CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Marginal Abatement Costs of Greenhouse Gas Emissions: A Meta-Analysis

Master's thesis

Author: Mgr. Alžběta Křížková
Study program: Economics and Finance
Supervisor: prof. PhDr. Tomáš Havránek, Ph.D.
Year of defense: 2022

Declaration of Authorship

The author hereby declares that she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 2, 2022

Alzbeta Krizkova

Abstract

This thesis uses up-to-date meta-analysis methods to produce a systematic summary of the literature on marginal abatement costs (MAC) of greenhouse gas emissions. It collects 242 MAC estimates for 2030 and 2050 from 59 studies. Besides the usual tests for publication bias, the study employs several modern non-linear tests, such as the TOP 10, the Kink method, the Stem method, and others. Subsequently, Bayesian model averaging is performed for the first time in MAC literature to reveal a mild negative publication bias for the MAC in 2050. The thesis reveals that newer studies provide higher estimates of MAC. Other factors influencing MAC estimation are the size of stabilisation targets, emissions baseline, utilising the LEAP model, the inclusion of other greenhouse gases besides carbon dioxide, and considering the long-run decision making. Several robustness checks are conducted along the way to confirm the selection of the dataset and the robustness of the BMA analysis (using weighted BMA, FMA, OLS). The true value of MAC in 2030 corrected for publication bias is around 32 EUR/tCO_2-eq, while for 2050, it is 59 EUR/tCO_2-eq.

JEL Classification	F12, Q54, Q52, Q43
Keywords	Meta-Analysis, Greenhouse Gas Mitigation, Marginal Abatement Costs, Publication Bias, Bayesian Model Averaging

Title	Marginal Abatement Costs of Greenhouse Gas Emissions: A Meta-Analysis

Abstrakt

Tématem této diplomové práce jsou mezní náklady snižování emisí skleníkových plynů. Studie sestavuje systematický přehled literatury, k čemuž využívá moderní metody metaanalýzy. Práce shromáždila 242 pozorování pro roky 2030 a 2050 z celkem 59 odborných studií. Kromě tradičně využívaných testů pro odhalení publikační selektivity pracuje s moderními nelineárními testy (například metody TOP 10, Kink, Stem a další). Následné bayesiánské průměrování modelů odhaduje vlastnosti modelování, které ovlivňují výslednou hodnotu mezních nákladů. Nejsilnější efekt se našel pro modely, které zpracovávají data v LEAP modelu. Další vlastnosti modelů, které ovlivňují výslednou hodnotu mezních nákladů, jsou zahrnutí jiných skleníkových plynů než CO_2 a předpoklad rozhodování v dlouhodobém horizontu. Pro potvrzení výběru správného datasetu a určení stability výsledků provádíme několik testů robustnosti. Diplomová práce našla důkaz pro mírnou negativní publikační selektivitu pro rok 2050. Hodnota mezních nákladů po opravení publikační selektivity je 32 eur/tCO_2-eq pro rok 2030 a 59 eur/tCO_2-eq pro rok 2050.

Klasifikace JEL	F12, Q54, Q52, Q43
Klíčová slova	Metaanalýza, Snížení emisí skleníkových plynů, mezní náklady na snížení emisí, Publikační selektivita, Bayesiánské Průměrování Modelů
Název práce	Mezní náklady na snížení emisí skleníkových plynů: Metaanalýza

Acknowledgments

I would like to thank my supervisor, prof. PhDr. Tomáš Havránek Ph.D., for his time, guidance, ideas, and valuable help with this thesis. Furthermore, I want to thank my family and friends for their continuous support and encouragement throughout my whole studies.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record
Krizkova, Alzbeta: *Marginal Abatement Costs of Greenhouse Gas Emissions: A Meta-Analysis*. Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 92. Advisor: prof. PhDr. Tomáš Havránek, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

BAU Business-as-usual

BE Between Effects estimator

CCS Carbon Capture and Storage

CO_2 Carbon Dioxide

CPI Consumer Price Index

DCAS Direct Capture and Storage

GHG Greenhouse Gas

IPCC Intergovernmental Panel on Climate Change

IV Instrumental Variable

MAC Marginal Abatement Cost

MRA Meta-Regression Analysis

OECD Organisation for Economic Co-Operation and Development

ppm parts per million

ppmv parts per million volume

RE Random Effects estimator

RCPs The Representative Concentration Pathways

UNFCCC United Nations Framework Convention on Climate Change

WLS Weighted Least Squares

Master's Thesis Proposal

Author	Mgr. Alžběta Křížková
Supervisor	prof. PhDr. Tomáš Havránek, Ph.D.
Proposed topic	Marginal Abatement Costs of Greenhouse Gas Emissions: A Meta-Analysis

Motivation Climate change crisis is an urgent matter that has to be dealt with as soon as possible. Even though, some information found in the public spheres do not seem to be supported by scientific research, thanks to civic engagement, various policies are now being proposed and often implemented. Nevertheless, it is crucial to continue researching the climate change topics in order to present up to date results obtained with credible scientific methods to the policy makers.

The usual rhetoric aims at mitigating the climate change by reducing the emissions of greenhouse gases. It has been a thoroughly discussed topic in the academic as well as political spheres. The topics of such political discussions are usually surrounding the costs of such mitigation. However, it can be difficult for politicians to make sense of the countless studies which seek to estimate various abatement costs. Available research papers work with a wide range of units and substances and often lead to unreliable conclusion. This study will aim to clarify such variation through the means of meta-analysis. More specifically, it will focus on marginal abatement costs (MAC) of greenhouse gas (GHG) emissions. Marginal abatement cost refers to the cost of eliminating a single unit of emissions.

Kuik, Brander & Tol (2009) already conducted a meta-analysis of marginal abatement costs of greenhouse gas emissions. In their study, they collected data from 26 studies published in 2006 and found the cost estimates to be sensitive to the stringency of the stabilisation target, the assumed emissions baseline, and other factors.

Hypotheses

Hypothesis #1: The MAC estimates are positively influenced by emissions baseline.

Hypothesis #2: The stabilisation target has a negative effect on MAC.

Hypothesis #3: The literature estimating marginal costs of greenhouse gas emissions is affected by publication bias.

Methodology The aim of my research is to examine the sensitivity of MAC estimates to the specifications and assumptions underlying these models. Specifically, I will be focusing on MAC of stabilisation targets, baseline emissions, the inclusion of other greenhouse gases in the emission target and other factors. Additionally, I will calculate MAC ranges for alternative stabilisation targets for GHG concentrations.

The first step in conducting a meta-analysis is the collection of primary research. I am going to be using all the studies examined by Kuik, Brander & Tol (2009) (I have already politely asked the authors for their dataset) and I will thoroughly research all relevant economic journals, as well as Google Scholar, to find the most appropriate studies. In each paper I need to carefully examine that standard errors (or other statistics from which standard errors can be computed) are included. In case of an absence of standard errors, I will follow the technique of Havranek, Irsova, Janda & Zilberman (2015).

After collecting all relevant studies, I am going to create my dataset. I will convert all the reported values to a common unit, clear the dataset from outliers and winsorise, subsequently. Regarding the publication bias, I am going to be utilizing the Ordinary Least Squares (OLS) technique, Fixed Effects, Between Effects and other suitable methods. I am going to be using cluster standard errors when possible and presenting confidence intervals. For dealing with heterogeneity, I am going to utilize Bayesian (baseline) and Frequentist model averaging.

Expected Contribution Using the methods of meta-analysis, I am going to conduct a quantitative survey of research papers estimating marginal abatement costs of greenhouse gas emissions. I am going to follow previously conducted meta-analysis by Kuik, Brander & Tol (2009). In addition to their model and methodology I am going to focus on publication bias using mixed-effects multilevel meta-regression. I am expecting to obtain different results after correcting for the bias. I am aware that the climate change research is developing at a rather fast pace. Therefore, I expect to obtain different values of costs than Kuik, Brander & Tol (2009) since their data collection is from 2006 and earlier. My findings can be directly used for climate change policy implications, because the results obtained will take into account all the factors that research papers in this field work with.

Outline

1. Introduction (Motivation) – There is already one meta-analysis on marginal abatement costs of greenhouse gas emission from 2009 (working with data from 2006), I am going to revisit this paper and bring the analysis up to date.

2. Literature Review – I will describe already existing literature on marginal abatement costs that I will be working with.

3. Methodology – I will present the technique of meta-analysis research methodology, including the necessary tests.

4. Data – I will describe the data and their sources.

5. Empirical results – I will discuss my findings, my baseline regressions and robustness checks.

6. Conclusion – I will summarize results of the research, their implications for policy and propose a course for future research.

Core bibliography

Kuik, O., Brander, L. & Tol, R.S.J., 2009. Marginal abatement costs of greenhouse gas emissions: A meta-analysis. Energy Policy, 37(4), pp.1395-1403. Available at: `https://www.sciencedirect.com/science/article/abs/pii/S0301421508007295` [Accessed September 07, 2020].

Havránek, T., Stanley, T.D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W.R., Rost, K. and van Aert, R.C.M. (2020), Reporting guidelines for meta-analysis in economics. Journal of Economic Surveys, 34: 469-475. doi:10.1111/joes.12363 [Accessed September 15, 2020].

Havranek, T., Irsova, Z., Janda, K., & Zilberman, D., 2015. Selective reporting and the social cost of carbon. Energy Economics, 51(C), 394–406. Available at `https://ideas.repec.org/s/eee/eneeco.html` [Accessed September 15, 2020].

Havranek, T., Herman, D., & Irsova, Z., 2018. Does Daylight Saving Save Electricity? A Meta-Analysis. The Energy Journal, International Association for Energy Economics, vol. 0 (Number 2). Available at: `https://ideas.repec.org/a/aen/journl/ej39-2-irsova.html` [Accessed September 07, 2020].

Barker, T., Köhler, J., Villena, M., 2002. The costs of greenhouse gas abatement: a meta-analysis of post-SRES mitigation scenarios. Environmental Economics and Policy Studies 5, 135–166. doi: 10.1007/BF03354027 [Accessed September 15, 2020].

Manne, A., Richels, R.G., 2006. The role of Non-CO2 greenhouse gases and carbon sinks in meeting climate objectives. The Energy Journal Special Issue on Multi- Greenhouse Gas Mitigation and Climate Policy, 393–404. Available at: https://www.jstor.org/stable/23297092 [Accessed September 15, 2020].

Chapter 1

Introduction

Greenhouse gas emissions play a fundamental role in climate change. One of the most profound global agreements aiming to reduce global warming, the Kyoto Protocol, established reducing greenhouse gas emissions as the key action to slowing down climate change. Public discussion is usually linked with the costs associated with emissions mitigation. However, scientific literature supplies different values, and each study works with different assumptions. How should we decide from the amount of scientific results, which ones to follow? This research seeks to explain the true effect of Marginal Abatement Cost (MAC) - the price of reducing one additional unit of emission. To find this true outcome, the thesis employs the method of meta-analysis and sorts through literature to find one true cost of mitigating greenhouse gas emissions. A meta-analysis is a methodical review and quantitative literature synthesis summarising (and explaining) the variation found among empirical results.

The study can be put alongside several meta-analyses estimating the true effect of the MAC. It directly extends the previous study conducted by Kuik *et al.* (2009) because it connects their dataset from 2006 with empirical results published since then. Other meta-analyses, for example Barker *et al.* (2006), Fischer *et al.* (2003) or Repetto & Austin (1997) focus on a relationship between the MAC and a specific aspect. By contrast, this study employs a wide range of characteristics collected from literature to reveal which one affects the MAC the most. On top of that, this study is the first to examine both publication bias and heterogeneity.

The analysis therefore consists of two main building blocks: publication bias analysis and heterogeneity analysis. Publication bias arises when the paper's publication depends on the significance of its results. In the heterogeneity analysis, we examine which model specifications influence the MAC estimate. We collect 242 observations of MAC for the years 2030 and 2050 from 59 primary studies. Due to a lack of uncertainty measures in primary studies, standard errors are approximated, following the method by Havranek *et al.* (2015). After a thorough examination, cleaning and adjusting of the data, the final variables are presented, together with their characteristics. Then, the meta-analysis can begin. First, the funnel plot is visually observed to reveal publication bias. The plot is relatively symmetrical, which does not signify publication bias in the literature. Meta-analysis regression should empirically validate this conclusion. We find a small publication bias after several linear and non-linear tests were employed. The true effect corrected for publication bias is relatively close to the sample mean, confirming this conclusion. The resulting MAC corrected for publication bias is 32 EUR/tCO_2-eq for 2030, while it almost doubles (59 EUR) for 2050.

In the BMA analysis examining model uncertainty, we find evidence of mild negative publication bias for the MAC in 2050. We reveal the size of the stabilisation target and emissions baseline to affect the MAC estimate negatively. The estimate of the MAC for 2030 is lower when a model employing the LEAP model works with overall GHG emissions or emissions from agriculture. On the other hand, the MAC 2050 is negatively affected by the number of regions, including intertemporal optimisation or multigas.

The thesis is structured as follows. Chapter 2 explains crucial concepts from climate change literature. Next, Chapter 3 presents key concepts for the MAC analysis and summarises previously conducted meta-analyses and their findings. Data collection, adjustments, and summary statistics are outlined in Chapter 4. The next chapter introduces the reader to meta-analysis and serves as an opening for the following two chapters. The inspection of publication bias is conducted in Chapter 6, and the heterogeneity analysis is in Chapter 7. Finally, we conclude the research in Chapter 8, together with limitations to our research and possible suggestions for future extension. Supporting materials are attached in Appendix A and Appendix B.

Chapter 2

Abatement of Greenhouse Gas Emissions

Greenhouse gas (GHG) emissions have been the centre of attention of public debate for a relatively long time now. Despite the apparent importance of the issue, previous research has not reached a consensus regarding the actual cost of reducing GHG emissions. Using the means of meta-analysis, we will seek to answer the following questions: What is the true effect of marginal abatement cost in empirical research? Are the reported effects subject to publication bias? To what extent does the research design (data, estimation methods, variables) systematically influence the reported results? We use both linear and non-linear methods to correct publication bias and deal with model uncertainty in the study using Bayesian model averaging (Steel, 2020).

The abatement of greenhouse gases is a broad topic requiring a certain level of understanding. Before starting with the meta-analysis, we first explain the crucial related concepts to understand the research better.

The theoretical background on climate change mitigation has been drawn mainly from the IPCC report *AR5 Climate Change 2014: Mitigation of Climate Change*, prepared by the Working Group III. The IPCC, or the Intergovernmental Panel for Climate Change, is a body of the United Nations responsible for aggregating knowledge on climate change. The report focuses on the literature discussing various aspects of climate change mitigation published between 2007 and 2014 (IPCC, 2014). An updated version of this report is scheduled for September 2022. It will focus on the literature published from 2014 to the present day and will most likely bring new evidence to the discussion. Working with a new version of the report (AR 6) could lead to more accurate conclusions and would make a suitable extension to the research presented here.

2.1 Climate Change

> *"It is unequivocal that human influence has warmed the atmosphere, ocean and land."* (IPCC, 2021)

2.1.1 Current Climate Change

The Intergovernmental Panel on Climate Change (IPCC) publishes reports on climate change which are backed by a synthesis of scientific data and accompanied by the likelihoods of certain statements (IPCC, 2022). The fifth assessment report (AR5) talks about observed changes in the climate system and the influence of greenhouse gas emissions on global warming.

In the last decades, the Earth's surface has been successively warmer than in any decade since 1850. The period 1983-2012 has likely been the warmest period in the last 1400 years for the Northern Hemisphere. Since the pre-industrial era, anthropogenic greenhouse gas emissions have increased, which has led to unprecedented atmospheric concentrations of carbon dioxide, methane, and nitrous oxide. The consequences of economic and population growth, together with other human activities on the Earth, were, according to the IPCC, detected in the climate system and are extremely likely "the dominant cause of the observed warming since the mid-20th century".

The evidence that human actions influence the climate system grows with each IPCC report published. The IPCC Fifth Assessment Report (AR5) states that "it is extremely likely that more than half of the observed increase in global average surface temperature from 1951 to 2010 was caused by the anthropogenic increase in GHG concentrations and other anthropogenic forcings together" (IPCC, 2014). Evidence of observed climate change can be shown in many regions on changing precipitation and melting ice and snow. On top of that, it can also affect the quantity and quality of water resources.

The risks of climate-related impacts are distributed unevenly and are more significant for disadvantaged people and communities in developing countries. Furthermore, even if emissions were mitigated, the impacts of climate change would continue for centuries. A detailed overview of potential future changes in the climate system can be found in the IPCC's report *AR5 Synthesis Report: Climate Change 2014*.

2.1.2 Future Climate Changes

The amount of anthropogenic CO_2 emissions is mainly influenced by population, namely its size, economic activity, energy use, land use pattern, technology, lifestyle, and climate policy. Scientific discourse presented in the IPCC's reports utilises the Representative Concentration Pathways (RCPs) to make projections of future scenarios based on the factors described above (IPCC, 2014). The RCPs describe four different pathways of the 21st century, dependent on GHG emissions and their atmospheric concentrations, air pollutant emissions, and land use. The pathways are consistent with the wide range of scenarios used in the literature.

Future climate change and its scale depend on current and future emissions as well as natural climate variability. The RCPs include one stringent mitigation scenario - RCP2.6, which aims to keep the global temperature rise below 2 °C above pre-industrial levels. There are two intermediate scenarios called RCP4.5 and RCP6.0, and one scenario with very high GHG emissions - RCP8.5. Additionally, there is a scenario for no further efforts in constraining emissions, the so-called 'baseline scenario'. The RCPs reflect consistent and robust evidence from the literature indicating that there is a linear relationship between cumulative CO_2 emissions and projected global temperature change up to the year 2100. The Figure 2.1 shows a graphical representation of the RCPs and the associated scenario categories.

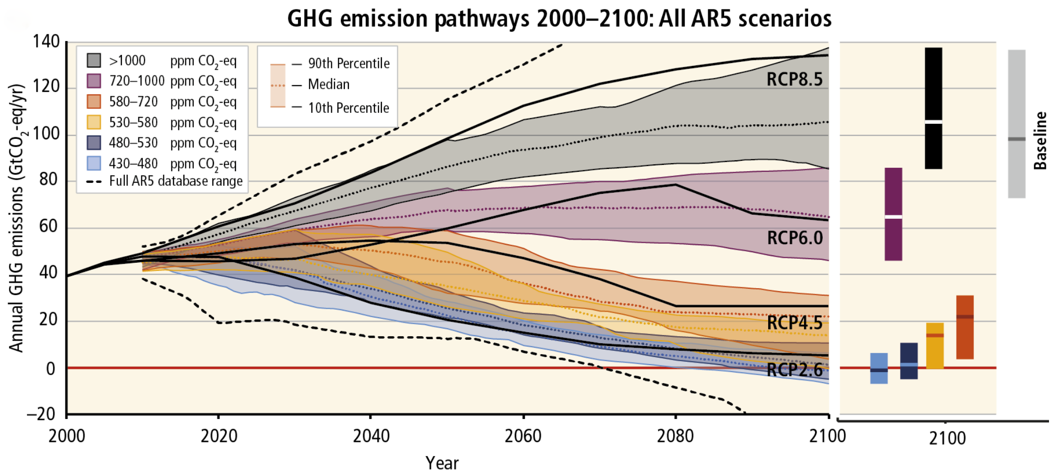


Figure 2.1: The Representative Concentration Pathways (RCPs), source: IPCC (2014, p. 21)

2.2 Stabilisation Targets

In order to accurately work with the different future scenarios of GHG emissions used in literature, we first need to understand their significance. Using stabilisation targets and RCPs is one way to consider the projected climate changes and impacts the GHG emissions stabilisation would bring. The targets analyse and employ information drawn from the scientific literature.

The climate stabilisation goals are often presented together with the global mean temperature change (in °C) and stabilised concentrations of carbon dioxide (in ppmv) (Council, 2011). When studying other GHG gases, the concentration of radiative forcing is usually expressed in terms of CO_2-equivalent. The pathways that would limit the warming below 2 °C relative to pre-industrial levels require substantial reductions of emissions over the following decades and near-zero emissions of CO_2 and other long-lived greenhouse gases by the end of the century (IPCC, 2014). With no additional effort to reduce the GHG emissions except for those in place today, emissions are expected to grow further. With the baseline scenario (with no additional interventions), the global mean surface temperature increase in 2100 is expected to range from 3.7 to 4.8 °C.

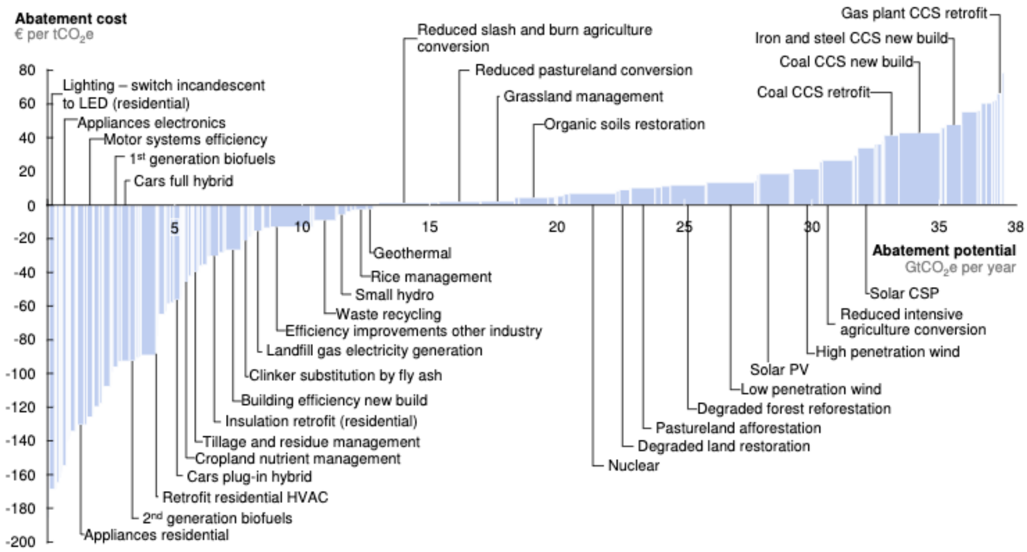
The studies included in the meta-analysis work with different stabilisation targets. To reasonably compare the results, the stabilisation targets were converted to a unified metric: concentration of greenhouse gases in the atmosphere - expressed in CO_2 equivalents (ppm/CO_2-eq). The other often used metrics include radiative forcing (W.m^{-2}), the concentration of the greenhouse gas CO_2 (ppm/CO_2), and global mean temperature ($°C$). Table 2.1 presents the classification by IPCC (2014) of various stabilisation targets divided into six categories (I-VI). The overview also serves as a conversion table between the metrics.

Category	Radiative forcing ($W.m^{-2}$)	CO_2 concentration (ppm)	CO_2-eq concentration (ppm)	Global mean temperature increase (°C)
I	2.5 - 3.0	350 - 400	445 - 490	2.0 - 2.4
II	3.0 - 3.5	400 - 440	490 - 535	2.4 - 2.8
III	3.5 - 4.0	440 - 485	535 - 590	2.8 - 3.2
IV	4.0 - 5.0	485 - 570	590 - 710	3.2 - 4.0
V	5.0 - 6.0	570 - 660	710 - 855	4.0 - 4.9
VI	6.0 - 7.5	660 - 790	855 - 1130	4.9 - 6.1

Table 2.1: Stabilisation targets and corresponding metrics, source: IPCC (2014)

2.3 Marginal Abatement Cost

To reduce the climate change risk, we need both adaptation and mitigation as complementary strategies. Each country has had a different past contribution to GHG emissions and has differing means and financial resources to address its adaptation and mitigation. Climate policy's design needs to reflect both individuals' and organisations' perceptions of risk and uncertainty. One of the methods to evaluate the risks from an economic point of view is marginal abatement cost (MAC) - how much would have to be paid to diminish one more unit of emission. In this way, we can use the MAC to describe the potential and cost of different abatement options (den Elzen *et al.*, 2007). For better understanding, the literature usually works with several mitigation options and constructs a marginal abatement costs curve - MACC.



Note: The curve presents an estimate of the maximum potential of all technical GHG abatement measures below €80 per tCO_2e if each lever was pursued aggressively. It is not a forecast of what role different abatement measures and technologies will play.
Source: Global GHG Abatement Cost Curve v2.1

Figure 2.2: Global GHG abatement cost curve beyond BAU – 2030, source: Enkvist *et al.* (2010, p. 7)

One of the most referenced works on abatement costs of greenhouse gases was developed in McKinsey. It was first published in 2007 and then revisited and updated in 2010. In the original report, the authors collected and highlighted the most beneficial way to abate GHG emissions. Rather than evaluating the science of climate change, the findings are aimed at policymakers to get oriented on the problem and offer a relatively simple assessment of the situation (Enkvist *et al.*, 2007).

The authors used data from the International Energy Agency (IEA) to establish a "business-as-usual" (BAU) projection for their comparison in 2010, 2020 and 2030. They focused on abatement costs below 40 euros per ton of CO_2, and their primary outcome was the global cost curve for GHG abatement. Three years later, the report was updated, motivated by the financial crisis in 2008 and its impact on previous estimates. In Figure 2.2, we can see the updated GHG abatement cost curve - Version 2.1 (Enkvist *et al.*, 2010).

The abatement cost curve plots possible ways to mitigate GHG emissions from the least expensive (on the left side) to the most expensive. For each abatement measure, there is an exact cost per ton of CO_2 emission reduced (y-axis) and a quantity of emissions available for reduction at that cost (x-axis). In most abatement curves, some measures show the negative abatement cost - meaning that money would be saved when choosing particular measures. However, this gap is usually explained as unaccounted-for costs in most literature. The curves, such as the McKinsey one, are based on engineering estimates and typically do not include behavioural considerations (Gillingham & Stock, 2018). These kinds of imperfections in the construction of abatement curves should be considered when concluding policy implications based on them.

Chapter 3

Key Concepts & Literature Review

The objective of this chapter is twofold: to present crucial literature and its empirical findings and introduce key concepts that have the potential to affect the MAC and will therefore be included in the analysis. Together with these concepts, we present previous empirical findings regarding particular aspects. The most recent paper directly followed is Kuik *et al.* (2009). Other crucial studies of the MAC include Barker *et al.* (2006), Fischer *et al.* (2003), Barker *et al.* (2002), and Repetto & Austin (1997).

The first meta-analysis that focused on the costs of mitigating climate change was conducted by Repetto & Austin (1997). Their paper *The costs of climate protection: A Guide for the perplexed*, undertaken at the World Resources Institute (WRI), works with 16 widely used models and explains "how key assumptions affect the predicted economic impacts of reaching CO_2 abatement targets". Their study was the first to reveal that only a few assumptions are important to affect the resulting estimate. The results show two main areas with the highest impact on climate-change mitigation. The CO_2 emission control should be instrumented by revenue-raising policies (carbon taxes, tradable permits), and these revenues should be used to reduce other burdensome taxes. This approach results in more expensive carbon-based fuels, implying higher costs throughout the economy. The authors suggest using the revenues to offset some of these higher costs and thus improve the economic impact. The final recommendation highlights the role of media in contributing to public understanding. It also includes advice on softening the impacts on the regions, industries, and communities that would be affected adversely as well as on negotiating international agreements to coordinate actions.

Target

The variable *Target* denotes the stabilisation targets introduced in the previous chapter. All the collected targets from studies were converted to ppm/CO_2-eq based on the conversion table 2.1. Fischer *et al.* (2003) claims that the more strict the stabilisation targets are, the less flexibility there is for alternative emission scenarios.

In response to the IPCC introducing the emissions scenarios, Barker *et al.* (2002) analysed the extent to which the mitigation costs in each scenario can be explained by the characteristics and the assumptions of the model. They combine the means of meta-analysis with scatter plots of the data to classify ranges of estimates rather than single values. They found a strong correlation between CO_2 reduction and GDP reduction as well as highly significant model characteristics which, when chosen correctly, can explain up to 70% of the variance.

Emissions Baseline

The variable *Baseline* demonstrates how technology development, economic growth, and industry structure influence the predictions in the model. The variable is expressed as a percentage of emissions in the future (2030 or 2050), where 1 (or 100%) denotes emissions in the baseline year. It essentially outlines how much the emissions are expected to grow in the future without any effort expended to mitigate them. Together with the stabilisation target, this variable shows the emissions mitigation effort.

Top-Down and Bottom-Up Models

The impact of climate change policies is modelled using two kinds of economic analyses - 'top-down' and 'bottom-up' models. The top-down model is an aggregate model of the economy as a whole "that represents the sale of goods and services by producers to households and the reciprocal flow of labour and investment funds from households to industries" (Repetto & Austin, 1997). The scope for technological substitution is deduced from the past. On the contrary, the bottom-up model considers the actual "technological options for energy savings and fuel-switching that are available in individual sectors of the economy, such as housing, transportation, and industry. Information on the costs . . . is then aggregated to calculate the overall cost" (Repetto & Austin, 1997). These models are usually more optimistic about the magnitude of cost-effective energy savings.

In order to capture which model was utilised in the research, we created a dummy variable called *Top-down*. When the employed model was top-down, the variable is equal to 1, while for the bottom-up, the variable is 0.

LEAP Model

There are several methodological approaches for estimating the MAC. In primary studies, we usually come across GAINS, AIM, MARKAL, TIMES or IMAGE simulation models. Here, we present the integrated model LEAP because it is used the most in our dataset, and we include it as a dummy variable. LEAP is a Windows-based tool for comprehensive analysis of GHG mitigation assessment, developed by the Stockholm Environment Institute with funding from the World Bank and the UNEP (UN Environment Programme). Countries worldwide utilise it to develop their Intended Nationally Determined Contributions (INDCs) - outlined steps they will take to tackle climate change. LEAP focuses on energy sector GHG emissions (but can be used across all sectors) and, apart from GHGs, examines local air pollutant emissions, energy security, economic costs, land-use change, and forestry (Hong *et al.*, 2016; Heaps, 2018). The variable *LEAP* equals 1 when the model employs the LEAP model for the estimation.

Induced Technological Change

Another significant concept when discussing climate change and possible pathways for policymakers is Induced technological change (ITC). Clarke *et al.* (2008) describe the term as "the alteration to the rate or direction of technological change in response to a particular policy or set of policies…; the concern is whether the sorts of policies that are considered in the climate context…might induce *additional* or *different* technological changes". In other words, the overall GHG mitigation policy and subsequent carbon price are directly connected with the direction and magnitude of progress in abatement technologies. For researchers and policymakers, a fundamental concern should be how much technological change would occur even without the climate policy and how much of it is a direct consequence of their policy. Kuik *et al.* (2009) therefore claim that "dynamic economic models should not take technical progress over time as given, but should explicitly model the interactions between policy and technical change."

The dummy variable *Induced Technological Change* reflects precisely that - when the model contains a specification induced technological change, the variable is equal to 1, and 0 when it does not.

Barker *et al.* (2006) in their meta-analysis focused on mitigation costs for global GHG with respect to induced technological change. They analysed the effect of technological change on the cost estimates, "measured as changes in welfare or gross world product, and of the required CO_2 tax rates and emission permit prices." The study acknowledged that induced technological change was a rather new topic in economic modelling literature in 2006, and the results they relied on were often controversial and experimental. Nevertheless, they concluded that even strict stabilisation targets can be achieved in 2030 without significantly affecting world GDP growth. The marginal abatement prices they found were \$15/tCO_2 for 550ppmv and \$50/tCO_2 for 450ppmv.

Intertemporal Dynamic Optimisation

Another dummy variable called *Intertemporal Dynamic Optimisation* reflects the time horizon of GHG emissions within the model. While some models assume long-living decision-makers who optimally decide on consumption, investments and abatement over an extended time period (variable equal to 1), other models consider only optimisation period by period (variable is 0). These different time profiles can affect the MAC in a particular year.

Fischer *et al.* (2003) looked into a wide range of estimates for marginal abatement costs, which led to undercutting the support for policies to reduce greenhouse gas emissions. They used four kinds of factors that could explain the differences in estimations of the MAC: emissions baseline, degree of flexibility allowed for emissions constraints, structural characteristics of the model, and characterisation of the benefits from pollution reduction (Fischer *et al.*, 2003).

They recommend that subsequent researchers of the MAC should carry these factors in mind when designing an analysis. The authors described two approaches to address the range of estimates. The first one is to match all the assumed policy systems and other relevant assumptions and, together with a mix of quantitative and qualitative techniques, reveal the differences in the models. The second approach is to define specific variables that are expected to explain the different factors described above.

Some variables they included in the model are similar to ours. They work with emissions baseline, number of regions, number of energy sources, CCS and technical change, a number of non-energy sectors, and dummy variables for perfect substitutes and international capital mobility.

Their analysis showed that "certain modelling choices have important effects on the estimated costs of reducing greenhouse gas emissions" (Fischer *et al.*, 2003). Lower MACs are reported in models that assume freer trade and more perfect substitution of goods across regions. On the other hand, MAC estimations are higher when assuming greater disaggregation in energy goods. Baseline scenarios only have a small influence on MAC.

Multigas

While carbon dioxide emissions are at the centre of climate change discussion, other greenhouse gases should also be considered when talking about mitigation. These GHGs are methane, nitrous oxide, ozone, water vapour, and fluoridated gases. There are several reasons why the main attention is directed at CO_2. Its emissions from fossil sources can be easily estimated from market data on fuel use, while for other GHGs, the estimation is more complicated. Also, the extensive research on energy markets, energy efficiency, and possible alternative energy supply technologies were motivated by the attempt to secure the supply and prices of fossil fuels (Reilly *et al.*, 2003). Finally, carbon dioxide accounts for more than half of the effect GHG emissions have on climate change (Stern, 2008). For an effective environmental and economic policy, one should address CO_2 as well as the other greenhouse gases.

If the research works with a multigas approach, the value of the variable *Multigas* is 1. When the research deals with single greenhouse gas, the variable equals 0.

Carbon Capture and Storage

The idea behind backstop technology is a belief that during the years of abating GHG emissions, there will be a "transition from one energy source to another" (Seo, 2021). This transition will lead to lower (or no) dependency on fossil fuels and other GHG emission sources and provide society with an almost inexhaustible energy source for a constant price (Seo, 2021).

Methods for capturing and storing carbon (such as "Direct Capture and Storage") are considered the first step towards a backstop technology. The process lies in capturing carbon dioxide directly from any air, as opposed to focused capture from a point source where a higher amount of CO_2 is present (e.g. biomass power plant, cement factory). This technology has already been implemented, and the world's largest plant for direct air capture opened in Iceland in September 2021. The DCAS technology has been called quasi-backstop because the economic model still determines the price but at least limits the price from the top.

Another dummy variable included in the meta-analysis is called *Carbon Capture and Storage (CCS)*. It is equal to 1 when the model acknowledges a possibility of carbon capture and storage technology as a partial solution to GHG emissions mitigation and 0 if there is no mention of this technology.

The remaining variables used in the meta-analysis are *Regions* and *Energy Sources*. Variable *Regions* indicates a number of regions the primary study works with and ranges between 1 and 162. *Energy Sources* stands for the number of primary energy sources in the model. The summary statistics of all variables can be found in Chapter 4.

Kuik *et al.* (2009) focused on the sensitivity of MAC estimates to the assumptions and specifications based on the models. On top of that, their goal was to predict the ranges of MAC for different stabilisation targets. The authors examined 26 models and collected 62 observations of MAC for the years 2025 and 2050. The variables included in their meta-analysis are similar to ours: stabilisation target, baseline emissions, number of regions, number of energy sources, and dummy variables for multigas, induced technological change, top-down approach, intertemporal optimisation, carbon capture and storage.

In addition to these variables, the authors focused on the scientific forum where the model was first presented. "A modelling forum is a meeting or a series of meetings of modelling groups that address a common research question and use a commonly agreed set of assumptions and a common reporting format." There were three modelling fora: EMF-21, IMCP and USCCSP. The authors found a difference in reporting for each forum: "Compared to the EMF-21 modelling forum, the models in the IMCP forum tend to report lower MAC, and the models in USCCSP forum tend to report higher values" (Kuik *et al.*, 2009).

In conclusion, they found that the MAC estimations were dependent on variables of stabilisation target, the emissions baseline, the intertemporal dynamic optimisation, multigas, the number of regions and energy sources, and (to a lesser degree) the modelling forum. When comparing their results to the policy currently in place in the UK and the EU, they "found that these policy-specific estimates are still on the low side if the ultimate aim of the policies is to meet very stringent long-term stabilisation targets" (Kuik *et al.*, 2009). Additionally, they recommended that economic models should focus on estimating the MAC for stabilisation scenarios below 500 ppm CO_2-eq.

Because we extended the dataset by Kuik *et al.* (2009), the last variable (*Kuik*) assigns 1 to studies that originated in their study. The variable equals 0 when the observation comes from the search query described in the next chapter. Because Kuik works with studies from 2006 while the new data comes from 2007 onward, the variable also serves as a time differentiation. Value 1 depicts a study from 2006, while 0 stands for newer data.

Chapter 4

Data

This chapter describes the requirements primary studies had to satisfy to be included in the dataset. We also outline the data collection process and obstacles we overcame to compile the data to a comparable set. Finally, the variables included in the dataset are presented along with their summary statistics and a list of studies used in the dataset.

The dependent variable we seek to explain is Marginal Abatement Cost (MAC). The literature offers estimations of the MAC for different points in time. After thorough consideration, we decided to collect the estimates for 2030 and 2050 because current literature works with these data points the most. The MAC is expressed as a price per ton of carbon dioxide (or equivalent) abated. We also had to standardise the various units used across the papers. We used 2020 Euros per tone of CO_2-equivalent as a unifying unit ($EUR_{2020}/tCO_2 - eq$). The selection of independent variables was made mainly on the previous meta-analysis by Kuik *et al.* (2009) but also reports from the IPCC and availability in examined papers.

4.1 Data Collection

We employed Google Scholar to find relevant primary studies. It is a well-known database with a powerful full-text search and unmatched scope of literature. After identifying a couple of relevant, heavily cited studies we wanted to include in the analysis, we built a search query in a way these studies appeared among the first search results. The final search query is as follows:

marginal abatement cost "curve" OR *"curves" greenhouse "gas"* OR *"gases emissions long-term.*

In addition, the query was restricted to the years 2007-2022. Kuik *et al.* (2009) collected their data in 2006; thus, we expected the final dataset to stretch over the last 17 years. Our data ended up spanning over 31 years as the oldest data included in the analysis were from 1990 and the newest from 2020. Within the search query, we went through the first 100 results and selected 72 studies to examine further - read their abstracts and see if they fit the requirements to be included in the analysis. We then repeated the search with studies no older than three years and added a few more studies to the list. Lastly, we used the snowballing method - inspecting references of the studies selected from the query. The search was terminated on 31 December 2021. The PRISMA diagram in Appendix A indicates specific numbers of papers added to the inventory in each step (Page *et al.*, 2021).

Subsequently, we examined the abstracts of the 76 selected papers and determined whether they could be included in our analysis. Some of the papers summarised findings of other papers and did not carry any new empirical results or worked with different assumptions. A substantial number of papers calculated the MAC for the current period, not the future. Some papers did not explicitly work with the marginal abatement costs, and some calculated the MAC for years other than 2030 or 2050. These types of studies were excluded from further examination. The dataset from Kuik *et al.* (2009) served as a template of variables we should be able to retrieve from the studies. Nevertheless, further adjustments had to be made to both datasets before merging them (these edits are described further in the text).

The papers considered for the meta-analysis were restricted to English-written to secure correct understanding. Due to the lack of uncertainty measures reported in papers, we considered limiting the selection to published papers. Studies published in peer-reviewed journals guarantee quality and avoid multiple inclusion of the same result. After carefully examining their methodology and assumptions, we decided to include papers published elsewhere to expand our dataset.

Finally, the dataset was compared and combined with the dataset by Kuik *et al.* (2009). We chose this paper for two main reasons: it is, to our best knowledge, the most current one in the field, and it combines several explanatory variables that have appeared in previous analyses. The paper is described in more detail in Chapter 2. After merging the two datasets, we had 135 estimates of MAC 2030 and 107 observations of MAC 2050. The Table 4.1 lists the 59 studies included in the analysis.

Aaheim *et al.* (2006)	Kesicki (2012)
Ahn & Jeon (2019)	Kesicki (2013)
Akimoto *et al.* (2012)	Kurosawa (2006)
Barker *et al.* (2006)	Löffler & Hecking (2017)
Beach *et al.* (2015)	Manne & Richels (2006)
Bernard *et al.* (2006)	de Oliveira Silva *et al.* (2015)
Böhringer *et al.* (2006)	Pellerin *et al.* (2017)
Bosetti *et al.* (2006)	Popp (2006)
Chen *et al.* (2020)	Purohit & Höglund-Isaksson (2017)
Chung *et al.* (2015)	Rao & Riahi (2006a)
Clarke *et al.* (2006)	Rao & Riahi (2006b)
Crassous *et al.* (2006)	Reilly *et al.* (2006)
Eide *et al.* (2011)	Sano *et al.* (2006)
Escobar Carbonari *et al.* (2019)	Sapkota *et al.* (2019)
Fawcett & Sands (2006)	Sapkota *et al.* (2021)
Fujino *et al.* (2006)	Smith & Wigley (2006)
Gerlagh (2006)	Sotiriou *et al.* (2019)
Gopal *et al.* (2018)	de Souza *et al.* (2018)
Hanson & Laitner (2006)	Subramanyam *et al.* (2017a)
Harmsen *et al.* (2019)	Subramanyam *et al.* (2017b)
Havlík *et al.* (2013)	Teng *et al.* (2019)
Hedenus *et al.* (2006)	Timilsina *et al.* (2017)
Jakeman & Fisher (2006)	Tol (2006)
Janzen *et al.* (2020a)	van Vuuren *et al.* (2006a)
Janzen *et al.* (2020b)	van Vuuren *et al.* (2006b)
Janzen *et al.* (2020c)	Vogt-Schilb *et al.* (2015)
Jiang *et al.* (2006)	Wagner *et al.* (2012)
Katta *et al.* (2019)	Xiao *et al.* (2014)
Katta *et al.* (2020)	Yue *et al.* (2020)
Kemfert *et al.* (2006)	

Table 4.1: Studies included in the meta-analysis

4.2 Data Adjustments

Before conducting the actual analysis, our data needed to be adjusted to be comparable. The first data alteration concerned the variable *Stabilization Target*. When investigating the primary studies, we could not uncover any estimate that could be used as a stabilisation target (the estimate was in three studies). The variable, however, appears in the dataset by Kuik *et al.* (2009). We decided to apply the best guess estimate based on the relative size of the *Baseline* variable and other characteristics stated in each paper to complete the observations. That way, we could at least approximately analyse its relationship with MAC estimates.

Next, we converted the data for *Stabilization Target* to one metric - CO_2-eq concentration, measured in ppm. Finally, the estimates of MAC needed to be normalised to one currency and one year - 2020 Euro (specifically, $EUR^{2020}/tCO_2 - eq$). For this transformation, we utilised market exchange rates from various currencies to EUR, consumer price index (CPI) from the OECD, and molecular weights to convert units to a common dimension CO_2-eq (Yahoo, 2022; OECD, 2022; Brander, 2021).

To accurately conduct a meta-analysis, we need to include some indicator of uncertainty (usually standard error) for the estimates we collect. Unfortunately, none of the primary studies included standard errors when presenting their results. There were no indicators of uncertainty in the dataset by Kuik *et al.* (2009); thus, we assume it is not a common practice in this area of study. To resolve the issue, we followed Havranek *et al.* (2015) and constructed a standard error approximation. Their technique works only for papers with more than one estimate, meaning we added a measure of uncertainty to 50 out of 59 studies. Furthermore, the technique works better the more estimates the study contains. Therefore, the resulting standard errors should be handled with caution since most of the studies in the dataset contained only two observations. To utilise this method, we assume that the estimates in each study are normally distributed. Then, we calculate the median of the estimates, and the difference between the 50$^{\text{th}}$ and the 16$^{\text{th}}$ percentile serves as an estimate for standard error. Even though this technique is initially meant to complete just a few missing observations rather than filling each value, this was the best option the literature offered. Weir *et al.* (2018) confirmed the validity of this technique when they concluded that approximating the missing standard deviations minimises loss of precision and overall performs better than omitting trials.

The need for another adjustment appeared when we merged the two datasets. While the new dataset worked with the years 2030 and 2050, as these were the years that appeared most in the search, Kuik et al. worked with the years 2025 and 2050. Intending to have a robust dataset, we decided to join the years 2025 and 2030 to one variable (MAC2030). We believe this alteration does not significantly affect the results. Later in the analysis, we conduct robustness checks to confirm this assumption.

Variable	Description
Response variables	
MAC 2030	Marginal Abatement Costs of Greenhouse Gases emissions in 2030
MAC 2050	Marginal Abatement Costs of Greenhouse Gases emissions in 2050
Study specific variables	
Publication Year	The year the paper was published
Google Citation	Number of citations in Google Scholar
No. of estimates	Number of estimates in a particular study
Kuik	Data from Kuik *et al.* (2009); time differentiation of studies (*dummy*)
Empirical setting	
GHG Emissions	Study analyses overall GHG emissions (*dummy*)
Agriculture	Study analyses emissions from agriculture (*dummy*)
Energy sources	Number of energy sources
Regions	Number of regions
Methodology	
LEAP model	Study utilises the LEAP model (*dummy*)
Top-down	Study utilises the top-down approach (*dummy*)
Technological Specific	
Intertemporal Optimisation	Study includes a specification of intertemporal optimisation (*dummy*)
Carbon Capture and Storage	Study includes a specification of carbon capture and storage (*dummy*)
Multigas	Study examines a multigas policy (*dummy*)
Induced Technological Change	Study includes a specification of induced technological change (*dummy*)
Target	Stabilisation target
Baseline 2030	Projected Baseline in 2030
Baseline 2050	Projected Baseline in 2050

Table 4.2: Description of the regression variables

The last adjustment included taking natural logarithms of the MAC 2030 and MAC 2050 variables. The advantage of this log-level transformation is that the coefficient resulting from regression can be interpreted as semi-elasticities; a one-unit change in the independent variable indicates the corresponding percentage change in the MAC. Before taking logarithms of MAC, we added 850 to each observation to correct for negative values. Further in this chapter, when plotting selected data characteristics, the MAC variables appear in both absolute values and logarithms. The choice was made to best illustrate the particular characteristic. Starting the next chapter, we will use the terms MAC, MAC elasticity, and log(MAC) interchangeably to address the logarithm of MAC. Nonetheless, the conclusions apply to MAC in absolute values, too. To obtain the true effect of MAC, we reverse the logarithm procedure and deduct 850.

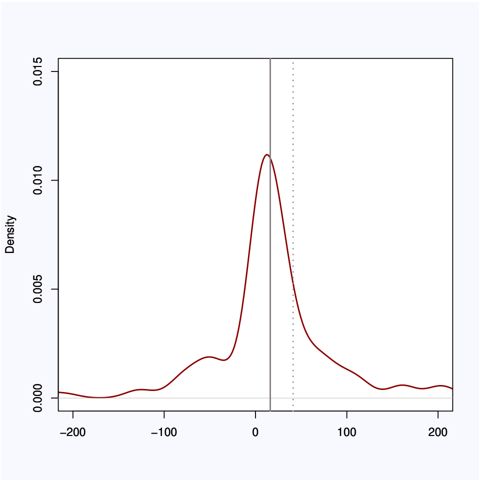
Before we could start investigating the data, we had to clean and scrutinise the whole dataset. We had to make sure all variables qualified to be included in the dataset, there were no typos, and we carried out winsorising. All the dummy variables have a decent variability (none of their means were close to 0 or 1). It should be noted that the variability of the dummy variable CGE comes mainly from Kuik's dataset since there were only a few papers from the last 17 years in the search query. The mean of the CGE variables from the dataset alone was 0.1, but after adding Kuik's data, the mean shifted to 0.47, so we decided to keep the variable in the dataset. Even after clearing the data, some outliers still remained, for which we utilised winsorisation - 2.5% from each side.

4.3 Summary Statistics

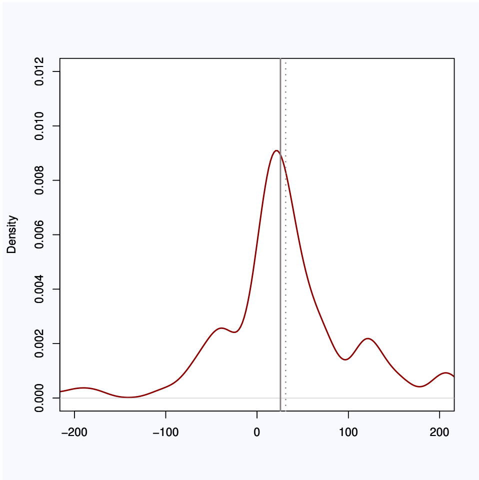
We obtained the estimates for MAC in 2030 and 2050 from each selected study and additional characteristics that serve as regression variables. As expected, most of the collected variables are dummy - gaining the value 1 if the characteristic is present, 0 otherwise. The final collection of variables used in the analysis can be found in Table 4.2.

MAC	2030	2050
No. of observations	135	107
Minimum	-266.20	-627.55
Median	16.16	25.64
Mean	41.21	26.42
Maximum	556.92	528.74
Standard Deviation	155.96	155.85

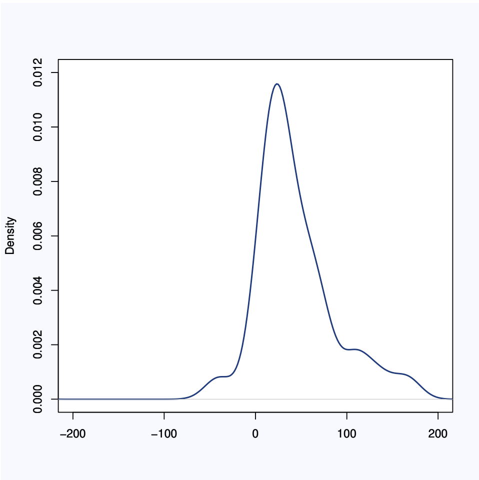
Table 4.3: Summary statistics of the MAC variables



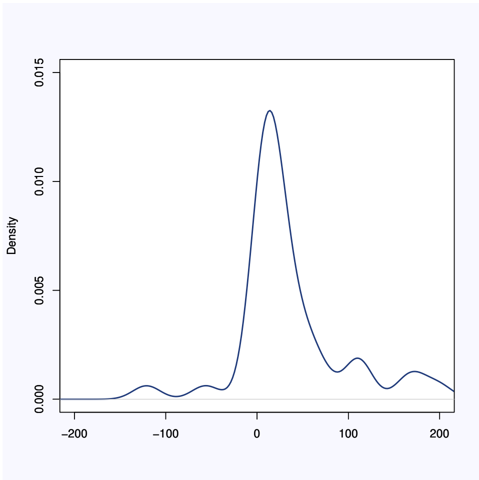
(a) MAC2030 All estimates



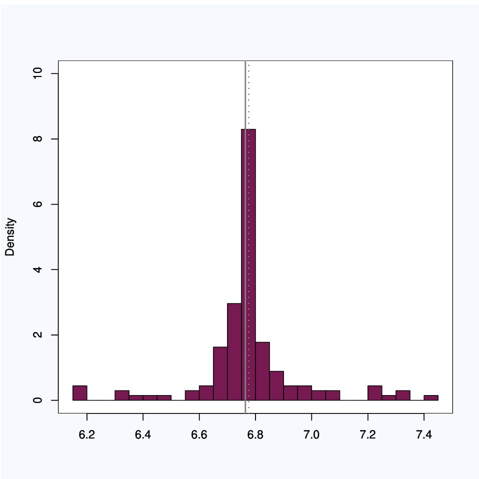
(b) MAC2050 All estimates



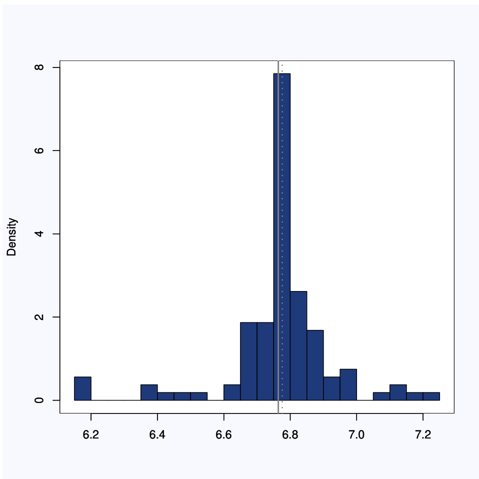
(c) MAC2030 Study level



(d) MAC2050 Study level



(e) log(MAC2030)



(f) log(MAC2050)

Figure 4.1: Kernel density plots and histograms for MACs

Note: The solid and dashed lines depict median and mean, respectively.

The Table 4.3 displays summary statistics for the MAC variables (in absolute values). The simple mean larger than the median, together with the kernel density plot in Figure 4.1, suggests skewness to the right for both MACs. For the regression, they will be transformed to logarithms to closer resemble a normal distribution. The histogram in Figure 4.1 proves that the skewness is reduced when working with logarithms. Also, the mean and median came closer together in both years.

Due to the lack of uncertainty measures reported in primary studies, this study distinguishes between two data groups. The first (all estimates) dataset includes all estimates collected from primary studies with the approximated standard error that is identical for all estimates from one study. The second dataset (study-level group) includes study-level medians and standard errors approximated using the method of Havranek *et al.* (2015). There are 43 study-level medians (and standard errors) for MAC 2030 and 34 medians for MAC 2050. We can see kernel density plots side-by-side for both MAC estimates (absolute values, no logarithm yet) versus study-level medians in the panels (a)-(d) in Figure 4.1. We can see that the study-level estimates are more centred around the mean for both years, and extremes are less distinctive. The reason for this redistribution closer to the mean is easily explicable. We need to remember that there were no uncertainty measures in primary studies. Therefore, this shift occurs because the standard errors (and median) could be obtained only from papers presenting more than one estimate. The more estimates in a paper, the closer the median gets to a 'middle' value. Even when a paper with more estimates contains an extreme value, others balance this one out and together appear moderate. Papers that only present one estimate have no way of balancing any deviations. Furthermore, since they do not fall in the study-level group, the all estimates dataset shows more extremes. Further in the analysis, we conduct a robustness check to confront these two data groups to reveal which dataset should be used for the analysis.

The Table 4.4 gives an overview of summary statistics for all explanatory variables. We collected 153 observations for most of the variables (especially the dummies). The variable *agriculture* shows the smallest variability from the dummies. We decided to keep it in the dataset since the mean of 0.1 still carries specific information and can lead to an insightful conclusion. Regarding the variables in an empirical setting, most studies focused on certain areas producing GHG emissions in one country (usually divided into a couple of regions). Only one study, Purohit & Höglund-Isaksson (2017) works with global emissions.

The diversity in the dummy variables indicates a great variety in the focus of primary studies. The pattern in papers providing two estimates (which is the majority of the papers) is the following: the authors first present the estimate for MAC unaffected by any of the specific factors, the second estimate is then influenced by one or more of these specific factors, such as CCS, Multigas, ITC, and others.

Variable	Observations	Mean	Standard Deviation
Publication Year	153	2013	5.90
Google Citation	153	65.42	85.10
No. estimates	153	4.85	3.85
Kuik	153	0.41	0.49
GHG Emissions	153	0.52	0.50
Agriculture	153	0.10	0.31
Energy sources	147	6.92	6.12
Regions	149	15.98	31.38
LEAP	153	0.21	0.41
Top-down	153	0.47	0.47
Intertemporal Optimisation	153	0.39	0.49
CCS	147	0.41	0.49
Multigas	153	0.52	0.50
ITC	153	0.24	0.43
Target	153	565.30	82.67
Baseline 2030	126	1.86	0.92
Baseline 2050	98	2.15	1.05

Table 4.4: Summary statistics of the explanatory variables

Forest plots in Figure 4.2 and 4.3 serve as a visual representation of collected estimates. The figures show how heterogeneous and different are the MAC estimates both within and across the studies. We can see that the older data from Kuik *et al.* (2009) are more compact and narrow than the newer data. Studies published after 2006 show noticeably more heterogeneity—the reason could be twofold. Primary studies from Kuik *et al.* (2009) worked with common research questions and assumptions. Additionally, all the papers were published in a single year, while the newly collected data covers 17 years of research and broader areas of GHG emissions.

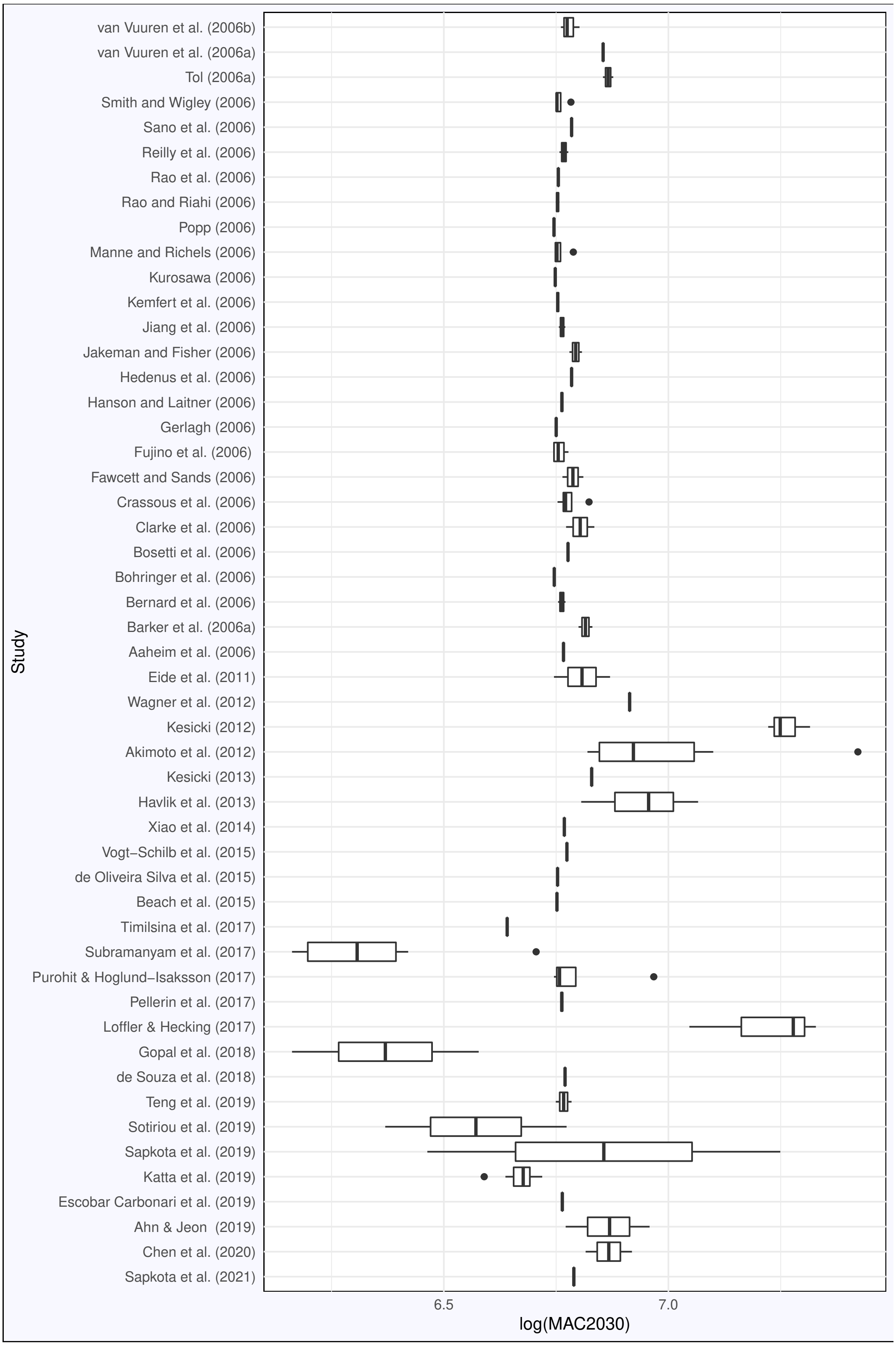


Figure 4.2: Forest plot of log(MAC2030)

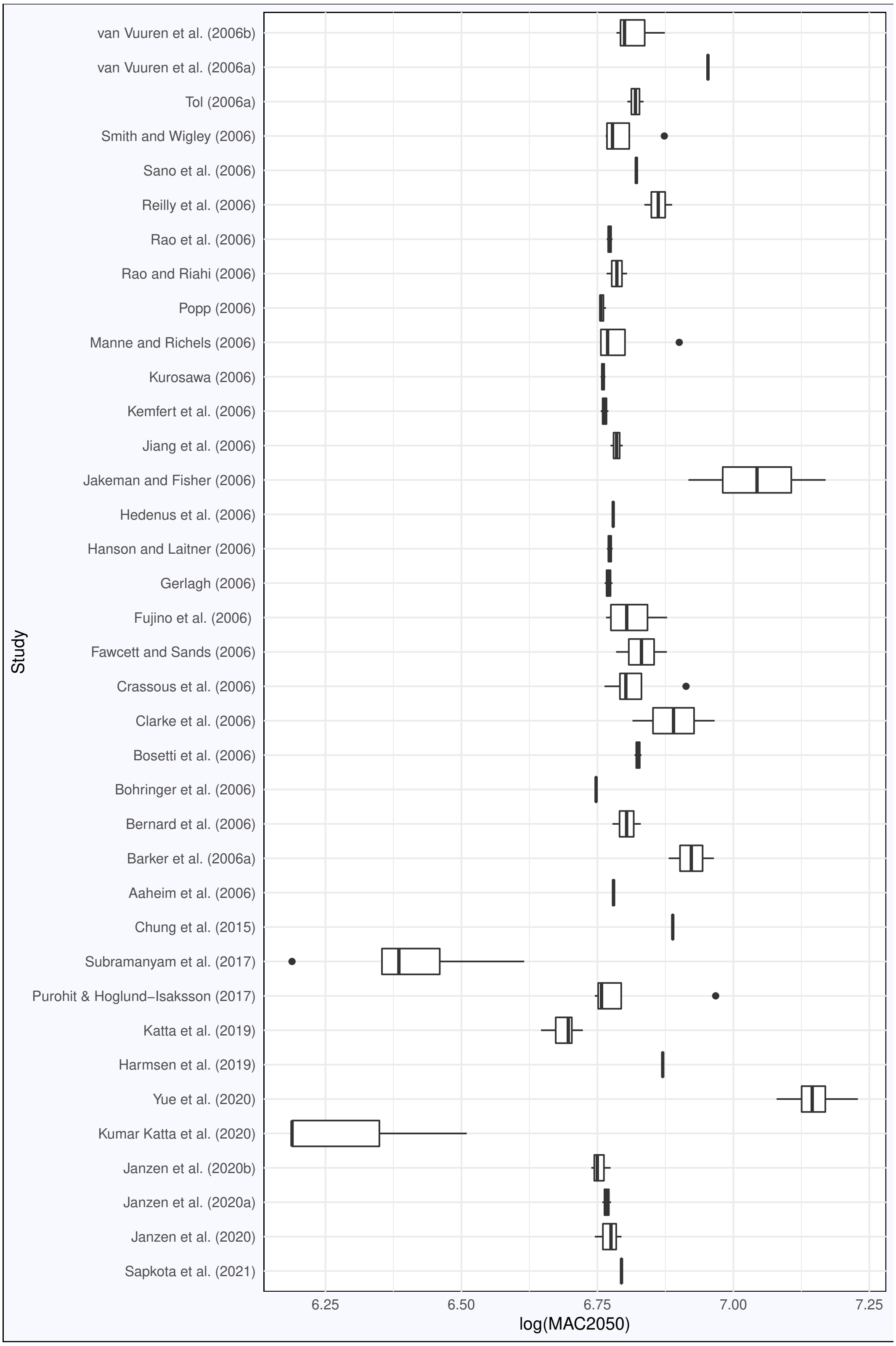


Figure 4.3: Forest plot of log(MAC2050)

Chapter 5

Meta-Analysis

Meta-analysis consists of rigorous quantitative methods to review, clarify underlying associations, evaluate, and draw valid statistical inferences from empirical research. For this purpose, we collected the estimates from selected primary studies and utilised a computing software R. We analysed various modelled estimates of the MAC and examined their dependence on the models' key assumptions and structural characteristics. To test this dependency, we constructed a meta-regression model with the MAC as a dependent variable. The MAC variable is assumed to be a linear function of a set of explanatory variables and a random error.

Generally speaking, meta-analysis should reveal an effect size which could be summarised across all primary studies. The effect size stands for a metric quantifying the relationship between two subjects. It captures both magnitude (absolute value) and direction (positive or negative) (Harrer, 2022). This chapter describes the general methodology when conducting meta-analysis and searching for the effect size. The following two chapters are dedicated to the meta-analysis methods.

5.1 Methodology

Conducting a meta-analysis essentially translates to evaluating policy from a more general perspective. The method systematically evaluates and compares additional information from previous research, experience, and quantitative results. Bergh *et al.* (1997) explain that in evaluating environmental costs, we are interested in "estimates of the monetary valuations of particular environmental costs". In such cases, the meta-analysis can be a valuable tool for identifying "indicators of central tendency in previous studies".

Havránek *et al.* (2020) provide us with guidelines for conducting meta-analysis. Aside from presenting us with a selection of studies employing modern techniques, they list all the necessary components that economic meta-analysis should contain. Generally, the meta-analysis regression analysis takes on the following form:

$$MAC_i = \beta_0 + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{5.1}$$

In our case, i = 2 for two MAC estimates, β_0 is constant, β_j represents a vector of coefficients, X_{ij} stands for a subset of all explanatory variables from n collected studies, and ϵ_i denotes a disturbance term.

The following two chapters present the empirical part of the meta-analysis. First, we examine publication bias using the Funnel Plot, the Meta-Analysis Regression, and the Caliper Test. Publication bias appears when the probability of a paper being published depends on the statistical significance of its results. Second, we focus on heterogeneity in the dataset, or other words, variation in outcomes between studies. We employ the Bayesian Model Averaging and the Frequentist Model Averaging for this analysis. Along the way, we present robustness checks that determine which data should be used for the meta-analysis and the strength of the results.

Compared to the thesis proposal, we decided to thoroughly examine which specifications affect the MAC estimate and find evidence for publication bias in the literature. Estimating the MAC ranges for alternative stabilisation targets as proposed did not align with the actual direction of the analysis. Nonetheless, with respect to how similar the results of our meta-analysis are to those conducted by Kuik *et al.* (2009), we expect the ranges to be approximately similar.

Chapter 6

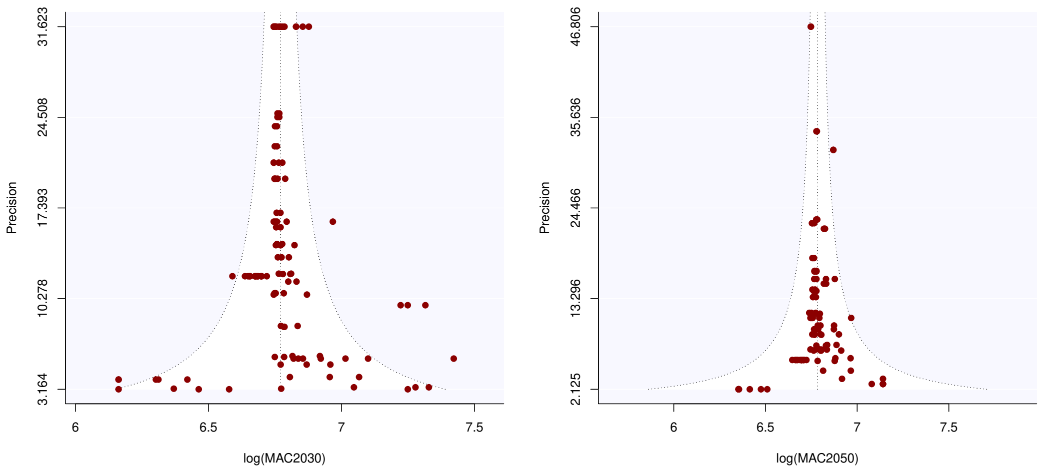
Publication Bias Analysis

Publication bias occurs when the chance of getting the study published (or even submitted) is affected by its results. In his "file drawer problem", Rosenthal (1979) introduces the idea that we cannot tell how many studies were conducted but never reported. The accepted view is that a study is more likely to be published when its results are statistically significant or confirm its hypothesis. However, even statistically insignificant studies contribute to the real outcome. It simply means that some conducted studies are inevitably missing in the dataset, and they are usually the ones with unfavourable findings. Even when there are meta-analysis techniques that help us find an unbiased estimate of the average effect size, if the dataset itself is distorted, we might not be able to find the 'true' effect reflecting reality.

We use several procedures to reduce (or at least reveal) the effect of publication bias (and other reporting biases). Unfortunately, the search for the so-called grey literature (studies not published in prominent journals, unpublished papers, dissertations, and others) goes beyond the scope of this diploma thesis. Nevertheless, this chapter presents statistical methods to examine the presence of publication bias. These procedures can not reveal the bias directly but may indicate it by specific properties in the data (Harrer, 2022).

6.1 Funnel Plot

A funnel plot provides a visual tool to inspect publication bias and heterogeneity. The assumption is that studies with large standard errors have larger effect sizes than those with lower standard errors (which have smaller effect sizes and might never be published). The funnel plot shows the observed effect sizes (usually elasticities) on the x-axis and their standard errors (or other precision measures) on the y-axis. If there is no publication bias in the sample, the data points should form an upside-down funnel (Anzures-Cabrera & Higgins, 2010).



((a)) log(MAC2030) ((b)) log(MAC2050)

Figure 6.1: Funnel plots

Since there were no precision measures in the primary studies and we had to approximate them, the dataset for the funnel plot contains only 126 elasticities for MAC2030 and 103 elasticities for MAC 2050. The plot for 2030 resembles an upside-down funnel; the points at the top are close to the sample mean, while at the bottom, the points are more spread out to both sides of the plot. Even though the plot is not strictly symmetrical, there are only a few outliers. The outliers come from two studies that report estimates that are close to each other. This causes the approximated standard errors to be relatively small, and the elasticities, therefore, do not fall within the funnel structure. Overall, the plot shows quite symmetrical patterns that do not indicate strong publication bias. Further investigation is needed to obtain more specific information about the bias.

The situation for MAC 2050 is quite similar - most points are at the bottom of the plot, and there are fewer points at the top. The points are mainly in the first third of the plot, well within the indicated borders. At the bottom, the data appear more on the right side than the left, which could signal publication selection. The pattern is symmetrical, which is not indicative of publication bias. Most estimates lie very close to the sample mean, and there are only a few points further on the sides of the plot. These few studies that deviate from the middle of the plot could later strongly affect our results, and we should be aware of them. Shapes of both funnel plots reveal that the approximated standard errors carry information about a study's precision.

6.2 Meta-Regression Analysis - FAT-PET test

Meta-regression analysis (MRA) represents a more quantitative and objective method of detecting publication bias than the previously presented funnel plot. Its primary aim is to "model, estimate, and understand the excess variation among reported empirical results" (Stanley, 2005). MRA reveals the sensitivity of reported estimates, estimation techniques, econometric models, and other issues. On top of that, it can be used to identify publication bias and its potential effects.

Stanley (2005) encapsulates the initial equation of MRA linear tests as an analysis of a relationship between the reported effect and its standard error:

$$MAC_{i,j} = \beta_0 + \beta_1 * SE(MAC)_{i,j} + \epsilon_{i,j}, \tag{6.1}$$

where i and j stand for the i^{th} observation in the j^{th} study. When publication bias is absent in the sample, the observed effects are expected to vary around the 'true' effect (β_0), independently of standard error. We assume the observed effects to be heteroskedastic, hence the error term ϵ_i. The coefficient β_1 provides information about the potential magnitude of a publication bias in the sample.

Two tests to quantitatively examine the publication bias are employed: the Funnel Asymmetry Test (FAT) and the Precision Effect Test (PET). The FAT tests whether the sample is affected by publication bias (the null hypothesis $H_0 : \beta_1 = 0$, meaning there is no publication bias), whereas the PET assesses whether there is a non-zero true effect once the publication bias is corrected ($H_0 : \beta_0 = 0$, the mean value after correcting for publication bias is zero). The tests use standard error as a proxy for the amount of selection needed to achieve statistical significance. Studies that report higher standard error need to find proportionally larger effect sizes to be significant. If this statistical significance cannot be achieved by re-estimation, different model specifications, or data adjustment, we assume the study is not published. Therefore, we expect greater publication bias for studies with larger standard errors (Doucouliagos & Stanley, 2013).

To address potential within-study correlation, we cluster the standard errors at the study level wherever possible. Inspired mainly by Alinaghi & Reed (2018) and Stanley (2005), we use the estimators presented in the following subsections. We divide the tests (and corresponding estimators) into three categories by their design - linear, study variation, and non-linear. Estimators introduced by authors other than the two previously mentioned are stated alongside the corresponding method.

6.2.1 Linear Tests

OLS Estimator

The OLS estimator given by Equation 6.1 calculates the arithmetic mean of MAC across studies. It serves as a benchmark to compare with other meta-analysis estimators. The OLS assumptions are violated in our case because the error term is not constant and varies among individual studies. This leads to heteroskedastic results, and we have to use clustering of the standard errors at the study level.

Weighted Least Squares Estimators

The following estimators work with the Weighted Least Squares (WLS) estimation. It allows for correcting heteroskedasticity in the baseline regression and puts more weight on results with smaller standard errors. To remove any remaining heteroskedasticity, we cluster standard errors by study level.

The WLS estimations work with Equation 6.1, where the weight (ω) is applied to each component. The list below presents all WLS estimations used to reveal publication bias:

1. The **Estimator weighted by the inverse standard error (Precision)** uses standard error as the weight. This is a baseline framework of WLS for tackling heteroskedasticity in the sample.

2. Another **estimator** uses the inverse of **the number of estimates reported per study** (n) as the weight. The weight is selected to give each study the same possibility to affect the result and does not handicap those with one or few estimates. The estimator is called **Study** in the text.

Study Variations

Second group of estimates allows for between- and within-study variation. Here, the error term breaks down into study-level random effects (ζ) and estimate-level disturbances (ν):

$$MAC_{i,j} = \beta_0 + \beta_1 * SE(MAC)_{i,j} + \zeta_j + \nu_{i,j}, \tag{6.2}$$

1. **The Fixed effects (FE) estimator** allows for variation across studies and captures the similarities the observations have within a study. It assumes one true effect size (weighted average) across all studies. Nevertheless, the standard error of FE can be too large for estimates within the study with a little variability, and the estimate is then based on studies with many observations (which could be our case).

2. **The Between-effects (BE) estimator** allows for variation between studies and thus should have balanced weights across studies.

3. **The Random effects (RE) estimator** recognises that the effect can differ between studies due to heterogeneity. It assumes the unobserved variables to be uncorrelated with the observed ones and uses a weighting matrix of both within- and between-study variance.

IV Estimator

Lastly, the **IV Estimator** takes the inverse of the square root of the number of observations $\frac{1}{\sqrt{n_{i,j}}}$ as an instrument for the standard error SE_i. Using Instrumental Variables (IV) offers another way to remove heteroskedasticity. This specific setting introduces the instrument correlated with the standard error but uncorrelated with the error terms. Another advantage of this method is that it tackles endogeneity, while the previous methods only assume no correlation between estimates and standard errors.

There are two conditions an IV needs to satisfy to be considered a strong instrument: validity and exogeneity. The instrument should exhibit a high correlation with standard error and a low correlation with the error term. Unfortunately, this is not the case with this instrument since we cannot reject the null hypothesis that the instrument is weak. Additionally, the F-statistic from the first stage equals 1.2, which further confirms that the instrument is weak. The following section, therefore, proposes another method - p-uniform* estimate. Unfortunately, this method is not suitable for our dataset. The following section presents other non-linear tests to reveal the publication bias.

6.2.2 Non-Linear Tests

Non-linear tests serve to check the validity of the previously presented tests. Unlike linear tests, these do not expect the publication bias to be a linear function of the standard error. We can observe if the linear relationship holds on the funnel plot. The points at the top of the funnel are less likely to be affected by publication bias because they have a very small standard error and high significance (Stanley, 2005). Again, the standard errors are clustered at the study level.

1. **The P-Uniform*** method was first introduced by van Aert & van Assen (2018) as an improvement of the original p-uniform method for detecting publication bias. The method assumes that p-values are uniformly distributed at the mean effect size. Therefore, the estimated coefficient should equal the 'true' effect when testing the hypothesis. Unfortunately, the collected data are not suitable for this technique since it works with variance among standard errors. Because we approximated standard errors, there is always one value per the study, no matter how many estimates the study provides.

2. **The Stem-Based method** is a non-parametric technique introduced by Furukawa (2019): "The estimate uses the studies with the highest precision, which correspond to the "stem" of the funnel plot, to estimate a bias-corrected average effect". The method selects only the most precise estimates that minimise the overall standard error. It is a relatively conservative procedure for detecting publication bias. The model was estimated using the R-code from the GitHub repository by Furukawa (2021).

3. **The TOP 10 method**, as the name suggests, takes only the top 10% of the most precise estimates. Then, this 10% is averaged and is considered the true effect, disregarding the rest of the dataset. The method is introduced in Stanley *et al.* (2010).

4. **The Weighted Average of Adequately Powered (WAAP)** by Ioannidis *et al.* (2017) selects estimates with suitable statistical power. It uses

a weighted average of these powerful estimates and returns an empirical lower bound for the bias.

5. **The Endogeneous Kink** introduced by Bom & Rachinger (2019) assumes that the most precise estimates have no bias. Then, it fits a linear regression of the primary estimates on their standard errors and adds a kink at the precision level where the studies are not likely to be reported.

6. **The Selection model** assumes that authors are less likely to publish their findings when the t-statistics are too high. Andrews & Kasy (2019) propose to divide the sample into several subsets bounded by the t-statistics thresholds. Subsequently, the selection model calculates how many studies in each subset are over- or under-represented in the primary literature and re-weights them.

6.3 Interpreting Results

Before we begin evaluating the FAT-PET tests, we present guidance on revealing publication bias from Doucouliagos & Stanley (2013). First, the FAT reveals whether the estimates are influenced by publication bias. The null hypothesis is that there is no publication bias. In other words, when β_1 is statistically significant, the publication bias is present. The extent of publication bias is as follows :

1. "If FAT is statistically insignificant or if $|\hat{\beta}_0| < 1$, then selectivity is 'little to modest'.

2. If FAT is statistically significant and if $1 \leq |\hat{\beta}_0| \leq 2$, then there is 'substantial' selectivity.

3. If FAT is statistically significant and if $|\hat{\beta}_0| > 2$, then there is 'severe' selectivity" (Doucouliagos & Stanley, 2013).

Consequently, the PET exposes the non-zero true effect of estimates when the publication bias is corrected. It tests whether $\beta_0 = 0$ and whether it is statistically significant. When β_0 is not significant, we can assume there is insufficient evidence to expect any empirical effect (Alinaghi & Reed, 2018). We can rely on Precision Effect Estimate with Standard Error (PEESE) when this situation arises. We can see from the test results that this further procedure will not be necessary.

6.3.1 Results for MAC 2030

The Table 6.1 shows that only one test detected publication bias at a 1% significance level - the estimator weighted by precision. The resulting estimate would imply that the publication bias is 'severe' in this case. All the other tests show non-significant FAT estimates, which indicates 'little to modest' selectivity. In addition, four of these are very close to zero in absolute value, so we assume no or mild publication bias to be present.

MAC 2030				
A: OLS, WLS				
	OLS	Precision	Study	
Standard Error	-0.32	2.19**	0.40	
(Publication bias)	(1.58)	(0.81)	(1.80)	
Constant	6.78***	6.78***	6.79***	
(Effect beyond bias)	(0.03)	(0.004)	(0.01)	
Observations	126	126	126	
Studies	43	43	43	
B: Study Variations, IV				
	FE	RE	BE	IV
Standard Error	-10.12	-0.37	-10.11	12.56
(Publication bias)	(35.6)	(1.71)	(35.58)	(32.13)
Constant	6.78***	6.78***	6.78***	6.54***
(Effect beyond bias)	(0.03)	(0.02)	(0.03)	(0.60)
Observations	126	124	126	126
Studies	43	41	43	43
C: Non-Linear Estimates				
	Stem	TOP10	WAAP	
Mean beyond bias	6.75***	6.83***	6.78***	
	(0.006)	(0.03)	(0.003)	
Observations	80	9	126	
	Kink	Selection		
Mean beyond bias	6.78***	6.78***		
	(0.003)	(0.02)		
Observations	126	125		

Significance codes: p<0.001 '***', p<0.01 '**', p<0.05 '*', p<0.1 '.'
Clustered standard error in parenthesis

Table 6.1: FAT-PET tests - results for log(MAC2030)

Looking further at constants from PET, we see very similar results for all estimators. The linear tests reveal a statistically significant estimate with a mean value of 6.75, revealing the true effect of the log(MAC2030). The value closely corresponds to the sample mean, 6.78.

The non-linear tests reveal similar outcomes. Even when inputting only a certain group of data sharing similar characteristics or statistical strength, the true effect of the log(MAC2030) lies between 6.75 and 6.83. These values are very close to the sample mean. Putting together all 'true' effect estimates from the tests, we arrive at the mean of 6.78. Thus, we conclude that publication bias is very small in the sample.

MAC 2050				
A: OLS, WLS				
	OLS	Precision	Study	
Standard Error	-1.16	0.27	-0.44	
(Publication bias)	(0.83)	(3.84)	(0.69)	
Constant	6.81***	6.78***	6.79***	
(Effect beyond bias)	(0.02)	(0.04)	(0.02)	
Observations	103	103	103	
Studies	34	34	34	
B: Study Variations, IV				
	FE	RE	BE	IV
Standard Error	2.97	-0.31	4.19	-3.85
(Publication bias)	(2.46)	(1.01)	(2.87)	(4.65)
Constant	6.77***	6.79***	6.77***	6.91***
(Effect beyond bias)	(0.02)	(0.02)	(0.02)	(0.14)
Observations	103	101	103	103
Studies	34	32	34	34
C: Non-Linear Estimates				
	Stem	TOP10	WAAP	
Mean beyond bias	6.76***	6.40***	6.77***	
	(0.03)	(0.08)	(0.004)	
Observations	86	8	103	
	Kink	Selection		
Mean beyond bias	6.76***	6.79***		
	(0.002)	(0.04)		
Observations	86	102		

Significance codes: p<0.001 '***', p<0.01 '**', p<0.05 '*', p<0.1 '.'
Clustered standard error in parenthesis

Table 6.2: FAT-PET tests - results for log(MAC2050)

6.3.2 Results for MAC 2050

Results for MAC 2050 in the Table 6.2 show a similar outcome. This is not a surprise since most of the observations come from the same papers, and we expect consistent behaviour from the authors. The FAT returned all estimates as non-significant, suggesting a 'little to modest' extent of publication bias in the sample. On the other hand, the true effect is statistically significant for all linear tests. This reveals an empirically backed effect of MAC 2050, the mean of all resulting estimates is 6.8, very similar to MAC 2030. Again, the non-linear tests further prove the true effect and its magnitude - the mean of all test results is 6.76. The sample mean for MAC 2050 is 6.77, which concludes that the publication bias in the sample is very small.

6.4 Wild Bootstrapping

We can use confidence intervals obtained via wild bootstrapping as a robustness check for OLS. This method is beneficial when conventional methods are unreliable due to violation of large-sample assumptions (in our case, a weak instrument) and does not assume identical and independent distribution of error terms. To obtain the results, we use the R package `fwildclusterboot` by Joshi *et al.* (2022) instead of Stata's `boottest`. Besides this practical change, the method stays the same as introduced by Roodman *et al.* (2019): we use a model estimated by OLS with a study-level cluster. The technique groups observations into several clusters, takes residuals from a null model (model fitted without any additional covariates), and randomly assigns weights constant within each cluster. These re-weighted residuals are then used to calculate new outcome variables, which then help to calculate test statistics for each cluster. The result of this procedure is a confidence interval for both standard error and a constant in the regression.

MAC 2030

The procedure returned the confidence interval [6.75,6.82] for the constant. Based on the results from other methods in the meta-regression, we can assume that similar intervals would apply to other estimators, too. On the other hand, the interval for standard error [-3.27,2.58] does not appear to be statistically significant since it passes over zero. This supports the results from the meta-regression analysis. When we reverse the logarithm and subtract 850 to reveal the interval for absolute value of the MAC resulting from wild bootstrapping, we obtain [3.21,61.42].

MAC 2050

Wild bootstrapping also confirms our findings for MAC 2050. The constant's confidence interval is [6.80,6.83], in line with the previous results. The interval for standard error is negative, [-2.17,-1.08]. Unlike MAC 2030, the whole interval lies below 0. Together, these intervals bring no further evidence of publication bias. The interval for the constant in absolute values is [44.00,74.17].

6.5 Caliper Test

In addition to the previously presented tests, we performed Caliper test by Gerber (2008). This test reveals the so-called "type II" publication selection - choosing statistically significant results, regardless of their direction. This leads to excess variation, resulting in large t-values being over-reported. Bruns *et al.* (2019) explain the logic of the test: "...the probability of a t-value being just above a given threshold or just below this threshold is 0.5 if the interval around the threshold is chosen sufficiently small." This approach relaxes the exogeneity assumption and eliminates the need for assumptions about the distribution of the t-values. The test reveals whether differences between statistical significance and insignificance affect the probability of reporting the study. Because the t-statistics for our data are distributed wider than is common in literature, we decided for caliper values -3.96 and 3.96 instead of the conventional -1.96 and 1.96. Similar modification can be found in Gechert *et al.* (2021). The modification allows more observations to enter the model, and the results are then more robust (at the cost of losing precision).

MAC 2030

For MAC 2030, the test starts revealing significant results from caliper of 0.4. The resulting ratios are consistent through narrowing the caliper sizes. The test reveals that less than half of the estimates are significant above the 5% threshold for the negative and positive estimates. The estimates around zero are not statistically significant until size 0.8 but show around 45% of estimates being significant. From the caliper test, we cannot confirm any preference of researchers for the direction of estimates, and there is a slight preference for non-significant results over the significant ones.

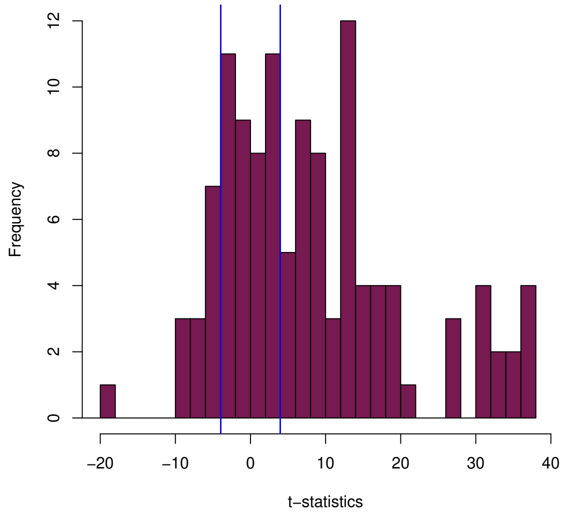


Figure 6.2: Distribution of t-statistics for log(MAC2030) estimates
(only a section is displayed for better illustration, the test was performed on the whole sample)

Caliper size	**0.2**	**0.4**	**0.6**	**0.8**
Threshold: **-3.96**	0.33. (0.17)	0.31* (0.13)	0.37** (0.11)	0.40*** (0.10)
Observations	*9*	*13*	*19*	*25*
Threshold: **0**	0.25 (0.25)	0.43 (0.20)	0.38. (0.18)	0.44* (0.18)
Observations	*4*	*7*	*8*	*9*
Threshold: **3.96**	0.56* (0.18)	0.44** (0.13)	0.40** (0.11)	0.38** (0.10)
Observations	*9*	*16*	*20*	*24*

The table shows the share of estimates that are above the critical value of t-statistic in 0.2, 0.4, 0.6, and 0.8 caliper; for example, the coefficient 0.40 means that 40% of the negative estimates are significant and 60% of negative estimates are non-significant. Standard error in parenthesis.
Significance codes: p<0.001 '***', p<0.01 '**', p<0.05 '*', p<0.1 '.'

Table 6.3: Caliper test - results for log(MAC2030)

MAC 2050

The caliper test confirms that non-significant values are preferred over the significant ones for the negative estimates. This result is statistically significant at a 5% level, even for the narrowest caliper size. Positive values repeat the same pattern, but the findings are slightly less significant. Again, non-significant values are preferred, which is visible on the jump after the 3.96 threshold on the histogram in figure 6.3. In contrast, for the values around 0, no conclusion can be made. There are not enough observations to provide statistically significant results.

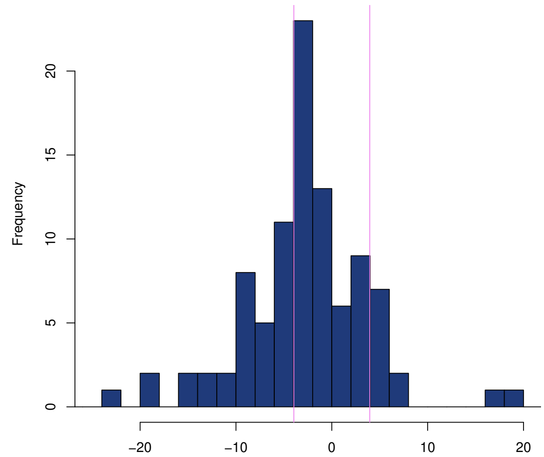


Figure 6.3: Distribution of t-statistics for log(MAC2050) estimates (only a section is displayed for better illustration, the test was performed on the whole sample)

Caliper size	**0.2**	**0.4**	**0.6**	**0.8**
Threshold: **3.96**	0.56* (0.18)	0.44** (0.13)	0.40** (0.11)	0.38** (0.10)
Observations	*9*	*16*	*20*	*24*
Threshold: **0**	1 (NA)	0.5 (0.5)	0.25 (0.39)	0.2 (0.37)
Observations	*1*	*2*	*4*	*5*
Threshold: **-3.96**	0.39** (0.12)	0.48*** (0.10)	0.32*** (0.08)	0.33*** (0.07)
Observations	*18*	*25*	*37*	*46*

The table shows the share of estimates that are above the critical value of t-statistic in 0.2, 0.4, 0.6, and 0.8 caliper; for example, the coefficient 0.40 means that 40% of the negative estimates are significant and 60% of negative estimates are non-significant. Standard error in parenthesis.
Significance codes: p<0.001 '***', p<0.01 '**', p<0.05 '*', p<0.1 '.'

Table 6.4: Caliper test - results for log(MAC2050)

Overall, we did not find strong evidence of publication bias in the sample for MAC 2030 or MAC 2050. The results show the true effects to be very close to the sample means for both years. The resulting MAC for 2030 in absolute value equals 32.32 EUR/tCO_2-eq, and for 2050 the cost almost doubles to 59.09 EUR/tCO_2-eq. Although the costs are similar to those that Kuik *et al.* (2009) worked with, we expected the newer studies to raise the value since they work with higher costs. Nevertheless, as we show later in the robustness check, the new studies suffer from higher publication bias than those from 2006. Another argument for expecting higher costs is CPI. The results are backed by the funnel plot, a selection of linear and non-linear FAT-PET tests, the wild bootstrapping, as well as the caliper test.

6.6 Robustness Check

The final part of this chapter is dedicated to robustness checks to validate the datasets used for the analysis. Robustness checks are based on various subsamples that yield different results, and we should be aware of this difference and examine it.

6.6.1 Study-Level Medians Dataset

The robustness check in this chapter concerns the MAC estimate, its median, and its standard error. As described in Chapter 4, the standard errors were approximated for studies which reported more than one observation. These constructed standard errors then lead to two ways to input the MAC into the meta-analysis. The first one is study-level, where we take the median from each study and the corresponding standard error. This dataset is significantly smaller than the original, but the values do not repeat, and each study is represented just once. The second approach is to put the constructed standard error next to each study's observation. Even though this results in a more extensive dataset, the standard errors are identical within a study. To assess the consistency between these two datasets, we rerun the FAT-PET tests (and baseline BMA in the following chapter) with the study-level dataset.

Tables B.1 and B.2 (in the Appendix B) reveal the FAT-PET test results for the study-level median estimates. The first thing worth emphasising is the size of this sample - for MAC 2030, there are 43 medians, compared to 126 observations in the default dataset. That is almost a three times smaller sample which probably results in less robust results. The situation is similar for MAC 2050 - there are 34 studies (compared to 103 in the default dataset). The number of observations is sufficient for linear tests. Nevertheless, for non-linear tests that work with a sub-sample of data (Stem, TOP 10), only a handful of observations pass the requirements to be included in the calculation. Besides the sample size, the results in the robustness check appear to be similar to the previously conducted tests. For MAC 2030, only the estimator weighted by precision reveals publication bias which would be deemed as 'severe'. All the other tests show insignificant estimates pointing at 'little to modest' selectivity in the sample. The true effect is highly significant for all tests, and the mean is slightly higher - 6.79, which is closer to the sample mean. Results for the study-level median of MAC 2050 show a similar outcome. The standard error estimates are not significant, providing little proof of publication bias in the sample. The effect beyond publication bias returns significant estimates with an average value of 6.75, which is again closer to the sample mean than the previous result. In conclusion, both datasets provide similar results - evidence for small publication bias for either year, and the true effect is very close to the sample mean. Each of them could serve as a primary dataset for our analysis. We chose the log(MAC) dataset for the size sample to get robust results from non-linear tests.

6.6.2 MAC 2025 vs. MAC 2030

As previously described, we came across a complication when joining the dataset of new data with the dataset from Kuik *et al.* (2009). While we collected data for 2030, Kuik *et al.* work with 2025. We decided to join them together because we do not expect a significant difference between the two. To prove this assumption, we conduct a meta-regression analysis and baseline BMA (in the next chapter). The results from the robustness check are attached in Appendix B. The meta-analysis regression returns similar outcomes. The non-significance of the coefficient for publication error remains, and the true effect stays significant. The mean of all true effects corrected for publication bias is slightly lower - 6.77.

Chapter 7

Heterogeneity Analysis

The last chapter investigates possible sources of heterogeneity among estimates. Using the explanatory variables collected from primary studies, we show how the true effect of MAC would change if the studies employed different study designs. In other words, we reveal which of the collected variables could be the source of heterogeneity and if they have a significant effect on the MAC estimate. From the widespread MAC estimates in both positive and negative values, we can assume that the assumption the primary studies impose reflects the size of the final estimate. First, we list the explanatory variables that could capture systematic differences. Second, we describe the techniques of Bayesian Model Averaging (BMA) since it helps us reveal the possible sources of heterogeneity. Next, we include a robustness check utilising BMA with weights and BMA without standard error, Frequentist model averaging (FMA), and OLS. Finally, we present our results and comment on the findings.

7.1 Explanatory Variables

All sixteen control variables collected from primary studies are listed in Table 7.1, accompanied by their means, standard deviations (SD), and variance inflation factors (VIF). For a detailed description, please refer to Chapter 4. The variables are divided into four categories: study-specific, empirical setting, methodology, and technology-specific.

The first step before the model computation is to treat collinearity. We assume it is likely present since we have a relatively small dataset and many variables. The VIF can serve as an indicator of collinearity between variables.

The convention in meta-analyses is to remove variables, to have the VIF under 10 (Viechtbauer, 2010).

Variable	**Mean**	**SD**	**VIF** 2030	**VIF** 2050
Study-Specific				
Publication Year	2013	5.90	17.23	91.32
Google Citation	65.42	85.10	1.89	2.86
Kuik	0.41	0.49	18.27	92.74
Empirical Setting				
GHG Emissions	0.52	0.50	4.00	12.15
Agriculture	0.10	0.31	2.21	5.37
Energy sources	6.92	6.12	3.21	1.94
Regions	15.98	31.38	2.32	2.32
Methodology				
LEAP	0.21	0.41	3.15	10.48
Top-down	0.33	0.47	3.72	4.30
Technology-Specific				
Intertemporal Optimisation	0.39	0.49	1.71	2.26
Carbon Capture and Storage	0.41	0.49	1.67	1.57
Multigas	0.52	0.50	1.72	1.61
Induced Technological Change	0.24	0.43	1.40	1.37
Target	565.30	82.67	1.67	1.61
Baseline 2030	1.86	0.92	1.77	NA
Baseline 2050	2.15	1.05	NA	2.45

Table 7.1: Summary statistics of the explanatory variables

When calculating the VIF for our dataset, we revealed that the variable *Publication Year* is highly correlated with other variables. After its removal, the maximal VIF of the data was 10.2 for the *GHG Emissions* variable in the MAC 2050 regression. Therefore, the number of explanatory variables decreased to 15 in each regression. Luckily, the variable *Kuik* still serves us as a time differentiation of studies (which was causing the collinearity).

7.2 Bayesian Model Averaging

The method of Bayesian model averaging (BMA) helps us indicate the best model to estimate which variables have a significant effect on the MAC outcome. The intuitive approach to regress the computed partial correlation coefficients on all the explanatory variables would be, in this case, very time consuming, and its results would be affected by inflated standard errors and missing true specifications of the model. The BMA helps us analyse heterogeneity even with the model uncertainty because it considers all possible models with different choices of covariates (Raftery, 1995).

First, we present the key equations of the BMA estimation. The term $p(M_k)$ stands for the prior probability that M_k is the true model. D represents the data, N the number of explanatory variables, and K the number of models, where $K = 2^N$. After observing the data, the posterior probability is derived from the extended Bayes' theorem:

$$p(M_k|D) = \frac{p(D|M_k)p(M_k)}{p(D)} = \frac{p(D|M_k)p(M_k)}{\sum_{m=1}^{K} p(D|M_m)p(M_m)}, \tag{7.1}$$

where the integrated likelihood of model M_k is as follows:

$$p(D|M_k) = \int p(D|\beta_k, M_k)p(\beta_k|M_k)d\beta_k. \tag{7.2}$$

The posterior model probability (PMP) indicates goodness-of-fit for model M_k, while the prior probability $pr(M_k)$ illustrates the author's prior beliefs regarding the probability of model M_k before observing the data (Zeugner & Feldkircher, 2015). Next, the BMA uses the posterior model probabilities to calculate the weighted posterior mean:

$$E(\beta_i|D) = \sum_{k=1}^{K} \hat{\beta}_{ik} pr(M_k|D), \tag{7.3}$$

and the weighted posterior variance (or standard deviation) for each explanatory variable:

$$Var(\beta_i|D) = \sum_{k=1}^{K} (Var(\beta_i|D, M_k) + \hat{\beta}_{ik}^2 pr(M_k|D) - E(\beta_i|D)^2. \tag{7.4}$$

Here, $\hat{\beta}_{ik}$ stands for the estimated regression coefficient for i^{th} variable in k^{th} model. Finally, Posterior Inclusion Probability (PIP) is calculated as the sum of the posterior model probabilities only of the models that include variable i:

$$PIP = \sum_{k=1}^{K} pr(M_k|\beta_i \neq 0, D). \tag{7.5}$$

The reader can find a detailed technical description of the BMA procedure in the documentation for the `BMS` package, which we utilised to work with BMA - Zeugner & Feldkircher (2015)

The PIP can be viewed as the probability that a given variable significantly affects the MAC variable. For example, the variable's PIP equal to 1 illustrates that all models utilise this variable. For the interpretation of PIP, we follow the structure by Kass & Raftery (1995). A variable has a significant effect when its PIP is above 0.5. Values of PIP above 0.5 are divided as follows:

- $0.5 < PIP < 0.75$ indicates a weak evidence of an effect,
- $0.75 < PIP < 0.95$ indicates a positive effect,
- $0.95 < PIP < 0.99$ indicates a strong effect,
- $PIP < 0.99$ indicates a decisive effect (Kass & Raftery, 1995).

The BMA method presented above is implemented in the following meta-regression equation:

$$MAC_{i,j} = \beta_0 + \beta_1 * X_i + \beta_2 * SE(MAC)_{i,j} + \epsilon_{i,j}, \tag{7.6}$$

where $MAC_{i,j}$ stands for the partial correlation coefficient, X_i represented matrix of explanatory variables, i and j stand for the i^{th} observation in the j^{th} study. The constant β_0 deems no interpretative power because it reflects the mean effect corrected for publication bias conditional on covariates X, while β_2 measures the magnitude and direction of publication bias.

Before the computation can begin, two computational issues need to be addressed. First is the challenge of computing integrals in the likelihood functions within the BMA process. Second is the excessively large model space that could be difficult to process for a personal computer. These obstacles can be solved by applying the Metropolis-Hastings algorithm of the Markov chain Monte Carlo method, which estimates only models with the highest PMP (Zeugner & Feldkircher, 2015). The BMA also requires the meta-analyst to specify the distribution priors over the parameter space (g) and the model space ($pr(M_k)$). The higher the parameter value prior, the more weight is put on data compared to prior beliefs. Since prior knowledge about the models is relatively small, we choose the so-called unit information prior (UIP), which sets g = N, as recommended by Hasan *et al.* (2016). We opted for a 'uniform model prior' (UMP) for the prior distribution, assigning an equal prior probability to each model. We do not use any weights for the baseline BMA, inclusion of weights in the BMA is in the robustness check.

7.3 Interpreting Results

The outputs of BMA analysis come in two forms - graphical and tabular. The figure illustrates the inclusion of variables in models. Different potential combinations of explanatory variables in columns are scaled by their PMP. The higher PIP for each variable, the higher the variables lies on the y-axis. The red-coloured cell (lighter in greyscale) indicates a negative coefficient direction in the regression, while the blue-coloured cell (darker in greyscale) illustrates a positive direction. The white cell means that the variable would be excluded from the particular model. The table provides precise results from the BMA analysis. It displays PIP for each variable that can be interpreted using the above boundaries.

Results for MAC 2030

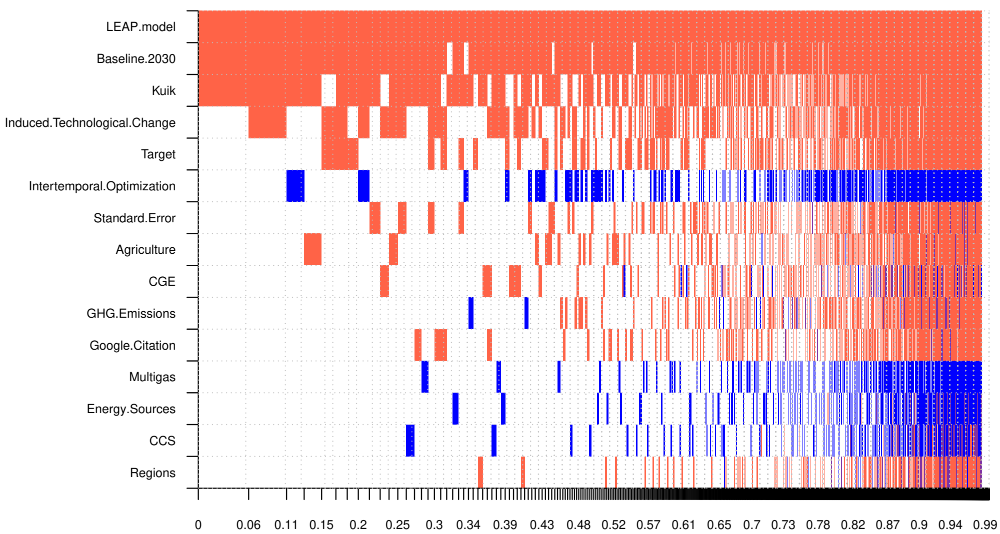


Figure 7.1: Model inclusion in Bayesian model averaging for MAC 2030
The response variable is the estimate of log(MAC) for 2030. The horizontal axis denotes cumulative PMPs and only shows the best 5000 models. We employ 1 million interactions and 3 million burn-ins to ensure convergence.

The baseline BMA output for log(MAC2030) reveals five variables that cross the threshold of PIP > 0.5: *GHG Emissions, Agriculture, LEAP model, Target,* and *Baseline*. A possible explanation and estimated effect are described in the section 7.5 that considers robustness checks.

Variable	PIP	Post. Mean	Post. SD
Standard Error	0.16	0.12	0.39
Study-Specific			
Google Citation	0.24	0.00	0.00
Kuik	0.18	-0.01	0.03
Empirical Setting			
GHG Emissions	**0.78**	-0.09	0.06
Agriculture	**0.59**	-0.07	0.07
Energy sources	0.18	0.00	0.00
Regions	0.17	0.00	0.00
Methodology			
LEAP model	**1.00**	-0.74	0.09
Top-down model	0.12	0.00	0.02
Technology-Specific			
Intertemporal Optimisation	0.29	0.02	0.03
Carbon Capture and Storage (CCS)	0.09	0.00	0.01
Multigas	0.11	0.00	0.01
Induced Technological Change (ITC)	0.23	-0.01	0.03
Target	**0.54**	0.00	0.00
Baseline 2030	**0.99**	-0.07	0.02

Variables with PIP > 0.5 in **bold** print.

Table 7.2: Coefficient estimates for log(MAC2030)

Results for MAC 2050

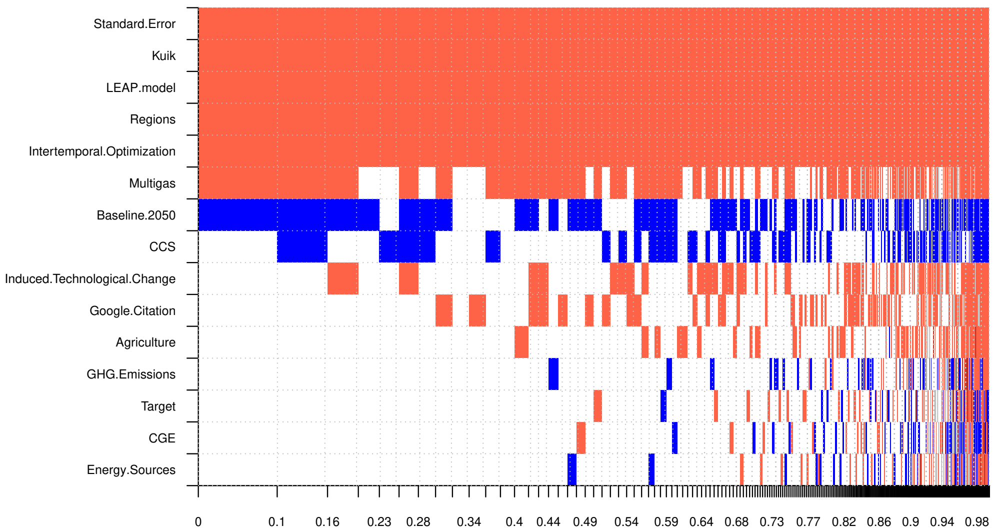


Figure 7.2: Model inclusion in Bayesian model averaging for MAC 2050
The response variable is the estimate of log(MAC) for 2050. The horizontal axis denotes cumulative PMPs, and only the best 5000 models are shown. We employ 1 million interactions and 3 million burn-ins to ensure convergence.

The BMA analysis for MAC 2050 produces seven variables explaining heterogeneity: *Standard Error, Kuik, LEAP, Intertemporal Optimisation, Regions, Multigas*, and *Baseline 2050*. It is worth noting that an analysis of MACs in two different years yields different variables that should explain the heterogeneity. We will further use a robustness check to verify whether these results hold. Further discussion on results is held after conducting robustness checks in part 7.6.

Variable	**PIP**	**Post. Mean**	**Post. SD**
Standard Error	**1.00**	-1.29	0.14
Study-Specific			
Google Citation	0.26	0.00	0.00
Kuik	**1.00**	-0.30	0.03
Empirical Setting			
GHG Emissions	0.11	0.00	0.01
Agriculture	0.15	-0.01	0.02
Energy sources	0.10	0.00	0.00
Regions	**1.00**	0.00	0.00
Methodology			
LEAP model	**1.00**	-0.43	0.05
Top-down model	0.10	0.00	0.01
Technology-Specific			
Intertemporal Optimisation	**1.00**	0.00	0.00
Carbon Capture and Storage (CCS)	0.39	0.01	0.02
Multigas	**0.67**	-0.03	0.02
Induced Technological Change (ITC)	0.30	-0.01	0.02
Target	0.11	-0.0002	0.00
Baseline 2050	**0.63**	0.01	0.01

Variables with PIP $>$ 0.5 in **bold** print.

Table 7.3: Coefficient estimates for log(MAC2050)

7.4 Robustness Check

The last section in this chapter is dedicated to robustness checks. We employ several methods to confront our results and assess their stability - BMA with different priors, BMA with weights, BMA with no Standard Errors, Frequentist Model Averaging, and OLS. The results are discussed in the next section.

BMA Variations

First, we run the BMA using the number of observations as weights. This approach gives each study equal importance in the analysis. The results from this modification are displayed in Table 7.4.

Next, we conclude a robustness check of running the BMA without standard errors. The construction of standard errors in our dataset poses a certain limitation for this study. We want to check how the results hold when the artificially created values are removed from the dataset. The results are displayed in Table 7.4.

Third, we change priors when running the BMA to 'BRIC' and 'random model prior', as recommended in Hasan *et al.* (2016). Results from this modification are displayed in the table B.5 in Appendix B.

Finally, we utilize R package `dilutBMS2` from Moser (2016) which contains a new prior for BMS which addresses collinearity. Concretely, it assigns more weights to models with no collinearity and vice versa. The results of the BMA using different priors are in Appendix B.

Frequentist Model Averaging

Frequentist Model Averaging (FMA) presents an alternative to baseline BMA since it addresses model uncertainty. The FMA analysis weights the variables with respect to their goodness-of-fit and parsimony and puts them in different combinations with each other. We follow the technique of model weighting by minimising Mallow's criterion presented in Gechert *et al.* (2021). The smaller this criterion, the higher weight assigned to the model. We utilise the orthogonalisation of covariate space as presented by Amini & Parmeter (2012). The drawback of this method lies in need to order the regressors before estimation - Hansen (2007) recommends ordering them by groups.

OLS

The last method is simple OLS weighted by the number of estimates but only includes variables with PIP higher than 0.5 (Matousek *et al.*, 2022). After a thorough analysis, we decided to include variables with significant effects from the weighted BMA instead of the baseline model.

Dataset Variations

The previous chapter introduced two robustness checks regarding the right data selection for the analysis. The different datasets are described in the section 6.6. We decided to analyse these different datasets with baseline BMA, too.

7.5 Results for MAC 2030

This section discusses the results from the baseline BMA analysis and robustness checks. Where applicable, our findings are contrasted with results from the literature, and the differences are explained.

Standard Error. The variable *Standard Error* and its (non)significance should reveal whether our conclusion about publication bias from the previous chapter holds even when we control for the context in which the model is estimated. From all model specifications, only weighted BMA marked standard error as significant. Again, this does not provide solid evidence of a strong publication bias. On the other hand, results for the new data indicate a strong negative publication bias. This suggests that the new data suffer more substantial publication bias than data from Kuik *et al.* (2009). For the whole dataset, however, we can confirm our findings from the previous chapter and assume mild publication bias for the MAC 2030.

GHG Emissions. The dummy variable *GHG Emissions* assigns value 1 for a study which works with overall GHG emissions rather than just one area (for example, traffic, agriculture, energy sector). The PIP of 0.78 indicates a positive effect, and the BMA results imply that models which work with overall GHG emissions report a lower MAC. This effect is confirmed by the BMA with different priors and the BMA addressing correlations. It could be because the studies working with overall GHG emissions do not scrutinise every

	BMA *weighted by no. of estimates*			BMA *no SE*			FMA			OLS		
	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	Coef.	SE	p-val.	Coef.	SE	p-val.
Standard Error	**1.00**	22.63	5.37	NA	NA	NA	-0.20	0.75	0.79	-1.39	0.71	0.06
Study Specific												
Google Citation	0.46	0.002	0.02	0.20	0.00	0.00	-0.0002	0.0002	0.32			
Kuik	0.12	0.03	0.12	0.40	-0.03	0.04	0.03	0.06	0.64			
Empirical Setting												
GHG Emissions	0.08	-0.005	0.07	0.43	-0.04	0.05	**-0.13**	0.05	0.02			
Agriculture	0.08	-0.004	0.08	0.19	-0.01	0.03	-0.11	0.06	0.07			
Energy sources	0.08	0.001	0.007	0.14	0.00	0.002	-0.002	0.004	0.61			
Regions	0.12	0.001	0.003	0.14	0.00	0.00	-0.001	0.001	0.32			
Methodology												
LEAP model	**1.00**	-1.7	0.28	**1.00**	-0.55	0.08	**-0.79**	0.09	0.00	**-0.25**	0.03	0.00
Top-down model	0.09	-0.01	0.09	0.17	-0.01	0.02	-0.01	0.05	0.82			
Technology-Specific												
Intertemporal Optimisation	0.11	0.02	0.08	0.4	0.03	0.04	0.05	0.04	0.21			
Carbon Capture and Storage (CCS)	0.09	-0.007	0.06	0.09	0.001	0.01	-0.03	0.03	0.46			
Multigas	0.18	0.04	0.02	0.09	0.001	0.01	0.02	0.03	0.46			
Induced Technological Change (ITC)	0.23	0.07	0.02	0.11	-0.003	0.02	-0.05	0.04	0.21			
Target	**1.00**	0.01	0.001	0.33	0.00	0.00	-0.0004	0.0002	0.05	**-0.001**	0.00	0.00
Baseline 2030	**1.00**	0.69	0.11	**0.83**	-0.04	0.02	**-0.07**	0.02	0.00	**-0.06**	0.02	0.01

Notes: The standard errors for FMA and OLS are clustered at the study level. Value 0.00 stands for number lower than 0.001. Variables with PIP > 0.5 in **bold** print.

Table 7.4: Potential sources of heterogeneity among MAC 2030

possibility of abatement. On the other hand, studies that focus on just one area have more space to dig deeper into the specific field and reveal every possible way of abatement. Therefore, papers studying GHG emissions report lower costs because they do not report all the possible ways of reduction.

Agriculture. The same logic holds for the dummy variable *Agriculture*, which applies when a primary study analyses GHG emissions originating from agriculture. The results imply that these studies report lower MAC elasticities by 7%. The results indicate that mitigating emissions from agriculture is not as costly as from other sectors. The reasons for that could be relative advancement of agriculture in mitigating emissions and subsidies that can help farmers transition faster to modern technology.

LEAP Model. Another dummy variable with significant PIP is the *LEAP model*. The baseline BMA analysis revealed that studies that work with the LEAP model report a MAC lower by 74%, and results from other model specifications suggest a similar value. Overall, the MAC estimates from the LEAP model are undervalued compared to estimates from other models. We should remember this when working with the LEAP model and its results in the future.

Target. This variable represents the stabilisation target for GHG emissions (the reader can find a detailed description of this variable in Chapter 2). The BMA suggests that studies with higher stabilisation targets tend to report a slightly lower MAC, but the magnitude of this relationship is almost negligible - an 0.02% higher MAC for an additional one unit in target, ceteris paribus. The direction of the effect of *Target* on the MAC is what we assumed in the hypothesis, but the value itself is smaller than we expected. Studies that expect the stabilisation targets to be higher do not have to lower the emissions as much, which logically corresponds to findings by Barker *et al.* (2006). Therefore, the costs to achieve a higher stabilisation target should be considerably lower than for lower targets. Nevertheless, the BMA addressing correlation confirmed both the direction and magnitude of the effect. Vogt-Schilb & Hallegatte (2014) found similar insufficient relationship between abatement costs and reduction targets.

Baseline 2030. The last variable that matters for heterogeneity is the *Baseline* describing time projections of emissions. Robustness checks confirm that models with a higher emissions baseline tend to report lower MACs. This result has the opposite direction of what we expected in our hypothesis. A higher emissions baseline means we expect the emissions to grow fast, and our mitigation efforts must be extensive (IPCC, 2014). The results suggest that the MAC does not have to be high, not even with a high emissions baseline. This is probably due to other factors prevailing.

Other variables were not deemed significant when explaining heterogeneity in the model. To our surprise, we found little evidence that the technology-specific variables are responsible for systematic differences in observed MAC elasticities.

7.6 Results for MAC 2050

The baseline BMA revealed seven variables that significantly explain heterogeneity in the model. We compare findings from baseline BMA with those from the robustness check and try to explain the resulting direction and magnitude of the effects found. Our results for the MAC 2050 closely resemble those from Kuik *et al.* (2009) - they also found emissions baseline, multigas, and the number of regions to be significant for model heterogeneity. In the dataset, there are more estimates from Kuik *et al.* (2009) for the year 2050 than 2030, which is probably why results for MAC 2050 follow theirs, while MAC 2030 does not.

Standard Error. Unlike MAC 2030, we found strong evidence for standard error explaining heterogeneity in the model. The baseline BMA suggests that studies with higher standard error report a lower MAC elasticity. Recalling the funnel plot, this can be caused by a handful of studies deviating from the mean (most of the studies lie very close to the sample mean). This demonstrates that negative publication bias is present in the literature even after controlling for estimation characteristics. The findings stay significant with all robustness checks. This result suggests that authors are more likely to report larger negative estimates than smaller ones and negates the findings for publication bias from the previous chapter.

Kuik. Following significant variable marks primary studies collected by Kuik *et al.* (2009). The findings hold for most robustness checks, which suggests that studies collected by these authors tend to report smaller MAC estimates.

	BMA *weighted by no. of estimates*			BMA *no SE*			FMA			OLS		
	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	Coef.	SE	p-val.	Coef.	SE	p-val.
Standard Error	**0.80**	-3.91	2.55	NA	NA	NA	**-1.24**	0.19	0.00	**-1.16**	0.17	0.00
Study Specific												
Google Citation	**0.99**	-0.004	0.0001	0.21	0.00	0.00	0.00	0.00	0.32	**-0.001**	0.00	0.01
Kuik	0.11	0.00	0.07	**1.00**	-0.17	0.04	-0.29	0.05	0.00			
Technological Setting												
GHG Emissions	0.30	-0.07	0.14	0.35	0.03	0.04	0.01	0.07	0.86			
Agriculture	0.19	0.05	0.15	0.37	-0.04	0.06	0.004	0.10	1.00			
Energy sources	0.29	0.01	0.01	**0.93**	-0.01	0.003	-0.001	0.002	0.57			
Regions	**0.86**	-0.01	0.004	0.16	0.00	0.00	**-0.001**	0.00	0.00	-0.001	0.00	0.00
Methodology												
LEAP model	**0.92**	-0.55	0.26	**1.00**	-0.39	0.06	-0.41	0.09	0.00	**-0.32**	0.04	0.00
Top-down model	**1.00**	-0.57	0.13	0.10	-0.00	0.01	-0.01	0.04	0.83	**-0.15**	0.03	0.00
Technology-Specific												
Intertemporal Optimisation	**0.60**	-0.14	0.14	**0.97**	-0.09	0.03	**-0.11**	0.03	0.00	-0.08	0.03	0.01
Carbon Capture and Storage (CCS)	0.34	-0.01	0.09	**0.76**	0.04	0.03	0.03	0.02	0.17			
Multigas	0.14	-0.01	0.05	**0.87**	-0.6	0.03	-0.05	0.02	0.02			
Induced Technological Change (ITC)	0.29	0.05	0.09	0.12	0.002	0.01	-0.03	0.02	0.13			
Target	**1.00**	0.01	0.001	0.16	0.00	0.00	-0.0004	0.0002	0.05	0.0001	0.00	0.11
Baseline 2050	0.29	0.03	0.01	0.23	0.004	0.01	**0.02**	0.12	0.00			

Notes: The standard errors for FMA and OLS are clustered at the study level. Value 0.00 stands for number lower than 0.001. Variables with PIP > 0.5 in **bold** print.

Table 7.5: Potential sources of heterogeneity among MAC 2050

We could explain this trend by the age of the studies. While studies from our dataset cover the last 15 years, all of their studies were published in 2006. Because the emissions mitigation efforts were not as successful as they seemed at the beginning of the millennium, these studies could have predicted lower baseline emissions and lower abatement costs (ECCP, 2003).

Region. The variable representing the number of regions was significant in three model specifications. The posterior mean suggests a slightly lower MAC for analyses with more regions. The value of the effect is almost negligible (-0.2% for an additional region), and its direction is expected. We suppose that the smaller the area the study focuses on, the more precise results it brings. Repetto & Austin (1997) reveal the same significant relationship between regions and MAC estimates.

LEAP model. The significance of the variable representing the LEAP model remains even for MAC 2050, and the direction stays negative, the same as for MAC 2030. The magnitude is smaller by almost a half. The studies that predict GHG emissions using the LEAP model overall estimate the MAC lower by 40%.

Intertemporal Optimisation. Finally, we found evidence that technological specification plays a significant part in model heterogeneity. Concretely, the variable represents models that employ intertemporal optimisation. The variable captures models that assume long-living decision-makers who establish consumption and investment by looking at the long term. The results suggest a negative direction of the relationship, which was expected. Models that employ this idea of forward-looking decision-makers present a lower abatement cost (by about 10%). The relationship holds with all robustness checks and aligns with findings of Vogt-Schilb *et al.* (2015).

Multigas. Another technological variable identifies studies that predict GHG emissions for other greenhouse gases besides CO_2. The BMA analyses suggest that models that work with multigas lead to lower MAC estimates. The effect holds only for two robustness checks, and the magnitude of the effect is relatively small. The results suggest that mitigating emissions of other greenhouse gases is less costly than CO_2 emissions.

Baseline 2050. The second variable which explains heterogeneity for both MAC 2030 and 2050 is *Baseline*. In this case, the direction is the opposite.

The analysis reveals that studies with higher emissions baselines tend to report a slightly higher abatement cost. This is the direction we expected for the MAC 2030, too. Studies with a higher baseline predict that the emissions would grow tremendously without any reduction. Therefore, we would expect that the costs to abate them would be bigger with higher emissions. Nevertheless, the effect it has on MAC is minimal, which corresponds to findings of Fischer *et al.* (2003) who found that emissions baseline explains only a small portion of differences in MAC estimates across studies.

Overall, the BMA analysis and its robustness checks found four variables explaining heterogeneity for MAC 2030 and seven for MAC 2050. The explanations and comparisons with findings from the literature are discussed above. We found evidence for negative publication bias for MAC 2050 but little evidence for publication bias in literature for MAC 2030. Besides mild publication bias, at least one variable from each category plays a role in explaining heterogeneity in the model. We expected the technological variables to play a more important role in explaining MAC estimates, just like Barker *et al.* (2006) found. On the other hand, Repetto & Austin (1997) also found technological specifications less significant.

Chapter 8

Conclusion

The presented master's thesis conducts a meta-analysis for marginal abatement costs (MAC) of greenhouse gas (GHG) emissions. The study collected 242 observations for MAC in the years 2030 and 2050 from 59 primary studies. The Funnel plot, the Meta-regression analysis, and the Caliper test were utilised to reveal publication bias in the literature. The Bayesian model averaging (BMA) analysis revealed what model specifications affect MAC estimates.

Publication bias analysis reveals little evidence of publication bias for MAC 2030 and MAC 2050. The funnel plots are mostly symmetrical, and FAT-PET tests reveal no significant selectivity. Additionally, estimates for true effect are statistically significant and close to the sample mean. For MAC 2030, the true effect corresponds to approximately 32 EUR/tCO_2-eq, while for MAC 2050, it is around 59 euro. The results are lower than we expected from the literature and show the abatement costs do not necessarily have to be high, especially when policymakers begin with low-cost abatement ways.

An analysis of heterogeneity reveals several factors that affect MAC estimates. For MAC 2030, all variables explaining heterogeneity lead to lower estimates. MAC is lower when the model employs the LEAP model and when the observation comes from older data, respectively, from studies published in 2006. Adding one unit in emissions baseline or stabilisation targets lowers MAC 2030 by around 6% or 0.02%, respectively. There were no significant findings for standard error regarding publication bias in the BMA.

For MAC 2050, the BMA revealed negative publication bias with a significant marginal effect. Other factors affect MAC 2050 negatively, except for the emissions baseline. Other characteristics lead to lower MAC estimates. Inclusion of the LEAP model, other GHGs, and observations from studies published in 2006 (as opposed to newer publications) would lower the MAC by around 80%, ceteris paribus. In the same way, the more regions a study examines, the lower the MAC is found. On the other hand, an increase in emissions baseline by a unit results in a 1% increase in the MAC estimate. The findings are supported by several robustness checks and correspond to the literature. We were able to confirm the hypotheses presented in the proposal.

The study can be considered a follow-up to the study by Kuik *et al.* (2009). Their results correspond more to mine for MAC 2050. This is likely because their studies mostly pursued the MAC in 2050. They found that the MAC is dependent on the stringency of the stabilisation target, the emissions baseline, the inclusion of intertemporal optimisation, the consideration of other greenhouse gases, and the number of regions and energy sources. This meta-analysis brings their dataset up-to-date and applies modern techniques to reveal further dependencies. This thesis is the first meta-analysis to examine both publication bias and model uncertainty for the MAC. The drawback of the presented study includes the construction of standard errors and data modifications. The absence of uncertainty measures in the primary literature results in constructing standard errors for studies with more than one estimate. However, there was no way to check the robustness of this process. Therefore, the analysis of publication bias stands on these constructed values. This could be why we found almost no evidence for publication bias, even though its presence was later revealed through BMA. Another inaccuracy could emerge when we used the best-guess estimate for the observations of the stabilisation target. We wanted to keep the information in the dataset but did not find any observation in primary studies. The resulting influence of the variable should therefore be taken indicatively. We believe that other data adjustments do not introduce more inaccuracy to the results since we conducted several robustness checks to confirm the selection of appropriate methods and datasets.

An interesting extension of this study would be comparing the results with projected policies in selected countries. After all, a meta-analysis should serve the general public and politicians to understand better the abundance of scientific works available on the subject. An update of this work could build on the IPCC report that is expected to be published in September 2022. We believe that the meta-analysis based on newer literature could bring new variables and uncover further relationships.

Bibliography

Aaheim, A., J. S. Fuglestvedt, & O. Godal (2006): "Costs Savings of a Flexible Multi-Gas Climate Policy." *The Energy Journal* **SI2006(01)**.

van Aert, R. C. M. & M. A. L. M. van Assen (2018): "Correcting for Publication Bias in a Meta-Analysis with the P-uniform* Method." https://osf.io/zqjr9 (Accessed: 2022-04-12).

Ahn, Y.-H. & W. Jeon (2019): "Power sector reform and CO2 abatement costs in Korea." *Energy Policy* **131**: pp. 202–214.

Akimoto, K., F. Sano, T. Homma, K. Wada, M. Nagashima, & J. Oda (2012): "Comparison of marginal abatement cost curves for 2020 and 2030: longer perspectives for effective global GHG emission reductions." *Sustainability Science* **7(2)**: pp. 157–168.

Alinaghi, N. & W. R. Reed (2018): "Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work?" *Research Synthesis Methods* **9(2)**: pp. 285–311.

Amini, S. M. & C. F. Parmeter (2012): "Comparison of model averaging techniques: Assesing growth determinants." *Journal of Applied Econometrics* **27(5)**: pp. 870–876.

Andrews, I. & M. Kasy (2019): "Identification of and Correction for Publication Bias." *American Economic Review* **109(8)**: pp. 2766–2794.

Anzures-Cabrera, J. & J. P. T. Higgins (2010): "Graphical displays for meta-analysis: An overview with suggestions for practice." *Research Synthesis Methods* **1(1)**: pp. 66–80.

Barker, T., J. Köhler, & M. Villena (2002): "Costs of greenhouse gas abatement: meta-analysis of post-SRES mitigation scenarios." *Environmental Economics and Policy Studies* **5(2)**: pp. 135–166.

Barker, T., M. S. Qureshi, & J. Köhler (2006): "The Costs of Greenhouse Gas Mitigation with Induced Technological Change." *Cambridge Centre for Climate Change Mitigation Research, Department of Land Economy, University of Cambridge, Working Papers* **89**.

Beach, R. H., J. Creason, S. B. Ohrel, S. Ragnauth, S. Ogle, C. Li, P. Ingraham, & W. Salas (2015): "Global mitigation potential and costs of reducing agricultural non-CO $_2$ greenhouse gas emissions through 2030." *Journal of Integrative Environmental Sciences* **12(1)**: pp. 87–105.

Bergh, J. C. J. M., K. J. Button, P. Nijkamp, & G. C. Pepping (1997): *Meta-Analysis in Environmental Economics*, volume 12 of *Economy & Environment*. Dordrecht: Springer Netherlands.

Bernard, A., M. Vielle, & L. Viguier (2006): "Burden Sharing Within a Multi-Gas Strategy." *The Energy Journal* **SI2006(01)**.

Bom, P. R. D. & H. Rachinger (2019): "A kinked meta-regression model for publication bias correction." *Research Synthesis Methods* **10(4)**: pp. 497–514.

Bosetti, V., C. Carraro, & M. Galeotti (2006): "The Dynamics of Carbon and Energy Intensity in a Model of Endogenous Technical Change." *The Energy Journal* **SI2006(01)**.

Brander, M. (2021): "Greenhouse Gases, CO2, CO2eq, and Carbon: What Do All These Terms Mean?" `https://ecometrica.com/assets/GHGs-CO2-CO2e-and-Carbon-What-Do-These-Mean-v2.1.pdf` (Accessed: 2022-01-21).

Bruns, S. B., I. Asanov, R. Bode, M. Dunger, C. Funk, S. M. Hassan, J. Hauschildt, D. Heinisch, K. Kempa, J. König, J. Lips, M. Verbeck, E. Wolfschütz, & G. Buenstorf (2019): "Reporting errors and biases in published empirical findings: Evidence from innovation research." *Research Policy* **48(9)**.

Böhringer, C., A. Löschel, & T. F. Rutherford (2006): "Efficiency Gains from 'What' - Flexibility in Climate Policy - an Integrated CGE Assessment." *Energy Journal* **Special Issue: Multi-Greenhouse Gas Mitigation and Climate Policy**: pp. 405–424.

Chen, L.-J., Z.-H. Fang, F. Xie, H.-K. Dong, & Y.-H. Zhou (2020): "Technology-side carbon abatement cost curves for China's power generation sector." *Mitigation and Adaptation Strategies for Global Change* **25(7)**: pp. 1305–1323.

Chung, Y., H. Kim, C. Paik, & Y. J. Kim (2015): "Marginal Abatement Cost Analysis for the Korean Residential Sector Using Bottom-Up Modeling." *Journal of Energy Engineering* **24(1)**: pp. 58–68.

Clarke, L., J. Edmonds, H. D. Jacoby, H. Pitcher, J. M. Reilly, & R. Richels (2006): "CCSP synthesis and assessment product 2.1." *Synthesis and Assessment Product 2.1 by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research* **2.1**.

Clarke, L., J. Weyant, & J. Edmonds (2008): "On the sources of technological change: What do the models assume?" *Energy Economics* **30(2)**: pp. 409–424.

Council, N. R. (2011): *Climate stabilization targets: emissions, concentrations, and impacts over decades to millennia.* Washington, D.C: National Academies Press. OCLC: ocn658581418.

Crassous, R., J.-C. Hourcade, & O. Sassi (2006): "Endogenous Structural Change and Climate Targets Modeling Experiments with Imaclim-R." *The Energy Journal* **SI2006(01)**.

Doucouliagos, C. & T. Stanley (2013): "Are All Economic Facts Greatly Exaggerated? Theory competition and selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

ECCP (2003): "Can we meet our Kyoto targets?" `https://ec.europa.eu/clima/system/files/2016-11/second_eccp_report_en.pdf` (Accessed: 2022-04-21).

Eide, M. S., T. Longva, P. Hoffmann, Endresen, & S. B. Dalsøren (2011): "Future cost scenarios for reduction of ship CO2 emissions." *Maritime Policy & Management* **38(1)**: pp. 11–37.

den Elzen, M., M. Meinshausen, & D. van Vuuren (2007): "Multi-gas emission envelopes to meet greenhouse gas concentration targets: Costs versus certainty of limiting temperature increase." *Global Environmental Change* **17(2)**: pp. 260–280.

Enkvist, P.-A., J. Dinkel, & C. Lin (2010): "Impact of the financial crisis on carbon economics: Version 2.1 of the global greenhouse gas abatement cost curve." https://sallan.org/pdf-docs/2009_mckinsey.pdf (Accessed: 2022-01-28).

Enkvist, P.-A., T. Nauclér, & J. Rosander (2007): "A cost curve for greenhouse gas reduction." https://www.mckinsey.com/business-functions/sustainability/our-insights/a-cost-curve-for-greenhouse-gas-reduction (Accessed: 2022-02-03).

Escobar Carbonari, D., G. Grosjean, P. Läderach, T. D. Nghia, B. O. Sander, J. McKinley, L. Sebastian, & J. Tapasco (2019): "Reviewing Vietnam's Nationally Determined Contribution: A New Perspective Using the Marginal Cost of Abatement." *Frontiers in Sustainable Food Systems* **3**: p. 14.

Fawcett, A. A. & R. D. Sands (2006): "Non-CO2 Greenhouse Gases in the Second Generation Model." *The Energy Journal* **SI2006(01)**.

Fischer, C., R. D. Morgenstern, C. Fischer, & R. D. Morgenstern (2003): "Carbon Abatement Costs: Why the Wide Range of Estimates?" *Discussion Papers 10537, Resources for the Future* .

Fujino, J., R. Nair, M. Kainuma, T. Masui, & Y. Matsuoka (2006): "Multi-gas Mitigation Analysis on Stabilization Scenarios Using Aim Global Model." *The Energy Journal* **SI2006(01)**.

Furukawa, C. (2019): "Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method." https://www.ssrn.com/abstract=3362053 (Accessed: 2022-04-12).

Furukawa, C. (2021): "Stem-based method to correct publication bias." https://github.com/Chishio318/stem-based_method (Accessed: 2022-04-17).

Gechert, S., T. Havranek, Z. Irsova, & D. Kolcunova (2021): "Measuring capital-labor substitution: The importance of method choices and publication bias." *Review of Economic Dynamics* .

Gerber, A. (2008): "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* **3(3)**: pp. 313–326.

GERLAGH, R. (2006): "ITC in a Global Growth-Climate Model with CCS: The Value of Induced Technical Change for Climate Stabilization." *The Energy Journal* **SI2006(01)**.

GILLINGHAM, K. & J. H. STOCK (2018): "The Cost of Reducing Greenhouse Gas Emissions." *Journal of Economic Perspectives* **32(4)**: pp. 53–72.

GOPAL, A. R., W. Y. PARK, M. WITT, & A. PHADKE (2018): "Hybrid- and battery-electric vehicles offer low-cost climate benefits in China." *Transportation Research Part D: Transport and Environment* **62**: pp. 362–371.

HANSEN, B. E. (2007): "Least Squares Model Averaging." *Econometrica* **75(4)**: pp. 1175–1189.

HANSON, D. A. & J. A. S. LAITNER (2006): "Technology Policy and World Greenhouse Gas Emissions in the AMIGA Modeling System." *The Energy Journal* **SI2006(01)**.

HARMSEN, J., D. P. VAN VUUREN, D. R. NAYAK, A. F. HOF, L. HÖGLUND-ISAKSSON, P. L. LUCAS, J. B. NIELSEN, P. SMITH, & E. STEHFEST (2019): "Long-term marginal abatement cost curves of non-CO2 greenhouse gases." *Environmental Science & Policy* **99**: pp. 136–149.

HARRER, M. (2022): *Doing meta-analysis with R: a hands-on guide.* Boca Raton: CRC Press.

HASAN, I., R. HORVATH, & J. MARES (2016): "What Type of Finance Matters for Growth? Bayesian Model Averaging Evidence." *The World Bank Economic Review* **32(2)**.

HAVLÍK, P., H. VALIN, A. MOSNIER, M. OBERSTEINER, J. S. BAKER, M. HERRERO, M. C. RUFINO, & E. SCHMID (2013): "Crop Productivity and the Global Livestock Sector: Implications for Land Use Change and Greenhouse Gas Emissions." *American Journal of Agricultural Economics* **95(2)**: pp. 442–448.

HAVRANEK, T., Z. IRSOVA, K. JANDA, & D. ZILBERMAN (2015): "Selective reporting and the social cost of carbon." *Energy Economics* **51**: pp. 394–406.

HAVRÁNEK, T., T. D. STANLEY, H. DOUCOULIAGOS, P. BOM, J. GEYER-KLINGEBERG, I. IWASAKI, W. R. REED, K. ROST, & R. C. M. AERT (2020):

"Reporting guidelines for meta-analysis in economics." *Journal of Economic Surveys* **34(3)**: pp. 469–475.

Heaps, C. (2018): "Using LEAP for GHG Mitigation Assessment." https://unfccc.int/sites/default/files/resource/LEAP%20UNFCCC%20Nov%202018.pdf.

Hedenus, F., C. Azar, & K. Lindgren (2006): "Induced Technological Change in a Limited Foresight Optimization Model." *The Energy Journal* **SI2006(01)**.

Hong, S., Y. Chung, J. Kim, & D. Chun (2016): "Analysis on the level of contribution to the national greenhouse gas reduction target in Korean transportation sector using LEAP model." *Renewable and Sustainable Energy Reviews* **60**: pp. 549–559.

Ioannidis, J. P. A., T. D. Stanley, & H. Doucouliagos (2017): "The Power of Bias in Economics Research." *The Economic Journal* **127(605)**: pp. F236–F265.

IPCC (2014): "Synthesis report: Summary for policy makers." https://www.ipcc.ch/site/assets/uploads/2018/02/AR5_SYR_FINAL_SPM.pdf.

IPCC (2021): "Summary for policymakers." https://www.ipcc.ch/report/ar6/wg1/#SPM.

IPCC (2022): "IPCC — Intergovernmental Panel on Climate Change." https://www.ipcc.ch/ (Accessed: 2022-04-22).

Jakeman, G. & B. S. Fisher (2006): "Benefits of Multi-Gas Mitigation: An Application of the Global Trade and Environment Model (GTEM)." *The Energy Journal* **SI2006(01)**.

Janzen, R., M. Davis, & A. Kumar (2020a): "An assessment of opportunities for cogenerating electricity to reduce greenhouse gas emissions in the oil sands." *Energy Conversion and Management* **211**.

Janzen, R., M. Davis, & A. Kumar (2020b): "Evaluating long-term greenhouse gas mitigation opportunities through carbon capture, utilization, and storage in the oil sands." *Energy* **209**.

Janzen, R., M. Davis, & A. Kumar (2020c): "Greenhouse gas emission abatement potential and associated costs of integrating renewable and low carbon energy technologies into the Canadian oil sands." *Journal of Cleaner Production* **272**.

Jiang, K., X. Hu, & Z. Songli (2006): "Multi-Gas Mitigation Analysis by IPAC." *The Energy Journal* **SI2006(01)**.

Joshi, M., J. E. Pustejovsky, & S. N. Beretvas (2022): "Cluster Wild Bootstrapping to Handle Dependent Effect Sizes in Meta-Analysis with a Small Number of Studies." *Research Synthesis Methods* p. jrsm.1554.

Kass, R. E. & A. E. Raftery (1995): "Bayes Factors." *Journal of the American Statistical Association* **90(430)**: pp. 773–795.

Katta, A. K., M. Davis, & A. Kumar (2020): "Development of disaggregated energy use and greenhouse gas emission footprints in Canada's iron, gold, and potash mining sectors." *Resources, Conservation and Recycling* **152**: p. 104485.

Katta, A. K., M. Davis, V. Subramanyam, A. F. Dar, M. A. H. Mondal, M. Ahiduzzaman, & A. Kumar (2019): "Assessment of energy demand-based greenhouse gas mitigation options for Canada's oil sands." *Journal of Cleaner Production* **241**: p. 118306.

Kemfert, C., T. P. Truong, & T. Bruckner (2006): "Economic Impact Assessment of Climate Change - A Multi-gas Investigation with WIAGEM-GTAPEL-ICM." *The Energy Journal* **SI2006(01)**.

Kesicki, F. (2012): "Intertemporal issues and marginal abatement costs in the UK transport sector." *Transportation Research Part D: Transport and Environment* **17(5)**: pp. 418–426.

Kesicki, F. (2013): "Marginal Abatement Cost Curves: Combining Energy System Modelling and Decomposition Analysis." *Environmental Modeling & Assessment* **18(1)**: pp. 27–37.

Kuik, O., L. Brander, & R. S. Tol (2009): "Marginal abatement costs of greenhouse gas emissions: A meta-analysis." *Energy Policy* **37(4)**: pp. 1395–1403.

KUROSAWA, A. (2006): "Multigas Mitigation: An Economic Analysis Using GRAPE Model." *The Energy Journal* **SI2006(01)**.

LÖFFLER, C. & H. HECKING (2017): "Greenhouse Gas Abatement Cost Curves of the Residential Heating Market: A Microeconomic Approach." *Environmental and Resource Economics* **68(4)**: pp. 915–947.

MANNE, A. S. & R. G. RICHELS (2006): "The Role of Non-CO2 Greenhouse Gases and Carbon Sinks in Meeting Climate Objectives." *The Energy Journal* **SI2006(01)**.

MATOUSEK, J., T. HAVRANEK, & Z. IRSOVA (2022): "Individual discount rates: a meta-analysis of experimental evidence." *Experimental Economics* **25(1)**: pp. 318–358.

MOSER, M. (2016): "dilutBMS2." `https://gitlab.com/matmo/dilutbms2` (Accessed: 2022-04-18).

OECD (2022): "Prices - Inflation (CPI) - OECD Data." `http://data.oecd.org/price/inflation-cpi.htm` (Accessed: 2022-04-23).

DE OLIVEIRA SILVA, R., L. G. BARIONI, T. Z. ALBERTINI, V. EORY, C. F. TOPP, F. A. FERNANDES, & D. MORAN (2015): "Developing a nationally appropriate mitigation measure from the greenhouse gas GHG abatement potential from livestock production in the Brazilian Cerrado." *Agricultural Systems* **140**: pp. 48–55.

PAGE, M. J., J. E. MCKENZIE, P. M. BOSSUYT, I. BOUTRON, T. C. HOFFMANN, C. D. MULROW, L. SHAMSEER, J. M. TETZLAFF, E. A. AKL, S. E. BRENNAN, R. CHOU, J. GLANVILLE, J. M. GRIMSHAW, A. HRÓBJARTSSON, M. M. LALU, T. LI, E. W. LODER, E. MAYO-WILSON, S. MCDONALD, L. A. MCGUINNESS, L. A. STEWART, J. THOMAS, A. C. TRICCO, V. A. WELCH, P. WHITING, & D. MOHER (2021): "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews." *BMJ* p. 71.

PELLERIN, S., L. BAMIÈRE, D. ANGERS, F. BÉLINE, M. BENOIT, J.-P. BUTAULT, C. CHENU, C. COLNENNE-DAVID, S. DE CARA, N. DELAME, M. DOREAU, P. DUPRAZ, P. FAVERDIN, F. GARCIA-LAUNAY, M. HASSOUNA, C. HÉNAULT, M.-H. JEUFFROY, K. KLUMPP, A. METAY,

D. Moran, S. Recous, E. Samson, I. Savini, L. Pardon, & P. Chemineau (2017): "Identifying cost-competitive greenhouse gas mitigation potential of French agriculture." *Environmental Science & Policy* **77**: pp. 130–139.

Popp, D. (2006): "Comparison of Climate Policies in the ENTICE-BR Model." *The Energy Journal* **SI2006(01)**.

Purohit, P. & L. Höglund-Isaksson (2017): "Global emissions of fluorinated greenhouse gases 2005–2050 with abatement potentials and costs." *Atmospheric Chemistry and Physics* **17(4)**: pp. 2795–2816.

Raftery, A. E. (1995): "Bayesian Model Selection in Social Research." *Sociological Methodology* **25**: p. 111.

Rao, S. & K. Riahi (2006a): "Importance of Technological Change and Spillovers in Long-Term Climate Policy." *The Energy Journal* **SI2006(01)**.

Rao, S. & K. Riahi (2006b): "The Role of Non-CO2 Greenhouse Gases in Climate Change Mitigation: Long-term Scenarios for the 21st Century." *The Energy Journal* **SI2006(01)**.

Reilly, J., M. Sarofim, S. Paltsev, & R. Prinn (2006): "The Role of Non-CO2 GHGs in Climate Policy: Analysis Using the MIT IGSM." *The Energy Journal* **SI2006(01)**.

Reilly, J. M., H. D. Jacoby, & R. G. Prinn (2003): *Multi-gas contributors to global climate change: Climate Impacts and Mitigation Costs of Non-CO2 Gases*. Pew Center on Global Climate Change Arlington, VA.

Repetto, R. C. & D. Austin (1997): *The costs of climate protection: a guide for the perplexed*. Washington, DC: World Resources Institute.

Roodman, D., M. Nielsen, J. G. MacKinnon, & M. D. Webb (2019): "Fast and wild: Bootstrap inference in Stata using boottest." *The Stata Journal: Promoting communications on statistics and Stata* **19(1)**: pp. 4–60.

Rosenthal, R. (1979): "The file drawer problem and tolerance for null results." *Psychological Bulletin* **86(3)**: pp. 638–641.

Sano, F., K. Akimoto, T. Homma, & T. Tomoda (2006): "Analysis of Technological Portfolios for CO2 Stabilizations and Effects of Technological Changes." *The Energy Journal* **SI2006(01)**.

Sapkota, T. B., F. Khanam, G. P. Mathivanan, S. Vetter, S. G. Hussain, A.-L. Pilat, S. Shahrin, M. K. Hossain, N. R. Sarker, & T. J. Krupnik (2021): "Quantifying opportunities for greenhouse gas emissions mitigation using big data from smallholder crop and livestock farmers across Bangladesh." *Science of The Total Environment* **786**: p. 147344.

Sapkota, T. B., S. H. Vetter, M. Jat, S. Sirohi, P. B. Shirsath, R. Singh, H. S. Jat, P. Smith, J. Hillier, & C. M. Stirling (2019): "Cost-effective opportunities for climate change mitigation in Indian agriculture." *Science of The Total Environment* **655**: pp. 1342–1354.

Seo, S. N. (2021): *Climate Change and Economics: Engaging with Future Generations with Action Plans*. Cham: Springer International Publishing.

Smith, S. J. & T. Wigley (2006): "Multi-Gas Forcing Stabilization with Minicam." *The Energy Journal* **SI2006(01)**.

Sotiriou, C., A. Michopoulos, & T. Zachariadis (2019): "On the cost-effectiveness of national economy-wide greenhouse gas emissions abatement measures." *Energy Policy* **128**: pp. 519–529.

de Souza, J., B. de Oliveira, J. Ferrer, & S. Pacca (2018): "Industrial low carbon futures: A regional marginal abatement cost curve for Sao Paulo, Brazil." *Journal of Cleaner Production* **200**: pp. 680–686.

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

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

Steel, M. F. J. (2020): "Model Averaging and Its Use in Economics." *Journal of Economic Literature* **58(3)**: pp. 644–719.

Stern, N. (2008): "The Economics of Climate Change." *American Economic Review* **98(2)**: pp. 1–37.

Subramanyam, V., M. Ahiduzzaman, & A. Kumar (2017a): "Greenhouse gas emissions mitigation potential in the commercial and institutional sector." *Energy and Buildings* **140**: pp. 295–304.

Subramanyam, V., A. Kumar, A. Talaei, & M. A. H. Mondal (2017b): "Energy efficiency improvement opportunities and associated greenhouse gas abatement costs for the residential sector." *Energy* **118**: pp. 795–807.

Teng, F., X. Su, & X. Wang (2019): "Can China Peak Its Non-CO2 GHG Emissions before 2030 by Implementing Its Nationally Determined Contribution?" *Environmental Science & Technology* **53(21)**: pp. 12168–12176.

Timilsina, G. R., A. Sikharulidze, E. Karapoghosyan, & S. Shatvoryan (2017): "Development of marginal abatement cost curves for the building sector in Armenia and Georgia." *Energy Policy* **108**: pp. 29–43.

Tol, R. S. (2006): "Multi-Gas Emission Reduction for Climate Change Policy: An Application of Fund." *The Energy Journal* **SI2006(01)**.

Viechtbauer, W. (2010): "Conducting Meta-Analyses in *R* with the **metafor** Package." *Journal of Statistical Software* **36(3)**.

Vogt-Schilb, A. & S. Hallegatte (2014): "Marginal abatement cost curves and the optimal timing of mitigation measures." *Energy Policy* **66**: pp. 645–653.

Vogt-Schilb, A., S. Hallegatte, & C. de Gouvello (2015): "Marginal abatement cost curves and the quality of emission reductions: a case study on Brazil." *Climate Policy* **15(6)**: pp. 703–723.

van Vuuren, D., B. Eickhout, P. Lucas, & M. den Elzen (2006a): "Long-Term Multi-Gas Scenarios to Stabilise Radiative Forcing - Exploring Costs and Benefits Within an Integrated Assessment Framework." *The Energy Journal* **SI2006(01)**.

van Vuuren, D. P., M. G. J. den Elzen, P. L. Lucas, B. Eickhout, B. J. Strengers, B. van Ruijven, S. Wonink, & R. van Houdt (2006b): "Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs." *Climatic Change* **81(2)**: pp. 119–159.

Wagner, F., M. Amann, J. Borken-Kleefeld, J. Cofala, L. Höglund-Isaksson, P. Purohit, P. Rafaj, W. Schöpp, & W. Winiwarter (2012): "Sectoral marginal abatement cost curves: implications for mitigation pledges and air pollution co-benefits for Annex I countries." *Sustainability Science* **7(2)**: pp. 169–184.

Weir, C. J., I. Butcher, V. Assi, S. C. Lewis, G. D. Murray, P. Langhorne, & M. C. Brady (2018): "Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review." *BMC Medical Research Methodology* **18(1)**: p. 25.

Xiao, H., Q. Wei, & H. Wang (2014): "Marginal abatement cost and carbon reduction potential outlook of key energy efficiency technologies in Chinas building sector to 2030." *Energy Policy* **69**: pp. 92–105.

Yahoo (2022): "Live Rate, Chart & News - Yahoo Finance." `https://finance.yahoo.com/` (Accessed: 2022-04-30).

Yue, X., J. Deane, B. O'Gallachoir, & F. Rogan (2020): "Identifying decarbonisation opportunities using marginal abatement cost curves and energy system scenario ensembles." *Applied Energy* **276**: p. 115456.

Zeugner, S. & M. Feldkircher (2015): "Bayesian Model Averaging Employing Fixed and Flexible Priors: The BMS Package for R." *Journal of Statistical Software* **68(4)**.

Appendix A

PRISMA Diagram

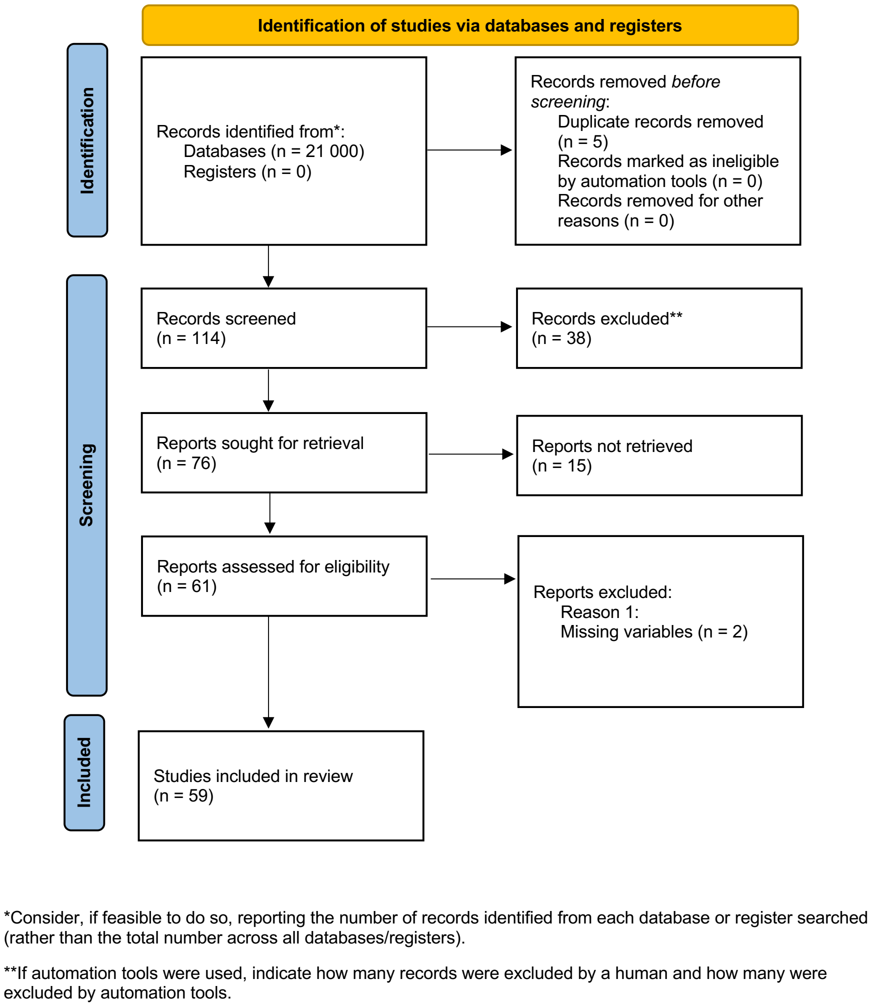


Figure A.1: PRISMA 2020 Diagram, template by Page *et al.* (2021)

Appendix B

Robustness Check

MAC 2030 - median				
A: OLS, WLS				
	OLS	Precision	Study	
Standard Error	-0.43	2.13*	0.39	
(Publication bias)	(1.22)	(0.93)	(1.31)	
Constant	6.79***	6.75***	6.78***	
(Effect beyond bias)	(0.02)	(0.004)	(0.01)	
Observations (Studies)	43	43	43	
B: Study Variations, IV				
	FE	RE	BE	IV
Standard Error	-23.48	-0.41	-23.48	5.21
(Publication bias)	(48.06)	(1.37)	(48.06)	(70.63)
Constant	6.79***	6.79***	6.79***	6.68***
(Effect beyond bias)	(0.03)	(0.02)	(0.03)	(1.36)
Observations (Studies)	43	41	43	
C: Non-Linear Estimates				
	Stem	TOP10	WAAP	
Mean beyond bias	6.96***	6.77***	6.79***	
	(0.12)	(0.05)	(0.006)	
Observations	3	4	43	
	Kink	Selection		
Mean beyond bias	6.79***	6.78***		
	(0.006)	(0.02)		
Observations	43	42		

Significance codes: p<0.001 '***', p< 0.01 '**', p< 0.05 '*', p< 0.1 '.',
clustered standard error in parenthesis

Table B.1: FAT-PET tests - results for median(MAC2030)

MAC 2050 - median				
A: OLS, WLS				
	OLS	Precision	Study	
Standard Error	-1.43	3.97	-1.26	
(Publication bias)	(1.04)	(7.13)	(1.30)	
Constant	6.82***	6.78***	6.82***	
(Effect beyond bias)	(0.02)	(0.05)	(0.02)	
Observations (Studies)	34	34	34	
B: Study Variations, IV				
	FE	RE	BE	IV
Standard Error	4.19	0.55	4.19	-0.12
(Publication bias)	(2.87)	(1.02)	(2.87)	(0.77)
Constant	6.77***	6.79***	6.77***	6.78***
(Effect beyond bias)	(0.02)	(0.01)	(0.02)	(0.02)
Observations (Studies)	34	32	34	34
C: Non-Linear Estimates				
	Stem	TOP10	WAAP	
Mean beyond bias	6.78***	6.27***	6.77***	
	(0.01)	(0.16)	(0.007)	
Observations	17	2	34	
	Kink	Selection		
Mean beyond bias	6.77***	6.79***		
	(0.008)	(0.06)		
Observations	34	34		

Significance codes: p<0.001 '***', p<0.01 '**', p<0.05 '*', p<0.1 '.'
Clustered standard error in parenthesis

Table B.2: FAT-PET tests - results for median(MAC2050)

MAC 2030				
A: OLS, WLS				
	OLS	Precision	Study	
Standard Error	-0.83	0.20	-1.36	
(Publication bias)	(2.31)	(1.91)	(1.82)	
Constant	6.81***	6.78***	6.85***	
(Effect beyond bias)	(0.03)	(0.02)	(0.06)	
Observations	65	65	65	
Studies	26	26	26	
B: Study Variations, IV				
	FE	RE	BE	IV
Standard Error	-0.25	-0.82	-0.26	3.21
(Publication bias)	(5.42)	(2.23)	(5.41)	(3.79)
Constant	6.78***	6.81***	6.78***	6.68***
(Effect beyond bias)	(0.02)	(0.07)	(0.02)	(0.14)
Observations	65	63	65	65
Studies	26	24	26	26
C: Non-Linear Estimates				
	Stem	TOP10	WAAP	
Mean beyond bias	6.75***	6.60***	6.78***	
	(0.06)	(0.04)	(0.006)	
Observations	18	6	64	
	Kink	Selection		
Mean beyond bias	6.78***	6.78***		
	(0.008)	(0.04)		
Observations	64	64		

Significance codes: p<0.001 ’***’, p<0.01 ’**’, p<0.05 ’*’, p<0.1 ’.’
Clustered standard error in parenthesis

Table B.3: FAT-PET tests - results for log(MAC2030), without data from Kuik *et al.* (2009)

	MAC 2030						MAC 2050					
	BMA *BRIC prior*			BMA *correlation*			BMA *BRIC prior*			BMA *correlation*		
	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD
Standard Error	0.12	0.12	0.40	0.14	0.11	0.37	**1.00**	-1.30	0.14	**1.00**	-1.3	0.14
Study Specific												
Google Citation	0.15	0.00	0.00	0.23	0.00	0.00	0.19	0.00	0.00	0.26	0.00	0.00
Kuik	0.20	-0.01	0.04	0.16	-0.01	0.03	**1.00**	-0.30	0.03	**1.00**	-0.30	0.03
Empirical Setting												
GHG Emissions	**0.55**	-0.06	0.06	**0.80**	-0.09	0.06	0.06	0.00	0.01	0.1	0.00	0.01
Agriculture	0.34	-0.04	0.06	**0.60**	-0.07	0.07	0.09	-0.004	0.02	0.15	-0.01	0.02
Energy sources	0.10	0.00	0.002	0.16	-0.001	0.002	0.06	0.00	0.00	0.10	0.00	0.001
Regions	0.09	0.00	0.00	0.16	0.00	0.00	**1.00**	-0.002	0.00	**1.00**	-0.002	0.00
Methodology												
LEAP model	**1.00**	-0.71	0.09	**1.00**	-0.74	0.07	**1.00**	-0.44	0.05	**1.00**	-0.43	0.05
Top-down model	0.07	-0.002	0.02	0.10	-0.001	0.02	0.06	0.00	0.01	0.10	0.00	0.01
Technology-Specific												
Intertemporal Optimization	0.16	0.01	0.03	0.28	0.02	0.03	**1.00**	-0.10	0.02	**1.00**	-0.10	0.02
Carbon Capture and Storage (CCS)	0.04	0.00	0.01	0.07	0.00	0.01	0.28	0.01	0.2	0.39	0.01	0.02
Multigas	0.04	0.00	0.01	0.09	0.001	0.01	0.47	-0.02	0.02	**0.67**	-0.03	0.03
Induced Technological Change (ITC)	0.13	-0.01	0.02	0.22	-0.01	0.03	0.18	-0.01	0.01	0.30	-0.01	0.02
Target	0.35	0.00	0.00	**0.55**	0.00	0.00	0.06	0.00	0.00	0.10	0.00	0.00
Baseline	**0.94**	-0.06	0.02	**1.00**	-0.07	0.02	0.47	0.01	0.01	**0.63**	0.02	0.01

Notes: The standard errors for FMA and OLS are clustered at the study level. Value 0.00 stands for number lower than 0.001. Variables with PIP $>$ 0.5 in **bold** print.

Table B.4: Additional specifications for BMA

	MAC 2030			**MAC 2050**			**MAC 2030**		
	Median			*Median*			*without Kuik et al.(2009)*		
	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD	PIP	Post. Mean	Post. SD
Standard Error	0.16	0.06	0.40	**0.95**	-1.34	0.52	**1.00**	-4.91	0.79
Study Specific									
Google Citation	0.15	0.00	0.001	0.19	0.00	0.001	0.13	0.00	0.00
Kuik	0.27	-0.02	0.04	**0.97**	-0.24	0.09	NA	NA	NA
Empirical Setting									
GHG Emissions	0.21	-0.01	0.03	0.24	0.02	0.07	0.32	-0.03	0.06
Agriculture	0.15	-0.005	0.03	0.22	-0.02	0.07	0.26	-0.03	0.06
Energy sources	0.15	0.00	0.002	**0.51**	-0.004	0.005	0.22	0.001	0.002
Regions	0.13	0.00	0.00	**0.62**	-0.001	0.001	**0.84**	-0.001	0.002
Methodology									
LEAP model	**0.98**	-0.37	0.12	**0.97**	-0.39	0.12	**1.00**	-0.31	0.10
Top-down model	0.17	-0.01	0.03	0.20	-0.01	0.03	0.17	-0.01	0.02
Technology-Specific									
Intertemporal Optimization	0.23	0.01	0.03	**0.92**	-0.11	0.05	0.42	-0.02	0.03
Carbon Capture and Storage (CCS)	0.16	0.004	0.02	0.30	0.01	0.03	0.13	-0.002	0.01
Multigas	0.21	0.01	0.03	0.36	-0.02	0.04	0.11	0.00	0.01
Induced Technological Change (ITC)	0.27	-0.02	0.04	0.17	-0.001	0.02	0.22	0.01	0.16
Target	0.15	0.00	0.00	0.21	0.00	0.00	0.33	0.00	0.00
Baseline	**0.84**	-0.05	0.03	0.24	0.01	0.02	**0.69**	0.02	0.01

Notes: The standard errors for FMA and OLS are clustered at the study level. Value 0.00 stands for number lower than 0.001. Variables with PIP > 0.5 in **bold** print.

Table B.5: Robustness Checks for BMA