance development of a country which performed below average (USA).

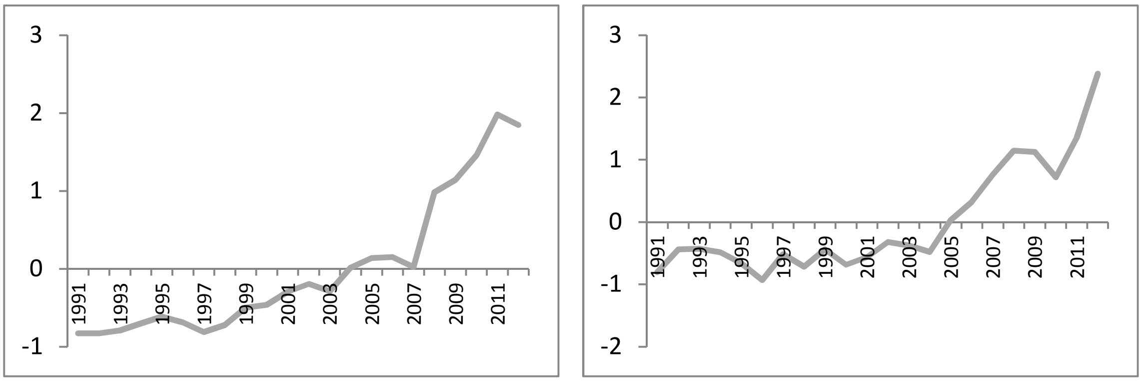


Figure 13: Development of the climate policy stringency over time in Germany (left) and the United States (right). (source: own illustration)

In both countries an upward sloping trend becomes apparent, indicating an increase in the stringency of climate policies in the observation period. In Germany the upward sloping

[45] Australia, Chile, Hungary, Ireland, Israel, Netherlands, Japan, Slovakia, and USA

trend started in 1997, which was the year of the adoption of the Kyoto Protocol. In the United States the upturn started a few years later: Until 2004, the CPMI curve is flat before it picks up in 2005. To also illustrate the development of the sub-indicators and their impact on changes in the composite index, Table 19 shows the development of the three sub indicators and the composite indicator for Germany and the United States over time. To facilitate readability Table 19 only includes selected years, which reflect turning points in the long-term development of the climate policy stringency of the two countries.

Country	Year	Composite Index	Taxes on energy products (38%)	Subsidies for renewable energy (53%)	Carbon pricing (9%)
Germany	**1991**	-0.8285123	-0.9308372	-0.8977651	
	1995	-0.6168438	-0.4803099	-0.8197761	
	2000	-0.4644463	-0.7292134	-0.355675	
	2004	0.0143099	-0.1261885	0.1167442	
	2005	0.1381289	0.0406204	0.2240553	0.0416831
	2006	0.1499256	0.2049598	0.1088615	0.1610118
	2007	0.0232697	0.5127357	-0.030784	-1.692772
	2012	1.845303	2.045779	2.143751	-0.7271517
USA	**1991**	-0.8034658	-1.106913	-0.7253025	
	1995	-0.6580722	-0.7324566	-0.7179818	
	2000	-0.6859618	-0.5138606	-0.9260562	
	2004	-0.4814844	-0.1317606	-0.8129148	
	2005	0.0361335	0.2220805	-0.0899931	
	2007	0.7607087	1.076566	0.6663826	
	2008	1.147765	1.656333	1.098572	-0.6758973
	2009	1.124437	0.8537427	1.637027	-0.7414953
	2010	0.7213451	1.200025	0.6358469	-0.7665585
	2012	2.379929	1.870009	2.98118	0.9897151

Table 19: Decomposition of the time series CPMI for Germany and the United States for selected years. (source: own illustration)

The first noticeable point in the German development of the CPMI is in 2004, when the indicator for subsidies on renewable energy lets the CPMI turn above average for the first time. The kink in the CPMI curve in 2007 has two reasons. First it was the last year of the first period of the EU ETS and the certificate price nearly dropped to zero. The second reason is a decrease in the expansion of renewable energy capacity. The latter might have been caused by investment uncertainties about future subsidies induced by discussions about an EEG novel in 2008. In the United States, the three indicators develop more congruent to each other than in Germany. The drop in 2010 is mainly caused by the introduction of the Regional Greenhouse Gas Initiative (RGGI) in nine northeastern states[46] in 2008, which started off with rather low prices of around 2$ per ton of CO_2. Another reason is the cutback in the expansion of renewable energy capacity. Germany and the United States are of course only a small fraction of the entire sample of OECD countries but they reflect nicely two general patterns which all countries in the sample exhibit. These are the upward sloping trend in the climate policy stringency and the fact that the most volatile measure in the CPMI is the sub-indicator on renewable energy subsidies measured by the annual growth of installed capacities of renewable energy. To check the dependence of the CPMI on the weighting factors of the three policy categories, the next sub-section presents some robustness checks.

3.5.3 Sensitivity Analyses

It should be stressed once again that the weighting factors are calculated based on estimates of the impact of the respective policy on the CO_2 emissions. Thereby the weighting factors of the CPMI are not subject to e.g. expert opinions on the importance of the individual policy categories, which is the case for many climate policy indices (see Table 15). However, not only weighting factors based on expert opinions, but also empirically founded weighting factors have to deal with uncertainty. Therefore it is necessary to conduct some robustness checks with respect to the weighting factors. The sensitivity analyses are illustrated using the most recent cross country results of the CPMI[47], since the weighting factors play a bigger role for the cross country comparison. The policy development of single countries is not so heavily influenced by the weighting factors as the overall trend of increasing stringency also applies to the development of the three sub-indicators. But for the comparison of countries the weighting of the three policy indicators has a larger impact as countries differ with regard to their policy focus. For instance, those countries which mainly focus on energy consumption taxes and not so much on subsidies for renewables are disadvantaged by the baseline weighting factors. In the baseline version of the CPMI, taxes on energy products have a share of 38%, subsidies on renewable energy of 53% and carbon pricing mechanisms of about 9%.

The first robustness check is to decrease the weighting factor of subsidies. This is done by using the lower bound[48] of the 95% confidence interval of the estimate for renewable subsidies. A decrease of the subsidy weighting factor of course simultaneously increases the

---

[46] Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New York, Rhode Island, and Vermont.

[47] The most recent result of the CPMI applies to the year 2012.

[48] Lower bound in absolute terms as the elasticities of the three policy indicators are negative.

shares of the other two policy categories. The modified shares look as follows: energy taxes 61%, renewable subsidies 24% and carbon pricing mechanisms 15%. While this weighting factor modification leads to a substantial shift from subsidies to energy taxes, the share of the carbon pricing mechanisms is still relatively small. To change that, a second robustness check is conducted by using the largest elasticity of the 95% confidence interval of the estimate for the carbon price impact on CO_2 intensity of GDP. The resulting weights are 56% for energy taxes, 21% for subsidies on renewables and 23% for the carbon pricing mechanisms.

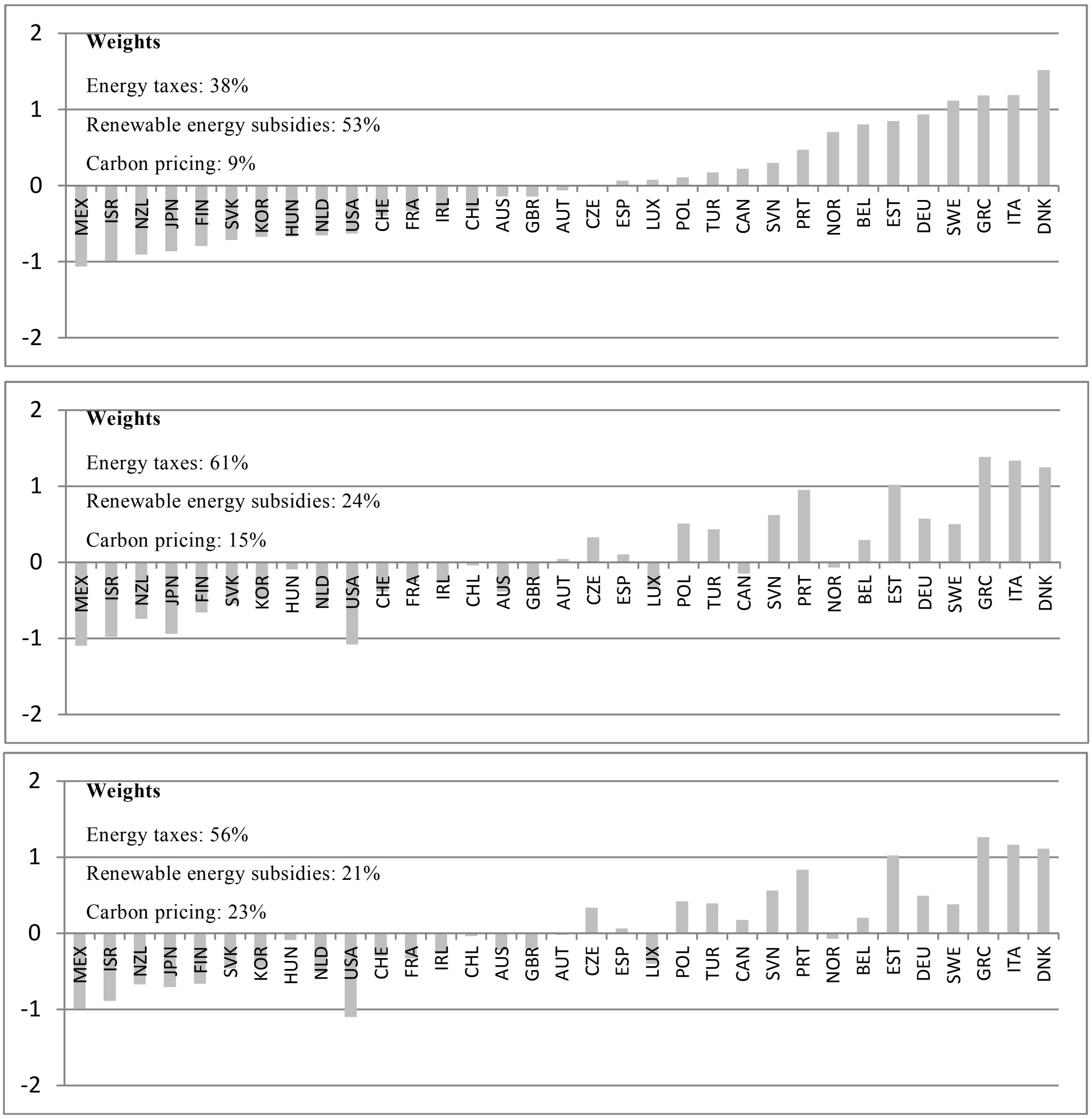


Figure 14: Results of the sensitivity analysis with respect to the weighting factor. The first chart depicts the reference CPMI. (source: own illustration)

Figure 14 depicts the results of these two robustness checks in relation to the baseline CPMI for the year 2012, which is shown in the first of the three charts. Again the order of the countries is chosen according to the 2012 baseline country ranking.

The best performer of the 2012 baseline CPMI is depicted on the far right. As expected the country ranking does change with changing weighting factors, but the overall picture of those countries whose climate policy stringency is above average and those whose stringency is below average turns out to be rather robust to changes in the weighting factors. This is because these countries have stringent policies across all three categories so that the change in the weights does not matter for them. The only exceptions are Austria, Canada, Luxembourg and Norway. These countries either switch from below average performance to above average performance (Austria) or vice versa (Canada, Luxembourg and Norway).

3.6 Conclusion

The CPMI measures the stringency of climate policy in OECD countries.[49] Stringency means the levels of the respective policies weighted by their effectiveness in reducing the emission intensity of production. It allows cross country comparisons per year as well as within country comparisons over time. The time period covered goes from 1991 until 2012. It is thereby the CPMI which provides index values for the longest time series compared to other climate policy indices. Also, it covers the most important period for climate policy as there has been hardly any policy measure before 1991 which was intended to reduce CO_2 emissions. The reason for 2012 being the current end of the observation period is data availability.

The major contribution of the CPMI, however, is that it is entirely data driven. This approach starts with the selection of variables as only quantitative policy measures are included. Thereby the CPMI does not have to rely on expert opinions or other indices. Also the construction of the composite index is purely data driven. The single policy variables are weighted according to each policy's scope. Taxes on fossil fuel consumption for example are weighted according to the consumption of the respective product. The aggregation of the sub-indicators is done using weighting factors from an empirical assessment of the impact of the respective policy measures on the reduction of CO_2 emissions. This data based construction guarantees objectivity of the index and enhances the transparency of the index's composition. Both are important for the informative purpose of the index as well as for the applicability in empirical analyses and in the calibration of Computable General Equilibrium (CGE) models.

---

[49] Except for Iceland which is missing due to data availability.

The CPMI consists of three sub-indicators which each contain several policy indicators: First, the indicator on taxes on energy products, second the promotion-indicator on renewable energy and third the indicator on carbon pricing mechanisms.

Next to the ranking of the countries according to their political effort to curb carbon emissions, the CPMI provides also two more general results. The within country comparison over time shows that in all countries observed there is a trend of increasing stringency of climate related policies. A second interesting finding stems from the cross-country comparison. There are seven countries (Denmark, Italy, Greece, Germany, Norway, Portugal and Spain) which perform consistently above average considering their effort in climate policies.

This study and the index resulting from this study should be understood as the first step into developing an index on the stringency of countries' climate policy. There is still scope for further modifications of the database as well as with regard to the methodology. The database would benefit from being extended regarding the policy measures as well as the countries. It would be interesting to include also the BRICS countries to relate the policy performances from the OECD countries to those from emerging economies.

3.7 References

Badura, Romina (2008), “A Survey of Composite Indices Measuring Country Performance: 2008 Update”, *UNDP/ODS Working Paper*, February 2008.

Bauer, Michael W. (2006), “Politikbeendigung als policyanalytisches Konzept“, *Politische Viertelejahresschrift* 47.

Birdsall, Nancy and David Roodman (2003), “The Commitment to Development Index: A Scorecard of Rich-Country Policies”, Center for Global Development.

EBRD (2008), “Securing sustainable energy in transition economies”, April 2008.

EEA (2015), “European Union Emissions Trading System (EU ETS) data from EUTL” (database), European Environment Agency. (Accessed on 21 November 2015)

EIA (2015a), “CO_2 Emissions”, International Energy Statistics (database). (Accessed on 17 November 2015)

EIA (2015b), “Carbon Intensity”, *International Energy Statistics* (database). (Accessed on 19 November 2015)

Freudenberg, Michael (2003), “Composite Indicators of Country Performance: A Critical Assessment”, *OECD Science*, *Technology and Industry Working Papers*, 2003/16, OECD Publishing.
DOI: http://dx.doi.org/10.1787/405566708255

Germanwatch (2013), “The Climate Change Performance Index. Background and Methodology”, November 2013.

Holzinger, Franziska (2006), “Methodological Pitfalls of Convergence Analysis”, *European Union Politics*, Vol. 7(2).

Hsu, A., L.A. Johnson, and A. Lloyd (2013), “Measuring Progress: A Practical Guide From the Developers of the Environmental Performance Index (EPI)”, New Haven: Yale Center for Environmental Law & Policy.

ICAP (2015), “Emissions Trading Worldwide”, *International Carbon Action Partnership (ICAP) Status Report 2015*.

IEA (2012), “Energy Prices and Taxes: Methodology Notes”, IEA 1st Quarter 2012.

IEA (2015a), “Addressing Climate Change”, *Policies and Measures Database* (database). (Accessed on 11 November 2015)

IEA (2015b), “End-use prices: Energy prices in national currency per toe”, *IEA Energy Prices and Taxes Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00440-en (Accessed on 06 October 2015)

IEA (2015c), “OECD - Electricity and heat generation”, *IEA Electricity Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00457-en (Accessed on 21 November 2015)

IEA (2015d), “OECD – Net capacity of renewables”, *IEA Renewables Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00467-en (Accessed on 20 November 2015)

IEA (2015e), “Emissions of CO2, CH4, N2O, HFCs, PFCs and SF6”, *IEA CO2 Emissions from Fuel Combustion Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00431-en (Accessed on 23 November 2015)

IEA (2015f), “CO2 emissions by product and flow”, *IEA CO2 Emissions from Fuel Combustion Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00430-en (Accessed on 23 November 2015)

IEA (2015g), “OECD product supply and consumption”, *IEA Oil Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00478-en (Accessed on 06 October 2015)

IEA (2015h), “OECD – Coal statistics”, *IEA Coal Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00554-en (Accessed on 13 October 2015)

IEA (2015i), “OECD - Natural gas supply and consumption”, *IEA Natural Gas Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00481-en (Accessed on 13 October 2015)

IEA (2015j), “OECD - Electricity/heat supply and consumption”, *IEA Electricity Information Statistics* (database).
DOI: http://dx.doi.org/10.1787/data-00462-en (Accessed on 06 October 2015)

IPCC (200), “Emission Scenarios”, *Special Report*, Cambridge University Press, UK.

Jevons, William Stanley (1865/1965), “The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-mines”, In: *Flux*, A.W. (Ed.), 3rd edition 1905, Augustus M. Kelley, New York.

Kaya, Y. (1990). “Impact of carbon dioxide emission control on GNP growth: Interpretation of proposed scenarios”, In: *IPCC Energy and Industry Subgroup, Response Strategies Workshop*, Paris, France.

Künkel, Nana, Klaus Jacob and Per-Olof Busch (2006), “Climate policies – (the feasibility of) a statistical analysis of their determinants”, Manuscript.

OECD (2015a), “Aggregate National Accounts, SNA 2008 (or SNA 1993): Gross domestic product”, OECD National Accounts Statistics (database).
DOI: http://dx.doi.org/10.1787/data-00001-en (Accessed on 14 October 2015)

Steves, Franklin and Alexander Teytelboym (2013), “Political Economy of Climate Change Policy”, *Smith School Working Paper Series*, Working Paper 13-06.

Surminski, Swenja and Andrew Williamson (2012), “Policy indexes – what do they tell us and what are their applications? The case of climate policy and business planning in emerging markets”, *Grantham Research Institute on Climate Change and the Environment Working Paper*, No. 88.

UNFCCC (1992), "United Nations Framework Convention on Climate Change", United Nations 1992.

World Bank (2014), "State and Trends of Carbon Pricing", Washington DC, May 2014.

3.8 Appendix

A1. Carbon Pricing Mechanisms in OECD countries

Country	**Carbon tax**	**ETS**
Australia		Carbon Pricing Mechanism (EPM)
Austria		EU ETS
Belgium		EU ETS
Canada	✓ (BC)	SPEDE (Québec)
Chile		
Czech Republic		EU ETS
Denmark	✓	EU ETS
Estonia		EU ETS
Finland	✓	EU ETS
France	✓	EU ETS
Germany		EU ETS
Greece		EU ETS
Hungary		EU ETS
Ireland	✓	EU ETS
Israel		
Italy		EU ETS
Japan	✓	Tokyo Cap and Trade Program/Target Setting Emissions Trading System in Saitama
Luxembourg		EU ETS
Mexico	✓	
Netherlands		EU ETS
New Zealand		NZ ETS
Norway	✓	EU ETS
Poland	✓	EU ETS
Portugal		EU ETS
Republic of Korea		
Slovakia		EU ETS
Slovenia		EU ETS
Spain		EU ETS
Sweden	✓	EU ETS
Switzerland	✓	Swiss ETS
Turkey		
United Kingdom	✓	EU ETS
United States of Ameri-ca		Regional Greenhouse Gas Initiative (RGGI) (CT, DE, ME, MD, MA, NH, NY, RI, VT), California Cap-and-Trade Program

Table 20: Overview of Carbon Pricing Mechanisms in OECD countries (source: World Bank (2014) and ICAP (2015), own illustration)

A2. Rankings of different climate policy indices including the CPMI

Country	CCPI	CDI	CLIMI	CPI	EPI	ISE	CPMI
Australia	21	21	22	22	2		10
Austria	13	11	13	11	6		13
Belgium	8	8	11	16	22		3
Canada	22	22	21	10	17		15
Czech Republic	16	5	12	21	3	4	7
Denmark	1	6	6	3	10		4
Finland	15	3	1	1	13		19
France	5	10	2	7	19		12
Germany	10	9	10	4	4	1	6
Greece	19	15	16	19	16		1
Hungary	6	2	18	17	20	5	22
Ireland	7	17	9	12	14		11
Italy	9	13	14	9	15		2
Japan	20	20	15	13	18		20
Netherlands	14	12	8	5	9	3	14
New Zealand	17	16	17	14	11		21
Norway	12	19	5	6	8		18
Portugal	2	4	19	20	12		5
Spain	11	7	4	18	5	2	8
Sweden	3	1	7	2	7		9
Switzerland	4	14	3	8	1		16
United States	18	18	20	15	21		17

Table 21: Overview of the rankings discussed including the CPMI. (source: own illustration)

A3. Development of the cross-country rankings over time

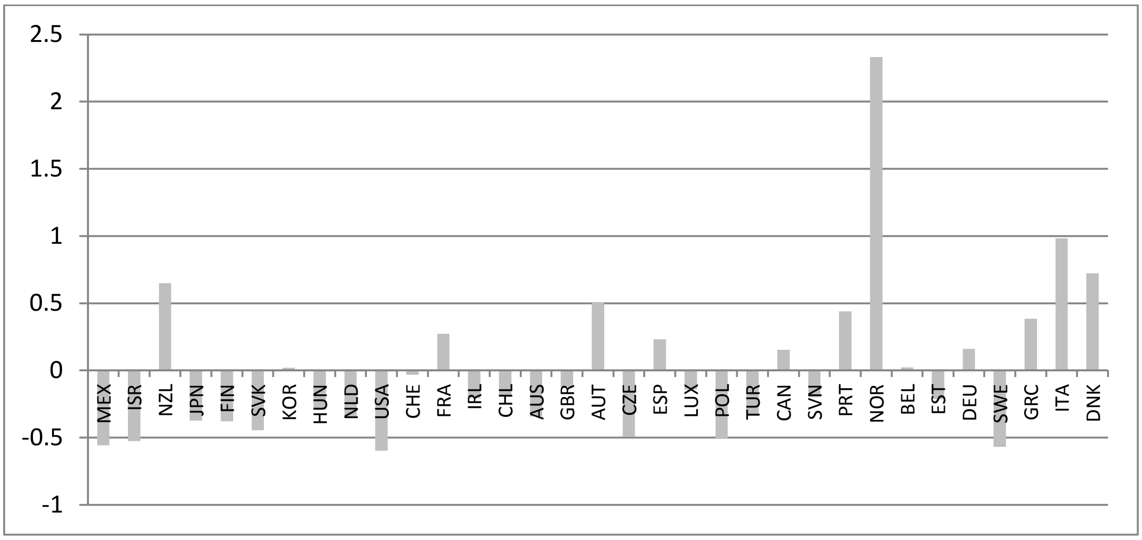


Figure 15: Cross-country comparison of the CPMI for the year 1995, which is the first year of the index in which index figures can be assigned to every country. (source: own illustration)

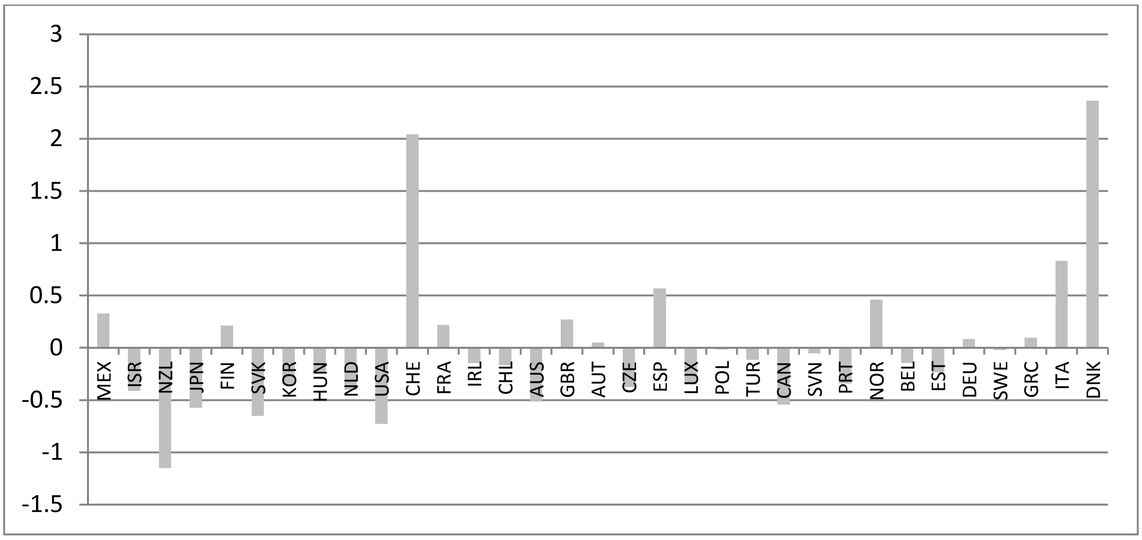


Figure 16: Cross-country comparison of the CPMI for the year 2000 (source: own illustration)

Curriculum Vitae

Julian Dieler
born on September 9, 1986 in Frankfurt am Main, Germany

05/2016 – present	Research Associate, Ifo Institute – Leibniz Institute for Economic Research at the University of Munich, Germany
10/2011 – 05/2016	Ph.D. Candidate, Munich Graduate School of Economics, University of Munich, Germany Junior Economist and Doctoral Student, Ifo Institute – Leibniz Institute for Economic Research at the University of Munich, Germany
04/2007 – 09/2011	Studies of economics (Diplom, equivalent M.A. in economics) University of Munich, Germany
09/2010 – 02/2011	Studies of economics Toulouse School of Economics, France
06/2006	Abitur (High School Diploma), Mariengymnasium Arnsberg, Germany