Tobias Eibinger

# Evaluating emission reductions in the transport sector: Empirical contributions

## Doctoral Thesis

to be awarded the degree of

Doctor of Business, Economics and Social Sciences/
Doctor rerum socialium oeconomicarumque (Dr. rer. soc. oec.)

at the University of Graz, Austria

Field of Economics

**Supervisor: Hans Manner**
Department of Economics, University of Graz

**Co-Supervisor: Karl Steininger**
Department of Economics and
Wegener Center for Climate and Global Change, University of Graz

Graz, February 2025

# Abstract

This thesis empirically examines emission reductions and the effectiveness of mitigation policies with a specific focus on transportation, the only EU-27 sector in which emissions have increased since 1990. Particular attention is paid to Austria and Luxembourg, both of which have experienced especially large increases in transport emissions relative to 1990 levels. Mitigating emissions in this sector is complicated by stock persistence and deeply ingrained social habits. The thesis is organized into four chapters, each complementing the others both thematically and methodologically to provide robust mitigation policy assessments. Chapter 1 offers empirical guidance on identifying and estimating major emission drivers in large-$N$, large-$T$ panels. It revisits studies in such settings and finds that not adequately accounting for this structure can result in incorrect inference and biased estimates. Chapter 2 moves to evaluate specific policy impacts, focusing on Austria's policy mix since 1950. It pinpoints the most effective transport-related policies, while acknowledging their interdependencies and dynamic effects. It finds that taxes on new cars and fuels were particularly effective. Chapter 3 turns to policies aimed at incentivizing public transport and thus reducing car usage. It uses Luxembourg's nationwide fare-free system as a quasi-experimental setting to measure its impact on transport emissions. It accounts for complexities such as large commuting inflows and COVID-19 related factors. An 8% reduction in road transport emissions is identified. Chapter 4 concludes and switches perspectives. Rather than asking which policies were most effective, it investigates the extent to which external factors – such as mild winters and overall economic activity levels – contributed to Austria's recent emission reductions. The findings suggest these external factors played only a relatively minor role.

# Abstract – German

In dieser Dissertation werden Emissionsreduktionen und die Wirksamkeit von Klimaschutzmaßnahmen empirisch untersucht, wobei ein besonderer Schwerpunkt auf den Verkehrssektor gelegt wird, der einzige Sektor der EU-27, in dem die Emissionen seit 1990 gestiegen sind. Besonderes Augenmerk wird auf Österreich und Luxemburg gelegt, die beide im Vergleich zu den Werten von 1990 besonders starke Zuwächse der Verkehrsemissionen verzeichnet haben. Die Reduktion von Emissionen in diesem Sektor wird durch die Persistenz der Bestände und tief verwurzelte soziale Gewohnheiten erschwert. Die Dissertation gliedert sich in vier Kapitel, die sich sowohl thematisch als auch methodisch ergänzen und zusammen eine fundierte Bewertung von Klimaschutzmaßnahmen ermöglichen. Kapitel 1 bietet empirische Anleitungen zur Identifikation und Schätzung wesentlicher Emissionstreiber in Panels mit großer Fallzahl ($N$) und langer Zeitreihe ($T$). Es greift Studien in diesen Settings auf und zeigt, dass eine unzureichender Berücksichtigung der Datenstruktur zu fehlerhaften Schlussfolgerungen und verzerrte Schätzungen führen kann. Kapitel 2 befasst sich mit der Bewertung spezifischer Maßnahmen und konzentriert sich dabei auf Österreichs Mix an Maßnahmen seit 1950. Es werden die effektivsten verkehrsbezogenen Maßnahmen identifiziert, wobie deren Interdependenzen und dynamische Auswirkungen berücksichtigt werden. Besonders effektiv erweisen sich dabei Steuern auf Neuwagen und Kraftstoffe. Kapitel 3 widmet sich Maßnahmen, die Anreize für den öffentlichen Verkehr schaffen und dadurch den PKW-Gebrauch reduzieren sollen. Als quasi-experimenteller Rahmen dient das flächendeckende Nulltarifsystem in Luxemburg, dessen Auswirkungen auf die Verkehrsemissionen untersucht werden. Dabei werden komplexe Einflussfaktoren wie die hohe Zahl an Pendlerströmen sowie COVID-19-bezogene Faktoren berücksichtigt. Es wird eine 8%-ige Verringerung der Straßenverkehrsemissionen identifiziert. Kapitel 4 schließt mit einem Perspektivenwechsel. Anstatt zu fragen, welche Maßnahmen am effektivsten waren, untersucht es, inwieweit externe Faktoren — wie milde Winter oder das gesamtwirtschaftliche Aktivitätsniveau — zu den jüngsten Emissionsrückgängen in Österreich beigetragen haben. Die Ergebnisse legen nahe, dass diese externen Faktoren nur eine vergleichsweise geringe Rolle spielten.

# Statement of co-authored work

The four chapters of my cumulative dissertation – “Evaluating emission reductions in the transport sector: Empirical contributions” – are co-authored. Chapter 3 is co-authored with another PhD candidate. The remaining chapters include at least one of my supervisors. My total contribution amounts to 240%. The authors and their contribution to each chapter are as follows:

- **Chapter 1**
  **Panel Data in Environmental Economics: Econometric Issues and Applications to IPAT Models**
  Tobias Eibinger (University of Graz) – 50%, Beate Deixelberger (University of Graz) – 40%, Hans Manner (University of Graz) – 10%.

- **Chapter 2**
  **Shifting Gears? Austria's Transport Policy Mix and CO2 Emissions from Passenger Cars**
  Tobias Eibinger (University of Graz) – 70%, Hans Manner (University of Graz) – 20%, Karl Steininger (University of Graz and Wegener Center for Climate and Global Change) – 10%.

- **Chapter 3**
  **Zero fare, cleaner air? The causal effect of Luxembourg's free public transportation policy on carbon emissions**
  Tobias Eibinger (University of Graz) – 50%, Sachintha Fernando (Martin Luther University Halle-Wittenberg) – 50%.

- **Chapter 4**
  **The Development of Austrian Greenhouse Gas Emissions since 2021**
  Tobias Eibinger (University of Graz) – 70%, Hans Manner (University of Graz) – 20%, Karl Steininger (University of Graz and Wegener Center for Climate and Global Change) – 10%.

# Acknowledgments

This thesis would not have been possible without the help and guidance of many people. I want to particularly thank my supervisor, Hans Manner, who always had an open door for questions, and while he provided valuable guidance, he always encouraged me to work independently on my research. I also want to express my gratitude to my co-supervisor, Karl Steininger, who was equally approachable and whose economic and political insights I greatly appreciated, as well as the discussions we had on these topics. In addition to my supervisors, I also want to thank my co-author Sachintha Fernando for the great collaboration and many insightful discussions.

In general, I want to thank many members of the Graz Economics and Public Economics departments as well as the Wegener Center for Cliamte and Global Change for fruitful discussions, both in internal seminars and on a bilateral basis. In this regard, I would particularly like to thank Riccarda Rosenball, Beate Deixelberger, and Jörn Kleinert for their support and friendship. I also want to extend my thanks to Richard Sturn, who supported and promoted my academic career in its earliest stages.

I had the opportunity to take a causal inference course at the CEU with Andrea Weber, which provided me with critical tools to conduct my research. I am very grateful for this opportunity as well as for the subsequent discussions and guidance she offered. I am also thankful for the chance to spend a research stay at the Spatial Economics department at VU Amsterdam and to Hans Koster for hosting me. The concentrated expertise in urban economics there created a very fruitful space for scientific exchange, and I am particularly thankful to Jos van Ommeren, Leon Bremer, and Yashvant Premchand for their insights and support in Amsterdam.

Of course, the journey of writing my thesis was not solely an academic one. Foremost, I want to thank Helene, my partner in love, for her unwavering support throughout this journey, which would not have been possible without her. Finally, I want to thank my parents, Annemarie and Peter, for their support and patience during challenging times.

I am certain I have forgotten to mention many others who also offered valuable support, and to them, I also extend my gratitude.

# Contents

# List of Figures

# List of Tables

# Synopsis

## Motivation

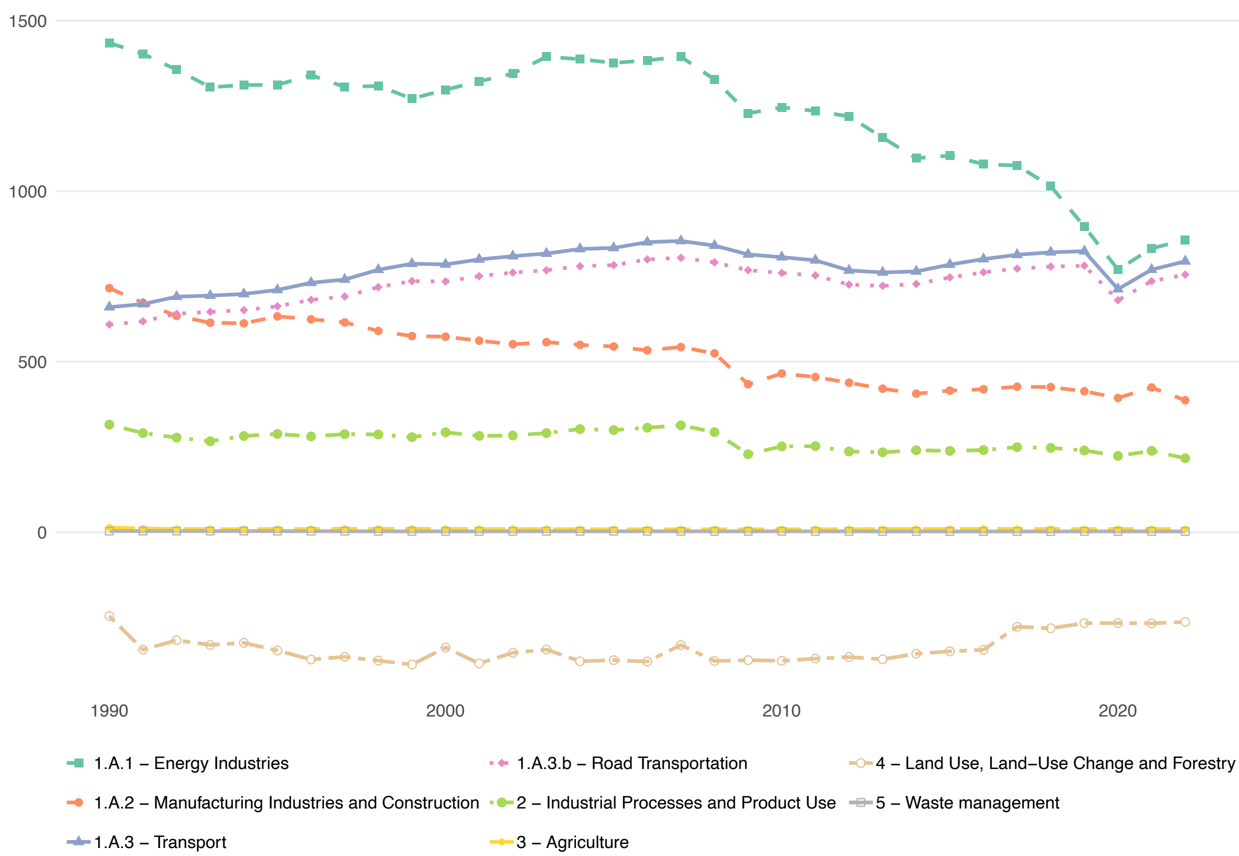
In 2024, the Copernicus Climate Change Service (2024) noted that the planet registered its highest-ever recorded temperatures, exceeding for the first time in a single year the long-term threshold set by the Paris Climate Agreement in 2015, in which 196 nations committed to limiting global warming to 1.5°C above pre-industrial levels (1850-1900). The Intergovernmental Panel on Climate Change (IPCC), which synthesizes research from thousands of scientists, has reached a strong consensus that climate change poses significant threats to both human and planetary health. Despite a comprehensive range of proposed interventions, current and planned policy measures have fallen short of achieving the ambitious emission reduction targets outlined in international agreements (IPCC, 2023). Policymakers frequently turn to scientific assessments for guidance, yet the efficacy of proposed interventions often remain a matter of both scientific and political debate.

This dissertation contributes to the scientific debate by delivering robust estimates of policy effectiveness, re-evaluating existing studies, and offering methodological insights. The research concentrates on the transport sector, a key contributor to overall emissions. Geographically, a focus is set on Europe, which is warming faster than any other continent, according to IPCC (2023) reports, with a specific focus on two Western European countries: Austria and Luxembourg. Both countries face distinctive policy and infrastructure challenges, allowing for a nuanced analysis of how various interventions perform under different national contexts.

The transport sector plays a crucial role in mitigating emissions. Figure 1 shows greenhouse gas (GHG) emissions in the EU-27 between 1990 and 2022, disaggregated by sector. Although industry remains the largest source of emissions, transportation is the only sector in which emissions increased from 1990 to 2023. Within the transport sector, road transportation causes the largest share of transport emissions. Overall, transport activities account for approximately 25% of the EU's total GHG emissions. Crucially, achieving the EU's intermediate and long-term climate goals requires significant emission cuts in this sector.

However, reducing emissions from transport is uniquely complex. Existing infrastructures, vehicle stocks, and social habits developed over decades create systemic inertia. Even if policymakers introduce ambitious measures, their impact on emission levels may manifest only gradually. For instance, changes in registration taxes or vehicle subsidies influence consumer choices at the point of purchase but may have delayed effects on the overall fleet composition. This thesis contributes to understanding these systemic delays and to identify which policies generate the most significant and sustained emission reductions.

**Figure 1:** Emissions (MT $CO_2$eq) by sector for EU-27




Note: This figure shows aggregated $CO_2$ emissions for EU-27 countries for selected sectors. Data is from the European Environment Agency (EEA), 2024.

Another key question in the thesis is, whether observed changes in transport emissions can be attributed primarily to policy interventions or to external factors such as weather patterns, economic growth, and global events, such as the COVID-19 pandemic. For instance, transport emissions fell during the COVID-19 pandemic, only to swiftly rise to pre-pandemic levels once travel restrictions were eased. At the same time, numerous policy measures aimed at curbing emissions were introduced or enhanced across EU Member States. To what extent did these policies meaningfully contribute to observed emission declines? This dissertation explores these questions in particular for the transport sector by analyzing both broader trends and specific national contexts.

## Austria and Luxembourg

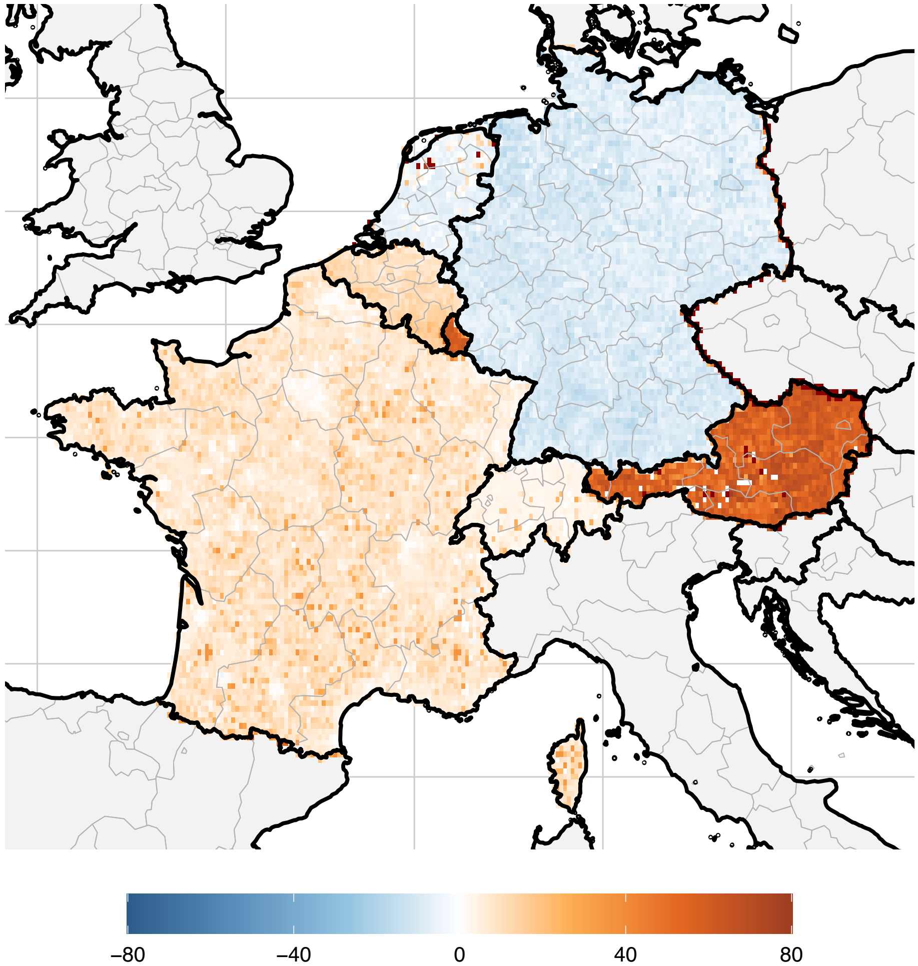
Two interesting national contexts shape the focus of the thesis: Austria and Luxembourg, each offering unique insights. Austria exemplifies a small, affluent economy in Europe with a long-standing cultural inclination toward individual mobility. Over the decades, Austria pursued an extensive range of policy instruments designed to limit transport emissions, including, for instance, speed limits, fuel taxes, and registration taxes. However, Austria's current policy framework is unequivocally insufficient to meet international climate goals to become carbon neutral by mid-century, let alone its own national goal to achieve this by 2040 (Anderl, Bartel, et al., 2021). One contribution of this dissertation is to synthesize Austria's transport policy mix from 1950 onward into a stringency index, allowing for a systematic estimation of policy efficacy while acknowledging potential interdependencies among these measures.

Luxembourg is the world's richest country (in terms of GDP per capita) and provides an unusually ambitious case to reduce transport related emissions. Most notably, it was the first country in the world to implement free public transportation for all users—an assertive "pull" measure intended to shift travel behavior away from private vehicles. Luxembourg's national context is complicated by high levels of cross-border commuting and a historically car-centric orientation.

Figure 2 underscores why Austria and Luxembourg stand out in a Western European context. It illustrates the percentage change in road transport emissions from 1990 to 2023 (a measure often used to define EU-wide emission reduction goals), highlighting that both countries experienced some of the region's largest proportional increases. In contrast, Germany and the Netherlands achieved minor reductions, while Belgium, Switzerland, and France showed only a slight emission growth. Investigating these two high-income countries – each with a distinct but evolving policy landscape – thus reveals valuable nuances about how different strategies function within the shared EU regulatory framework.

Focusing on smaller countries like Austria and Luxembourg is particularly relevant within the European Union's climate policy apparatus. Under the Effort Sharing Regulation (ESR) and related carbon budget allocations, each Member State must meet specific emission-reduction targets for non-ETS sectors, including transport. Although larger countries often receive greater attention, the collective contributions of smaller nations are critical for reach-

**Figure 2:** Percentage Change in $CO_2$ Emissions (2023 vs 1990) in Western Europe




Note: This figure shows percentage changes in road transport emissions for 2023 relative to 1990. Emissions are shown in a spatial distribution, data is from Crippa et al., 2022.

ing EU-wide goals. Political changes may alter policy details but do not supersede supranational mandates. According to the 2023 ESR, Luxembourg must reduce its emissions by 50% by 2030 relative to 2005 levels, while Austria is required to cut its emissions by 48% within the same time frame (European Commission, 2024a).

## Related Literature

This dissertation contributes to a substantial body of literature examining greenhouse gas emissions and the effectiveness of policies designed to mitigate them, with a particular focus on policy impact evaluations within the passenger transport sector. The related literature can be structured in a general-to-specific framework. At a general level, accounting identities can be used to decompose emissions into its main sources. For instance, the IPAT identity (Ehlrich & Holdren, 1971) relates an environmental *Impact*, such as greenhouse gas emissions, to *Population*, *Affluence*, and *Technology*. Such an identity is often used to motivate empirical model specifications to analyze drivers of emissions, including economic activity, technological advances, and population growth (Dong et al., 2018; Opoku et al., 2022; Rafiq et al., 2016; Zheng et al., 2023). Some contributions look at transport emissions specifically, exploring how urbanization, infrastructure investments, and fuel mixes contribute to either decoupling economic growth from emissions or exacerbating them (Andrés & Padilla, 2018; Georgatzi et al., 2020; Guo et al., 2022; W.-Z. Wang et al., 2021; Xu & Lin, 2016; Zhang et al., 2017). The first chapter of this dissertation contributes to this literature by providing methodological guidance and reassessments of related studies.

While these studies offer valuable insights into the key drivers of emissions, researchers and policymakers are often interested in identifying the most effective policies for mitigating emissions and influencing these drivers. For such analyses, more granular research designs are usually required. This dissertation contributes to the literature on transport policy by focusing specifically on policies targeting passenger transport emissions. A substantial body of research examines various policy interventions in this domain. To contextualize this thesis, it is useful to categorize these approaches. Policies aimed at adoption try to accommodate consequences of global warming, while mitigation policies aim to reduce emissions and thus attenuate global warming. This dissertation falls within the broader category of mitigation measures.

Such measures are commonly classified using frameworks such as the Avoid–Shift–Improve (ASI) model (Creutzig et al., 2018). "Avoid" strategies aim to reduce travel demand through spatial planning and urban development. "Shift" strategies promote less emission-intensive transport modes, often through regulations, bans, or taxes that discourage car use while encouraging public transport or non-motorized alternatives. Finally, "Improve" strategies focus on technological advancements, primarily through the adoption of battery electric vehicles (BEVs). However, policies often span multiple ASI categories. For example, a carbon tax not only discourages fossil fuel vehicle use by making it more expensive ("Shift") but also reduces overall mileage ("Avoid") and can incentivize improvements in vehicle efficiency over time ("Improve"). Given these overlaps, it is useful to classify policies based on their immediate outcomes or primary mechanisms rather than strictly within the ASI framework.

Policies that increase the cost of driving and incentivize reduced car usage, such as carbon taxes and congestion charges, have been shown to achieve modest yet significant reductions in transport emissions. For instance, Andersson (2019b) finds that Sweden's carbon tax led to moderate but statistically significant reductions in transport-sector emissions, while

Pretis (2022b) reports notable decreases in transport emissions in British Columbia following similar measures. Additionally, Gerlagh et al. (2018) demonstrate that $CO_2$-sensitive vehicle registration taxes can improve fleet efficiency, indirectly reducing emissions. In Austria, Koch et al. (2022) identify that higher fuel taxes and truck tolls introduced around 2005 were effective in curbing emissions. Regulatory measures, such as standards, limits, and bans – including low-emission zones and speed restrictions – impose direct constraints on driving and vehicle emissions (Davis, 2017; Gallego et al., 2013; Sarmiento et al., 2023; Wolff, 2014). These policies directly target emissions-intensive behaviors and can be highly effective in urban environments. Chapter 2 of the thesis relates to these studies by evaluating the efficacy of Austria's transport policy mix since 1950, including both taxes and regulatory measures.

Policies designed to enhance the attractiveness of alternative transport modes, such as public transport improvements and fare reductions, show mixed results in reducing car reliance (Aydin & Kürschner Rauck, 2023; Borsati et al., 2023; Bull et al., 2021; Cats et al., 2017; Gohl & Schrauth, 2024; Liebensteiner et al., 2024). For example, Cats et al. (2017) find that fare reductions in Tallinn (Estonia) led to increased public transport ridership but did not significantly reduce car usage. Similarly, Bull et al. (2021) report that free-fare vouchers in Santiago (Chile) primarily induced additional leisure trips rather than replacing car trips. Evaluations of Germany's "9-Euro ticket" indicate modest pollution reductions in areas with well-developed transit networks (Aydin & Kürschner Rauck, 2023; Gohl & Schrauth, 2024), while Liebensteiner et al. (2024) find limited shifts away from car use. Chapter 3 of this dissertation builds on this literature by examining the impact of free public transport on road transport emissions. This analysis contributes to the ongoing debate on the effectiveness of fare-based policies, offering new evidence on whether and under what conditions such measures can contribute to emission reductions.

There is overarching consensus in the literature that policy mixes are required to achieve optimal emission reduction potential and to be politically feasible (Koch et al., 2022; Kuss & Nicholas, 2022; Winkler et al., 2023). However, empirical evaluations have predominantly focused on isolated policies, limiting the understanding of how multiple policies interact to influence emissions. These require discussion of specific contexts and potential interactions with other existing policies or economic conditions. This has spurred uncertainty about what policy measures are most effective. The dissertation incorporates this critique into the policy analyses presented in Chapters 2 and 3.

Various methodologies exist to examine the effects of environmental policies. Two prominent examples include computable general equilibrium (CGE) models and econometric analyses. CGE models can account for great economic complexity and handle feedback effects between different economic sectors (Robinson et al., 1999). For specific policy analyses, these models can simulate future counterfactual scenarios to estimate economy-wide policy effects. However, CGE models rely heavily on theoretical assumptions about behavioral responses and equilibrium conditions. Econometric methods, by contrast, exploit observed data to identify empirical patterns and test policy impacts, while also allowing counterfactual simulations akin to CGE models, e.g. with difference-in-differences designs (Carbone et al., 2020).

This dissertation adopts an empirical approach to evaluate the effects of environmental policies, with a particular focus on transport policies. Various empirical strategies are employed to analyze how policies interact and shape environmental outcomes. Time-series and panel data methods are used to model long-term relationships between policy measures and environmental indicators, capturing potential interdependencies. Quasi-experimental research designs often focus on one specific policy that is treated as an exogenous event and thus assigned a causal interpretation. A fuller discussion of these methods and their relation to the chapters of the thesis is given in the Methodology Section.

# Contributions

This dissertation contributes to environmental economics and transport policy evaluation by addressing a central question: How can policy instruments be better designed and assessed to achieve meaningful and sustained reductions in transport emissions? It combines empirical methods – most notably time-series and panel econometrics – with in-depth policy analysis and is organized around four key contributions, each constituting a chapter in the cumulative dissertation:

## 1) Panel Data in Environmental Economics: Econometric Issues and Applications to IPAT Models

This paper tackles methodological challenges in environmental economics when using large-$N$, large-$T$ panel data. Such a setting is commonly encountered in cross-country analyses with annual observations, including a large time span (e.g., 30 years) as well a cross-sectional dimension of approximately equal size (e.g., 30 countries). In such macropanel data, traditional econometric techniques often fail to account for issues like nonstationarity and cross-sectional dependence, jeopardizing the validity of standard inference. Ignoring these complexities can lead to distorted confidence intervals and biased estimates that may even reverse the inferred direction of policy impacts. By illustrating these complexities, this contribution provides a roadmap for applied researchers to conduct more rigorous policy analyses.

In an environmental context, such large-$N$, large-$T$ structures are frequently encountered in empirical assessments of IPAT identities. Beyond offering a methodological guide for applied researchers, this contribution replicates and re-evaluates existing IPAT-based studies, revealing that certain previously reported findings do not hold under more rigorous estimation procedures. For instance, the anticipated negative effect of human development on emissions is not statistically confirmed, and the influence of green technology appears weaker than originally suggested. These outcomes underscore the importance of robust econometric methods in delivering reliable insights for policymakers. To facilitate best practices, the chapter concludes with a step-by-step guide – including Stata code – on how to accurately estimate IPAT-type (and related) models under macropanel conditions.

## 2) Shifting Gears? Austria's Transport Policy Mix and $CO_2$ Emissions from Passenger Cars

This study develops a novel policy stringency index for the Austrian transport sector covering the period from 1950 to 2019. This long-range historical perspective makes it possible to trace how tax policies and regulatory measures have jointly evolved. In this, we recognize that policies may interact with each other, amplifying or muting their efficacy. Over such an extensive horizon, policies may themselves be influenced by changing emission levels—for instance, higher emissions can spur new or stricter measures. This study explicitly models these interdependencies in a multivariate econometric time-series model that allows policies and other variables to interact almost freely. This is a novelty compared to many existing studies on single transport policies.

The newly constructed index distinguishes between policies targeting investment decisions (e.g., buying a new vehicle) and those addressing usage behavior (e.g., fuel taxes or speed limits). Findings indicate that investment-focused policies, particularly Austria's engine-related insurance tax and its registration tax on new vehicles, have been notably effective in reducing passenger-car emissions over time. These measures produce statistically significant emission reductions roughly five years after their implementation or when their stringency is increased. This delayed impact aligns with the economic rationale that adjustments to entrenched transport infrastructure and vehicle stock naturally unfold over an extended period.

## 3) Zero fare, cleaner air? The causal effect of Luxembourg's free public transportation policy on carbon emissions

This paper provides a quasi-experimental evaluation of Luxembourg's move to eliminate fares on public transport in March 2020. Although the scheme is financed by a national tax, residents, commuters, and tourists alike pay no fare to ride buses, trains, or the tram. Fully free public transport programs are rare and typically restricted to specific cities or small regions, often on a partial basis. Indeed, Luxembourg is the first country in the world to provide free fares for everyone. It thus presents an exceptionally interesting, but also a challenging context because of its strong commuter inflows, fuel tourism, and the COVID-19 pandemic.

To extract a causal effect of the policy on road transport emissions, we adapt and extend a new difference-in-differences type of methodology, proceeding in two main stages. First, we compare Luxembourg's emission trends with matched other European regions before the policy intervention. These other regions are chosen such that they follow Luxembourg's trends closely prior to the policy intervention. A policy effect is then usually extracted by assuming that these regions would have continued to follow comparable paths in the absence of a policy intervention (free fares). As this is not straightforwardly interpretable in this case, this brings us to the second main step. We control for COVID-related shifts that may have caused these parallel trends to diverge. Although this would only be an issue insofar as COVID-19 shifted mobility patterns in Luxembourg differently compared to the control regions. We carefully consider and control for such potential changes, including, for instance, commuting inflows, working from home, and COVID-19 cases. We find evidence that Luxembourg was not affected differently compared to other regions.

This approach enables a robust estimate of the policy's causal effect on road transport emissions, which is estimated at around an 8% reduction. We show that this figure is statistically significant, robust, and its magnitude reasonable. The novelty of this contribution is twofold: (1) To provide a causal estimate of the world's first national free public transport scheme, and (2) offering a methodological framework to address confounding factors, such as the COVID-19 pandemic. Other studies, for instance on Germany's 9-Euro, find smaller policy effects. However, direct comparisons to fully free fares are limited by differences in both scope and duration. Unlike short-term partial discounts, a permanent elimination of fares also removes the inconvenience of ticket purchasing (even at low cost) and might suggest a stronger potential for sustained shifts in travel behavior.

### 4) The Development of Austrian Greenhouse Gas Emissions since 2021

Austrian greenhouse gas emissions declined substantially from 2021 to 2023, reaching a level 14% below 1990 levels. This memo-style contribution offers a timely overview of the factors underpinning these reductions and examines how short-term economic and meteorological developments intersect with longer-term policy dynamics. Large-scale macroeconomic models, while capable of providing detailed insights, are often time-consuming to implement. In contrast, the memo applies transparent dynamic time-series models to preliminary emissions data, offering a rapid yet credible assessment of the key drivers behind the observed decline.

Specifically, we study to what degree these emission cuts can be attributed to external events, such as mild winters, economic trends, and energy savings due to price increases, and to structural shifts following policy interventions and increases in energy shares of renewables. We find that these reductions were only partially influenced by a milder winter and weaker economic performance. The majority of the emission reductions appears to stem from increases in renewable energy shares, where the data also suggests that rising energy prices played a notable role in incentivizing this shift. This work thus underscores the complementary importance of quickly deployable analytical tools in providing early evidence to guide policy decisions.

## Methodology

The diverse data structures inherent in transport policy analysis necessitate a versatile methodological framework. Time-series analysis enables the examination of trends and dynamic interactions over time, as well as capturing the interdependencies among multiple policy variables and economic indicators. Macro panel data, which combines cross-sectional and time-series dimensions, is essential for addressing broader patterns and heterogeneities across regions or countries. Difference-in-differences (DiD) approaches are instrumental in identifying causal effects by comparing treated and control groups before and after policy interventions, thereby mitigating the influence of confounding factors.

This dissertation employs a combination of these empirical techniques to provide a comprehensive evaluation, focusing on methodological rigor and applications to transport policies in Austria and Luxembourg. The following section gives a brief overview of the applied techniques, where each of the sub-sections is related to a specific contribution of the dissertation.

## Time Series

A dynamic time-series framework focuses on a single outcome variable measured over time, often to assess the effect of a policy intervention that may be continuous (e.g., a tax rate). A general specification is:

$$Y_t = \alpha + \beta D_t + \gamma Y_{t-1} + \delta X_{t-1} + \epsilon_t,$$

where $Y_t$ is the outcome at time $t$, $D_t$ denotes the policy variable, $X$ represents additional relevant variables, and $\epsilon_t$ is the error term. Lagging both $Y$ and $X$ allows the model to capture dynamic relations, including inertia in the outcome and delays in the impact of policy measures or external factors. The aim is to consistently estimate the main policy effect, given by the parameter $\beta$. A key challenge is nonstationarity, where persistent trends can lead to spurious results. If policy and outcome variables share a long-run equilibrium relationship (cointegration), an error correction model (ECM) can distinguish short-run fluctuations from long-run adjustments.

Chapter 4 of the dissertation builds on and applies such a dynamic framework to explain recent emission reductions in Austria, assessing the roles of exogenous factors (such as economic performance and weather conditions) and structural changes.

The single-equation time-series model can be extended into a multiple-equations setting to account for more complex interactions among multiple variables. A Vector Autoregression (VAR) framework models a system of equations:

$$\mathbf{Y}_t = \mathbf{A} + \mathbf{B}\mathbf{D}_t + \mathbf{\Phi}\mathbf{Y}_{t-1} + \epsilon_t,$$

where $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{kt})'$ is a vector of $k$ endogenous outcome variables, $\mathbf{D}_t$ is a vector of policy variables, and $\mathbf{A}$ is a vector of intercepts.

In this model, each variable influence all other variables to reflect complex interdependencies, such as those between different policies. Tracing the dynamic effects of conotemporaneous policy changes is not straightforward because they follow highly nonlinear paths. Chapter 2 of the dissertation utilizes an extended VAR-type model to estimate the dynamic effects of Austria's transport policies dating back to 1950.

## Panel Data

Panel data structures combine cross-sectional observations on units $i$ over multiple time periods $t$. A basic representation is given by:

$$Y_{it} = \alpha + \beta D_{it} + \gamma X_{it} + \mu_i + \lambda_t + \epsilon_{it},$$

where $\mu_i$ captures unobserved, time-invariant characteristics specific to unit $i$ (individual fixed effects), while $\lambda_t$ accounts for unobserved factors that vary over time but are constant across units (time fixed effects). By controlling for both individual and time fixed effects, the model effectively isolates the impact of the policy from confounding factors that are either constant within units or across time periods. These panel models can further be extended to include dynamic effects, similar to the pure time-series cases.

The relative sizes of cross-sectional and time dimensions influence the importance of issues like nonstationarity and cross-sectional dependence. Chapter 1 deals with estimation strategies tailored to settings where these dimensions are of comparable magnitude, which introduce such complexities. Econometric theory on robust estimation in these cases is still not fully matured and applied research is consequently prone to biased estimation. The chapter revisits specific studies building on this methodology.

## Difference-in-Differences (DiD) Settings

DiD approaches exploit panel data to identify causal effects by comparing treated and untreated groups before and after policy interventions. Suppose $Y_{it}(0)$ denotes the potential outcome for unit $i$ without a specific policy of interest (e.g., free public transport), and $Y_{it}(1)$ denotes the potential outcome with the policy. Presence of a specific policy is indicated by $D_i = 1$ and its absence by $D_i = 0$.

In the simplest setup with two periods (pre-treatment $t = 0$ and post-treatment $t = 1$) and two groups (treated with $D_i = 1$ and control with $D_i = 0$), the average treatment effect on the treated can then be computed by simple comparison of two means:

$$ATT = \underbrace{E[(Y_{i,1}(1) - Y_{i,0}(1)) \mid D_i = 1]}_{\text{Change in outcomes for treated group}} - \underbrace{E[(Y_{i,1}(0) - Y_{i,0}(0)) \mid D_i = 0]}_{\substack{\text{Change in outcomes for untreated group} \\ \text{(counterfactual)}}}.$$

The first term on the right-hand side represents the observed change in outcomes for the treated group, while the second term represents the observed change for the control group. The difference between these two quantities provides an estimate of the causal impact of the policy intervention.

This simple approach can be extended to a regression framework, which allows computation of uncertainty bands:

$$Y_{it} = \alpha + \delta_t + \gamma_i + \beta D_{it} + \epsilon_{it},$$

where $D_{it}$ indicates the presence or absence of a policy, $\delta_t$ captures time fixed effects, $\gamma_i$ accounts for individual fixed effects, and $\beta$ estimates the treatment effect of the policy. A critical assumption underpinning the DiD approach is the parallel trends assumption, which posits that in the absence of the policy, the average change in outcomes for the treated and control groups would have followed parallel paths. Under this assumption, the DiD estimator consistently estimates the average treatment effect $\beta$.

Chapter 3 of the dissertation applies an extended DiD framework to evaluate the causal effect of free fares on transport emissions in Luxembourg. Finding a single region (or multiple regions) that are comparable to Luxembourg such that the parallel-trends assumption holds is difficult. To find more similar regions in such cases, a weighted average of controls can be computed. Such approaches are called Synthetic Control (SC) methods. However SC typically requires close matches between regions. To overcome these limitations, Chapter 3 uses a synthetic DiD framework, which integrates both DiD and SC approaches. It, for example, allows for level differences between regions, weighs time periods optimally, and can provide estimates of dynamic policy effects over time, called event-study type estimates. We extend these event-study estimates to control for COVID-19 related impacts and other confounders, which appear novel in the literature.

# Concluding remarks

The four chapters of this thesis are structured to complement each other both thematically and methodologically. They all contribute to evaluating mitigation policies with a focus on transport, where reducing emissions remains challenging because policy impacts often materialize with delays because of the persistence of the existing vehicle stock. More scientific evidence is needed that considers this persistence and recognizes the interdependencies among policies to determine the most effective measures. Robust methods are crucial in such studies for adequately assessing policy effects to derive reliable estimates.

Chapter 1 discusses econometric approaches for large-$N$, large-$T$ panel settings and offers researchers practical guidance. It identifies the main drivers of emissions at an aggregate level and revisits existing studies, illustrating how methodological choices can significantly alter policy conclusions. Future work might expand on these findings both empirically and methodologically. Replication studies that re-evaluate earlier conclusions can enhance policy assessments, and a more detailed understanding of the primary drivers of emissions can enable more refined policy analysis. Moreover, the methodological insights of this chapter extend beyond environmental studies to other research areas with similar data structures.

Chapter 2 focuses on Austria's historical transport policy mix, accounting for vehicle stock persistence and policy interdependencies to estimate their dynamic effects on car emissions. Consistent with Chapter 1, it confirms that overall economic activity is a major driver of Austria's historical growth in passenger car emissions, while specific policies have played only a minor role. However, some—especially emission-based taxes on new cars and insurance taxes—have shown statistically significant effects and can generate "double dividends" by reducing emissions and raising fiscal revenue. In this vein, Estonia recently introduced emission-based registration taxes that will likely take several years to show measurable effects. Austria's policy mix appears to have had only a limited impact, especially given the national net zero target for 2040.

More stringent, additional measures may be necessary, including subsidizing and promoting public transport. Only after the study period ended did Austria introduce a nationwide climate ticket to encourage public transport use and a carbon tax. Extending the time frame or geographic scope of the Austrian analysis could clarify how different policy mixes function elsewhere. Because the carbon tax has a regressive aspect offset by a progressive climate bonus, these social and behavioral dimensions present an intriguing research opportunity. The abolition of the bonus currently discussed in ongoing governmental debates adds another layer of complexity for future studies that could explore resulting behavioral responses when a tax only retains its regressive characters.

Chapter 3 examines zero-fare public transport in Luxembourg, applying a new methodology to refine policy evaluation. The findings indicate a significant effect on passenger transport emissions. Other initiatives, such as Germany's "9-Euro ticket", have been estimated to mainly increase ridership but yield more mixed outcomes on emissions. This could be partly because the ticket required purchase and was time-limited. Understanding how fully free-fare systems that are not time-limited and require no ticket purchases induce long-term shifts from cars to public transport would be highly informative. Such research could be extended to other free public transport programs, whether entirely free or targeted at specific groups or modes.

Chapter 4 adopts a broader perspective by analyzing whether external factors contributed to recent emission reductions in Austria. It moves beyond the transport sector and examines overall emissions in 2022 and 2023, distinguishing exogenous influences, such as overall economic activity and changes in renewable energy prices, from direct policy effects. The findings indicate that short-term reductions may not always align with long-term climate objectives. The methodology used here can be applied to quick initial assessments of policy outcomes or the causes of emission trends whenever long-term studies are not feasible.

Methodologically, the thesis applies a range of econometric methods, each suited to different policy contexts. For example, identifying main drivers of emissions, evaluating a particular policy intervention, or making a quick assessment of external factors each call for different analytical approaches. Specifically, the chapters draw on dynamic time-series models, (macro)panel data analyses, and quasi-experimental methods to evaluate various policies in different contexts. For specific policy analyses, for example, Chapters 2 and 3 show how controlling for confounding events, such as the COVID-19 pandemic, or incorporating stringency measures and policy interdependencies can refine policy impact estimates. Such techniques can be readily applied to other settings, where external shocks and methodological choices similarly influence how policies are assessed.

Overall, this dissertation makes significant methodological, empirical, and policy-relevant contributions to understanding how transport-sector emissions can be effectively reduced. Finally, the aim of the dissertation is to enhance both academic understanding and practical policy development and to thus contribute to more effective strategies for emission mitigation policies and to address climate change.

# Chapter 1

# Panel Data in Environmental Economics: Econometric Issues and Applications to IPAT Models

This paper addresses econometric challenges arising in panel data analyses related to IPAT (environmental Impact of Population, Affluence and Technology) models and other applications typically characterized by a large-$N$ and large-$T$ structure. This poses specific econometric complexities due to nonstationarity and cross-sectional error correlation, potentially affecting consistent estimation and valid inference. We provide a concise overview of these complications and how to deal with these with appropriate tests and models. Moreover, we apply these insights to empirical examples based on the IPAT identity, offering insights into the robustness of previous findings. Our results suggest that using standard panel techniques can lead to biased estimates, incorrect inference, and invalid model adequacy tests. This can potentially lead to flawed policy conclusions. We provide practical guidance to practitioners for navigating these econometric issues.

## 1.1 Introduction

The Paris Agreement in 2015 marks the commitment of the international community to a carbon-free society by setting the target to stay well below +2°C of global warming (IPCC, 2018). In 2021, the EU adopted its new Climate Law, in which it transformed the goals set in its Green Deal to become climate neutral by 2050 (EC, 2021). To realize these goals, a comprehensive understanding of the factors driving Greenhouse Gas (GHG) emissions in different settings is paramount. This has spurred numerous studies to leverage the IPAT identity (Ehlrich and Holdren, 1971), which puts environmental impact (I) in relation with population (P), affluence (A), and technology (T). Even though the baseline IPAT model is a simple setting, it serves as the basis for many empirical studies in environmental economics.

As in all empirical research, findings and their policy implications often depend strongly on the specific statistical methods being used, which should always be carefully chosen. This appears to be a particularly pervasive issue in panels with large $N$, and $T$ structures, termed as *macropanels*. Studies based on the IPAT model, but also other models relying on similar data sources, often do not use adequate methods that can deal with the specific properties of the data being analyzed, which are mainly the presence of nonstationarity and cross-sectional dependence. The objective of our paper is to provide a concise overview on these econometric complications and provide guidance for practitioners navigating these challenges, which is illustrated using different empirical applications. We put specific emphasis on both static as well as dynamic versions of the common correlated-effects (CCE) model, first proposed by Pesaran (2006), due to its generality and relative simplicity. We demonstrate the significance of the chosen methodology by showcasing empirical applications that exemplify the sensitivity of empirical results and policy implications. Furthermore, the possibility of endogeneity is often a delicate issue in empirical analyses, and IPAT models are no exception. Indeed, the problem of endogeneity extends to the broader scope of this study. Accordingly, we provide specific attention to this issue within the methodological frameworks and empirical applications we cover.

IPAT models are typically estimated with panel data at the country or regional level with annual frequency. Estimation procedures for such data structures have to abandon the usual "small $T$", "large $N$" assumption in micropanels in favor of "large $T$", "large $N$" asymptotics. This warrants greater consideration of time-series specific issues in such panel data models. Ignoring these issues can lead to inconsistent estimation, invalid inference, and invalid test statistics. In particular, cross-sectional dependence, nonstationarity, and cointegration must be investigated carefully. These issues are not always adequately considered in the related literature. Ignoring nonstationarity can lead to coefficient estimates that may not be meaningful in the absence of a cointegrating relation. If cointegration holds, estimates may be biased or even inconsistent if regressors correlate with the error. Transforming the data to be stationary is a popular reaction to this issue, but this approach is not necessarily satisfactory. Often, valuable information that could otherwise be used to distinguish between short-run and long-run effects is lost, this is particularly relevant in the presence of cointegrating relations. Additionally, the estimation precision can be limited if information in the data is removed. This holds even when variables are assumed to be cross-sectionally independent, an assumption that appears unrealistic in many cases, especially so in macropanels. Sufficiently strong cross-sectional dependence can lead to inconsistent coefficient estimates even under stationarity. While micropanels often assume cross-sectional independence of units, this assumption is likely violated in macropanels, in which units are usually aggregated at a level that units can impact each other. We can think of such interactions at the country level, where cross-sectional dependence can be assumed whenever economic theory suggests that countries are large enough to affect another country. While the distorting effect of this dependence structure is generally recognized in many empirical contributions, it is done so insufficiently. Often, the distorting effect on inference is accounted for, while resulting invalidity of test statistics for unit roots and cointegration are less often considered. Moreover, potential inconsistency of parameter estimates receives even less attention.

We focus our contribution and the empirical examples on models based on the IPAT identity, although the context to which our discussion applies extends far beyond. We do this because the IPAT model is particularly simple and thus well suited for illustrative purposes. Moreover, there is an active literature estimating IPAT-type models with panel data in different contexts. The results are commonly used to inform policy implications and are as such particularly relevant from a practical point of view. Empirical IPAT-related studies on EU countries include Andrés and Padilla (2018), who study the determinants of GHG emissions from transport for the EU-28 in the period 1990-2014. They study a particularly comprehensive list of drivers, including the composition of transport modes and fuel usage. González et al. (2019) study the relationship between CO2 emissions from passenger cars in 13 EU countries over the period 1990-2015. Their emission drivers include fuel efficiency, other technological improvements, economic activity, share of diesel-powered cars, and motorization.

More recently, Georgatzi et al. (2020) studied the effect of environmental policy stringency, climate change mitigation technologies, share of value added by different transport modes, and infrastructure investments on CO2 emissions in 12 EU countries from 1994-2014. Xu and Lin (2016) investigate the drivers of transport-related CO2 emissions in 30 Chinese provinces over the period 2000-2013. The authors consider GDP per capita, energy intensity, urbanization, cargo turnover, and private vehicle inventory. Guo et al. (2022) analyze the environmental Kuznets hypothesis in the transport sector for 30 Chinese provinces from 1998-2017. Overall, they find evidence of an inverted U-shape relationship between economic activity and transport emissions. Zhang et al. (2017) find an inverted U-shape relation between urbanization and transport emissions in a study of 141 OECD countries over the period 1961-2011. Hashmi and Alam (2019) study the effect of environmental regulation and innovation on CO2 emissions for 29 OECD countries from 1999-2014. They show that larger environmental taxes and green patents can reduce emissions. W.-Z. Wang et al.

(2021) study the impact of urbanization of transport-related CO2 emission for 33 OECD countries from 1960-2014. They find that urbanization weakly decreases emissions from transport. Opoku et al. (2022) find a significant relation between human development and environmental sustainability for 33 OECD countries from 1996–2016.

These studies provide valuable contributions to quantify the effect of determinants on emissions, and thus to evaluate the efficacy of various policy options to mitigate emissions. However, most studies do not fully consider the above-mentioned fact that the underlying data are typically characterized by unit roots and cross-sectional dependence. We are aware only of handful of studies that correctly consider these aspects. Rafiq et al. (2016) explore the impact of urbanization and trade openness on emissions in 22 emerging economies from 1980-2010 and find that both factors contribute to increased emissions. Dong et al. (2018) investigate the effect of population size, economic growth, and renewable energy on CO2 emissions in a panel of 128 countries, finding that the first two increase emissions, while the latter can reduce them. Zheng et al. (2023) highlight the significance of energy-efficiency as an important factor for environmental sustainability in G-7 countries form 1990-2020. Finally, Pablo-Romero et al. (2017) analyze the environmental Kuznets hypothesis for the transport sector in 27 EU countries from 1995-2009 and find evidence of a concave relation between economic activity and transport emissions.

We illustrate the sensitivity of results and policy conclusions to the chosen methodological framework with several empirical examples. Throughout these applications we contrast carefully specified models with one-way fixed-effects models, the workhorse of many empirical contributions, as well as with two-way fixed-effects models. Additionally, we consider extensions with instrumental variable (IV) techniques. We start off with a simple IPAT model for CO2 emissions in the EU transport sector and then go on to re-visit three recent contributions based on more general IPAT models. Throughout, we guide the reader through a carefully conducted empirical analysis highlighting the methodological issues we intend to emphasize. The simple specification we start with allows us to utilize both static and dynamic versions of the CCE model to distinguish between short and long-run effects as well as to extend the models to instrumental variables. We intend to provide particularly extensive guidance for practitioners in this context.

Following up on this exposition, we apply adequate models to a few selected extended specifications of the IPAT identity based on recently published studies in this literature. These include studies on the effects of human development, regulation, transport modes, and energy sources on emissions. Our findings indicate that neglecting nonstationarity and cross-sectional dependence can impact the results, ranging from small differences to losses in statistical significance and even reversed signs in some cases. The consequences for policy implications are potentially great. More specifically, we cannot replicate a significant relation between human development and emissions. The effectiveness of environmental regulation on emissions can be confirmed, while our estimates do not indicate significant interactions between green innovation (proxied by green patent applications) and emissions. Regarding transport modes and energy sources, we find that switching from road to water transport as well as from oil to electricity as an energy source significantly reduces emissions, which seems intuitive. However, these results cannot be replicated by standard fixed-effects models.

The rest of this paper is structured as follows. section 1.2 provides an overview of the econometrics for nonstationary panel data. We discuss adequate tests for weak cross-sectional dependence, unit roots, and cointegration in such settings. Moreover, we discuss a recommendable approach for such situations, namely static and dynamic common correlated effects models, which are flexible enough to accommodate various macropanel characteristics. section 1.3 applies the recommended methods to empirical examples based on the IPAT identity. Finally, section 2.5 concludes. The appendix contains econometric details on the methods discussed, a practical step-by-step guide, a list of useful STATA commands and a table pointing to some key references.

## 1.2 Econometrics for Nonstationarity and Cross-Sectionally Dependent Panels

The basic model setup that we are referring to throughout our paper is motivated by the IPAT identity. In a regression context, this model is referred to as the STIRPAT (Stochastic Impact by Regression on Population, Affluence, and Technology) model, which is proposed by Dietz and Rosa (1997). It starts with the IPAT identity and transforms the variables into natural logarithms. Coefficients and an error term are added to obtain the regression model

$$log(I) = \alpha + \beta_1 log(P) + \beta_2 log(A) + \beta_3 log(T) + u.$$

The model states that the environmental impact (I) is determined by population (P), affluence (A), and technology (T). Environmental impact is measured by GHG emissions, affluence by real GDP per capita, and technology by energy intensity. This model is typically estimated using panel data. The issues we discuss below are not exclusive to this model class, but they apply to all applications that are analyzed using macro-panels, i.e., panels that typically consist of a limited number of units/countries (say 20-50) over number of time periods featuring medium to long term behaviour (typically between 20 and 100, depending on the available frequency).

We provide an overview of the specific econometric challenges that arise when modeling nonstationary panel data, which frequently underlie such regression models, and provide guidelines on adequate testing and modeling approaches. We start by discussing the properties of OLS estimators when the panel data series are characterized by nonstationarity. We then discuss the implications of cross-sectional dependence. One consequence is that sufficiently strong dependence can lead to invalid inference and inconsistent estimates even when variables are stationary. This must at least be addressed by using robust inference as proposed in Driscoll and Kraay (1998), but it is in fact preferable to use models that explicitly account for this feature. Furthermore, tests for unit roots and cointegration have to be adapted to account for cross-sectional dependence. We discuss some of these tests. An appropriate model that allows for cross-sectional dependence as well as unit roots and cointegration under weak conditions is the common correlated effects model, which will be discussed in 1.2.4 and 1.2.4. This is our recommended modeling approach as it is a flexible model that works in a wide range of situations.

One issue that is of paramount importance for any regression analysis where one is interested in establishing causal relationships is endogeneity and the IPAT model is no exception to this. Firms can be expected to react to policies aimed at reducing GHG emissions, which is likely to feedback into output and energy efficiency and policies may directly affect GHG emissions, which may cause endogeneity. While this important issue is not the main focus of our paper, it cannot be completely ignored, as is often the case in the literature. The methods we present can be extended to instrumental variable (IV) estimation and in section 1.2.4 below we revisit this issue. However, to keep the exposition tractable and to focus on the points we wish to emphasize, we abstract from endogeneity problems in the general discussion.

The treatment in this section attempts to avoid technical issues and focuses on summarizing the main conclusions from the related literature and giving clear recommendations. Important generalizations and extensions are briefly presented at the end of the specific subsections and in 1.2.4. In Appendix A.1 we provide additional formulas and some econometric details. Furthermore, in Appendix A.2 a clear step-by-step procedure for our recommended approach is provided along with STATA commands for those methods to allow for easy implementation by researchers. It also contains a table with some key references on the main methods and some of their extensions for readers interested in more econometric details and extensions.

### 1.2.1 Properties of OLS Estimators for nonstationary Data

In the following, we provide a brief overview of asymptotic theory in nonstationary panel data models with large $N$ and $T$. Much of this theory is based on Phillips and Moon (1999), who develop the asymptotic theory for such panel settings. Throughout this subsection, we assume that errors are cross-sectionally independent. This assumption will be discussed in the following subsection. Generally, asymptotic results for nonstationary panels with large $N$ and $T$ differ from their pure time-series counterparts. It is useful to distinguish two cases to discuss asymptotic results. 1) Spurious regression case, in which the residuals from an OLS regression involving two nonstationary variables are themselves nonstationary and 2) cointegration case, in which the residuals are stationary. OLS will consistently estimate a long-run average relation in the case of spurious regression, in contrast to the time-series case, in which the estimate will be inconsistent. When the variables exhibit a homogeneous cointegrating relation, consistency of the estimator depends on the correlation between the regressor and the error term.

We will briefly outline these two cases, while the technical details can be found in Appendices A.1.1 and A.1.2. Bias-corrected versions can allow for endogeneity and provide consistent estimators in this case. Examples include fully-modified OLS (FM-OLS) and dynamic OLS (D-OLS); see, for example, Kao and Chiang (2001) and Pedroni (2001). Recall that the clear advantage of cointegrating regressions is the fact that long-run relationships can be established, convergence rates of the estimators are faster, and that using appropriate models long-run and short-run relations can be distinguished.[1]

For simplicity, we focus our discussion on the properties of pooled OLS estimation, as the associated asymptotic properties are most commonly and thoroughly studied in the related literature. Moreover, the presented results remain largely unaltered for more complex models that include individual specific effects, in particular fixed effects, which we discuss at the end of this section. We start with a setting where panels are nonstationary but not cointegrated. Consider the simple panel regression model

$$y_{it} = \beta' x_{it} + u_{it}, \tag{1.1}$$

where the cross-sectional units and time periods are denoted by $i = 1, ..., N$ and $t = 1, ..., T$, respectively. The variables $y_{it}$ and $x_{it}$ are assumed to be nonstationary in the sense of being $I(1)$ processes. The pooled OLS estimators of equation (1.1) is consistent. This is shown formally for model (1.1) in Appendix A.1.1. Phillips and Moon (1999) show that

[1] Error correction models are the most prominent example, which are also available in the nonstationary panel case.

the estimator is also unbiased. This result is in stark contrast to the usual time-series asymptotics, where the OLS estimator converges to a random variable and therefore is not consistent. The intuition behind this result is that the cross-sectional dimension provides additional information that can be used to obtain an estimator that does not converge to a random variable. Interpretation of the estimation results still has to be conducted carefully, as it is not apparent how the estimated relation should be interpreted. Although the estimator in the panel setting consistently identifies a long-run relation between $y_{it}$ and $x_{it}$, this relationship is not necessarily meaningful (Breitung and Pesaran, 2008).

We now turn to the case in which $y_{it}$ and $x_{it}$ exhibit a cointegrating relation. Consider again the two $I(1)$ random vectors $y_{it}$ and $x_{it}$, and assume the following data generating process:

$$
\begin{aligned}
y_{it} &= \beta' x_{it} + e_{it}, \\
x_{it} &= x_{i,t-1} + u_{it},
\end{aligned}
\tag{1.2}
$$

where $w_{it} = (e_{it}, u'_{it})'$ is stationary. Consequently, $y_{it}$ and $x_{it}$ are cointegrated. Note that this model assumes a homogeneous cointegrating relation. The pooled OLS estimator is $T\sqrt{N}$-consistent when $e_{it}$ and $u_{is}$ are uncorrelated. This is formally stated in Appendix A.1.2. When this condition is violated, i.e., the regressors are endogenous, a persistent second-order bias is introduced and the estimator is inconsistent (Kao & Chiang, 2001; Phillips & Moon, 1999). This stands in contrast to the pure time series asymptotics under similar conditions. Intuitively, this is due to differing convergence rates in the two cases. In a pure time-series context, the estimator is superconsistent, i.e., $T$-consistent, while the bias grows with rate $\sqrt{T}$. In the panel setting, the bias grows with the same rate as the asymptotic rate of convergence, i.e., order $\sqrt{N}$. Note that this directly motivates the use of instrumental variable techniques. When the cointegrating relation is heterogeneous, the usual small-sample bias due to temporal error correlation (which induces endogeneity) is present, but the estimator remains consistent.

Models (1.1) and (1.2) can be extended to include individual specific effects. These can take the form of unit-specific effects, common time trends, and individual-specific time trends. Unit-specific and common time-specific effects can be easily accommodated by transforming the raw data before they are plugged in to the formulas for estimation. Unit-specific effects can be removed by subtracting time-averages from the raw data, while common time-effects are eliminated by subtracting cross-sectional averages. Phillips and Moon (2000) propose to model individual-specific trends by estimating time trends and using these to de-trend the data in a first step and estimating the de-trended data with pooled OLS in a second step. Most commonly, unit-specific effects are included to capture unobserved heterogeneity. When the panels are nonstationary, unit-specific effects in (1.1) can be interpreted as individual specific deterministic trends, while in (1.2) they can be seen as individual effects in the cointegrating relation. Phillips and Moon (2000) show how to extend models (1.1) and (1.2) to unit-specific effects as well as heterogeneous time trends, Kao and Chiang (2001) study a model with homogeneous cointegration including unit-specific effects, and Mark and Sul (2003) study multiple variations of this model, including unit-specific effects, individual-specific time trends, and common time trends. All these studies find results regarding small-sample biases and consistency comparable to those for the pooled estimator discussed above.

### 1.2.2 Testing for Cross-Sectional Dependence

A highly relevant issue when working with macropanels is the potential presence of cross-sectional dependence (CSD). While this is usually not an issue in micropanels, individual units in macropanels need not be cross-sectionally independent. The presence of CSD has consequences for testing and estimation even under stationarity. Estimators can be inconsistent if the source of CSD is correlated with the regressors and inference may be invalid when CSD is sufficiently strong, requiring the use of the robust standard errors proposed in Driscoll and Kraay (1998). Furthermore, panel unit-root tests that do not explicitly account for CSD are invalid under CSD; for details see Chudik and Pesaran (2015a). Similar conclusions hold for panel cointegration tests that do not account for CSD (Breitung and Pesaran, 2008). A particularly easy solution to this problem is the inclusion of time fixed-effects. However, this only adequately accounts for cross-sectional dependency structures when all units are commonly impacted by the source of CSD. This may be a very restrictive assumption. Particularly in data structures usually found in the context of IPAT models, cross-sectional units are sufficiently aggregated (often at the country level) that they can impact each other. We will discuss models that account for these more general forms of CSD in section 1.2.4.

To determine whether there is cross-sectional dependence in the data, the test proposed by Pesaran (2015) can be applied. The null hypothesis of weak cross-sectional dependence is tested against the alternative hypothesis of strong cross-sectional dependence. Sufficiently weak CSD does not pose serious problems for conventional estimation and strict cross-sectional independence is likely an unrealistic assumption for most real-world data. The test statistic of the Pesaran CSD test was originally designed to test residuals of panel data models. However, if interest lies in testing the presence of CSD in the time series of the dependent and explanatory variables as to decide on an appropriate unit root and cointegration test, the Pesaran CSD test can be applied as well. The original test statistic applied to regression residuals $\hat{u}_{it}$ is given by

$$CSD = \sqrt{\frac{2T}{N(N-1)}} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{p}_{ij} \right), \tag{1.3}$$

where $\hat{p}_{ij} = \sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} / (\sum_{t=1}^{T} u_{it}^2)^{1/2} (\sum_{t=1}^{T} \hat{u}_{jt}^2)^{1/2}$ are the estimated pairwise correlation coefficients. Under the null hypothesis, the statistic is asymptotically distributed normal with $CSD \overset{a}{\sim} N(0,1)$.

Juodis and Reese (2021) show that the CSD test proposed by Pesaran (2015) diverges when applied to residuals from a model that already corrects for CSD in some form, including models with time-fixed effects (e.g., TWFE) and common-correlated effects models (CCE). They propose a weighted CSD (CSDw) test with pairwise correlation coefficients given by

$$\hat{p}_{ij} = \sum_{t=1}^{T} w_i \hat{u}_{it} w_j \hat{u}_{jt}, \tag{1.4}$$

where the weights $w$ are identically and independently Rademacher distributed (i.e. $w_i$ takes on the values 1 and -1 with probability 1/2). Under the null hypothesis, the weighted statistic is again asymptotically normal, i.e., $CSDw \overset{a}{\sim} N(0,1)$.

### 1.2.3 Testing for Unit Roots and Cointegration under CSD

We have established that nonstationary series can impact estimation results and interpretation when not adequately accounted for. However, nonstationary series can convey additional information that can improve precision of results and policy implications when adequately accounted for. Consequently, it is important to test for nonstationarity of the series and to test their order of integration. Such tests are common and well established procedures in the pure time-series context, while they are less common in micropanels, where the cross-sectional dimension is large, but only a few time periods are observed. When $N$ and $T$ dimensions are both large, these topics gain in relevance. However, standard testing procedures can provide misleading inference in the presence of cross-sectional dependence. Testing procedures that recognize this complication are referred to as second-generation tests. We provide a brief overview of unit-root and cointegration tests that specifically model CSD. Our focus is on the most popular tests that are commonly available in econometric software. That being said, we also briefly discuss further tests that accommodate further econometric complexities.

**Panel Unit-Root Tests**

A number of unit-root tests for panel data have been suggested in the literature. Those that do not account for CSD are referred to as first-generation tests (e.g. Hadri, 2000; Im, Pesaran, and Shin, 2003; Levin, Lin, and Chu, 2002; Maddala and Wu, 1999). Pesaran (2007), for example, shows in Monte Carlo simulations that panel unit-root tests that assume cross-sectional independence can be severely biased in the presence of sufficiently strong CSD. The tests usually over-reject in this case. Simply de-meaning the series before applying first-generation tests is not guaranteed to resolve this issue. Consequently, panel unit-root tests that do explicitly account for CSD have been developed. They are typically denoted as second-generation tests (Bai and Ng, 2004; Pesaran, 2007).

A popular panel unit-root test that accounts for cross-sectional dependence is an augmented version of the test suggested by Im, Pesaran, and Shin (2003), proposed by Pesaran (2007). This is a second-generation unit-root test, which is able to control for cross-sectional dependence. The procedure for this test is based on augmenting the usual augmented Dickey-Fuller (ADF) regression for each series with the lagged cross-sectional mean and its first-difference to capture the cross-sectional dependence. The individual ADF statistics are then averaged. The null hypothesis of homogeneous nonstationarity (unit root) is tested against the heterogeneous alternative. The test statistic is formally given in Appendix A.1.3.

More recently, second-generation panel unit-root tests have been extended to incorporate additional econometric complexities. Cavaliere (2005) shows that augmented DF-type tests can be distorted when errors are heteroscedastic. Westerlund (2014) proposes a simple test that allows for unconditional heteroscedasticity in the errors while putting minimal restrictions on the data-generating process. The Lagrange-multiplier-type test is thus very general and simple to implement. Monte Carlo results show good small-sample properties. Pesaran et al. (2013) extend the CIPS test from Pesaran (2007) to account for multiple common factors. Lee et al. (2016) extends this test to allow for structural breaks. However, small-sample properties are only satisfactory for $T$ larger than fifty.

A more general concern regarding panel unit-root tests is discussed by Pesaran (2012). In a pure time-series context, the alternative hypothesis to a unit-root test is clearly defined. But the alternative can be heterogeneous with respect to the unit $i$ in a panel context. Consider the following two extreme cases. $H_1^a$: Each series is stationary and $H_1^b$: At least one series is stationary. The first one has a clear interpretation but is very restrictive. Moreover, a test against $H_1^a$ will have power typically when not all series are stationary. Therefore, a rejection of $H_0^a$ is not very convincing. The second alternative is only appropriate when

$N$ is finite and lacks power in large $N$, large $T$ settings. Pesaran (2012) notes that in such settings, it is more appropriate to consider an alternative between these two extremes. $H_1^c$: $f(N)$ series are stationary, where $f$ is an increasing function wrt $N$. Indeed, this specification is used in the CIPS test from Pesaran (2007). If $T$ is relatively small ($\sim$15), these tests are only powerful in some average sense, indicating whether a significant fraction of the series are stationary. The exact proportion of these series can only be extracted when $T$ is sufficiently large. Ng (2008) and Hanck (2013) provide strategies to identify the stationary fraction in a panel.

**Panel Cointegration Tests**

Cointegration tests can again be separated into two different groups. One group of tests, referred to as first-generation panel cointegration tests does not explicitly account for cross-sectional dependence. Examples include Kao (1999), Pedroni (1999, 2004), and Westerlund (2007). Not accounting for CSD can distort these test statistics (Banerjee et al., 2004). Some of these tests allow to account for CSD by de-meaning the series before testing, which can alleviate problems caused by specific forms of CSD. However, small sample-properties of these testing procedures remain suspect (Westerlund & Edgerton, 2008). The second group of tests explicitly accounts for cross-sectional dependence and is referred to as second-generation panel cointegration tests.

The second-generation error-correction-based cointegration test proposed by Westerlund (2007), for example, is appealing because it can handle possible cross-sectional dependence and is easily implemented. Westerlund (2007) suggests mean-group and panel statistics for the null hypothesis of no cointegration. The mean-group statistics test whether there is at least one cross-sectional unit that is characterized by a cointegrating relation, while the panel statistics test whether the panel as a whole is cointegrated. Based on Monte Carlo simulation results from Westerlund (2007), the $G_\tau$ test statistic seems to work best under cross-sectional dependence among the mean-group statistics. The panel cointegration test from Westerlund (2007) and the $G_\tau$ test statistic are formally given in Appendix A.1.4.

Such error-correction-based cointegration tests can be very data intense and thus restrictive in the number of variables that can be included in a large $N$, large $T$ setting. Other residual-based approaches are much less demanding in this respect. A very general test of this type is proposed by Westerlund and Edgerton (2008). The test can accommodate CSD, heteroscedasticity and serial correlation in the errors as well as unknown structural breaks in the cointegrating relation. The null of the test is no cointegration in the panel.

Cointegration tests usually restrict the series to be tested to follow an $I(1)$ process. More recently, Trapani (2021) proposed a residual-based test that allows for a mix of $I(1)$ and

$I(0)$ series to enter the cointegration test. Trapani (2021) notes that erroneously imposing slope homogeneity on a model can by construction introduce nonstationary residuals. To distinguish the source of nonstationarity of the residuals, the authors therefore propose two test statistics to test for slope homogeneity and cointegration, respectively. Furthermore, testing against mixed alternatives can also be relevant for cointegration tests. Hanck (2012) provides a note on mixed signals in panel cointegration tests.

### 1.2.4 Estimating Panel Data Models Under CSD

In this section, we provide a discussion on the estimation under cross-sectional dependence. We primarily focus on common-correlated effects (CCE) type of models in the spirit of Pesaran (2006). We choose to focus on this type because it is among the most popular in the literature, robust to various specifications, and easy to implement, thus making it attractive for empirical applications. Therefore we propagate the use of these models and estimate different versions thereof in the empirical examples provided in the next section. The original static stationary version from Pesaran (2006) has since been extended to a static nonstationary as well as dynamic stationary and nonstationary cases. It has also been extended to incorporate instrumental-variable techniques to accommodate endogeneity. We refer to small-sample properties of the various CCE-type estimators whenever possible. Overall, these are known to have good small-sample properties in their static versions. Dynamic models require a more nuanced discussion and vary among models. In 1.2.4 we discuss the static version of the model, in 1.2.4 the dynamic one, in 1.2.4 we discuss how to handle endogeneity using IV estimation, and we dedicate a short discussion in 1.2.4 to alternative estimation procedures and results under CSD.

#### Static Common Correlated Effects

The panel common-correlated effects (CCE) model was originally proposed by Pesaran (2006) to accommodate cross-sectional dependence in stationary models. Asymptotic results for this model have later been extended to the nonstationary case by Kapetanios, Pesaran, and Yamagata (2011). The CCE model is a particularly flexible specification that is easy to implement and can cope with the econometric challenges that may arise when working with nonstationary macropanels. It allows for cross-sectional dependence in the errors, nonstationarity and possible cointegration. It is thus among the most general static models in the literature and especially well suited for the subsequent empirical analysis.

The CCE model regresses the endogenous variable on individual-specific observed regressors as well as observed and unobserved common factors:

$$y_{it} = \alpha_i' d_t + \beta_i' x_{it} + \lambda_i' f_t + u_{it}, \qquad i = 1, 2, \ldots, N;\ t = 1, 2, \ldots T, \tag{1.5}$$

which is a linear heterogeneous panel data model, where $d_t = (d_{1t}', d_{2t}', d_{3t}')'$. $d_{1t}'$ is a vector of deterministic components, $d_{2t}'$ is a vector of observed common effects assumed to have a unit root and $d_{3t}'$ is a vector of stationary observable common effects. $x_{it}$ is a vector of individual-specific regressors, $f_t$ is a vector of unobserved common effects, and the errors $u_{it}$ are assumed to be $i.i.d$ across $i$. The unobserved common factors may be correlated with $(d_t, x_{it})$ and the regressor $x_{it}$ is modelled as

$$x_{it} = A_i' d_t + \delta_i' f_t + v_{it}, \tag{1.6}$$

where $A_i'$ and $\delta_i'$ are factor loading matrices with fixed components. $v_{it}$ is assumed to follow general covariance stationary processes and can be interpreted as the components of $x_{it}$ distributed independently of the common effects. The observed common factors, $d_t$, as well as the unobserved common factors, $f_t$, may contain $I(1)$ components. In the case that either of these contain a unit root, $y_{it}$, $x_{it}$, $d_t$, and $f_t$ may be cointegrated. The authors suggest using cross-sectional averages of $y_{it}$ and $x_{it}$ to proxy unobserved common factors. It is worth noting that this common factor model encompasses both the FE and TWFE specifications as special cases when specific restrictions on the unobserved factors are imposed. It is then clear that the common factor model is less restrictive compared to these special cases by allowing for multiple unobserved effects in multiplicative form.

Pesaran (2006) proposes two alternative estimators for the model given in equation (1.5). The pooled version pools the cross-sectional observations and assumes (potentially incorrectly) parameter homogeneity, but the loading of the common factor may differ between units. The resulting CCE estimator is then similar to a fixed-effects estimator of a model that allows for cross-sectional dependence by extending it with cross-sectional means. The other version assumes parameter heterogeneity. This mean-group estimator is based on Pesaran and Smith (1995) and gives a cross-sectional mean of the individual-specific estimates. Efficiency gains can be achieved with the pooled version of the estimator. The two estimators are formally given in Appendix A.1.5.

One may wonder whether this flexible approach is useful for the relatively small sample sizes available for estimating IPAT models. Pesaran (2006) studies small sample properties for both CCE estimators in Monte Carlo exercises. Both estimators show satisfactory results both for models that assume parameter homogeneity and heterogeneity. These small-sample properties are particularly important for empirical applications that in the context of macropanels are usually characterized by cross-sectional and time dimensions of around 20. Even for such a small sample, both estimators show very small root mean-squared errors in the Monte Carlo experiments. The mean-group estimator generally performs better when the true data-generating process features heterogeneous parameters, while the pooled version does better under parameter homogeneity. Furthermore, the pooled version even shows slightly better properties in very small samples (with $N$ and $T$ equal to 20) when the true specification features heterogeneous parameters. Simulation results from Kapetanios et al. (2011) show that these small-sample properties hold even under more general conditions, in particular when the unobserved factors are allowed to follow unit root processes. Hence, we can conclude that CCE models can be estimated reliably for studying IPAT equations based on yearly data and a limited number of countries.

Inference can be conducted based on the asymptotic distribution of the estimators. However, these asymptotic results may not hold sufficiently well in small samples and significance levels may be biased. Bootstrapped standard errors and confidence intervals may lead to more precise inference in such cases. Gonçalves and Perron (2014) propose cross-sectional bootstrap methods to improve small-sample inference. Westerlund et al. (2019) shows that many of the desired results of the CCE estimators can be extended to large $T$ panels with good small-sample properties. They also propose a cross-section based bootstrap method to improve small-sample inference.

**Dynamic Common Correlated Effects**

Many models that empirical researchers are interested in are better described by dynamic compared to static structures. Autoregressive distributed lag (ARDL) models are popular and flexible models to estimate dynamic effects. These specifications add lagged values of the dependent and independent variables to the regression. The model can be consistently estimated even when variables consist of a mix of $I(0)$ and $I(1)$ series and the dependent and independent variables are jointly determined. This is particularly useful when feedback effects from the dependent variable (e.g., emissions) on some explanatory variables (e.g., energy intensity, GDP) cannot be ruled out. Chudik et al. (2016), however, mention that consistent estimation fails when the error term $u_{i,t}$ contains unobserved common factors that correlate with the regressors. There are two CCE-type estimators that provide consistent estimation results in this setting. One is based on the usual ARDL approach augmented with cross-sectional averages. This cross-sectionally augmented model is referred to as CS-ARDL and is proposed by Chudik and Pesaran (2015b) for the stationary case. Cao and Zhou (2022) recently showed that the CS-ARDL estimator remains consistent when unobserved common factors are nonstationary. The general procedure of ARDL-type models is to first estimate the short-run coefficients and then compute the long-run coefficients based thereon. The second approach is to directly estimate the long-run coefficients. This can be accomplished by transforming the ARDL model into a distributed lag structure. The resulting specification is referred to as the cross-sectionally augmented distributed-lag (CS-DL) model and is proposed in Chudik et al. (2016), and consistent estimates can be obtained regardless of whether variables are $I(0)$ or $I(1)$. We briefly compare the properties of the two approaches and then outline the specific models.

The CS-DL has several advantages over the CS-ARDL model. It is robust under the following scenarios: misspecification of lag orders, both slope homogeneity and heterogeneity, unit roots in the unobserved common factors and regressors, and breaks in the errors. The main drawback of the model is that it is inconsistent under feedback effects, i.e., when lagged values of the dependent variable correlate with the regressors in $t$. Contrary, the CS-ARDL does not require strict exogeneity of the regressors and allows for such feedback effects, but it is sensitive to lag misspecification. Chudik et al. (2016) compare the small-sample properties for estimates of the long-run coefficients for two models for combinations of $N$ and $T$ ranging from 30 to 200. The CS-DL shows good properties even for small $T$ (<50), while the CS-ARDL shows substantial small-T biases. Chudik and Pesaran (2015b) study small-sample properties of the short-run estimates of the CS-ARDL model. Estimates of the regressors show good properties for small $T$, while the coefficient of the lagged dependent variable shows strong small-sample biases. The authors propose bias corrections, but they are unable to completely eliminate it.

The CS-ARDL with $p = 1$ and $q = 1$ is given by:

$$y_{it} = \mu_i + \varphi_i y_{i,t-1} + \beta_{0i} x_{it} + \beta_{1i} x_{i,t-1} + \sum_{\ell=0}^{p_{\bar{z}}} \zeta'_{i\ell} \bar{z}_{t-\ell} + \epsilon_{it}, \tag{1.7}$$

where $\bar{z}_t = (\bar{y}_t, \bar{x}_t)'$ are the cross sectional averages of the dependent and independent variables, $\zeta_{il} = (\zeta_{yil}, \zeta_{xil})'$ are the estimated coefficients of the cross-sectional averages (which are generally treated as nuisance parameters) and $\epsilon_{it} = \lambda'_i f_t + u_{it}$. The long run equilibrium effect is captured by $\theta_i$, which is here defined by

$$\theta_i = \frac{\beta_{0i} + \beta_{1i}}{1 - \varphi_i}, \tag{1.8}$$

where $\beta_{0i}$ represents the short run effect of $x_{it}$ on $y_{it}$. The model can be estimated by the pooled mean group estimator proposed by Pesaran et al. (1999). The mean-group estimator gives the average over the individual long-run estimates: $N^{-1} \sum_i^N \theta_i$. According to Chudik and Pesaran (2015b), there are two conditions for the estimator to be valid: First, there must be a sufficient number of cross-sectional average lags included in individual equations of the panel.[2] Second, the number of cross section averages must be at least as large as the number of unobserved common factors.

Ditzen (2021) notes that the CS-ARDL can be transformed into an error-correction model (ECM). The CS-ECM model in this case is:

$$\Delta y_{it} = \mu_i - \phi_i[y_{i,t-1} - \theta_{1i}x_{it}] - \beta_{1i}\Delta x_{it} + \sum_{\ell=0}^{p_{\bar{z}}} \zeta'_{i\ell}\bar{z}_{t-\ell} + \epsilon_{it}, \tag{1.9}$$

where $[y_{i,t-1} - \theta_{1i}x_{it}]$ is the error correction term. The error correction speed of adjustment parameter $\phi_i$ is defined by $\phi_i = (1 - \varphi_i)$.

The cross-sectionally augmented distributed lag (CS-DL) model developed by Chudik et al. (2016) transforms the CS-ARDL specification such that no pre-determined values of the outcome variable appear on the right-hand side:

$$y_{it} = \mu_i^* + \theta_i x_{it} + \sum_{\ell=0}^{p_{\bar{x}}-1} \delta_{i\ell}\Delta x_{i,t-\ell} + \sum_{\ell=0}^{p_{\bar{y}}} \eta_{i\ell}\bar{y}_{i,t-\ell} + \sum_{\ell=0}^{p_{\bar{x}}} \omega_{i\ell}\bar{x}_{t-\ell} + u_{it}^*, \tag{1.10}$$

where $\bar{x} = N^{-1}\sum_i^N x_{it}$, $\bar{y} = N^{-1}\sum_i^N y_{it}$ and $p_{\bar{y}} = 0$.[3] The long-run coefficients are then directly given by $\theta_i$ and the individual long-run estimates can again be averaged to give the mean-group estimator. A pooled version of the estimator is also available. Cross-sectional dependence can be accounted for by augmenting the regression by cross-sectional averages of unobserved common factors. A more detailed derivation of this transformation and the associated estimators are given in Appendix A.1.6.

### Endogeneity

As mentioned above, endogeneity is a likely issue when estimating IPAT models. In summary, estimation of CCE models can be done by IV in a fairly straightforward way. Due to the availability of panel data, lags of the endogenous variables are natural candidates for the instruments and their validity can typically be argued for. We will discuss specific approaches to incorporate IV estimation in a CCE framework below. We want to discuss two general issues that render a discussion on endogeneity in the context of this contribution more nuanced. Before we discuss specific IV approaches proposed in the literature we want

[2] Chudik and Pesaran (2015b) suggest to set the number of lags for the cross-sectional averages, $p_{\bar{z}}$, at $T^{1/3}$. The choice of this parameter ultimately depends on the researcher and might have to be weighed against a loss in degrees of freedom in small samples.

[3] Chudik et al. (2016) propose to set $p_{\bar{x}}$ equal to the integer part of $T^{1/3}$.

to point out that the models discussed above may naturally mitigate endogeneity in the following two ways. First, CCE accounts for general sources of cross-sectional dependence, which might attenuate correlation between the regressors and the residuals. Second, dynamic models can account for feedback effects from the dependent variable to the regressors. ARDL-type models in a pure time-series context are robust against feedbacks from the dependent variable to the regressors. In fact, these models allow for joint determination of the variable, comparable to a VAR-type setting. This property carries over to the panel setting. The drawback of these models is that they are sensitive to the lag-selection. Distributed-lag type models are more robust in this regard, but are more sensitive towards feedbacks. The models can therefore be seen as complements to each other. That being said, endogeneity should be considered in basically any model based on the IPAT-identity. This includes dynamic specifications, which can still be subject to endogeneity outside feedback effects. While panel data naturally lends itself for selecting lagged variables as instruments, the validity of this choice has to be considered carefully. Including instrumental variables in dynamic CCE models requires more careful consideration about the endogeneity and pre-determinedness of the variables included in the model that is being estimated. These models may include lagged variables as well as differences of these. As these moments are already used in the estimation, they cannot function as instruments as effectively.

CCE can be extended to include instrumental variables in several ways. Harding and Lamarche (2011) expand the CCE model from Pesaran (2006) to allow for endogenous regressors in homogeneous panels, whereas Forchini et al. (2015) explore this aspect in heterogeneous panel data by utilizing reduced form equations instead of instrumental variable (IV) estimation. Baltagi et al. (2019) extend the CCE approach by allowing for endogenous regressors and unknown common structural changes in slopes and error factor loadings. Neal (2015) extends both the static and dynamic CCE estimators of Pesaran (2006) and Chudik and Pesaran (2015b) to incorporate instrumental variables when coefficients are assumed to be heterogeneous. Particularly, lags of the endogeneous regressors are suggested as the instrument set, and OLS estimation is simply replaced by an IV-approach. This procedure is outlined in Appendix A.1.7. For consistency of the associated estimator, instruments should be exogenous, linearly independent, sufficiently correlated with the regressors, and satisfy the order condition for instruments. Furthermore, all assumptions for consistent estimation of CCE-type estimators except for exogenous regressors are still required to hold. Norkutė et al. (2021) develop two IV estimators for large dynamic panel data models with exogenous covariates and a multifactor error structure one for models with homogeneous slope coefficients and another for models with heterogeneous slope coefficients. Their approach is to project out the common factors from the exogenous covariates of the model, and to construct instruments based on defactored covariates. These estimators are linear, computationally robust and inexpensive, do not need to seek for instrumental variables outside the model and require no bias correction. Their method performs well in finite samples even for small $N$ and $T$. This procedure is implemented in STATA and therefore easily accessible. Ditzen (2018) implemented an IV version of the CCE estimators in STATA. The approach is in the spirit of Neal (2015) and uses a set of instruments with which OLS is replaced by IV estimation. The endogenous regressors are thereby regressed on all exogenous variables and instruments in a first stage and the instrumented variables substitute the endogenous regressors in a second stage. The packages are described in Table A.2.2 in the Appendix.

**Alternative Approaches**

There are alternative approaches to deal with cross-sectional dependence. We give a brief overview of this topic to provide the interested reader with important references in that literature. A popular alternative to the CCE approach is the principal components (PC) approach, which goes back to Coakley et al. (2002) and is used by Bai (2009). This method utilizes PC analysis extracting the factors to account for cross-sectional dependence in the estimation of equation (1.5). Song (2013) extends the model from Bai (2009) to the dynamic case with heterogeneous slopes. In their extensive study on panel data models featuring multifactor error structures, Karabiyik et al. (2019) compare the CCE method with the PC method. In general, the PC approach operates under the assumption that the number of factors is known. Therefore, one must first obtain a consistent estimator for the number of factors to implement this method. An advantage of the CCE method in comparison to the PC approach is its independence from prior knowledge of the number of unobserved factors. However, it is crucial to note that the rank condition, which constrains the number of factors, plays a pivotal role in determining the validity of some of the results presented by Pesaran (2006). An additional benefit of the CCE method over the PC approach lies in its avoidance of iterations. As a result, it offers computational simplicity compared to the PC approach. However, both approaches assume that the regressors are correlated only with $f_t$, and they assume stochastic independence between $x_{it}$ and $u_{it}$. Additionally, Westerlund and Urbain (2015) also give a comparison of PC and CCE. Sarafidis and Wansbeek (2012) find that the CCE approach outperforms the iterative PC method in all cases, except for when when the rank condition is not satisfied. Another advantage of the CCE estimator is that it can be extended to fixed $T$ mircropanels and retains much of its desired properties and robustness. The asymptotic theory of this extension is studied by Westerlund et al. (2019), who also find and good small-sample properties of the CCE estimators in such settings. Similar to Gonçalves and Perron (2014) for the large $N$, large $T$ case, the authors propose cross-section based bootstrap methods to improve small-sample inference. The extension to micropanels stands in contrast to principal-component based estimators that perform poorly in such settings.

Another issue that is worth mentioning is polynomial cointegration. This is particularly relevant in empirical studies of the environmental Kuznets curve, which studies the relation between economic growth and environmental impact, usually is taken in log-forms. It is often assumed that such models follow a cointegrating relation. An inverted U-shaped relation is assumed and model specifications therefore usually include polynomials of GDP, which proxies economic growth. In the simplest case, log of GDP and the square of log of GDP are included. Wagner (2015) highlights that when e.g. log of GDP follows an $I(1)$ process, powers of this integrated process may not follow the same integrated process. Regression models that aim to estimate a cointegrating relation including powers of regressors are termed cointegrated polynomial regressions (CPR). Recently, Wagner and Reichold (2023) provide theory for regressions that aim to estimate such specifications in the context of CSD. They propose a group-mean fully-modified OLS estimator that is consistent under CPR settings. Moreover, the estimator is consistent in large $N$, large $T$ settings under cross-sectional dependence and the stochastic regressors may be endogenous.

## 1.3 Empirical Applications

In this section, we apply the methodological approach outlined in section 1.2 to a simple model based on the IPAT identity. We provide a main application in the context of the EU transport sector, where we guide the reader through a careful econometric analysis of a baseline IPAT model. Additionally, in section 1.3.5 we apply our insights to three specifications related to the IPAT identity that have recently been applied in the literature and discuss their sensitivity to the applied econometric techniques. Throughout this section, we provide results from a fixed-effects as well as a two-way fixed-effects specification as our benchmarks. These will be compared against various versions of common-correlated effects type estimators.[4] The specific testing procedure pre-estimation as well as post-estimation will be discussed and a concise summary of these steps is provided in the form of a step-by-step guide in Appendix A.2.1 together with a table in Appendix A.2.2, which lists the relevant STATA packages that we used.[5] To start the empirical application, we first describe the data and then go on to discuss the baseline model that motivates the main drivers of transport-

[4]For CCE-type estimators, we use asymptotic standard errors. Note that bootstrapped versions can also be computed. We found that results generally remain unchanged.

[5]All subsequent testing and estimation procedures have been carried out in STATA 17.

related GHG emissions. We test the panel data for cross-sectional dependence, unit roots, and cointegration. In accordance with the test results, we apply the CCE model to the data (static and dynamic), discuss the results, and contrast these with fixed-effects models that do not explicitly account for nonstationarity and cross-sectional dependence. Furthermore, we apply the instrumental-variable technique to account for possible endogeneity in the regressors.

### 1.3.1 Data

We consider data for 22 EU countries over the period 1990-2019: Austria, Belgium, Bulgaria, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Netherlands, Poland, Portugal, Romania, Slovakia, Spain, Sweden, and United Kingdom. Data on GHG emissions, population, and energy consumption are taken from Eurostat. Data on real GDP per capita are obtained from The World Bank. Statistics on transport volume are from the Odyssee-Mure database (Odyssee-Mure, 2021). The panel is unbalanced and the observed time series span from a minimum of 25 to a maximum of 30 years. GHG emissions are in million tonnes of CO2 equivalent and include carbon dioxide (CO2), nitrous oxide (N2O), methane (CH4), hydrofluorocarbons (HFC), perfluorocarbons (PFC), sulfur hexafluoride (SF6) and nitrogen trifluoride (NF3). Population (Pop) is average total population in a given year. GDP per capita is in real terms and in constant 2015 USD. We compute energy intensity (EI) as final energy consumption divided by total transport activity (defined as the sum of passenger and freight activity measured in gross tonne-kilometres).

### 1.3.2 Model Framework

The IPAT model applied to our example on the EU transport sector takes the following form:

$$log(GHG_{it}) = \alpha_{it} + \beta_1 log(P_{it}) + \beta_2 log(GDP_{it}) + \beta_3 log(EI_{it}) + u_{it}, \tag{1.11}$$

where the country and year are denoted by $i$ and $t$, respectively, with $i = 1, ..., 22$ and $t = 1990, ..., 2019$. Different assumption can be made concerning the $\alpha_{it}$ such as one-way or two-way fixed effects or common correlated effects as discussed above. $GHG$ denotes the emissions from the transport sector, $P$ denotes average total population, $GDP$ is real GDP per capita in constant 2015 USD, $EI$ stands for energy intensity. These variables are all taken in natural logarithms. The coefficients can be interpreted as elasticities, i.e., the percentage change in expected GHG due to a percentage change in the regressors.

### 1.3.3 Test Results and Model Selection

The resulting data has comparable sizes in the cross-sectional and time dimensions. As outlined in section 1.2, such data structures can be characterized by nonstationarity and cross-sectional dependence. Recall the implications from theory that nonstationarity can lead from spurious interpretations to inconsistent estimates, depending on whether the panel is cointegrated and endogeneity of regressors. Sufficiently strong CSD can have serious consequences even under stationary conditions. It can lead to invalid standard errors and inconsistent estimates when the source of CSD is correlated with the regressors. To find an adequate econometric model, we therefore apply a battery of tests to select an adequate model. The procedure has the following structure. 1) Test all panel series for weak CSD, 2) apply adequate unit root and cointegration tests, 3) choose an appropriate estimation procedure and test residuals after estimation to check model adequacy.

Recall from section 1.2.3 that testing procedures for unit roots and cointegration depend on the degree of CSD in the panel series. This is the reason why we have to test for cross-sectional dependence in the panel of each variable in a first step. We perform the test on the panel series in log form as well as in log of first-differences, because we want to apply unit-root tests on the same series to test the order of integration. Table 1.1 shows the results from the Pesaran (2015) test in equation (1.3) for weak CSD against the alternative of strong cross-sectional dependence. We find strong evidence that all variables except for $log(Pop)$ are strongly cross-sectionally dependent in levels and first differences according to conventional significance levels. In levels, population is strongly cross-sectionally dependent at the 10% significance level. In first differences, it is only weakly cross-sectionally dependent.

**Table 1.1:** Test for weak cross-sectional dependence

| | Levels | | First Differences | |
|---|---|---|---|---|
| | Statistic | p-value | Statistic | p-value |
| log(GHG) | 30.28 | 0.000 | 24.47 | 0.000 |
| log(Pop) | 1.82 | 0.069 | -0.59 | 0.552 |
| log(GDP) | 69.19 | 0.000 | 45.47 | 0.000 |
| log(EI) | 12.87 | 0.000 | 4.91 | 0.000 |

Note: Test statistic by Pesaran (2015) given in equation (1.3). $H_0$: weak cross-sectional dependence against $H_1$: strong cross-sectional dependence.

To test for unit roots we apply the Pesaran (2007) second-generation CIPS test as discussed in section 1.2.3 (see equation (A.4) in Appendix A.1.3). The results are shown in Table 1.2. We cannot reject the null of homogeneous nonstationarity for any of the variables for a test specification with a constant and with or without a trend. Moreover, we find sufficient evidence to reject the null in favor of heterogeneous stationarity for variables in first differences, i.e., the variables are $I(1)$.

We next apply the cointegration test from Westerlund (2007) as discussed in section 1.2.3 (see equation (A.5) in Appendix A.1.2). Given testing results that all variables are $I(1)$, we can include all variables in the cointegration test specification, i.e. we include $log(GHG)$, $log(Pop)$, $log(GDP)$, $log(EI)$. To control for cross-sectional dependence, we compute robust p-values with 300 bootstrap replications. We conduct the test both for a specification with a constant and a trend as well as a constant only. Both specifications reject the null hypothesis of no cointegration in all panels at the 5-% significance level according to the $G_t$ mean-group test statistic (see Appendix A.1.2, equation (A.6)). We take this as sufficient evidence in favor of a cointegrating relationship.

**Table 1.2:** Second-generation unit-root test (CIPS)

| | Levels | | First Differences |
|---|---|---|---|
| | Constant | Constant & Trend | Constant |
| log(GHG) | -1.238 | -2.167 | -4.548$^{***}$ |
| log(Pop) | -0.738 | -2.506 | -2.317$^{**}$ |
| | (3.734) | (-0.639) | (-3.745)$^{***}$ |
| log(GDP) | -1.69 | -2.500 | -3.44$^{***}$ |
| log(EI) | -2.03 | -1.959 | -4.949$^{***}$ |

Note: Test statistic by Pesaran (2007) given in equation (A.4). $H_0$: Homogeneous nonstationarity. Lag selection is based on F-Test from 0 to 4 lags. Critical values were simulated according to the time-series structure of the unbalanced panel with 10,000 replications. The critical values at the 5% (1%) significance level are for the specification with 1) constant only: -2.200 (-2.381), 2) constant and a trend: -2.75 (-2.951). Statistics in parentheses are for the first-generation unit-root test from Im, Pesaran, and Shin (2003), lag selection based on AIC from 0 to 4 lags. Statistical significance is indicated by: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

The CCE estimators from section 1.2.4 are well suited to deal with the specific data characteristics such as CSD, nonstationarity, and cointegration. The static CCE is consistent even if the regressors and/or factors are nonstationary provided that the variables are cointegrated. In fact, Kapetanios et al. (2011) show that consistency of the estimates only require the residuals from the CCE estimation to be stationary. The dynamic CCE model explicitly accounts for nonstationary and cointegrated variables. These estimators distinguish between a short-run and long-run relations. All versions of the CCE estimator are available in a pooled version that assumes homogeneous coefficients and a mean-group version that assumes slope heterogeneity. We apply the test from Pesaran and Yamagata (2008) that compares a weighted fixed-effects regression with unit-specific cross-sectional OLS regressions under the alternative. We can clearly reject the null hypothesis of homogeneous slope coefficients at any meaningful significance level. Following these results, we will apply the mean-group versions of CCE estimators in the empirical application.

Small-sample properties of the static CCE estimator as well as tests on the model improve if the number of unobserved common factors is lower or equal to the number of regressors plus regressand (Kapetanios et al., 2011). The dynamic CCE models even require the number of cross-sectional averages to be at least as large as the number of unobserved common factors for consistency. We thus test the number of unobserved common factors. Test statistics from Ahn and Horenstein (2013) and Onatski (2010) report between two to three common factors. By including cross-sectional averages of both the regressand and regressors in the CCE estimations, we exceed the number of estimated factors.

### 1.3.4 Estimation Results

Table 1.3a shows the estimation results for a model with unit fixed-effects as well as one with unit and time fixed-effects as benchmarks. We compare these to those of the static mean-group common correlated effects estimation. Recall that the fixed-effects regression may be biased if variables are nonstationary. Moreover, cross-sectional dependence may even lead to inconsistent estimates under stationarity. It is, therefore, interesting to compare these results with those from a CCE regression. We note that both models are static in nature, which we will address later in this section.

As both the dependent and independent variables are in log-form, the coefficient estimates can be interpreted as elasticities. The absolute values of all coefficients are below one, implying an inelastic relation with respect to GHG emissions. In the CCE specification, for example, a 1% increase in real GDP per capita is estimated to increase transport GHG emissions by about 0.58%, all else kept constant. Similarly, a 1% decrease (i.e., improvement) in energy intensity is estimated to decrease emissions by 0.57%. The coefficient estimate for population is positive but statistically insignificant. We now compare these results to the FE and TWFE specifications. The coefficient estimates from the two models are very similar. The coefficient estimates of $log(GDP)$ and $log(EI)$ are positive and significant, as expected. However, population is estimated to have a negative effect on emissions, although the effect is statistically insignificant. Nevertheless, the result is contrary to what we would expect. The CCE estimates are more in line with our expectations. GDP is estimated to have a slightly higher effect at around 0.65 compared to 0.58 in the CCE specification. Energy intensity is estimated to have a slightly smaller effect compared to CCE. The latter provides an estimate at around 0.57 compared to 0.5 and 0.43 from the FE and TWFE models, respectively.

The results are broadly in line with what we would expect from the discussion in section 1.2.1. Since the panel series show evidence in favor of cointegration, we expect consistent estimates in case that the regressors do not correlate with the error term, but a small-sample bias can be induced. Estimates will be inconsistent under endogeneity of the regressors, an issue which we discuss below. However, first we study the model adequacy of the various specifications without IV extensions, for which we run a battery of tests on the regression residuals. Test results are shown in Table 1.5. Pesaran's (2015) test for weak CSD clearly indicates remaining CSD in the residuals from the FE model. The weighted CSD (CSDw) test from Juodis and Reese (2021) indicates that the null hypothesis of weak CSD cannot be rejected for the TWFE and CCE models. Test results for the Im, Pesaran, and Shin (2003) unit-root test for the residuals show that the null hypothesis of a unit root cannot be rejected for the FE model, but can be rejected for the TWFE at the 10% level, and even at the 1% significance level for the CCE model. Finally, we test the residuals for the number of remaining unobserved common factors following the test criterion proposed by Gagliardini et al. (2019), henceforth referred to as GOL criterion. We find evidence that all common factors are successfully removed.

We recognize that energy improvements may be driven by political decisions that in turn depend on the evolution of GHG emissions. This may also be relevant in the transport sector, in which emissions increased dramatically between 1990 and 2019. We therefore allow for endogeneity of $log(EI)$. We apply the instrumental-variable (IV) technique in the fixed-effects and static CCE models. The instrumental-variable estimation for the CCE specifications follows the implementation from Ditzen (2018). This is a particularly straightforward procedure that uses a two-stage least-squares approach instead of OLS estimation in the CCE framework. The endogenous regressors will be regressed on the exogenous variables and instruments in a first-stage, and the instrumented variables can then be used in a second-stage regression instead of the endogenous regressor. The approach is in spirit close to Neal (2015). In all IV-variants of the static models, $log(EI)$ is instrumented by three lags of itself. The choice of the amount of lags is a balance between information gain and loss in precision. Given our sample, we found three lags to be a good balance between these two criteria. We test for weak and valid instruments in the FE and TWFE models.[6] The Sargan-Hansen overidentification test cannot reject validity of all instruments and the Cragg-Donald F-statistic indicates that instruments are sufficiently strong. We additionally applied a Hausman test for endogenetiy. We compared the CCE specification without IV to the one instrumenting $log(EI)$ by three lags of itself and do not find evidence that the coefficient estimates systematically differ, i.e., there is no statistical evidence of endogeneity.

[6]Standard test statistics for weak instruments and instrument validity are invalid for mean-group models such as the estimated CCE models in this section. We therefore rely on results from the fixed-effects models to infer properties of instruments.

Nevertheless, we find it instructive to discuss and compare the estimation results from the IV specificatons given in Table 1.3b. That being said, the choice and validity of instruments should always be justified using well grounded arguments. In this spirit, we may assume that EI is pre-determined because of the following reasons. The data we study is at annual frequency and the country level, it is highly aggregated. GHG policies in the transport sector are unlikely to lead to behavioral changes of a magnitude that meaningfully influences energy intensity in a forward looking manner. This argument is strengthened by the high persistence of transport-related variables and investment decisions, such as vehicle life times and new car purchases.

In the FE and TWFE specifications, the results appear similar to the estimates without IV. The effect of population is again estimated to be negative and insignificant. The coefficient estimate for $log(EI)$ is a bit larger compared to those without IV. It is now statistically significant at the 1% level in the FE-IV estimation. In the CCE-IV specification, all coefficient estimates are larger compared to their non-IV counterparts. The estimated effect of population on transport GHG emission is almost double the size, but still statistically insignificant. The estimated coefficients of $log(GDP)$ and $log(EI)$ remain highly significant and are only slightly larger compared to the CCE specification without IV. We note that the effect of energy intensity is estimated to be larger in all IV estimations compared to their non-IV counterparts. Overall, the results appear to be quite robust against possible endogeneity in energy intensity. Table 1.5 shows that the FE-IV is again the only among the three that still shows evidence of remaining CSD after estimation and where we cannot reject the null of a unit root at any meaningful significance levels. The other models seem to model CSD adequately. We can reject the null of a unit root in the TWFE residuals barely at the 10%-level, while the test on the CCE residuals clearly rejects. We find no remaining common factors for all models.

In addition to possible endogeneity issues discussed, we recognize that policies that contemporaneously impact emissions from transport may lead to omitted variables bias. This is one reason to estimate dynamic specifications including lags of the dependent variable, which may capture much of potentially omitted variables as a natural proxy. EU countries with higher emissions may face higher pressure to reduce emissions. This hinges on the assumption that current emissions depend on past emissions, which seems a valid assumption given the persistence of transport-related variables. Another more pragmatic reason for dynamic models is that the true model is generated by dynamic interactions. Moreover, we can use additional information coming from the nonstationary series that appear to be cointegrated. In Table 1.4, we report results for a distributed-lag model augmented with cross-sectional averages (CS-DL), see model (1.10), and an error-correction specification augmented with

**Table 1.3:** Estimation results for the FE, TWFE, and CCE models

**(a)** Static models

| | FE | | TWFE | | CCE | |
|---|---|---|---|---|---|---|
| | Estimates | Std.Errors | Estimates | Std.Errors | Estimates | Std.Errors |
| log(Pop) | $-0.151$ | (0.2244) | $-0.203$ | (0.2816) | 0.3606 | (0.5340) |
| log(GDP) | $0.637^{***}$ | (0.0739) | $0.653^{***}$ | (0.0933) | $0.5839^{***}$ | (0.0841) |
| log(EI) | $0.502^{**}$ | (0.1824) | $0.434^{**}$ | (0.1864) | $0.5697^{***}$ | (0.0526) |
| Num. obs. | 624 | | 624 | | 624 | |
| N | 22 | | 22 | | 22 | |
| T | $25-30$ | | $25-30$ | | $25-30$ | |

**(b)** Static models with instrumental variables (IV)

| | IV-FE | | IV-TWFE | | IV-CCE | |
|---|---|---|---|---|---|---|
| | Estimates | Std.Errors | Estimates | Std.Errors | Estimates | Std.Errors |
| log(Pop) | $-0.2539$ | (0.1988) | $-0.1687$ | (0.1978) | 0.6854 | (0.5463) |
| log(GDP) | $0.6093^{***}$ | (0.0767) | $0.6862^{***}$ | (0.0890) | $0.6246^{***}$ | (0.0953) |
| log(EI) | $0.5754^{***}$ | (0.2036) | $0.4557^{**}$ | (0.2111) | $0.6417^{***}$ | (0.1189) |
| Num. obs. | 558 | | 558 | | 558 | |
| N | 22 | | 22 | | 22 | |
| T | $22-27$ | | $22-27$ | | $22-27$ | |

Note: Estimation outputs are based on the IPAT model given in equation (1.11). The dependent variable is $log(GHG)$. Estimates for the constant are not reported. Part (a) displays results for a one-way fixed effects estimation (FE), a two-way fixed effects estimation (TWFE), and a static mean-group CCE model based on equation (1.5). Part (b) shows estimation results using instrumental variables (IV) for each of the models. $log(EI)$ is instrumented by three lags of itself. Standard errors are in parentheses. Fixed effects models show heteroscedasticity and autocorrelation robust standard errors clustered at the country level. Standard errors for CCE-type models are based on Pesaran (2006). Standard errors are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

cross-sectional averages (CS-ECM), see model (1.9). Among the two, the CS-DL model is more robust against misspecification of dynamics and serial correlation of errors, but sensitive to feedback effects from the dependent variable to the regressors (Chudik et al., 2016). The CS-ECM model has the additional advantage that it estimates the error-correction speed of adjustment. Endogeneity should also be considered in dynamic specifications. Although ARDL-type models allow for joint determination of variables and are thus robust against feedback effects from emissions to the explanatory variables, including energy intensity. The same does not hold for the distributed-lag specification. Nevertheless, the regressors in the ARDL-type models can still correlate with the error term and endogeneity issues cannot be completely ignored. We test for endogeneity by comparing the estimators from the dynamic

**Table 1.4:** Estimation results for dynamic CCE models

| | CS-DL | | CS-ECM | |
|---|---|---|---|---|
| | Estimates | Std.Errors | Estimates | Std.Errors |
| log(Pop) | 0.3112 | (0.4245) | 0.0417 | (0.4381) |
| log(GDP) | $0.5257^{***}$ | (0.1155) | $0.6322^{***}$ | (0.1428) |
| log(EI) | $0.5266^{***}$ | (0.0840) | $0.5609^{***}$ | (0.0988) |
| L.log(GHG) | | | $-0.7319^{***}$ | (0.0792) |
| d.log(Pop) | 0.6712 | (1.4702) | 0.7002 | (1.5646) |
| d.log(GDP) | $0.1519^{*}$ | (0.0916) | $0.2551^{**}$ | (0.1067) |
| d.log(EI) | 0.0000 | (0.0651) | $0.1781^{***}$ | (0.0588) |
| Num. obs. | 602 | | 602 | |
| N | 22 | | 22 | |
| T | $24-29$ | | $24-29$ | |

Note: Estimation output is based on dynamic specifications of the IPAT model given in equation (1.11). Estimators are of mean-group type. The augmented distributed-lag specification (CS-DL) is based on model (1.10) and the augmented error-correction specification (CS-ECM) is based on model (1.9). $L.log(GHG)$ is the error-correction speed of adjustment. The negative coefficient sign indicates that deviations from the cointegrating relation adjust at a speed of around 0.73 per period after deviation. The first three rows in the Table show the long-run coefficient estimates, while the bottom three rows show short-run estimates, where $d$ indicates first differences, and $L$ is the lag-operator. The dependent variable is $log(GHG)$. Estimates for the constant are not reported. Standard errors in parantheses are based on Pesaran (2006). $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

**Table 1.5:** Test on residuals from regression models

| Model | CSD | p-value | IPS | p-value | # Factors |
|---|---|---|---|---|---|
| FE | 18.02 | 0.000 | 0.8216 | 0.7944 | 0 |
| TWFE | 0.77 | 0.443 | -1.5840 | 0.0566 | 0 |
| CCE | 0.90 | 0.366 | $-13.3867$ | 0.000 | 0 |
| FE IV | 17.95 | 0.000 | 0.8165 | 0.7929 | 0 |
| TWFE IV | 0.79 | 0.428 | $-1.3013$ | 0.0966 | 0 |
| CCE IV | 0.41 | 0.679 | $-12.1010$ | 0.000 | 0 |
| CS ECM | $-0.13$ | 0.900 | $-21.6990$ | 0.000 | 0 |
| CS DL | $-0.03$ | 0.976 | $-19.4605$ | 0.000 | 0 |

Note: Models with unit fixed effects only (i.e., FE and FE IV) are tested for weak CSD with the test from Pesaran (2015). Models that control for some form of CSD are tested with the weighted test for weak cross-sectional dependence from Juodis and Reese (2021). Both versions test $H_0$: weak cross-sectional dependence against $H_1$: strong cross-sectional dependence. 50 replications to reduce dependence on draws from Rademacher distribution. $H_0$ for Im et al. (2003) unit-root test: All panels contain unit roots against $H_1$: some panels are stationary. Lag length chosen according to AIC from lags between 0 and 2. Remaining number of unobserved common factors are estimated according to GOL criteria from Gagliardini et al. (2019).

CCE models without IV to their respective counterparts using instrumental variables. The latter again instrument energy intensity by lags of itself.[7] The Hausman test does not provide any evidence that the estimates systematically differ, indicating a lack of endogeneity.

Qualitatively, the results from the dynamic specifications match well. Both the DL and ECM specification distinguish between short-run and long-run coefficient estimates. The long-run effects of $log(GDP)$ and $log(EI)$ are estimated to be positive and significant. As in the static models, the coefficients are below one, implying an inelastic relation with emissions. The coefficients from the CS-ECM model are somewhat larger compared to the CS-DL specifications for all variables. The long-run effect of $log(Pop)$ is positive but insignificant in both models. The short-run effects of population on GHG emissions are also insignificant in the two specifications. However, their magnitude is considerably larger compared to the long-run estimates. Short-run estimates for $log(GDP)$ are positive and significant in both specifications. The short-run estimate for $log(EI)$ in the DL model is estimated to be zero, while the effect in the ECM Model is positive and significant. The error-correction speed of adjustment in the ECM model indicates that departures from the long-run cointegrating relation adjust at a speed of around 70% per period. Compared to the FE and TWFE models, we find that the long-run estimates from the dynamic specifications confirm that GDP might be slightly overestimated while EI appears to be underestimated. The difference is larger compared to the TWFE model. The dynamic models estimate the effect of GDP between 0.53 to 0.63 compared to around 0.65 from the TWFE model. The estimate for energy intensity ranges from 0.53 to 0.56 compared to around 0.43 in the TWFE estimation. All residual tests for the dynamic specifications are satisfactory.

### 1.3.5 Further Applications

We now extend our analysis to IPAT-related model specifications that have recently been published in the literature. The examples have been chosen because they fit particularly well to our baseline model from above and the data and code to replicate results is available, for which we highly credit the authors. It is not our intention to criticise any particular papers. Econometric complexities surrounding nonstationarity and cross-sectional dependence seem common in this literature and the following examples are by no means an exception. Rather, they show that ignoring the points we raised in this article can have significant consequences for conclusions drawn from the analyses. As we stated earlier, this need not necessarily

[7] The dynamic specifications with IV are based on the models shown in Table 1.4. The log of energy intensity is again instrumented by three lags of itself. We additionally instrument the difference of log energy intensity by three lags of itself. Estimation results are in line with the specifications without IV. Results are not reported here but are available upon request.

be the case but depends on the specific data structure of each application. Consequences can range from being negligible to drawing conclusions that are the opposite of what the data actually support. In the examples below, the data is characterized by nonstationarity and strong CSD. Our analysis contrasts the estimation approaches in the original papers to estimation procedures we recommend.

**Human Development**

In the first application, we re-investigate the study from Opoku et al. (2022), who examine the effect of human development on environmental sustainability. The authors argue that the related literature predominantly focuses on the nexus between human capital and emissions. Estimation results on this relation are on a wide spectrum, ranging from an estimated positive relation (increasing emissions) to an estimated negative linkage (reducing emissions) and no significant effect at all. For a comprehensive overview of the literature and specific citations, we refer to Opoku et al. (2022). They note that human development encapsulates more dimensions than only human capital. Their study uses the UN Human Development Index (HDI) to proxy human development and to capture a more expansive array of dimensions, including health, knowledge, and economic power. They study the relation between HDI and several measures of environmental sustainability (including CO2 emissions) and show a consistently significant negative effect, implying that higher human development is associated with a reduction in emissions. This finding is consistent across several model specifications for 33 OECD countries over the period 1996-2016.

We use both the model and data from Opoku et al. (2022) to re-estimate the following model:

$$\begin{aligned} log(CO2_{it}) =& \alpha_{it} + \beta_1 HDI_{it} + \beta_2 log(Inc_{it}) + \beta_3 Trade_{it} + \\ & \beta_4 log(Tec_{it}) + \beta_5 log(Pop_{it}) + u_{it}, \end{aligned} \tag{1.12}$$

where $log(CO2_{it})$ is log of CO2 emissions in tonnes, $HDI$ is the UN Human Development Index and proxies human development, $Inc$ stands for income and is measured by log of real (constant, 2010 USD) GDP, $Trade$ measures trade openness as the total trade as a percentage of GDP, $log(Tech_{it})$ measures technology and is given by the log of the total expenditure on research and development, and $log(Pop_{it})$ is the log of total population.

Opoku et al. (2022) estimate a FE model as a benchmark, but they also consider that the explanatory variables may be endogenous. To account for this, they estimate an IV-GMM model that instruments all regressors by one lag of each regressor. A specific discussion on potential sources of endogeneity and validity of instruments is missing. However, Opoku et al. (2022) report that an overidentification test indicates that all instruments are valid. The associated estimation results from the two models are shown in Table 1.6 in columns (1) and (2). The coefficient estimates match remarkably well and statistical significance is also similar. The estimation results based on these two models suggest that a one percentage-point increase in the $HDI$ significantly decreases CO2 emissions by about 3%. However, the Pesaran (2015) CSD test finds evidence of strong CSD for all variables. Further, we conduct the Pesaran (2007) CIPS test and find mixed evidence regarding nonstationarity. Panel unit-root tests suggest that CO2 and HDI have a unit root although the results are not perfectly robust. In addition, we implement the Westerlund (2007) cointegration test and find no evidence in favor of a cointegrating relation. Even though Opoku et al. (2022) use Driscoll-Kraay standard errors (Driscoll and Kraay, 1998) to account for CSD, the parameter estimates from the FE and IV-GMM may not be robust due to the indicated lack of cointegration.

To accommodate CSD, we estimate three specifications. First, a TWFE model that accounts for constant contemporaneous correlation between countries. The estimates from this model are shown in column (3). The effect of $HDI$ is now estimated to be statistically insignificant. Additionally, the sign is flipped and a positive coefficient is estimated. The effect of population is now positive and significant. The estimates for income, trade openness and technology are comparable to the original specifications. As discussed earlier, imposing constant correlations may be overly restrictive. We therefore estimate a CCE model as a second specification in column (4). The pooled version was chosen according to the test from Pesaran and Yamagata (2008).[8] The CCE model is more flexible in accommodating cross-sectional correlations between the panel units compared to the TWFE specification. To explore the robustness of the CCE specification to endogeneity, we estimate the model with the same IV specification as in Opoku et al. (2022). That is, we include one lag of each regressor as the set of instruments and treat all regressors as endogenous. Estimation results for this IV-CCE specification are given in column (5). The weighted CSD test from Juodis and Reese (2021) as well as the test for remaining unobserved common factors from Gagliardini et al. (2019) indicate that both CCE specifications accommodate CSD sufficiently well.

The estimates of the CCE and IV-CCE are very similar. The CCE model estimates the coefficient of $HDI$ at around $-0.34$, while the IV-CCE specification gives an estimate at around $-0.58$, both of which we find to be statistically insignificant. These estimates are markedly lower compared to the statistically significant coefficient of $-3$ estimated by Opoku

[8]The test result is a borderline case that can motivate both a pooled and mean-group version. However, the pooled version generally shows better small-sample properties and was chosen therefore.

Table 1.6: **Revisiting Opoku et al. (2022)**

| | (1) FE | (2) IV-GMM | (3) TWFE | (4) CCE | (5) IV-CCE |
|---|---|---|---|---|---|
| HDI | -3.034*** | -3.215*** | 0.608 | -0.343 | -0.581 |
| | (0.503) | (0.397) | (1.045) | (0.394) | (0.429) |
| log(GDP) | 0.461*** | 0.574*** | 0.286** | 0.557*** | 0.586*** |
| | (0.145) | (0.085) | (0.127) | (0.081) | (0.095) |
| Trade | -0.003*** | -0.003*** | −0.002** | 0.001* | 0.001** |
| | (0.001) | (0.000) | (0.001) | (0.000) | (0.000) |
| log(Tech) | 0.138*** | 0.116*** | 0.176*** | -0.001 | -0.014 |
| | (0.047) | (0.031) | (0.040) | (0.023) | (0.025) |
| log(POP) | 0.122 | -0.141 | 0.872*** | 0.706*** | 0.911*** |
| | (0.253) | (0.144) | (0.230) | (0.155) | (0.160) |
| $N$ | 33 | 33 | 33 | 33 | 33 |
| $T$ | 21 | 20 | 21 | 21 | 20 |

Note: Estimation output is based on equation (1.12). The dependent variable is $log(CO2)$. Estimates for the constant are not reported. Column (1) shows a one-way FE model with Discroll-Kraay robust standard errors and column (2) shows IV-GMM with Discroll-Kraay robust standard errors, where each variable is instrumented by one lag of itself. Results in the first two columns are taken from Opoku et al. (2022). Column (3) shows a TWFE model with robust standard errors that are robust against heteroscedasticity and autocorrelation, they are clustered at the country level, column (4) shows the estimation output using a pooled CCE model, and column (5) shows estimation results of pooled CCE using instrumental variables (IV), where each regressor is treated as endogenous and the instruments set includes one lag of each regressor. Standard errors for CCE specifications are robust against heteroscdasticity and autocorrelation. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

et al. (2022). While the FE and IV-GMM specifications estimate significant negative effects for trade openness and significant positive effects for technology, the CCE specifications find contrary signs for both. The CCE model and the CCE model using IV find some statistical significance for trade openness, whereas they do not imply statistical significance for technology. Our estimated impact of income on emissions is similar in comparison to the FE and IV-GMM specifications. Notably, we observe a positive significant effect of population, which Opoku et al. (2022) find to be statistically insignificant. This results resonates with our findings from our baseline model above.

Overall, our findings suggest a potential overestimation of the effect of human development on environmental sustainability within the framework of the model specified in (1.12). We want to emphasize that this finding is contingent upon the specific model configuration and sample studies. The outcome may not be replicable under alternative specifications or a different dataset. Recall that HDI integrates measures related to health, knowledge, and economic power. The latter dimension is independently captured by GDP as a separate explanatory variable in the model. This warrants a more nuanced interpretation of HDI when explicitly accounting for economic power. HDI may also capture the thought that higher human development leads to technological advancements that may be more environmentally friendly. Technology is another explanatory variable in the model that can control for this effect. The remaining explanatory power of HDI may therefore not be related to all three dimensions that HDI captures equally. Opoku et al. (2022) acknowledge this complexity and provide additional estimation results in which they substitute HDI with education in one specification and human capital in another model. They again find significant negative coefficient estimates for the two new models. We also re-estimated the model with these specifications and the results are again negative but statistically insignificant.[9] Given our results and the thoughts outlined, we caution against taking this re-estimation exercise as clear evidence that HDI has no significant effect on environmental sustainability. Rather, additional research on this topic is required.

### Environmental Regulation and Innovation

For the second application, we revisit Hashmi and Alam (2019), who investigate the impact of environmental regulation and innovation on CO2 emissions across 29 OECD countries over the period 1999-2014. The authors note that OECD countries incorporate environmental taxation and green innovation in their national strategies to reduce GHG emissions and meet international emission reduction targets. They argue that it is therefore imperative to empirically assess the relation between these two market-based policy options and their impact on emissions. Their findings suggest that both market-based policies have the potential to significantly reduce CO2 emissions. However, they also note that this beneficial effect can be counteracted by a comparatively larger increase in emissions due to non-green innovation, measured as non-green patents. The authors also find that emissions are mainly driven by population and affluence, such that the effect of environmental taxation and green innovation is small in comparison. To re-visit this study, we draw on the model and data from Hashmi and Alam (2019). The model is given by:

$$\begin{aligned} log(CO2_{it}) =& \alpha_i + \gamma_t + \beta_1 log(Pop_{it}) + \beta_2 log(EnTax_{it}) + \\ & \beta_3 log(EnPat_{it}) + \beta_4 log(NoEnPat_{it}) + u_{it}, \end{aligned} \tag{1.13}$$

where CO2 emisions are the total CO2 emissions in kt that stem from the burning of fossil fuels and the manufacturing of cement. Pop is total population, GDP is measured in per capita terms, EnTax are environmentally related tax revenues per capita, EnPat are envi-

[9]Results are available upon request.

ronmental patent counts an NoEnPat are non-environmental patent counts. All variables are given in natural logarithms. Table 1.7 shows the estimation results based on the analysis and data from Hashmi and Alam (2019). The results from the TWFE and IV-GMM from the authors are given in column (1) and column (2), respectively. Environmental taxes and green patents mitigate emissions, while non-green patents augment them. All effects are estimated to be statistically significant.

We implement the Pesaran (2015) CSD tests and find strong CSD in all variables. It remains ambiguous whether the TWFE can accommodate this CSD sufficiently well. While the GOL criterion from Gagliardini et al. (2019) does not find remaining unobserved common factors after estimation, the weighted CSD statistic from Juodis and Reese (2021) indicates remaining CSD. We then go on to test the variables for stationarity. The Pesaran (2007) CIPS test with a trend shows that all variables are nonstationary except for environmental patents, while the test without a trend provides evidence that all variables are nonstationary. The implementation of a Westerlund (2007) cointegration test indicates that there is no cointegration in the data. In column (3) of in Table 1.7, we estimate a CCE model. The diagnostic test on the residuals from these models proposed by Juodis and Reese (2021) find no remaining CSD and the GOL criteria from Gagliardini et al. (2019) find zero remaining unobserved common factor. The pooled version was chosen according to the test from Pesaran and Yamagata (2008).

Hashmi and Alam (2019) estimate a TWFE model as their main specification. They additionally discuss potential endogeneity and estimate an IV-GMM specification. Specifically, it is argued that regulation and technology (i.e., patents) may be endogenous. These variables are treated as endogenous regressors, and lags of these variables are used as instruments. We again follow in the spirit of the original contribution and re-estimate our CCE specification with IV accordingly. We are, however, constrained in the sample size and cannot accommodate all three variables as endogenous regressors in one specification. Therefore, we decided to run two separate IV-CCE regressions. One only treats regulation as endogenous and instruments this variable by one lag of itself. The estimates are shown in column (5). The second IV-CCE specification treats the technology terms as endogenous. The instrument set is populated by lags of these regressors and the results are given in column (4). Hausman tests for endogeneity comparing the IV-CCE specifications to the CCE model reject, implying that indeed endogeneity is an issue here.

All model specifications provide qualitatively similar results. Particularly, the effect of regulation ($EnTax$) is estimated to significantly reduce emissions. The TWFE and all CCE specifications provide relatively consistent estimates ranging from $-0.03$ to $-0.05$, while the IV-GMM provides an estimate at $-0.2$. The effect of GDP is estimated quite consistently

**Table 1.7:** Revisiting Hashmi and Alam (2019)

| | (1) TWFE | (2) IV-GMM | (3) CCE | (4) IV-CCE (Pat) | (5) IV-CCE (Tax) |
|---|---|---|---|---|---|
| log(POP) | 1.503*** | 0.965*** | 1.952*** | 1.091*** | 0.907*** |
| | (0.119) | (0.020) | (0.309) | (0.291) | (0.290) |
| log(GDP) | 0.493*** | 0.580*** | 0.595*** | 0.642*** | 0.589*** |
| | (0.068) | (0.050) | (0.117) | (0.114) | (0.113) |
| log(EnTax) | -0.0298** | -0.196*** | -0.0348*** | -0.039*** | -0.051*** |
| | (0.011) | (0.026) | (0.0122) | (0.015) | (0.016) |
| log(EnPat) | -0.017** | -0.069* | -0.0110** | -0.006 | -0.007 |
| | (0.008) | (0.037) | (0.00545) | (0.006) | (0.005) |
| log(NoEnPat) | 0.113*** | 0.116*** | 0.00582 | -0.011 | 0.024 |
| | (0.017) | (0.039) | (0.0195) | (0.021) | (0.020) |
| $N$ | 29 | 29 | 29 | 29 | 29 |
| $T$ | 16 | 16 | 16 | 15 | 15 |

Note: Estimation output is based on equation (1.13). The dependent variable is $log(CO2)$. Estimates for the constant are not reported. Column (1) shows the TWFE model with Driscoll–Kraay robust standard errors and column (2) shows the IV-GMM model with Driscoll–Kraay robust standard errors. Results in the first two columns are taken from Hashmi and Alam (2019). Column (3) shows the estimation output using a static pooled CCE specification. Column (4) shows estimation results of pooled CCE using IV for the variables $log(EnPat)$ and $log(NoEnPat)$, which are instrumented each by one lag of itself. Column (5) shows estimation results of pooled CCE using IV for the variable $log(EnTax)$, instrumented by one lag of itself. Standard errors for CCE specifications are robust against heteroscdasticity and autocorrelation. Standard errors are given in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

across all specification, ranging from 0.5 to 0.65. Coefficient estimates for population vary slightly more, but are nonetheless comparable, ranging from 0.9 to 1.5. It is interesting to note that all IV specifications provide smaller estimates compared to the non-IV counterparts. Finally, we find larger differences for the estimates regarding patents. While the sign of both environmental and non-environmental patents estimates is consistent throughout all models, they vary in magnitude and statistical significance. The TWFE and CCE models provide similar estimates for environmental patents at around $-0.17$ and $-0.011$, respectively. The IV-CCE models estimate the effect at around half this magnitude and report no statistical significance. Contrary, the IV-GMM model estimates a significantly larger effect. The effect of non-environmental patents is estimated at similar magnitude in the TWFE and IV-GMM models at 0.11. All CCE specifications provide statistically insignificant results.

Overall, we can confirm some main findings and conclusions from Hashmi and Alam (2019). Namely, that environmental taxation appears to be an effective market-based policy options to reduce CO2 emissions. Contrary to the authors, we do not find strong evidence of a statistically significant effect of green innovation. Aside from this inference, we can confirm the finding from the authors that the effect of regulation is estimated to be stronger compared to green innovation. Non-green innovation is again estimated to be statistically insignificant. This, however, confirms the authors note that the emissions reduction of regulation policies is not significantly counteracted by non-environmentally friendly patents. Finally, we agree with the authors in the assessment that the main driving forces for emissions in the model are population and affluence, where we estimate these effects to be even larger compared to the authors' results.

### Transport Modes and Energy Sources

For the third application, we follow Andrés and Padilla (2018), who study the drivers of emissions from the transport sector for EU countries. The authors employ the IPAT model in this context and decompose the technology term into transport activity and energy intensity. These terms are additionally disaggregated into different modes of transport (road, rail, water, aviation) and energy sources (oil, electricity, renewable, gas), respectively. The authors estimate a FE model, a specification with panel corrected standard errors (PCSE), and a feasible generalized least-squares (FGLS) model. Small-sample properties and consistency of the resulting estimators may be affected by nonstationarity and CSD, which is not accounted for adequately. The main findings of Andrés and Padilla (2018) are that a switch from road to rail and inland water transport as well as from oil to electricity, gas, and renewables (which are basically biofuels) significantly reduces emissions. Due to retrospective changes in the dataset used by Andrés and Padilla (2018), we have to restrict our database to some degree and cannot replicate the original study exactly.[10] While Andrés and Padilla (2018) studied 25 EU countries from 1990-2014, we have to restrict our sample to 12 countries, but can extend the time period to 1990-2019. We will see that our analysis can qualitatively support the main results from the authors. However, we note that results from a FE and TWFE estimation can provide misleading results.

The resulting model is a particularly extensive model among IPAT specifications. It is given by:

$$\begin{aligned} log(GHG_{it}) =& \alpha_i + \beta_1 log(GDP_{it}) + \beta_2 log(Pop_{it}) + \beta_3 log(TA_{it}) + \beta_4 log(EI_{it}) \\ &+ \sum_{j=1}^{J-1} \mu_j M_{jit} + \sum_{k=1}^{K-1} \Omega_k S_{kit} + u_{it}, \end{aligned} \tag{1.14}$$

[10] The data on transport activity from Odysee-Mure initially contained many zero values for some variables in some countries that have since been updated to NA values.

where $i$ indexes the country, $t$ time, $j$ transport mode, and $k$ energy source. Note that this is based on the same data and a similar but extended IPAT-model as in our baseline application above. Environmental impact is measured by GHG emissions from transport (road and freight), Population (POP) is average total population, and Affluence is proxied by real GDP per capita. The technology term includes Energy Intensity (EI) and Transport Activity. The former is defined as final energy consumption divided by total transport activity (defined as the sum of passenger and freight activity measured in gross tonne-kilometres). Transport activity (TA) is defined as per capita freight activity (measured in tonne-kilometers per 1000 capita). The technology term is further decomposed into transport modes (road, rail, water, aviation) and energy sources (oil, gas, electricity, renewable). All variables that are not shares are taken in logarithmic form. Data is taken from the OECD and Odyssee-Mure databases. As a benchmark, we estimate the model using the FE approach (column 1, Table 1.8), which we choose over the RE model according to the Hausman (1978) specification test.[11] As an additional benchmark specification, we estimate a two-way fixed-effects model in column (2).

The results from the FE and TWFE specifications are very similar. The effect of variables that are not shares can be interpreted as elasticities. We find similar positive and significant effects of those variables in both specifications. Shares can be interpreted similarly with respect to their reference category. Shifts in transport modes are interpreted relative to road transport, while shifts in energy sources are interpreted relative to oil. Specifically, we find that rail transport leads to reduced emissions relative to road transport, while water and aviation contribute to increased emissions. A shift from road to water transport is statistically insignificant in both models, but the TWFE model estimates a negative effect as expected, while the FE model estimates a positive effect. A shift from oil to renewables and gas reduces emissions, while a switch to electricity increases them, which appears counterintuitive.

The counterintuitive results might be a consequence of CSD and nonstationarity of the data. We test the data for CSD with the test from Pesaran (2015) and find significant evidence of strong CSD for all variables except for $rail\_share$ and $water\_share$. Next, we find all variables except for $renew\_share$ to be nonstationary according to the CIPS test from Pesaran (2007). Testing for cointegration in this application is complicated by the fact that shares are naturally bounded between 0 and 1[12]. We consequently conduct a cointegration test for the variables that are not shares and find no evidence to support a cointegrating

[11] Despite differences in samples, our FE results do not substantially differ from the results from Andrés and Padilla (2018). An exception concerns *water_share*, which they estimate to be negative and significant, while we find a positive but insignificant effect in the FE specification.

[12] Additionally, the model exceeds the maximum number of variables that can be tested with the specification from Westerlund (2007) in STATA 17.

**Table 1.8:** Revisiting Andrés and Padilla (2018)

| | (1) FE | (2) TWFE | (3) CCE-P | (4) CCE-P_Diffs |
|---|---|---|---|---|
| log(GDP) | 0.289*** | 0.189* | 0.283*** | 0.187** |
| | (0.060) | (0.101) | (0.083) | (0.085) |
| log(POP) | 0.743** | 0.567 | 1.289*** | 0.986*** |
| | (0.272) | (0.320) | (0.062) | (0.301) |
| log(TA) | 0.478*** | 0.483*** | 0.308*** | 0.355*** |
| | (0.038) | (0.038) | (0.034) | (0.032) |
| log(EI) | 0.832*** | 0.860*** | 0.598*** | 0.628*** |
| | (0.090) | (0.100) | (0.057) | (0.053) |
| rail_share | -0.640*** | -0.613*** | -0.613*** | -0.435*** |
| | (0.177) | (0.175) | (0.149) | (0.134) |
| water_share | 0.0511 | -0.353 | -3.252*** | -1.514*** |
| | (0.826) | (1.139) | (0.739) | (0.474) |
| aviation_share | 1.265* | 1.296** | -0.0794 | 2.081 |
| | (0.630) | (0.533) | (1.397) | (1.789) |
| elec_share | 2.761*** | 2.527** | -1.671* | -2.604*** |
| | (0.854) | (0.952) | (0.966) | (0.896) |
| renew_share | -1.104*** | -1.327*** | -1.231*** | -1.053*** |
| | (0.232) | (0.234) | (0.089) | (0.122) |
| gas_share | -0.513*** | -0.592*** | -0.954*** | -1.625*** |
| | (0.146) | (0.160) | (0.281) | (0.248) |
| $N$ | 12 | 12 | 12 | 12 |
| $T$ | 25-30 | 25-30 | 25-30 | 24-29 |

Note: Estimation output is based on equation (1.14). The dependent variable is $log(GHG)$. Estimates for the constant are not reported. Standard errors in parentheses are robust against heteroscedasticity and autocorrelation, they are clustered at the country level for fixed effect specifications. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All variables that are not shares are in natural logarithms.

relation. We run the diagnostic tests on the residuals from the FE and TWFE models from Pesaran (2015) and Juodis and Reese (2021), respectively. The tests indicate remaining CSD in the residuals from both models. To account for these complications, we estimate both a CCE model and a CCE model in differences. For both specifications the pooled version was chosen according to the test from Pesaran and Yamagata (2008). We want to note that while endogeneity can play a role in this application, this is an issue not discussed in the original contribution. Unfortunately, the sample size is limited to such a degree that lagged variables as instruments are not feasible. The same reason prevents us from estimating dynamic specifications.

Estimation results for the two CCE specifications are presented in columns (3) and (4) of Table 1.8. We first compare results for variables that are not shares. The effect of GDP is comparable to FE and TWFE results. Population is estimated to have a larger effect. The FE specification estimates this effect at around $-0.7$, while the CCE and CCE in differences give estimates of around $-1.3$ and $-1$, respectively. Contrary, transport activity is estimated to have a smaller effect. The FE and TWFE model provide estimates of about $-0.5$, while the CCE models show coefficients at around $-0.3$. Similarly, energy intensity is estimated to have smaller effects. The two CCE specifications give estimates at around $-0.6$ compared to the FE and TWFE models with coefficients at around $-0.8$. Regarding shifts in the transport mode relative to road, we find that water transport now significantly lowers emissions, while the FE and TWFE models report the effect to be statistically insignificant. $aviation\_share$ is insignificant in the CCE specifications.[13] Finally, regarding shifts in energy sources relative to oil, we estimate that a shift from oil to electricity significantly reduces emissions, which is more in line with expectations, while the FE and TWFE models find the effect to be positive and significant. Neither of the two CCE specifications exhibits remaining CSD, affirming the expected result: both a shift from road to water transport and a move from oil to electricity contribute to emission reductions. Our estimates thus confirm the main arguments from Andrés and Padilla (2018) that switching from road and oil to alternative transport modes and energy sources reduces transport emissions significantly. Note, however, that the FE and TWFE specifications we estimated cannot replicate these results. Specifically, the effect of a switch from oil to electricity is estimated to be positive and significant in these specifications.

## 1.4 Conclusion

We discuss the econometric challenges that arise in the setting of nonstationary panel data where the cross-sectional and time-series dimensions are of comparable size, which is the type of data often used to study drivers of GHG emissions. Nonstationarity should not simply be ignored and such panels should always be tested for unit roots and cointegration. Transforming the data into first differences can be a solution for nonstationary data, but valuable information can be lost that leads to a loss in precision of results and policy implications. This is particularly relevant when variables exhibit a cointegrating relation. In particular, when variables are taken in differences, it is no longer possible to differentiate between short-run effects and those that alter the long-run equilibrium relation of the estimated model. The latter seems to be a crucial aspect for policy advice related to emission reduction. In addition to complications surrounding nonstationary data, cross-sectional dependence has to be considered. Consequences of ignoring this dependency structure can

[13] We attribute this insignificance as well as the positive point estimate of the CCE in differences to the fact that inland aviation plays a minor role in total transport emissions for many of the observed countries.

cause invalid test statistics, incorrect inference and inconsistent estimates. We recommend using common correlated effects (CCE) models that have desirable properties even in small samples in a variety of cases: under stationarity, unit roots, cointegration and often even under its absence. Dynamic and ECM versions of the model are also available.

We take the insights regarding the econometric methodology to empirical applications related to the IPAT identity and carefully guide the reader through the steps in the econometric analysis. We start by investigating a simple version of the IPAT model applied to the EU transport sector. We apply the static as well as dynamic versions of the CCE model to distinguish short-run and long-run effects. We contrast these results with fixed effects models that do not explicitly account for nonstationarity and cross-sectional dependence. Additionally, we compare IV-versions of fixed-effects and CCE estimators to account for possible endogeneity of the regressors. Using the appropriate methods, some difference in the estimates arise. These crucially depend on data characteristics, including, nonstationarity, cross-sectional dependence, cointegration, and endogeneity. Using the appropriate methods, some difference in the estimates arise. The insights from our contribution are applicable to a wide range of literature related to empirical analyses of GHG emissions, but also to other applications that involve panel data. In particular, they are relevant in all models that may be subject to cross-sectional dependence and nonstationarity. Many models related to the IPAT identity are examples thereof.

We thus additionally estimate three of such specifications that have recently gained attention in the literature. These include studies on the effect of human development, regulation, and transport modes and energy sources on emissions. We find that conventional methods ignoring nonstationarity and CSD lead from small differences in estimation results, possibly caused by using biased estimators, to (apparently) wrong coefficient signs and potentially incorrect standard errors. Some differences in policy implications are discussed. To be specific, we could not find sufficient evidence to conclude that higher human development decreases emission significantly. Our estimates suggest that we can confirm that environmental regulation significantly reduces emissions, while we estimate weaker and statistically not significant effects from green technology on emissions. We find evidence that suggest that switching from road to water transport and from oil as an energy source to electricity both significantly reduce emissions. We could not replicate these results with fixed-effects models.

Overall, we think that these findings confirm the importance of careful econometric estimation procedures. In practice, we recommend the following procedure to practitioners dealing with nonstationary panel data, more details on which can be found in the appendix along with STATA commands. 1) Adequately test the data for CSD, 2) apply correct unit root and cointegration tests, 3) if data are cointegrated the static CCE remains consistent, but it is preferable to apply a dynamic CCE model that captures the additional information provided (if possible), 4) if data are not cointegrated, static and dynamic CCE specifications can still consistently estimate long-run relations under certain conditions; CCE models in first-differences can additionally be applied, 5) apply battery of tests on residuals to confirm model adequacy.

# Chapter 2

# Shifting Gears? Austria's Transport Policy Mix and CO2 Emissions from Passenger Cars

Policy makers likely have to resort to a differentiated mix of complementary policy measures to achieve global targets on carbon-neutrality. To help policy makers design effective measures, we analyse the effect of environmental policies on CO2 emissions from passenger cars in Austria from 1965-2019. In a first step, we propose a novel environmental policy stringency index tailored to the Austrian transport sector for the period 1950-2019. In a second step, we analyse the effect of different policies on transport-related CO2 emissions. We use a structural vector autoregressive model to allow for interdependencies between the policies and remaining endogenous variables. We find that policies targeting the investment decision to buy new cars reduced emissions in Austria more significantly than policies targeting the usage of cars. The engine-related insurance tax quantitatively shows the strongest impact on emissions, while the standard fuel consumption tax shows the strongest statistical significance.

## 2.1 Introduction

Understanding the impact of past transport policies is crucial for guiding future policy designs to meet both national and EU climate goals. In this paper, we analyze the effectiveness of transport-related policies in Austria over the period 1950–2019 within a dynamic econometric framework that accounts for systemic delays and interdependencies inherent in the transport sector. As a first step, we develop a novel policy stringency index for the Austrian transport sector by carefully selecting policies in collaboration with experts from the Austrian Environmental Agency. This index is then integrated into our econometric model to identify which policies had a significant impact on transport emissions.

Several characteristics of our study make it a unique contribution to the literature. First, to the best of our knowledge, a transport-specific policy stringency index does not exist for any country. Second, by covering a long time period from 1950 to 2019, we exploit stronger variations in policy stringencies than most other studies. Third, we incorporate the index in a dynamic econometric framework that can deal with possible endogeneity and interdependencies. The model allows us to study the dynamic diffusion of the effect of policy bundles as well as single policies over time while acknowledging interdependencies. Overall, such a comprehensive analysis is novel in the literature.

Our research is highly relevant in both the national and European context. The European Union (EU) has set an ambitious climate-neutrality target for 2050 (EC, 2021) and is currently implementing more stringent emission targets with its "Fit-for-55" package to achieve this goal. The transport sector will play a crucial role in the transition towards a carbon neutral society. While overall greenhouse gas (GHG) emissions in the EU decreased by 28% during 1990-2019, emissions from transport increased by 20% and in 2019 accounted for about a quarter of the EU's total GHG emissions. The largest share of transport emissions (mostly CO2) stems from road transport (EEA, 2021). One major pillar of the "Fit-for-55" package to reduce transport emissions is the Effort Sharing Regulation (ESR), under which each member state has to fulfill binding emission targets through implementation of national policies.

We focus our investigation on Austria, because it poses a particularly interesting case for analysis within the EU. It was one of the first countries to ratify the Paris Agreement, and it set itself the ambitious goal to become carbon-neutral by 2040 (BKA, 2020). However, between 1990 and 2019, GHG emissions from transport in Austria increased by 74.4%. In 2019, the transport sector accounted for 30% of total GHG emissions and 19% of total emissions were emitted by road passenger transport alone (Anderl, Bartel, et al., 2021). Policy instrument packages to meet Austria's environmental target are yet to be implemented and existing policy measures are not expected to achieve a significant reduction in motorized individual transport emissions (Anderl, Gössl, et al., 2021).

We recognize that the transport sector is characterized by systemic delays, partly due to the relatively long lifetime of vehicles. Policies aimed at influencing the existing vehicle stock may have delayed effects on emissions, which must be considered when devising effective measures. Our dynamic econometric approach allows us to capture these time lags and address likely endogeneity among determinants of emissions. While numerous studies have examined individual transport policies, comprehensive analyses of policy mixes and their interdependencies over extended periods are scarce. Existing research often focuses on specific policies or short time frames, limiting the understanding of long-term policy effectiveness. By covering a long period from 1950 to 2019 and constructing a unique transport-specific policy stringency index, our study fills this gap in the literature.

The importance of policy mixes in addressing transport emissions is notably emphasized by Dugan et al. (2022), who analyse a range of policy packages that comprise different policies for Austria with a computable general equilibrium model. They show that a balanced policy package can mitigate negative effects associated with single policies. Winkler et al. (2023) focus on London as a case study and argue that meeting carbon-budgets that are compatible with meeting the Paris agreement will require significant reductions in car use in addition to changes in vehicle design. Koch et al. (2022) use a break detection method for time series on transport emissions and attribute breaks to policy changes, finding no effective transport-related policies for Austria. Stechemesser et al. (2024) use a similar but refined method to study over 1,500 policies in 41 countries. For Austria, they find a single significant policy effect around the year 2005, which they attribute to a combination of increases in the fuel tax and truck tolls. Gerlagh et al. (2018) examines the effect of fiscal policies on vehicle efficiency, including in Austria, finding that more CO2 sensitive registration taxes reduced new vehicle emissions.

Studies on individual transport policies are more prevalent than those examining policy mixes and interdependencies between policies. For example, Ostermeijer et al. (2019) investigate the impact of residential parking costs on car ownership in the Netherlands, indicating significant variances in parking costs and their effect on car ownership rates. Adler and van Ommeren (2016b) explore the congestion relief benefits of public transit during transit strikes in Rotterdam, finding substantial reductions in car congestion. Hintermann et al. (2022) provide insights into Pigovian road pricing through a large-scale field experiment in Switzerland. They find that the group that received a pricing treatment significantly reduced external transport costs.

Andersson (2019b) explores the effect of a carbon tax in Sweden and finds a modest reduction in transport emissions, while Pretis (2022a) studies a carbon tax in British Columbia and finds that it led to a decrease in transport emissions only, with no significant impact on overall emissions. Berger et al. (2022) estimate that the impact of speed limit policies can significantly reduce greenhouse gas emissions and improve road safety in Austria. Kuss and Nicholas (2022) conduct a meta-analysis of measures aimed to reduce car usage in European cities, identifying a dozen effective strategies, including congestion charges, parking and traffic control, and limited traffic zones.

We find that the observed increase in the stringency of policies that target the investment decision to buy cars has proven more effective than those observed targeting the usage of cars. The engine-related insurance tax is found to have had the strongest long-run impact on emissions in our study. But the standard fuel consumption tax - an emission sensitive tax on new vehicles - shows the strongest statistical significance in reducing emissions. The effect of both policies comes with a time delay as it takes time for more efficient cars to significantly impact fleet emissions. That being said, the magnitude of the effect of any policy that we consider is very limited. Austria will need to drastically increase these policies in stringency and implement additional measures to meet its policy targets.

The remainder of the paper is organised as follows. section 2.2 is dedicated to the computation of the policy stringency index. It starts with a short literature overview, discusses transport-related policies in Austria, and then explains the construction of the index. section 2.3 establishes the econometric model used to analyse the determinants of transport related CO2 emissions and provides empirical results. section 2.4 provides a policy discussion, and section 2.5 concludes.

## 2.2 Policy Stringency Index

In this section, we develop the policy stringency index for the Austrian transport sector. We first provide a brief literature overview on the construction of indexes and the incorporation of policy stringency measures, such as indexes, in econometric analyses. Then, we go on to discuss transport-related policies that Austria introduced over the period 1950-2019. We group these policies into two categories, one related to the purchase behavior of new vehicles and one related to the usage of vehicles. The policies and categories have been established with the help of policy experts from the Austrian Environmental Agency. Finally, we describe how we assign stringency scores to the different policies for every year and discuss the resulting stringency index.

### 2.2.1 Literature

Indexes that aim to quantify the stringency of policies face the problem of multidimensionality (e.g. Brunel and Levinson, 2016; Galeotti et al., 2020). Countries can draw on a diverse toolkit comprising different measures to achieve policy goals. These may include taxation, subsidies, and regulation. Individual policies can be characterized by disparate levels of effectiveness and metrics. An index has be constructed such that these different policies are comparable on a common scale. Several approaches to devise such indexes have been proposed in the literature. Among them, composite indexes have more recently gained popularity. These indexes aim to aggregate individual indicators by simply counting the number of regulations or the use statistical and data-driven techniques to create the index.

A prominent index of this type is the OECD environmental policy stringency index developed by Botta and Koźluk (2014). The index is composed of several market-based and non-marked-based policies. Numerous studies have used this index since its introduction in 2014. For instance, Georgatzi et al. (2020) study (among other variables) the impact of increased environmental policy stringency on CO2 emissions in 12 EU countries from 1994-2014 using panel cointegration techniques. Yirong (2022) used the index within a nonlinear autoregressive-distributed-lag (ARDL) model to analyze policy effects on CO2 emissions in high-polluting economies from 1990-2019. K. Wang et al. (2020) analysed the effect of stricter environmental policies on air quality for a panel of OECD countries for the period 1990-2015 using system generalized-method-of-moments (GMM) to account for endogeneity. Corrocher and Mancusi (2021) studied OECD and BRICS countries for the period 1995-2014 and found that higher discrepancies in the stringency of the index hinders international collaboration on energy-related technologies.

A different measure of environmental policy stringency can be found in Probst and Sauter (2015), who use a count-based indicator to study the effect of policy stringency on CO2 emissions for 46 countries over the period 1990-2010. To account for endogeneity issues, they apply a vector autoregressive model, which is similar in spirit to our econometric analysis. Neves et al. (2020) proxy policy stringency by counting market-based instruments in EU countries. To control for endogeneity, they use an ARDL type model and find that environmental regulations reduced CO2 emissions EU countries during 1995-2017 in the long run.

Hille and Möbius (2019) use shadow prices to compute environmental policy stringency and estimate their effect on air emission in OECD countries over 1996-2009. They use a system GMM approach to account for endogeneity and find that carbon-related policies significantly reduced air emissions. Hashmi and Alam (2019) proxy policy stringency by environmental tax revenue. They found that larger environmental taxes reduced CO2 emissions in OECD countries during 1999-2014.

### 2.2.2 Transport-Related Policies in Austria

The policies under consideration for our index as well as their categorizations have been established in accordance with experts from the Austrian Environmental Agency. We identified two broad categories: 1) Taxes mainly affecting the investment decision to buy a new car and 2) measures affecting the usage of cars. Table 2.1 outlines this structure and lists the individual indicators (policies) for each category, which are explained in more detail below. We focus our analysis on policies that directly impact combustion engine vehicles. We exclude subsidies on electric vehicles, because our sample-period ends in 2019 and until very recently electric vehicles in Austria were almost nonexistent relative to combustion-engine cars, making an analyses of the effect of policies directly promoting the switch to electric vehicles in our framework infeasible.

We focus our analysis on national policies and did not include EU regulations explicitly. However, our econometric model controls for EU directives, such as the fleet regulation and the biofuels directive, indirectly. The EU fleet regulation sets emissions limits on newly registered cars, while the biofuels directive set minimum shares for the use of biofuels and other renewable fuel in the transport sector. Both policies directly impact the efficiency of vehicles and as a result CO2 emissions. We capture this mechanism in our model described in section 2.3 by directly including an indicator for energy-efficiency in our analysis. This indicator freely interacts with all other policies directly captured in the index as well as with remaining endogenous variables in the model.

**Table 2.1:** Categorized Policy Instruments

| Invest | Usage |
|---|---|
| Standard Fuel Consumption Tax | Excise Duty on Mineral Oils (Fuel Tax) |
| Engine-Related Insurance Tax | Temporary Speed Limits |
| | Car-Free Days |
| | IG-L |

Note: The composite index can be disaggregated into the two main categories Invest and Usage. These can further be disaggregated into their sub-components.

The Standard Fuel Consumption Tax is a tax on new cars. It is commonly referred to as NoVA - Normverbrauchsabgabe and was introduced in 1992 (Normverbrauchsabgabegesetz, 1991). It is a direct successor to the Luxury Tax introduced in 1978, which put a tax on the purchase of luxury goods, including cars (2. Abgabenänderungsgesetz, 1977). The NoVA was calculated based on fuel consumption from 1992 to 2013, and based on CO2 emissions from 2014 onwards. The Engine-Related Insurance Tax is a yearly tax covering all registered vehicles. It was calculated based on engine size from 1952 to 1992 (Kraftfahrzeugsteuergesetz, 1952), and from 1993 onwards based on engine power (Kraftfahrzeugsteuergesetz, 1992). Both the NoVA and the Insurance Tax can be quite high for large cars such a SUVs and drive up the costs for purchase and maintenance significantly.

The Excise Duty on Mineral Oils is a fuel tax that differentiates between petrol and diesel fuels. It was already in force in 1950 and is thus the oldest policy in the index

(Mineralölsteuergesetz, 1949). In its current version, the law on the mineral oil tax was implemented in 1995 (Mineralölsteuergesetz, 1994). The Air Pollution Control Act (IG-L) allows provincial governors to enact speed limits in areas with strong air pollution since 1997 (Immissionsschutzgesetz – Luft, 1997). Austria enacted a temporary speed limit of 100 km/h from November 1973 to March 1974 as a fuel-saving initiative in response to the oil crisis(Geschwindigkeitsbeschränkungs-Verordnung, 1973). In 1974, Austria additionally implemented car-free days (Änderung des Bundesgesetzes über Verkehrsbeschränkungen zur Sicherung der Treibstoffversorgung, 1974).

### 2.2.3 Computing the Stringency Index

Our review of the literature on stringency indexes in section 2.2.1 highlights the diverse methods for synthesizing policies into an index, each with its own set of trade-offs. To establish a robust methodological foundation, we align our approach with the widely adopted OECD environmental policy stringency index Botta and Koźluk (2014). This choice allows us to adopt a proven framework and mitigate potential critiques of index construction. In our composite index, both policy categories ($Invest$ and $Usage$) will contribute equally to the composite index. The individual policies within each category are also weighted equally. This ensures that the effect of a given measure is not a priori influenced by different predetermined weights.[1] Following the OECD Stringency Index, we adopt a 7-step scale for the index. It spans from from 0 (indicating absence) to 6 (indicating highest stringency).

To assign a stringency score to each policy in each year, we first calculate the associated impact (cost) of a policy. The cost of the Fuel Tax is given in Eurocents and thus straightforwardly available. The remaining policies in the $Usage$-category (Temporary Speed Limits, Car-Free Days, IG-L) are of a qualitative nature. The stringency of such policies is constant over time. Upon implementation, these instruments are captured by a dummy indicator that is set equal to one if the policy was in force throughout the year, otherwise it is weighted according to the time it was in force in a given year.

The cost of policies in the $Invest$-category (Standard Fuel Consumption Tax and the Engine-Related Insurance Tax) depend on the characteristics of vehicles (see section 2.2.2 for details). To calculate coherent effective policy costs for this category, we resort to constant attributes of cars.[2] Data on characteristics of vehicles has been gathered from the National

[1] Obviously, this assumes that each policy is of equal significance. Below we also consider a decomposition of the indices to analyse the effect of individual policy types to address the concern.

[2] Alternatively, we could use average attributes of a car in a given year, but this approach would lead to changes in the index even if policy measures did not change. Characteristics of cars evolve are not constant but evolve over time, which could potentially result in a reduction in the level of stringency, even in the absence of any changes to the relevant policy measures.

Inventory Reports from the Austrian Environmental Agency as well as from "Verkehr in Zahlen" from the German BMDV (2019).[3]

To compute the effective cost of the Standard Fuel Consumption Tax, we construct the average newly registered vehicle over the period 1970-2019, as this tax only applies to new cars.[4] We calculate the associated cost of the Engine-Related Insurance Tax for the average car over the period 1970-2019 considering the entire fleet (contrasted to considering only newly registered vehicles).

To compute the effective cost of the Standard Fuel Consumption Tax, we construct the average newly registered vehicle over the period 1970-2019, as this tax only applies to new cars.[5] Let $x_t$ be the realization of a specific characteristic of a newly registered car (e.g. power, fuel consumption, emissions, price) in year $t$. Let $P_t^{new}$ and $D_t^{new}$ stand for the number of newly registered diesel and petrol vehicles in a given year, respectively. The characteristics of the average newly registered car are given by[6]:

$$\overline{x}^{new} = \frac{\sum_t \left( D_t^{new} \cdot \overline{x}_{d,t}^{new} + P_t^{new} \cdot \overline{x}_{p,t}^{new} \right)}{\sum_t \left( D_t^{new} + P_t^{new} \right)},$$

where $\overline{x}_{d,t}^{new}$ and $\overline{x}_{p,t}^{new}$ are the average characteristics of newly registered diesel and petrol cars in a given year $t$, respectively, with:

$$\overline{x}_{d,t}^{new} = \frac{1}{D_t^{new}} \sum_d x_{d,t},$$

and

$$\overline{x}_{p,t}^{new} = \frac{1}{P_t^{new}} \sum_p x_{p,t},$$

where $d = 1, \ldots, D_t^{new}$ and $p = 1, \ldots, P_t^{new}$, i.e., we sum over all cars of type diesel or petrol, respectively.

We calculate the associated cost of the Engine-Related Insurance Tax for the average car over the period 1970-2019 considering the entire fleet (contrasted to considering only newly

[3]We do not take a stance on whether characteristics from cars driven in Germany proxy attributes of cars driven Austria well. For the construction of the index it is important to compute policy costs based on constant average vehicle characteristics.

[4]This period has been chosen according to data availability on vehicle characteristics.

[5]This period has been chosen according to data availability on vehicle characteristics.

[6]The relevant characteristics of the average newly registered car are: 7.1 l/100km, 173 gCO2/100km, 18,650 EUR net price.

registered vehicles). We compute this average car similarly as above[7]:

$$\overline{x}^{fleet} = \frac{\sum_t \left( D_t^{fleet} \cdot \overline{x}_{d,t}^{fleet} + P_t^{fleet} \cdot \overline{x}_{p,t}^{fleet} \right)}{\sum_t \left( D_t^{fleet} + P_t^{fleet} \right)},$$

where $P_t^{fleet}$ and $D_t^{fleet}$ stand for the number of registered diesel and petrol vehicles in a given year, respectively. $\overline{x}_{d,t}^{fleet}$ and $\overline{x}_{p,t}^{fleet}$ give the average registered diesel and petrol car in a given year $t$, respectively, with:

$$\overline{x}_{d,t}^{fleet} = \frac{1}{D_t^{fleet}} \sum_d x_{d,t},$$

and

$$\overline{x}_{p,t}^{fleet} = \frac{1}{P_t^{fleet}} \sum_p x_{p,t},$$

where $d = 1, \ldots, D_t^{fleet}$ and $p = 1, \ldots, P_t^{fleet}$.

Once the associated costs of all policies have been calculated, we can assign stringency scores to the policies. We employ a data-driven approach and first compute the inter-percentile range between the 90th and 10th percentile of the distribution of a policy cost over the years and then segment it into five equally sized bins with width:

$$w = \frac{(p_{90} - p_{10})}{5}.$$

The cost of a non-qualitative policy in a given year is then matched with these bins and assigned the corresponding score. Table B.1 in Appendix D.2 shows the score assignment and the associated thresholds for policies for which a direct cost can be calculated (Fuel Tax, SFC Tax, Insurance Tax). Table B.2 in Appendix D.2 provides a full list of costs assigned to a policy alongside the associated scores for each year.

When each policy has been assigned a score in each year, we simply aggregate the score for every policy into the two main categories ($Invest$ and $Usage$) for each year, where every policy receives equal weight. The composite index can reach a maximum value of 6, implying that each of the two main categories can contribute a maximum of 3 to the composite index. We compute:

$$Invest_t = \frac{1}{2} \left( \frac{1}{2} SFC\ Tax_t + \frac{1}{2} Ins\ Tax_t \right),$$

where the weights inside the brackets attach equal weights to the two policies $SFC\ Tax$

[7] The relevant characteristics of the average car over the entire fleet are: 1660 ccm, 69 kW

and $Ins\ Tax$. The weight outside the brackets attaches equal weight to the $Invest$ and the $Usage$ categories. $SFC\ Tax_t$ and $Ins\ Tax_t$ can thus contribute a maximum of 1.5 to the composite index each and $Invest_t$ can contribute a maximum of 3. Similarly, we can compute:

$$Usage_t = \frac{1}{2}\left(\frac{1}{4}\mathit{Fuel\ Tax}_t + \frac{1}{4}\mathit{Speed\ Limit}_t + \frac{1}{4}\mathit{Car\text{-}Free\ Days}_t + \frac{1}{4}\mathit{IG\text{-}L}_t\right)),$$

where the weights inside the brackets again attach equal weights to the four policies in the $Usage$ category and the weight outside the brackets signals equal weights for the two main policy categories ($Invest$ and $Usage$). Each of the four policies can thus contribute a maximum of 0.75 to the composite index and $Usage_t$ can overall contribute a maximum of 3. We can then further aggregate the two categories into the overall composite index with equal weights for each of the two policy categories:

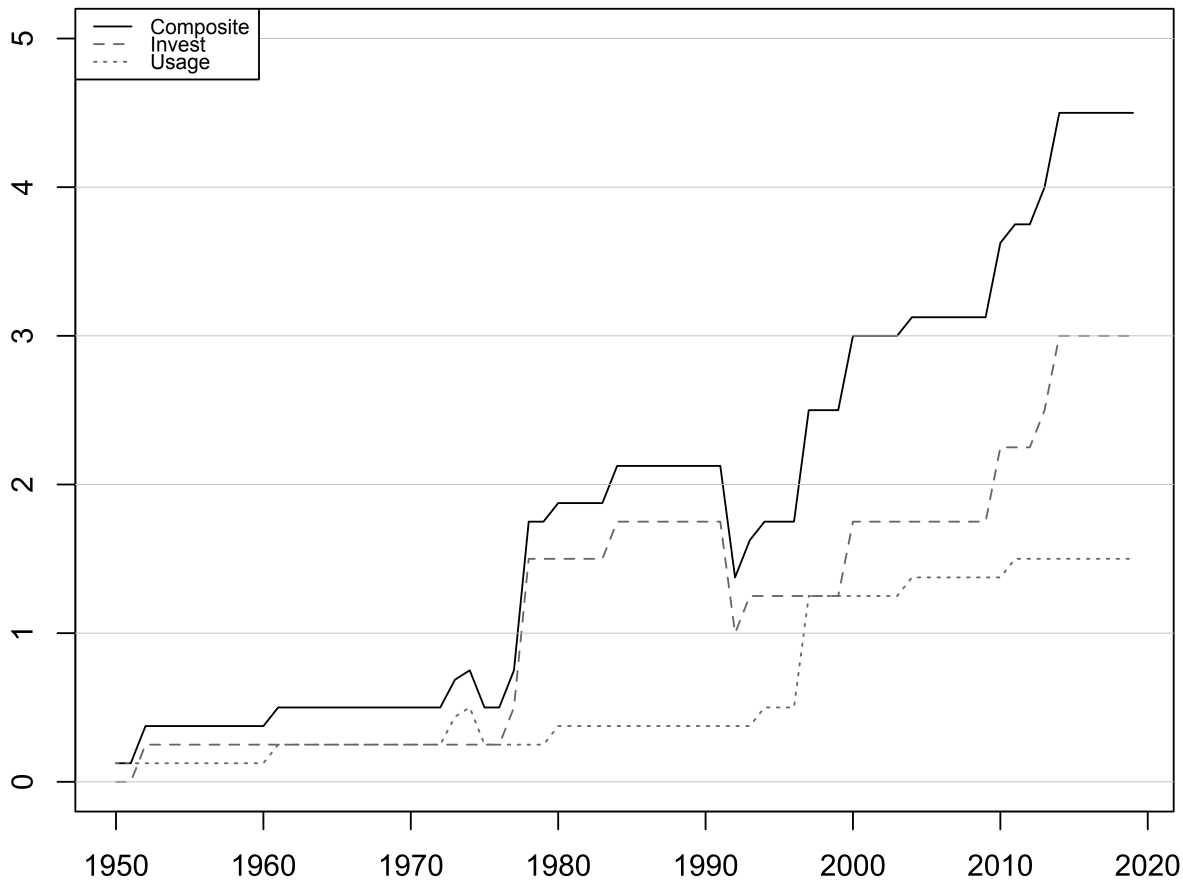
$$Comp_t = Invest_t + Usage_t,$$

where $Comp_t$ can take a maximum value of 6.

The final index is shown in Figure 2.1. The composite index is indicated by the solid line. The sub-index based on policies affecting investment decisions is given by the dashed line and measures affecting the usage of cars by the dotted line. The largest observed level of the composite index is 4.5, which is below the theoretical maximum of 6. This discrepancy arises because the stringency indicated by the index in a given year is calculated in relation to the observed most stringent value of all policy measures over the entire sample period. While some measures are at their most stringent level in 2019, this is not true for all instruments. The sub-index for the policy category $Invest$ reaches its maximum value of 3 in 2014. While the sub-index for $Usage$ only reaches a maximum of 1.5 with a theoretical maximum of 3. This is because the temporary speed limits and car-free days were temporary measures in the 1970s (each of them can take a theoretical maximum value of 0.75).

Usage-related policies spike in 1973 to 1974, indicating car-free days and temporary speed limits. The policies do not achieve their theoretical maximum value of 3 because they were not in effect over an entire given year. A particularly steep increase can be noticed in 1997, which can be attributed to the implementation of the the Air Pollution Control Act. Other small increases are due to increases in the fuel tax. Measures affecting the investment category sharply increased in 1978, when the luxury tax on new cars (which later transformed into the NoVA) was introduced. The tax was restructured and based on fuel usage from 1992

**Figure 2.1:** Passenger Transport Policy Stringency Index for Austria, 1950-2019




on. The effective tax rate on the average car dropped markedly.[8] In 2000, the insurance tax increased sharply. Increases in 2010 and 2014 can be attributed to an increased stringency in the NoVA.

## 2.3 Econometric Analysis

In this section we go on to study the effect of the instruments embodied in the transport related environmental policy stringency index on CO2 emissions from passenger transport. As policies aimed at influencing emissions from the transport sector are characterized by interdependencies, direct or indirect ones, endogeneity issues have to be considered. To address this, we employ a vector autoregressive (VAR) model that, by construction, treats all variables as endogenous. Aside from the index variables and CO2 emissions, the VAR

[8]The tax rate would have dropped equally for the average car in 1992.

model should include other determinants of transport emissions.

### 2.3.1 Data

Data on CO2 emissions, energy intensity, and the fleet composition has been provided by the Austrian Environmental Agency. The data have been extracted from their Network and Emissions Model (NEMO), developed by Dippold et al. (2012). CO2 emissions are measured in 1000t, data on the vehicle fleet contain the total fleet of petrol and diesel powered cars (including hybrid and plug-in hybrid vehicles) in a given year, and data on the energy intensity of vehicles are given by gCO2/100km.

NEMO simulates fleet composition based on annual registration data and survival probabilities, using predefined vehicle segments (e.g., gasoline, diesel, electric) to model emissions from road-specific driving patterns. NEMO is known to provide good model accuracy and, moreover, can abstract from fuel tourism. This allows us to focus on emissions actually generated within Austria - a distinct advantage in contrast to fuel-consumption based approaches. Finally, we extract population statistics from Statistik Austria (2021).

For the econometric analyses, we use CO2, vehicle fleet, and energy intensity in per capita terms. Data on GDP is measured in real GDP and has been taken from the Austrian Economic Chamber (WKO, 2021). Oil prices are composed of WTI prices up to 1986 and Brent (Europe) from 1987 onwards. Both time series were extracted from the FRED Economic Data base (U.S. Energy Information Administration, 2022a, 2022b). To calculate GDP/CAP and oil prices in real terms, we used the Austrian consumer price index, which we extracted from OENB (2022).

Clean data for all mentioned variables are available for the period 1965-2019. Table B.1 summarizes and describes the variables. Table B.2 presents the summary statistics of these variables, time series plots are shown in Figure B.1. CO2/CAP, Fleet/CAP, and GDP/CAP all show a clear upward trend. The financial crisis around 2009 is clearly discernable in the time series of GDP and CO2 per capita. The energy intensity has a decreasing trend, i.e., cars got more efficient, although the efficiency did not improve much prior to the 1980s. By inspecting the time series on international oil prices, one can clearly see a stark increase in prices during the first and second oil crises, starting in 1973 and 1979, respectively.

### 2.3.2 General Model Framework

Our baseline model is given by:

$$\begin{aligned} CO2/CAP_t =& \alpha + \beta_1 Comp_t + \beta_2 EI_t + \beta_3 Fleet/CAP_t \\ &+ \beta_4 GDP/CAP_t + \beta_5 Oil_t + u_t, \end{aligned} \tag{2.1}$$

where $Comp_t$ stands for the composite index containing the policies outlined in Table 2.1. This model forms the basis for our specification. We explicitly treat all variables except for GDP per capita and oil prices as interdependently determined in our econometric framework, which will be outlined in section 2.3.3.

We can disaggregate $Comp_t$ to get model specifications that allow for a more fine-grained policy analysis. In a first step, we can disaggregate the composite index into the two main sub-indexes: $Invest$ and $Usage$. We modify the structural baseline model motivated by equation (2.1) to obtain:

$$\begin{aligned} CO2/CAP_t =& \alpha + \beta_1 Invest_t + \beta_2 Usage_t \\ &+ \beta_3 EI_t + \beta_4 Fleet/CAP_t + \beta_5 GDP/CAP_t + \beta_6 Oil_t + u_t. \end{aligned} \tag{2.2}$$

In a next step, we can disaggregate the two main policy categories. We modify the baseline model to obtain the two following models:

$$\begin{aligned} CO2/CAP_t =& \alpha + \beta_1 Insurance\ Tax_t + \beta_2 SFC\ Tax_t + \beta_3 Usage_t \\ &+ \beta_4 EI_t + \beta_5 Fleet/CAP_t + \beta_6 GDP/CAP_t + \beta_7 Oil_t + u_t, \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} CO2/CAP_t =& \alpha + \beta_1 Fuel\ Tax_t + \beta_2 Use\ Qual_t + \beta_3 Invest_t \\ &+ \beta_4 EI_t + \beta_5 Fleet/CAP_t + \beta_6 GDP/CAP_t + \beta_7 Oil_t + u_t. \end{aligned} \tag{2.4}$$

In equation (2.3), we disaggregate the policy category $Invest$ into its subcomponents: engine-related insurance tax ($Insurance\,Tax$) and standard fuel consumption tax ($SFC\,Tax$). The policy category $Usage$ contains too many individual policies to fully disaggregate it. In equation (2.4), we thus disaggregate the policy category $Usage$ into the mineral oil tax ($Use\ Tax$) and gather the remaining policies in the new category qualitative usage ($Use\ Qual$).

All of these specifications share the same drawbacks that (i) they are likely to suffer from endogeneity and (ii) they are static models, whereas we are interested in dynamic effects. The next section describes our modeling approach to accommodate these complications.

### 2.3.3 VAR Analysis

Most variables in equation (2.1) are likely to be endogenous. These include the composite index and CO2/CAP, EI, and Fleet/CAP. Real GDP/CAP and real oil prices are likely to be determined outside this system and we treat them as exogenous. Vector autoregressive (VAR) type models with exogenous variables are an appropriate model class for our analysis (Sims, 1980, Lütkepohl, 2005). Due to potential nonstationarity of most variables, we have to test the variables for unit roots and for cointegrating relations in order to establish which model form is most suitable.

Figure B.1 clearly shows that the variables included in our model exhibit some kind of trend. For the econometric analysis, it is important to establish whether the variables are characterized by a stochastic trend (i.e. a unit root) or a deterministic one. Several tests have been proposed to test the presence of a unit root, but many unit root tests suffer from low power when applied to near-unit processes; see, e.g., Kilian and Lütkepohl (2017). Elliott, Rothenberg, and Stock (1996) propose a unit root test that dominates other tests in terms of small sample properties and power. It is based on the Augmented-Dickey-Fuller test (ADF) and tests the null hypothesis of a unit root. We apply the test to the variables in equation (2.1).

The resulting test statistics are shown in Table B.1 in Appendix D.2. The results for the variables in levels and first differences based on models with a constant only as well as a constant plus trend specification are given. The results reveal that the null hypothesis of a unit root cannot be rejected at the 10% significance level in the tests with only a constant as well as a constant and trend for all variables in levels. The series in first differences appear to be stationary, as the null can be rejected at the 5% level for both models (trend and constant as well as constant only). We can thus conclude that the variables in level form are I(1).

Next, we test for a cointegrating relation between the endogenous variables. The results of the Johansen cointegration trace test for CO2/CAP, EI, and Fleet/CAP are shown in Table B.2 in Appendix D.2[9]. The test cannot reject the null hypothesis of a cointegration rank of zero (i.e. no cointegration) at the 10% level. We further confirm this result by analyzing all pairwise cointegrating relations, where we find no evidence for cointegration (results available upon request). We thus conclude that there is no evidence in favor of a

[9]We do not include the policy stringency variables in the test because these are naturally bounded.

cointegrating relation between the variables.

Consequently, we adopt a (structural) VAR model for the first differences of the variables for our analyses. The model considers all variables to be endogenous and each variable is determined by lagged values of all other variables. As mentioned above, we include GDP per capita and international oil prices as exogenous variables in the model as they are likely important drivers of CO2 emissions. Such a VARX model with $p$ lags of the endogenous and $q$ lags of the exogenous variables in its structural form is given by:

$$\boldsymbol{B}_0\boldsymbol{y}_t = \boldsymbol{\mu} + \textstyle\sum_{i=1}^{p} \boldsymbol{B}_i\boldsymbol{y}_{t-i} + \sum_{j=0}^{q} \boldsymbol{\vartheta}_j\boldsymbol{x}_{t-j} + \boldsymbol{u}_t, \tag{2.5}$$

where $t = 1, \ldots, T$, $\boldsymbol{y}_t$ is a $K \times 1$ vector containing the endogenous time series and $\boldsymbol{x}_t$ is an $M \times 1$ vector containing the exogenous time series. $\boldsymbol{\mu}$ is a vector of intercepts, $\boldsymbol{B}_0$ is a $K \times K$ parameter matrix containing the contemporaneous interactions, $\boldsymbol{B}_i$ are $K \times K$ matrices containing the coefficients of the lagged endogenous variables, $\boldsymbol{\vartheta}_j$ are $M \times K$ matrices containing the coefficients of the exogenous variables, and $\boldsymbol{u}_t$ is the $K \times 1$ vector of structural errors, which are assumed to be independent of each other. Note that without prior restrictions the model is not identified.

Applying the model to the variables in equation (2.1) and taking first differences we obtain: $\boldsymbol{y}_t = [\Delta Comp_t, \Delta EI_t, \Delta Fleet/CAP_t, \Delta CO2/CAP_t]'$ and $\boldsymbol{x}_t = [\Delta log(GDP/CAP_t), \Delta log(Oil_t)]'$. Note that the endogenous variables are taken in first differences, whereas the endogenous ones are taken in log-differences. The estimation of the model is based on its reduced form:

$$\boldsymbol{B}_0^{-1}\boldsymbol{B}_0\boldsymbol{y}_t = \boldsymbol{B}_0^{-1}\boldsymbol{\mu} + \textstyle\sum_{i=1}^{p} \boldsymbol{B}_0^{-1}\boldsymbol{B}_i\boldsymbol{y}_{t-i} + \sum_{j=0}^{q} \boldsymbol{B}_0^{-1}\boldsymbol{\vartheta}_j\boldsymbol{x}_{t-j} + \boldsymbol{B}_0^{-1}\boldsymbol{u}_t, \tag{2.6}$$

which can be rewritten more compactly as:

$$\boldsymbol{y}_t = \tilde{\boldsymbol{\mu}} + \textstyle\sum_{i=1}^{p} \boldsymbol{\phi}_i\boldsymbol{y}_{t-i} + \sum_{j=0}^{q} \boldsymbol{\theta}_j\boldsymbol{x}_{t-j} + \boldsymbol{v}_t, \tag{2.7}$$

where $\tilde{\boldsymbol{\mu}} = \boldsymbol{B}_0^{-1}\boldsymbol{\mu}$, $\boldsymbol{\phi}_i = \boldsymbol{B}_0^{-1}\boldsymbol{B}_i$, $\boldsymbol{\theta}_j = \boldsymbol{B}_0^{-1}\boldsymbol{\vartheta}_j$, and $\boldsymbol{v}_t = \boldsymbol{B}_0^{-1}\boldsymbol{u}_t$.

The reduced form can be estimated by simple OLS, and a specific structure can be imposed on $\boldsymbol{B}_0$ to recover the structural parameters and interpret the results.[10] In order to identify the system, we place specific short run restrictions on the coefficient matrix $\boldsymbol{B}_0$ as shown in Table 2.2. The columns contain the shocks to each variable, and the rows indicate which variables are affected by this shock. The identification is justified as follows. It is reasonable to assume that policies influence CO2 emission contemporaneously, but higher

[10] The most popular structure is a triangular one based on the Cholesky decomposition. However, this implies that the outcomes are contingent upon the ordering of the variables and the restrictions are not always economically meaningful.

emissions may translate into stricter policies with a delay. Similarly, this holds also for EI and Fleet/CAP. EI can influence both the fleet and emission contemporaneously, whereas the fleet only has an immediate effect on emissions. We exclude contemporaneous interactions among the policy categories, as these may be difficult to order and justify. Additionally, we postulate that policies affect energy intensity with a delay.

**Table 2.2:** Identification of VARX(1,1) model with non-recursive short-run restrictions.

| | Comp | EI | Fleet/CAP | CO2/CAP |
|---|---|---|---|---|
| Comp | 1 | 0 | 0 | 0 |
| EI | 0 | 1 | 0 | 0 |
| Fleet/CAP | * | 0 | 1 | 0 |
| CO2/CAP | * | * | * | 1 |

Note: The * indicates a possible contemporaneous interaction, whereas a 0 stands for a restriction, i.e. a coefficient of zero.

One drawback of the VAR framework is its high data intensity. Therefore, the length of the lags of the variables have to be chosen carefully. To choose this optimally, we run a series of specification tests. For these tests, we specify a VAR model with one lag for the endogenous as well as exogenous variables (i.e., a VARX(1,1) model), which we choose to maximizes the degrees of freedom. The series of model adequacy tests confirm that a lag selection of one for both the endogenous and exogenous variables is valid. Various test statistics for lag-length selection shown in Table B.3 in Appendix B.3 select a lag length of 1 for the endogenous variables. The statistics include the Akaike information criterion (AIC), the Schwarz criterion (SC), the Hannan-Quinn (HQ) information criterion, and the final prediction error (FPE). While these criteria do not explicitly select a lag-length for the exogenous variables, the remaining adequacy tests show positive results for a VARX(1,1) specification.

The autocorrelation properties of the residuals of the VARX(1,1) model are shown in Table B.4 in Appendix B.3. It shows the results from the test proposed by Edgerton and Shukur (1999). The test is based on a VAR model of the error vector and tests the null hypothesis of no residual autocorrelation, i.e., all coefficients of the $h$ orders of the VAR process are equal to zero. The results show that we are not able to reject the null at any meaningful significance level. We also test for ARCH effects in the residuals with a multivariate LM-type test from Doornik and Hendry (1997). Test results are shown in Table B.5 in Appendix B.3 and show no sign of ARCH effects in the residuals. Given the test results, we model a structural VARX(1,1) with the chosen short run restrictions. The empirical results will be discussed in the following section.

### 2.3.4 Impulse Response Analysis and Dynamic Multipliers

Due to the interdependence of the variables, the coefficients of the VAR are difficult to interpret directly. Therefore, other concepts have been proposed to analyse such a system. One popular type of analysis for such models is the study of impulse response functions (IRFs). The basic idea of an impulse response analysis is to consider the vector moving average representation of the VAR to express model in terms of past shocks, specifically its structural errors $u_t$. This enables us to study how the system responds to structural shocks (impulses) related to the individual endogenous variables. Responses to shocks to exogenous variables, in contrast, can be studied with dynamic multipliers (DMs), which follow a standard ceteris-paribus interpretation. The interpretation is still similar to those of IRFs but without a dynamic feedback mechanism.

In the next subsection, we study the responses of CO2/CAP to shocks to the composite index. Then we go on to break down the index into its two main categories: $Invest$ and $Usage$. We then further disaggregate these and study specific policies contained in these sub-indices in more detail.

#### A Shock to the Composite Index

We start by analyzing the effect of a shock to the composite index on CO2 emissions from passenger cars. The VARX(1,1) in this setting is motivated by the model defined in equation (2.1). Figure 2.2 shows the cumulated impulse response of $CO2/CAP$ from passenger cars to a structural shock to the composite index ($Comp$), energy intensity ($EI$), and fleet/cap ($Fleet/CAP$) over time (years). The solid curves show the IRFs over time, the dashed curves provide a bootstrapped 90% confidence interval (CI), and the solid lines are plotted at zero to distinguish significant responses. The effect of a shock to a specific variable on $CO2/CAP$ is considered statistically significant at the 90% CI whenever both confidence bands are either below or above the zero line. The labels on the y-axis indicate the minimum and maximum values of the lower and upper CIs, respectively. Additionally, the estimated long-run responses are given.

The shock to the composite index is of size 6. This equals the maximum stringency the index can take if every quantitative policy is at its most stringent level and every qualitative policy is in effect. In other words, it is a 100% increase in the maximum stringency the index can theoretically take. Recall that the stringency index shown in Figure 2.1 reaches a maximum stringency of 4.5. The shock is thus of substantial size. As the variables in the structural VARX model are taken in first differences, the associated (cumulated) impulse responses to those shocks tend towards a long-run equilibrium. The long-run response of

$CO2/CAP$ to a shock to $Comp$ of size 6 settles at $-0.49$ kt. This is a significant effect, given that the maximum amount of emissions in our data is at 1.365 kt (see Table B.2). Converted into percentages this means that a 100% increase in the theoretical maximum stringency of the index reduces passenger transport CO2 emissions per capita by around 36% relative to its highest value.

The impulse responses of $CO2/CAP$ to shocks to $EI$ and $Fleet/CAP$ are also shown in Figure 2.2. We consider negative shocks to these variables, meaning an improvement in energy intensity and a decrease in the degree of motorization. For the shock sizes we decided to use economically/technically meaningful values, in contrast to the usual choice of one standard deviation shocks. Qualitatively, both shocks decrease emissions as to be expected. The shock size to $EI$ is set to $-25$ gCO2/100km. This equals around 15% of the minimum (most efficient) value of $EI$ over our sample period. The long-run effect on passenger transport CO2 emissions per capita to this shock settles at a about $-0.14$ kt. In the long-run, a 15% improvement in $EI$ relative to its minimum reduces CO2 emission per capita by around 10% relative to its maximum. The effect stays statistically significant for around seven years.

The shock to $Fleet/CAP$ is set to $-50$ vehicles per 1000 person. This relates to roughly a 9% reduction relative to its highest value. The reaction of $CO2/CAP$ is quite stark: it decreases by about $-0.90$ kt (around 66% relative to its highest value). The effect is statistically highly significant for the entire period. We attribute this strong response to the dynamic feedback mechanism in the VAR system. The shock to the fleet can lead to a reinforcing dynamic that further reduces the fleet in the following periods. If this effect is strong enough, this can justify the impulse response.

Figure 2.3 shows the cumulated dynamic multipliers (DMs) of $CO2/CAP$ to shocks to the exogenous variables, $GDP/CAP$ and $Oil$. The shocks are of unit size and constant, and we see that the effects are highly significant. The exogenous variables are taken in log-scales. The interpretation thus follows a level-log model: an increase in the exogenous variable by 100% leads to a unit change in $CO2/CAP$ as given by the solid line in Figure 2.3. Therefore, an increase in $GDP/CAP$ by 100% then leads to an increase in $CO2/CAP$ by 0.7153 kt. The effect is thus quite strong. Higher international oil prices are associated with a decrease in (per capita) passenger transport CO2 emissions, but the effect is markedly weaker compared to $GDP/CAP$. Increasing $Oil$ by 100% in the long-run reduces $CO2/CAP$ by around 0.085 kt, which amounts roughly to a 6% decrease in emissions relative to their highest value.

**Figure 2.2:** Cumulated impulse responses for $CO2/CAP$ (1965-2019). Responses to a shock of 6 to $Comp$, $-25$ to $EI$, and $-50$ to $Fleet/CAP$. Hall's percentile intervals are at 10% significance level with 1000 bootstrap replications.

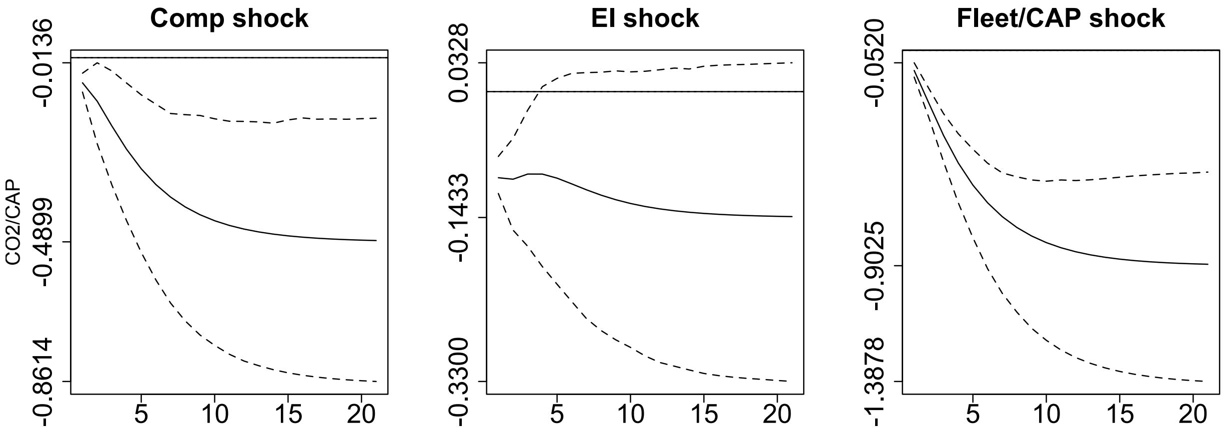
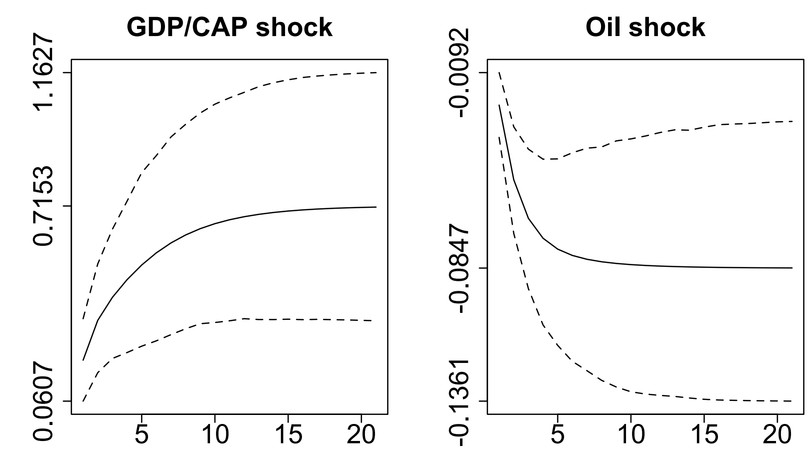



**Figure 2.3:** Dynamic multipliers for $CO2/CAP$ (1965-2019). Response to a 100%-shock to $GDP/CAP$ and $Oil$. Hall's percentile intervals are at 10% significance level with 1000 bootstrap replications.



### A Shock to the Sub-Indexes

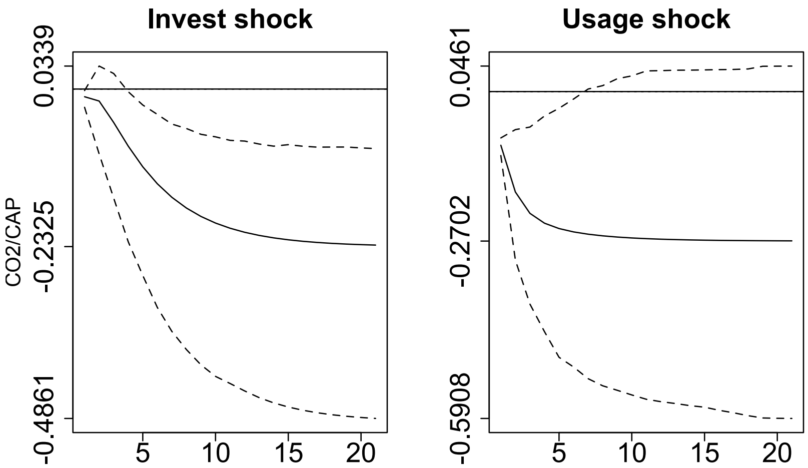
We now go on to provide a more fine-grained analysis of the disaggregated policy categories. We start by decomposing the composite index into its two main sub-indexes: $Invest$ and $Usage$. To estimate the impact of shocks to these policy categories, we model a VARX(1,1) based on equation (2.2). Figure 2.4 contains the cumulated impulse responses to structural shocks to the two main policy categories. Shocks to the policy variables are chosen such that they represent a 100% increase in the maximum stringency (i.e. 3). Qualitatively, a shock to each of the policy categories shows a negative effect on per capita passenger transport CO2 emissions.

The response of $CO2/CAP$ to a shock to $Invest$ settles at $-0.2325$ kt. Thus, a 100% increase in the maximum stringency of the category $Invest$ in the long-run reduces passenger transport CO2 emissions per capita by around 17% relative to its highest value. The effect of a shock to $Invest$ starts close to the zero-line and gets statistically significant only after around 5 years. This seems intuitive given that policies in this category affect the purchase behavior of new vehicles. Purchases may either not be undertaken or altered towards more efficient vehicles. Either way, it will take time for this effect to materialize in a significant reduction in emissions over the entire fleet. A shock to $Usage$ reduces emissions by about 0.2702 kt in the long-run (around 20% relative to their highest value). The effect thus seems to be a bit stronger than that of a shock to the invest category. The effect of a shock to $Usage$ significantly reduces emissions from period 0 on and remains statistically significant for around 7 years.

Next, we disaggregate the $Invest$-category into its sub-components: $Insurance\ Tax$ and $SFC\ Tax$. To analyze the effect of a shock to these policies on per capita CO2 emissions from passenger cars, we study a VARX(1,1) based on to equation (2.3). Figure 2.5 shows the corresponding impulse responses. It depicts the response of $CO2/CAP$ to shocks to the insurance tax and standard fuel consumption tax, which together make up the $Invest$ category. Shocks are again chosen to double the maximum stringency of each policy (i.e. 1.5).

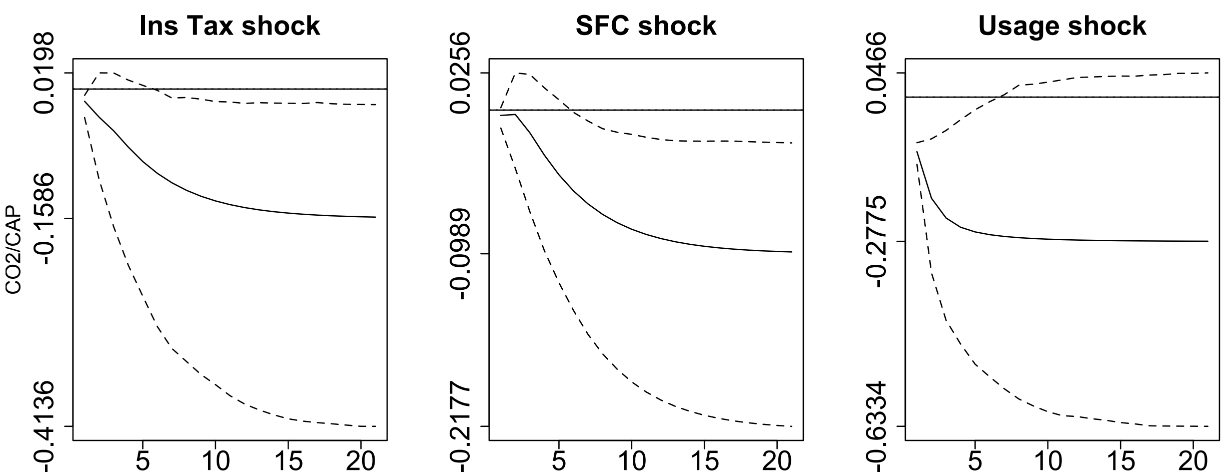
Qualitatively, we can see that both policies exhibit a negative effect on CO2 emissions per capita. The shocks become significant after a few years, which seems consistent with the overall effect of $Invest$ shown in Figure 2.5. Quantitatively, a shock to $Insurance\ Tax$ reduces emissions by approximately 0.1586kt, which amounts to a 12% reduction relative to the highest emission value. A shock to $SFC\ Tax$ reduces emissions by about $-0.0989$ kt per 1000 persons in the long-run (around a 7% reduction relative to the highest emission value). Both policies significantly reduce emissions in our sample period. While the effect of $Insurance\ Tax$ is stronger, the effect of $SFC\ Tax$ overall shows stronger statistical

**Figure 2.4:** Impulse responses for $CO2/CAP$ (1965-2019). Responses to a shock of 3 to $Invest$ and $Usage$. Hall's percentile intervals are at 10% significance level with 1000 bootstrap replications.



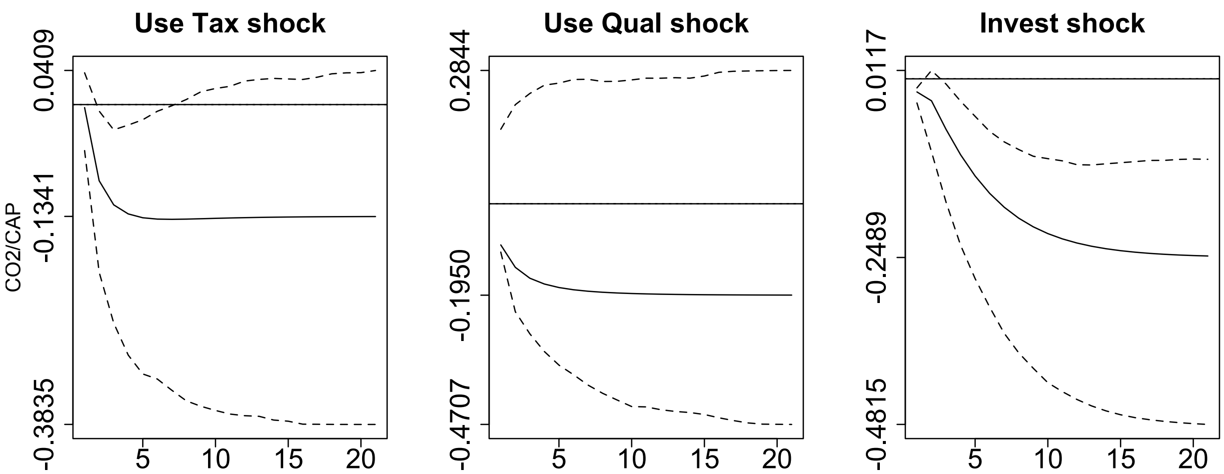

significance.

**Figure 2.5:** Impulse responses for $CO2/CAP$ (1965-2019). Responses to shocks of size 1.5 to $Ins\ Tax$ and $SFC$ (doubling the maximum stringency). Hall's percentile intervals are at 10% significance level with 1000 bootstrap replications.




In a final step, we disaggregate the sub-index $Usage$ into its components. Impulse responses corresponding to equation (2.4) are shown in Figure 2.6. It shows the response of $CO2/CAP$ to shocks to the mineral oil tax ($Use\ Tax$) and the remaining qualitative usage-related policies ($Use\ Qual$). Together these two categories constitute the $Usage$ sub-index. Shocks are again chosen to increase the maximum stringency of each policy by 100%. Each policy within this category can contribute a maximum of 0.75 to the index. This means that $Use\ Tax$ is shocked with 0.75, while the remaining three policies of the usage-related category are gathered in $Use\ Qual$ and are shocked with 0.75 each. $Use\ Qual$ is thus shocked by $0.75 \cdot 3 = 2.25$.

**Figure 2.6:** Impulse responses for $CO2/CAP$ (1965-2019). Responses to shock of size 0.75 to $Use\ Tax$ and 2.25 to $Use\ Qual$ (doubling the maximum stringency of each). Hall's percentile intervals are at 10% significance level with 1000 bootstrap replications.




Qualitatively, we can see that both policies are less significant compared to the investment-related polices. $Use\ Tax$ is only briefly significant, whereas qualitative usage-related measures appear to be highly insignificant. Quantitatively, a shock to $Use\ Tax$ reduces emissions by approximately 0.1341kt, which amounts to a 10% reduction in emissions relative to their highest value. A shock to $Use\ Qual$ reduces emissions by about $-0.1950$ kt per 1000 persons in the long-run (around a 14% reduction relative to their highest value).

### Robustness

Equations (2.3) and (2.4) are disaggregated forms of equation (2.2). Equation (2.3) disaggregates $Invest$ into its subcomponents. The remaining variables in equations (2.2) and (2.3) remain the same. Similarly, equation (2.4) disaggregates $Usage$ in (2.2) and the remaining variables are again the same in the two models. We can utilize this structure for robustness checks. Ideally, a shock to the same variables in the three models should show very similar responses in $CO2/CAP$ across all three models. For example, we can see from Figure

2.5 (corresponding to equation (2.3)) that the effect of a shock to $Usage$ is very similar to the effect shown in Figure 2.4 (corresponding to equation (2.2)). By comparing Figure 2.6 (equation (2.4)) to Figure 2.4, we see very similar effects of a shock to $Invest$ on $CO2/CAP$. These results are reassuring and add credibility to the robustness of the different specifications motivated by equations (2.2)-(2.4). Additionally, remaining impulse responses of $EI$ and $Fleet/CAP$) as well as the remaining dynamic multipliers of $GDP/CAP$ and $Oil$ show very similar results. They are qualitatively identical to those reported above and quantitatively they differ only marginally, adding further credibility to the robustness of our model specifications.

The policy stringency index is calculated from nominal values whenever policy stringency is determined by price changes. The reasoning is that if we calculated the index based on real terms, it would simply change in stringency whenever prices change - even though the policy itself did not change. In this, we follow the methodology of the OECD environmental policy stringency index. However, to address this issue, we re-calculate the impulse-responses based on equations (2.2)-(2.4) extended by the consumer price index as an additional exogenous variable. This approach ensures that the stringency index remains unaffected by price changes. At the same time, the effect of changing price levels is controlled for. The IRFs remain consistent across all specifications.

## 2.4 Policy Discussion

In this section, we provide a policy discussion that focuses on three key areas. First, we discuss the impact of changes in policy stringency on CO2 emissions based on the results of the econometric analysis from section 2.3.4. Second, we put these results in a wider context to explore policy options for Austria to achieve its national target to achieve net-zero GHG emissions by 2040 considering legal and economic aspects. Third, we assess the external validity of our insights for other EU countries with similar or differing policy and economic environments.

### 2.4.1 Emissions reacting to Changes in Policy Stringency

Our first interest is in the comparative effectiveness of different observed policies on reducing CO2 emissions. Table 2.3 concisely summarizes the effect of changes in the policy stringency index on passenger transport CO2 emissions in Austria based on the impulse response functions (IRFs) from section 2.3.4. The first column shows different aggregation levels of the index. Composite is the overall index. Invest and Usage are the two main sub-indexes, which can be further disaggregated into their respective components. The level of aggregation is indicated by a slight indentation of the index components. Columns 2-4 are associated with increases in policy stringency. Column 2 shows by how many index-points an index component is increased. All stringency increases are chosen to represent a 100% increase in policy stringency. Whenever possible, this increase is stated in monetary or percentage terms for single policies. Columns 5-7 show the effect of the associated stringency increase. Column 5 shows the reduction in kt CO2 per capita from passenger cars. Column 6 shows the reduction in these emissions relative to 2019 emissions in percentage terms. The statistical significance of the emission impacts due to stringency increases is shown in Column 7.

**Table 2.3:** Changes in policy stringency and effects on passenger transport CO2 emissions

| | **Policy Stringency Increase** | | | **Effect on CO2 Emissions** | | |
|---|---|---|---|---|---|---|
| | (2) | (3) | (4) | (5) | (6) | (7) |
| | Index-Points | Change | Single policy | kt CO2 | Relative to 2019 | Statistically significant? |
| Composite | 6 | 100% | | −0.49 | −36% | yes |
| Invest | 3 | 100% | | −0.23 | −17% | after 5 years |
| Ins. Tax | 1.5 | 100% | 272 EUR | −0.16 | −12% | after 5 years |
| SFC | 1.5 | 100% | 14 p.p. | −0.10 | −7% | after 5 years |
| Usage | 3 | 100% | | −0.27 | −20% | up to 7 years |
| Fuel Tax | 0.75 | 100% | 0.45 EUR | −0.13 | −10% | btw. 1-6 years |
| Use Qual | 2.25 | 100% | | −0.19 | −14% | no |

Note: Results are based on the impulse response function (IRF) results reported in section 2.3. Statistical significance is based on a 90% confidence interval. The policy stringency increase is reported in monetary terms or percentage points when possible, with associated values based on 2019 levels of stringency. Statistical significance remains throughout the rest of the time horizon studied, i.e., up to 20 years, if not stated otherwise.

Consider, for example, the first row of Table 2.3. The composite index is increased by 6 index points. This is equivalent to a 100%-increase in its theoretical maximum value, which can only be reached when all policies are in effect and at their most stringent level ever chosen within the time period of analysis. The composite index actually takes on a maximum value of 4.5, which is lower than its theoretical maximum value of 6. Increasing the stringency of the composite index by 6 (100%) thus represents a stronger change in policy stringency than actually observed between 1965 and 2019. This change would reduce CO2/CAP from passenger cars by ∼0.49 kt, which is a reduction of ∼36% relative to 2019 emission levels. The effect is highly statistically significant throughout its contemporaneous effect (time period 0) up to 20 years after the change in stringency.

Overall, we find that policies affecting the investment decision to buy new cars have been implemented at stringency levels that were more effective in Austria than the ones of instruments affecting the usage of vehicles. The engine-related insurance tax showed the strongest quantitative impact. It increased the price of emission-intensive vehicles and is charged on a yearly basis, which can incentivise individuals to shift to more efficient vehicles. In this regard, it may also be viewed as a push measure to promote the electrification of the vehicle fleet. In this framework, the standard fuel consumption tax can have similar effects. We find this effect to be smaller compared to the insurance tax, possibly because the standard fuel consumption tax is only targeted at new registrations, while the insurance tax is levied on the entire fleet. We further find that the second main policy category - limiting the usage of combustion-engine vehicles - was an effective policy category to reduce emissions in the short run. This category includes the fuel tax and qualitative measures (speed limits and car-free days). However, we find these latter measures to be statistically insignificant.

These results are in line with the notion that the transport sector is characterized by persistence. Especially policies that address the acquisition of cars and thus the development of the fleet composition will result in emissions reacting with a time lag. We can see this effect in the responses of CO2 emission to changes in the stringency of the standard fuel consumption or engine-related insurance tax. Conversely, policies targeting the usage of vehicles have been shown to have an imminent effect, but this effect seems to get watered down in the long-run. This may be attributable to people getting accustomed to the increase in, for example, fuel prices and revert to old driving habits with a time lag. Another possible explanation is that monetary stringency increases are usually not indexed to inflation, which reduces the real price of the tax increase over time. Considering their divergent time impact profile the two policy categories, focusing on influencing investment versus usage decisions, may thus well complement each other.

### 2.4.2 Implications for Austrian Policies

Austria set the ambitious goal to become carbon-neutral by 2040. For the passenger transport sector this could be ensured in the long term by a sufficiently early enforced requirement in new car registrations of zero-emission vehicles only, and by a complementary set of measures to guide the development in the short and medium term. As the transformation to a carbon-neutral transport system is not the only objective to be tackled, the policy package will need to be more comprehensive. It needs to simultaneously address issues such as local air pollution, noise, health and safety, urban sprawl, and affordability (Dugan et al., 2022; Jochem et al., 2016; Santos et al., 2010; Steg & Gifford, 2005). The switch in engine technology from combustion to electric alone,i.e., the requirement of zero-emissions engine technology, will not suffice to address all of them.

To achieve such GHG emission reduction targets in passenger car transport over time, as discussed above, countries have instruments at their discretion that either work via incentives for or regulation of the use of cars ("usage" indicators) or via incentives or regulation of what type of car (and engine) the users acquire and thus what cars make up the national fleet, including the dynamics of the fleet development over time ("invest" indicators). Austrian transport policies between 1965-2019 included a range of instruments of both types. A structured empirical analysis of the short term and long term effects on passenger transport CO2 emissions of each of these policies at their respective stringency levels revealed that "invest" policies were the ones with stronger emission reduction implications.

In achieving a zero-emission requirement in the future itself, Austria, as a Member State of the European Union, has to act according to European law. With the European Union's 'Fit for 55' package – aiming at reducing GHG emissions by 55% by 2030 – the EU has effectively abolished the admission of passenger and light duty vehicles powered by combustion engines from 2035 onwards. The Council in March 2023 has further specified that for the use of e-fuels (i.e. climate-neutral synthetic fuels) combustion engines are still admissible. An EU wide date in admission restrictions from 2035 onwards, however, would not ensure carbon neutrality of the private vehicle transport sector by 2040, given an average lifetime of passenger cars in Austria of 15 years (EAA, 2019).

Could Austria unilaterally restrict registrations earlier? At first sight the answer is no. EU Member States are required to permit market access to all vehicles as specified by the harmonized European Union wide regulation. Otherwise they would interfere with the "right to the free movement of goods", a fundamental principle of the EU Treaty (Article 28) (Steininger, Posch, et al., 2024). However, a Member State could invoke another article of the Treaty on the Functioning of the EU, which allows Member States to take environmental action, if the environmental problem is ´specific´ to the respective country (Art. 114 (5)). Austria would need to prove that its conditions are particular and different than in EU Member States in general, requiring such a unilateral measure. One could think of arguments, such as a particularly high sprawl in settlement structures or its alpine topography, but it would remain very uncertain whether these arguments would be sufficiently strong to justify such a unilateral intervention.

The transition in actual registrations, however, need not to be governed by registration bans of combustion engines, but could be enhanced and especially started earlier by respective economic incentives, e.g., with registration fees differentiating between combustion and other engines. Equivalently, the standard fuel consumption tax - as found effective in our analysis, see section 2.4.1 - could be differentiated much more significantly by engine type. If such differentiation is sufficiently strong, registrations will shift (BMK, 2021). EU Member

States do have own decision power in this field. Achieving an adequate fleet composition in the net zero target year is not the only concern of countries. For the example, in the EU, national greenhouse gas emission targets are set for both specific points in time (in particular 2030 and 2050) and for the path to get there. Up to 2030, a target path of linear emission reduction is specified for each member state for the emissions in the "effort sharing" system, i.e., all emissions outside the EU Emissions Trading System (ETS) that is operated at the overall European level. Exceeding this path is sanctioned by asking countries to acquire emission reductions of countries that lower their emissions more than they are required to.

Consequently, nations are not only concerned with the year to reach carbon neutrality, but also with emission reductions throughout the period until they reach carbon neutrality. Therefore, it is not the admission regulation alone that is of interest, but also other complementary policy measures influencing transport emissions well before the year carbon neutrality is sought to be achieved. The broad range of policies reflected in Sections 2.2 and 2.3 and synthesized in Table 2.3 thus remains crucial also throughout the period of transition to carbon neutrality, with e.g., economic instruments complementing technological standard setting, in particular for EU member states as they do have more individual leeway with the former than with some variants of the latter.

In our analysis of the "Usage"-type instruments, fuel taxes turned out effective. As an instrument in that spirit, Austria in 2021 has implemented a national carbon tax, effective also for the transport sector. The feature that the rate is increasing by 10 €/ton each year, addresses our finding that without such a rise the emission-reduction effect would fade out. Given that this instrument is already established, the current discussion in Austria in the context of its update of the National Energy and Climate Plan to be submitted to the European Commission focuses on other additional usage instruments: both kilometre based road pricing and reducing the speed limit are found significantly effective (Steininger, Riahi, et al., 2024). The former instrument works very similar to the fuel tax analysed above (but in this case also for alternative fuel vehicles, such as electric vehicles). While speed limits turned out not statistically significant in the past in our analysis above, the much stricter and more widely applied levels now discussed (30/80/100 kilometres/hour for municipal/countryside/highway) are broadly assessed to be of significant impact (Steininger, Riahi, et al., 2024).

### 2.4.3 External Validity

Austria presents an interesting case to study because it implemented a wide variety of policies over the past decades. These include policies targeted at the purchasing behavior of vehicles and the usage of vehicles. For the case of Austria, we find that policies targeting the purchasing behavior (the household permanent consumption good, i.e., investment decision), have been introduced in a way that they were more effective than policies addressing usage, but that this stronger effect comes with a time lag. The explanation for this stronger relevance may be a historic one. Environmental regulation started out as a field of foremost legal administration, largely resorting to command and control instruments. While in the Anglo-Saxon regions economic environmental policy instruments addressing usage were present at least over the last half century, countries such as Austria started to expand their regulatory toolbox to include them much later, and still make use of them on a comparatively smaller scale.

We believe our results externally validate because of the microeconomic principle that price changes and other policy measures generally impact consumer utility and thereby influence more sustainable transport choices. This economic rationale reinforces the broad applicability of our analysis in achieving environmental objectives through targeted policy interventions to different countries. The exact degree of applicability, however, depends on macroeconomic parallels as well as the respective policy environment of these countries over time. We will discuss these aspects in more detail below. Furthermore, we believe that our results validate not only between countries but also for specific policy goals not historically targeted by the policies we study. An example can be an accelerated switch to a more environmentally friendly engine technology, such as battery electric vehicles.

Macroeconomic factors that influence the applicability of our results to other countries include GDP per capita, population density, the geographical landscape, and energy sources. These elements collectively shape the feasibility and public reception of specific policy interventions. Economic strength may impact the viability of specific policy measures as well as their public acceptance. Population density and the distribution of population centers affects commuting habits and distances driven. Austria's geographical landscape allows it to harness renewable energy from hydroelectric power, which can be vital for supporting battery electric vehicles and to reduce emissions from a transition to more sustainable engine technologies. Countries comparable to Austria in these respects include several countries in Scandinavia - namely Sweden, Norway, Finland, and Denmark. However, Belgium, Czechia, Slovenia, Germany, and Switzerland also share numerous characteristics with Austria, making them relevant for the extrapolation of our results.

Regarding the policy environment, the Austrian experience is one of a balanced policy approach, combining both investment and usage strategies that the literature indicates to be crucial in general. Our findings on Austria's transport policies first offer insights relevant and valid for countries with similar policy environments. For example, France, Ireland, Belgium, Denmark, and Sweden introduced taxes on new vehicle registration and ownership that are directly or indirectly based on CO2 emissions and have similar or higher fuel taxes. Our results indicate that such a policy combination can be particularly effective in accelerating a transition away form combustion-engine vehicle towards zero-emission vehicles as well as other modes of transport. The latter three countries share similarities with Austria in terms of economic size and structure, further strengthening our believe of external validity in these cases.

Second, our results offer insights for countries such as Bulgaria, Croatia, Czechia, Estonia, and Poland who all lack CO2-based vehicle taxes and have lower fuel taxes. Among these, Czechia shows the closest macro-level resemblance to Austria. Nevertheless, we believe our results possess a degree of external validity for all mentioned countries, which is anchored in the micro-foundation previously discussed. The Austrian experience across more than five decades analysed here shows that both types of policies could well complement each other for achieving the now increasingly more relevant carbon neutrality target in passenger transport. Especially instruments that affect investment decisions could introduce the more long-term effective component to the national transport policy instrument package. Results also indicate that policy types do not substitute for each other, but each have their merits and contribute to achieving the emission targets. Thus we consider our results to underscore the importance of the broader principle of balancing investment and usage policies in national transport policy design.

## 2.5 Conclusion

In this paper, we introduce a new environmental policy stringency index targeted to the Austrian transport sector for the period 1950-2019. The index encompasses two main policy categories: Investment-related policies and policies that directly or indirectly limit the usage of vehicles. We then incorporate this index in an econometric model to study the efficacy of these policies in reducing CO2 emissions in the Austrian passenger transport sector. Our results can help policy makers design balanced policy packages to accelerate the transition towards a carbon-neutral transport sector. Moreover, our results are not only relevant to the Austrian transport sector, but provide policy conclusions and recommendations for a broader set of countries.

We find that for the Austrian case stringent taxes affecting the investment decision to buy a new car show the strongest effect on passenger transport CO2 emissions out of the two main policy categories. Among these policies, the engine-related insurance tax quantitatively shows the strongest impact. Doubling the stringency reduces emissions by about 0.16 kt per

1,000 persons and year. Whereas the standard fuel consumption tax (an emission-based tax on newly registered vehicles) shows the strongest statistical significance and a 100% increase in its stringency (that is the maximum divergence in stringency observed bewteen any two points in time within the period of analysis) reduces emissions by about 0.10 kt per 1000 persons and year. Both policies take a few years to show a significant impact on emissions. This is to be expected given that it takes time for newly registered vehicles to disseminate broader and only then to impact fleet emissions.

Targeting the usage significantly reduces emissions only in the short run. Among these policies, the mineral oil tax (a tax on fuel consumption) is found to be most effective. Doubling the stringency of this policy reduces $CO_2$ emissions from passenger cars in the long-run by about 0.13 kt per 1000 persons and year. Note that its impact is smaller in magnitude compared to the investment-related policies and it is only statistically significant for a short period. The remaining usage-related policies are found to be statistically insignificant during our sample period.

Our study opens several avenues for future research. The policy stringency index for the transport sector, tailored to a specific country, is a novel approach. The proposed methodology can be applied to a range of other countries to assess the external validity of our results. Another avenue is the development of theoretical models and empirical study of micropanels to provide a microfoundation for policy effects and a more detailed understanding of these at the individual level. Another extension is an updated dataset that includes more recent information regarding the adoption of electric vehicles and associated policies, such as subsidies. Finally, policies could further be disaggregated to study their specific effect on different propulsion technologies, e.g., petrol and diesel. This approach might offer a clearer view of how different policies influence the transition to more sustainable transportation options.

## Funding information

This chapter is part of a project funded by the Austrian Climate Research Programme (ACRP) project B960266 – KR18AC0K14626 under the title “The role of persistence in tackling Austria’s climate target: Policies for the transport sector”.

# Chapter 3

# Zero fare, cleaner air? The causal effect of Luxembourg's free public transportation policy on carbon emissions

In March 2020, Luxembourg became the first country to make public transport free. We use this unique setting to evaluate the policy's impact on carbon emissions. Synthetic difference-in-differences allows us to identify a suitable control group. We use spatial emissions data to construct a panel of NUTS 2 control regions in the EU from 2016 to 2021. Our estimates indicate an average reduction of around 8% in road transport emissions. We account for potential confounders, such as the COVID-19 pandemic, shifts in commuting behaviors and advancements in vehicle technologies. Robustness checks support the credibility of our results.

**Acknowledgements:**
We thank Christoph Wunder, Amelie Wuppermann, Melanie Krause, Mike LaForest-Tucker, Johannes Gessner, Jules Linder, Davide Cerruti, and Andrea Weber, as well as seminar audiences at Eureka (VU Amsterdam), Doctoral Colloquium (University of Graz), MEEP 2024 (Hamburg), EAERE 2024 (Leuven), Applied Econometrics Workshop 2024 (Leipzig), AURÖ Workshop 2024 (Karlsruhe), Applied Economist Workshop 2024 (Hannover), and EEA (2024) for helpful feedback. We also thank Damian Clark, Asjad Naqvi, and Raian Kudashev for sharing insights on the methods, COVID data, and institutional settings, respectively.

## 3.1 Introduction

In March 2020, Luxembourg became the first country in the world to abolish fares on all modes of public transit, including buses, trains, and trams, throughout the country to mitigate transport-related externalities (Research Luxembourg, 2021). The provision of affordable and efficient public transport is often discussed as an effective way of reducing carbon ($CO_2$) emissions from the transport sector (Federal Transit Administration, 2010; International Transport Forum, 2020). Accessible, affordable, and efficient public transit can encourage a shift from private motorized transport to more environmentally friendly modes. However, despite these benefits, fully free public transport policies are scarce.

We leverage this quasi-experimental setting in Luxembourg to causally identify and quantify the impact on $CO_2$ emissions in the road transport sector. To evaluate the effect of this policy, we use the recently introduced synthetic difference-in-differences (SDID) method to construct a credible counterfactual for Luxembourg and compare the post-intervention outcomes against it (Arkhangelsky et al., 2021). This allows us to isolate the policy's effect from other confounding factors to achieve robust causal inference.

Luxembourg stands out from other European Union (EU) countries in many ways. It has the highest Gross Domestic Product (GDP) per capita, the highest motorization rate, and the highest per capita $CO_2$ emissions from transport. These unique characteristics pose challenges in finding comparable regions for constructing a counterfactual scenario. To overcome this, we conduct our analysis at the Nomenclature for Territorial Units for Statistics (NUTS) 2 level, as Luxembourg itself constitutes a NUTS 2 region.[1] While entire countries may not serve as suitable comparison units for Luxembourg, other NUTS 2 regions such as Brussels, Amsterdam, or Paris offer more appropriate benchmarks. This level of analysis ensures a more meaningful comparison of emission trajectories.

To enhance the robustness of our identification, we employ SDID, which combines elements of traditional difference-in-differences (DID) and synthetic control (SC) approaches while overcoming their limitations in our context. The uniqueness of Luxembourg's case makes it less plausible that the parallel trends assumption required for DID estimation will hold. SC methods require a donor pool of units similar in predictors of the outcome to the treated unit – a requirement that is unlikely to be met in our setting. In contrast, the SDID method, combines elements of both DID and SC and allows us to construct a counterfactual $CO_2$ emission trajectory for Luxembourg from a pool of donor regions without relying on matches in absolute levels at any stage of the procedure – which is essential to draw causal

[1] NUTS is an EU classification system that divides countries into three levels. These classifications are used for collecting, developing, and harmonizing European regional statistics, conducting socio-economic analyses, and framing EU regional policies.

inferences about the policy's impact in our specific setting.

Moreover, we address potential confounding factors related to the COVID-19 outburst. The pandemic likely caused variations in mobility patterns that are unrelated to the free public transportation policy. However, this only complicates identification insofar as mobility behavior in Luxembourg changed differently compared to the control regions. To examine this, we draw on data on working from home and commuting inflow for Luxembourg. We find that Luxembourg's mobility patterns in response to the pandemic were largely consistent with those observed in other EU regions. We account for these patterns in our models to enhance the accuracy of our identification strategy. To control for regional variation in pandemic response, we additionally control for daily regional COVID-19 cases in our estimations.

The potential donor pool for constructing Luxembourg's counterfactual comprises all other European regions at the NUTS 2 level over the period 2016-2021. From this pool, we exclude regions that have implemented any form of public transportation subsidy during the study period (this is elaborated in Section 3.4). After ensuring a balanced sample, our final donor pool includes 137 NUTS 2 regions and 822 region-time observations. Using this dataset, we estimate that the free public transport policy in Luxembourg led to an average treatment effect on the treated (ATT) of around -0.083, i.e., to a reduction in $CO_2$ emissions from the road transport sector by 8.3%.

Our results are significant at the 95% confidence level. We conduct an event study analysis to verify that parallel trends hold in the pre-treatment period. We conduct various robustness and sensitivity tests, including a placebo test by backdating the policy to 2019, iteratively leaving out regions and countries from the donor pool, a specification that accounts for fuel tourism effects, and analyzing a more restricted sample of NUTS 2 regions. We also examine the sensitivity of our results to different model specifications. We also carry out an SDID analysis on the $CO_2$ from energy use in the building sector to detect if our effect is purely driven by the impact of the COVID-19 pandemic. Our findings remained consistent across all these tests. Reassuringly, our estimates closely align with survey-based assessments of Luxembourg's free transit policy. We additionally extend our estimates to 2022 and find an increasing effect, but are cautious in interpreting these findings due to data quality issues and confounding factors (further discussed in Section -3.7).

We contribute to the literature by providing the first causal assessment of a free public transport policy on $CO_2$ emissions. Methodologically, we employ novel approaches to address the unique challenges presented by Luxembourg's distinct characteristics and the concurrent COVID-19 pandemic. Additionally, this study offers a framework for addressing COVID-19 as a potential confounder in similar research contexts. To the best of our knowledge, there is only one other study that directly looks at Luxembourg's free public transportation policy. Bigi et al. (2023) use an agent-based modeling approach and indicate that the policy significantly contributed to a modal shift from private vehicles to public transport. Our findings contribute to this narrative by providing a causal ex-post evaluation of the policy's impact on $CO_2$ emissions.

The existing literature on the effects of *free* public transport on $CO_2$ emissions is still scarce, largely because such policies were relatively uncommon. Tallinn (Estonia) introduced free public transit in 2013 and extended it since. Descriptive work by Cats et al. (2017) found that this policy is associated with an increase in public transport usage, but had no significant effect on car usage. Bull et al. (2021) randomly assigned free public transport vouchers to workers in Santiago (Chile), which were primarily used during off-peak hours. This suggests that the vouchers were more often utilized for leisure activities rather than reducing car usage.

Our paper links to a larger body of literature that ex-post evaluates transport policies designed to decrease reliance on motorized vehicles. Policies aimed at mitigating transport emissions can be categorized into three main types. The first one examines policies intended to directly reduce or restrict the use of motor vehicles by making driving more costly or less convenient. These include initiatives such as driving restrictions (Davis, 2008, 2017; Gallego et al., 2013), low-emission zones (Sarmiento et al., 2023; Wolff, 2014), road pricing (Gibson & Carnovale, 2015), and tax-based instruments (Andersson, 2019a; Pretis, 2022b).

The second type of policies promotes a shift to public transport, mainly by subsidizing public transit systems or improving infrastructure. This body of literature is particularly relevant to our study, as we also investigate the effects of improved public transport, specifically through enhanced access. Despite the apparent overlap between free transit and subsidized transit programs, we maintain that it is useful to distinguish between the two. While one might consider free transit merely a specific type of subsidy, factors such as user perceptions, convenience, and behavioral responses can diverge markedly when no fare is charged. Consequently, a fare-free system might lead to ridership patterns that differ from those observed in more traditional, partially subsidized scenarios.

Recent research on subsidized transit demonstrates mixed evidence regarding environmental outcomes. For instance, Aydin and Kürschner Rauck (2023) and Gohl and Schrauth (2024) evaluate the impact of Germany's 9-Euro ticket, introduced in 2022, and both report a decrease in air pollution, particularly in regions with robust public transit networks. However, contrasting findings are presented by Liebensteiner et al. (2024), who observe that while the 9-Euro ticket led to a significant increase in train rides during leisure hours, it only marginally reduced car usage. Similarly, Borsati et al. (2023) find no significant effect on air quality from a four-month public transport subsidy in Spain.

Research on public transit infrastructure provides additional insights. Li et al. (2019), Lalive et al. (2018), and Chen and Whalley (2012) show that expanding subway and rail services in China, Germany, and Taipei, respectively, improves air quality. Gendron-Carrier et al. (2022) find no average effect from subway openings across 58 cities, but reductions in pollution in more polluted cities. Overall, these studies suggest public transit investments can improve air quality, though outcomes vary by local context.

Some studies indirectly measure the effects of public transport in the absence of explicit policy interventions, using transit strikes to assess substitution between public and private transport. For instance, Anderson (2014), Adler and van Ommeren (2016a), and Bauernschuster et al. (2017) find significant increases in congestion following transit strikes in Los Angeles, Rotterdam, and Germany's five largest cities, respectively.

Policies related to the third type aim to improve the energy and fuel efficiency of vehicles through regulations such as gasoline content standards (Auffhammer & Kellogg, 2011). While most studies focus on individual policies, some jointly examine multiple interventions (Koch et al., 2022; Kuss & Nicholas, 2022; Winkler et al., 2023).

The rest of the paper is organized as follows. Section 3.2 briefly introduces Luxembourg's free public transport policy. The Data used is detailed in section 3.3. The identification strategy is discussed in section 3.4. The empirical strategy, including the SDID procedure, is detailed in section 3.5. section 3.6 provides our empirical results and section 3.7 our robustness tests. The results and potential mechanisms are discussed in section 3.8. Finally, section 3.9 provides concluding remarks.

## 3.2 Background: Luxembourg and the policy

Luxembourg is a small country in Western Europe and spans an area of about 2,586 km$^2$, making it one of the smallest countries in the EU. In the NUTS statistical classification, Luxembourg is treated as a single region at all levels. The country hosts several EU institutions, with its economy primarily driven by banking and finance. Despite its small size and population, Luxembourg has the highest GDP per capita among EU countries, at approximately 140,000 USD. The economic hub is concentrated in Luxembourg City, the capital, located in the south. The country experiences a significant daily inflow of commuters from neighboring Belgium, Germany, and France, with around 200,000 people commuting daily, representing a substantial portion of its population of approximately 660,000. Luxembourg has the highest per capita $CO_2$ emissions from transport among EU member states, at around 8,200 kg. It also has the highest car density in the EU, with about 700 cars per 1,000 inhabitants. These characteristics set the country quite far apart from other EU countries.

On March 1, 2020, Luxembourg became the first country in the world to offer free public transport nationwide, available to all residents and visitors regardless of age and income group. Tickets are only required for 1st class travel. This initiative was part of the broader mobility strategy, “Modu.2.0” that aimed at improving the sustainability of the mobility system (Ministère du Développement Durable et des Infrastructures, 2018). Luxembourg designed this policy with the aim of reducing car usage to counter its high car density and significant congestion problems. Before the implementation of this policy, annual revenue for ticket sales in Luxembourg amounted to about 41 million euros, which accounted for approximately 8% of the annual cost of transport system maintenance (Ministère du Développement Durable et des Infrastructures, 2018).

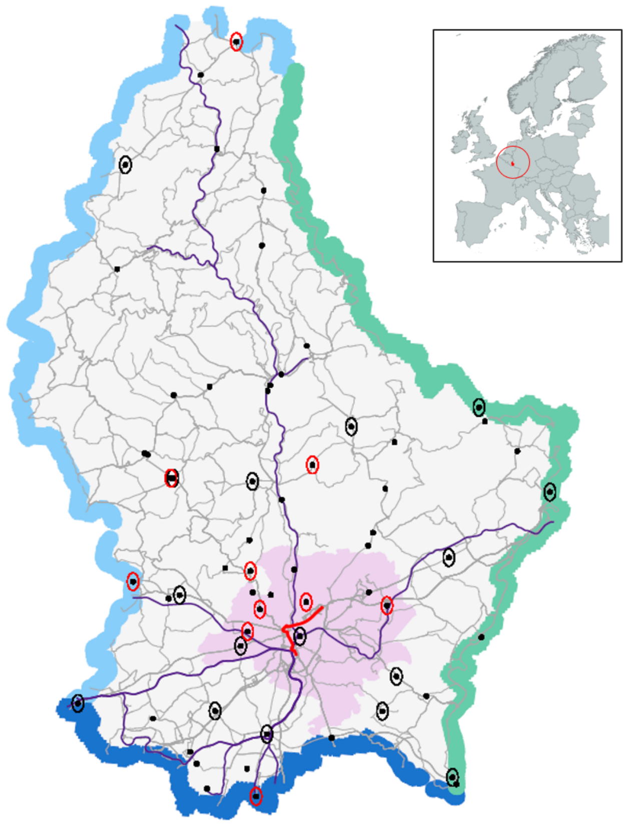
The existing public transportation infrastructure forms the backbone of the policy initiative and comprises buses, trams, and trains. The public transit network is sketched in Figure 3.1, where bus lines are shown in grey, train lines are in purple, and the tram line in red. Buses are the predominant mode of public transportation in Luxembourg and offer quite a comprehensive coverage across the entire country. They connect different localities as well as cross-border lines (Ministère du Développement Durable et des Infrastructures, 2020). Altogether about 400 bus lines are running through Luxembourg, connecting the entire country (Administration des transports publics, 2024). Trains additionally cover the country in a star-like network, originating in Luxembourg City and connecting it to cross-border connections (Département de la mobilité et des transports, 2020).

The city of Luxembourg is additionally served by the only tram line in the country, which covers around 10 km through 17 stations (Département de la mobilité et des transports, 2024). Before the implementation of the free public transportation policy, Luxembourg charged differentiated public transport fares based on the duration and length of travel. Special rates for children and the elderly were available, as outlined in the Ministerial Regulation of July 14, 2017 - *Règlement ministériel du 14 juillet 2017 fixant les tarifs des transports publics* (Le Ministre du Développement durable et des Infrastructures, 2017). Short-term tickets, valid for a maximum of 2 hours from validation were priced at 2 euros. Long-term tickets, valid for 1, 2, and 3 days, ranged from 4 to 12 euros, while annual network subscriptions were priced at 440 euros.[2]

It is worth noting that the free public transit policy was complemented by enhancements in the transportation infrastructure, notably through the strategic expansion of the national rail network's capacity and extensions in the tram line coverage. In 2017, Luxembourg

[2] A detailed schedule of public transport fares is available at (Le Ministre du Développement durable et des Infrastructures, 2017).

**Figure 3.1:** Luxembourg public transport network and traffic camera posts



Note: The map shows Luxembourg's borders with Belgium (light blue), Germany (green), and France (dark blue). Black dots indicate traffic posts, with circled ones showing a drop in traffic from 2019 to 2021; red-circled dots mark the top 10 largest declines. Light grey lines represent regional buses, dark purple lines are national rail, and the red line marks the tram. Luxembourg City is shaded in light pink. Public transport networks shown as of 2018; data is from Luxembourg's open data portal.

introduced a tram line traversing Luxembourg City, initially connecting 8 stations. The following year saw the line's expansion by 3 more stops. December 2020 marked another extension, enlarging the network by 2 kilometers and incorporating 4 additional stations. By September 2022, the tram network further expanded with the addition of 2 new stations.

The most recent extension lies outside our sample period. Since the 2020 extension coincides exactly with the introduction of the free transit policy, it is impossible to disentangle their individual effects directly. However, we show in a robustness test that the 2017 expansion did not lead to significant reductions in emissions, suggesting that the comparatively minor 2020 extension is unlikely to have had a notable effect either. We will return to the latter aspect in more detail in section 3.6.

Currently, the tram stretches over 10 kilometers, serves 17 stations, and includes 6 major interchanges (Département de la mobilité et des transports, 2024). Three more tramlines are planned to be completed by the end of 2035 (Luxtoday, 2022). Luxembourg also improved parking availability, particularly near border areas for its cross-border commuters. Additionally, through negotiations with neighboring transport networks, fares for cross-border transport have been lowered (Ministry of Mobility and Public Works, 2020). Consequently, the new scheme is designed to benefit not only residents but also commuters from neighboring countries. The strategic objective for 2025 is to reduce congestion during peak hours while transporting 20% more people than in 2017.

Figure 3.1 also illustrates traffic posts in Luxembourg measuring bi-directional car travel volume. The traffic volume data is compiled by the Administration des Ponts et Chaussées (Luxembourg Bridges and Roads Administration) and includes daily traffic counts. We map the points for which we obtain an uninterrupted time series over the period 2018-2021. The traffic posts circled all experienced a decrease in annual bi-directional car traffic volume compared to 2019, and the ten red circles experienced the largest drop. The circled traffic posts are largely situated in the vicinity of Luxembourg City and mostly close to public transport networks. Overall, traffic volume increased annually up to 2019 and stagnated after 2019, on average.

## 3.3 Data

We combine the following data to estimate the causal effect of Luxembourg's free public transport policy on $CO_2$ emissions from road transport. Data on the outcome variable, per capita $CO_2$ emissions from the road transport sector, are constructed by combining spatial road transport $CO_2$ emissions extracted from the European Emission Database for Global Atmospheric Research (EDGAR) v8 (Crippa et al., 2022) with population data from Eurostat's (2024) regional statistics. We select emissions as our primary outcome variable not only because they directly relate to a core policy goal, but also because emissions data are consistently available as a panel dataset across regions. This availability is essential for conducting a robust causal evaluation in this context. In contrast, data on transit ridership or vehicle mileage remain scarce and are not always measured uniformly, making them less suitable for a systematic assessment of policy impacts.

To control for other factors that may influence $CO_2$ emissions from road transport, we include several covariates. Controls related to the pandemic include daily COVID-19 cases at the NUTS 2 level, sourced from Naqvi (2021), as well as data on working from home and commuting inflows, obtained from a special extraction from the EU Labor Force Survey (EU-LFS). Other controls encompass fuel prices, which we source from the European

Commission's (2024b) weekly oil bulletin. Energy intensity is taken from EEA (2024) and captures changes in efficiency of cars. Data on loaded goods is included to capture the effect of freight transport emissions and obtained from Eurostat's (2024) regional statistics. Finally, we use data on real GDP per capita from the regional statistics to control for overall differences in economic development.

We drop regions with missing data and regions that experienced methodological breaks or data-quality disruptions in data generation to ensure a coherent balanced panel.[3] After additionally dropping bad controls (see Section 3.4), we are left with 19 EU countries (including Luxembourg) and a total of 138 regions over the sample period 2016-2021, giving a total of 828 region-year observations. The following subsections discuss in more detail the outcome variable, $CO_2$ emissions from road transport, and the COVID-19 related controls used in our analysis for this remaining set of countries.

### 3.3.1 $CO_2$ emissions data

Road transport emissions are categorized under the Intergovernmental Panel for Climate Change (IPCC) 1996 sector category 1.A.3.b. Emissions are calculated as the product of fuel consumption times the associated IPCC emission factors. The EDGAR database provides annual sector specific grid maps expressed in ton substance with a spatial resolution of 0.1 degrees $\times$ 0.1 degrees. We aggregate these grid cells to the corresponding NUTS 2 regions for the following 19 located in Europe: Austria, Belgium, Croatia, Czech Republic, Denmark, Finland, France, Greece, Hungary, Ireland, Italy, Latvia, Luxembourg, Netherlands, Poland, Portugal, Slovakia, Slovenia, and Spain. The NUTS 2 regional borders are extracted from the Eurostat database (European Commission, 2022).

We present the evolution of $CO_2$ emissions from road transport for Luxembourg and other NUTS2 regions over time in Figure 3.2.[4] Panel (a) shows the evolution of the log of annual $CO_2$ emissions from road transport over the period 2016-2021. Luxembourg is indicated by the solid black line, while other NTUS 2 regions are shown in gray. The impact of COVID-19 can be seen in a drop in emissions from 2019 to 2020 across all regions. In 2021, an increase in emissions can be observed. However, both the drop and subsequent increase vary across regions.

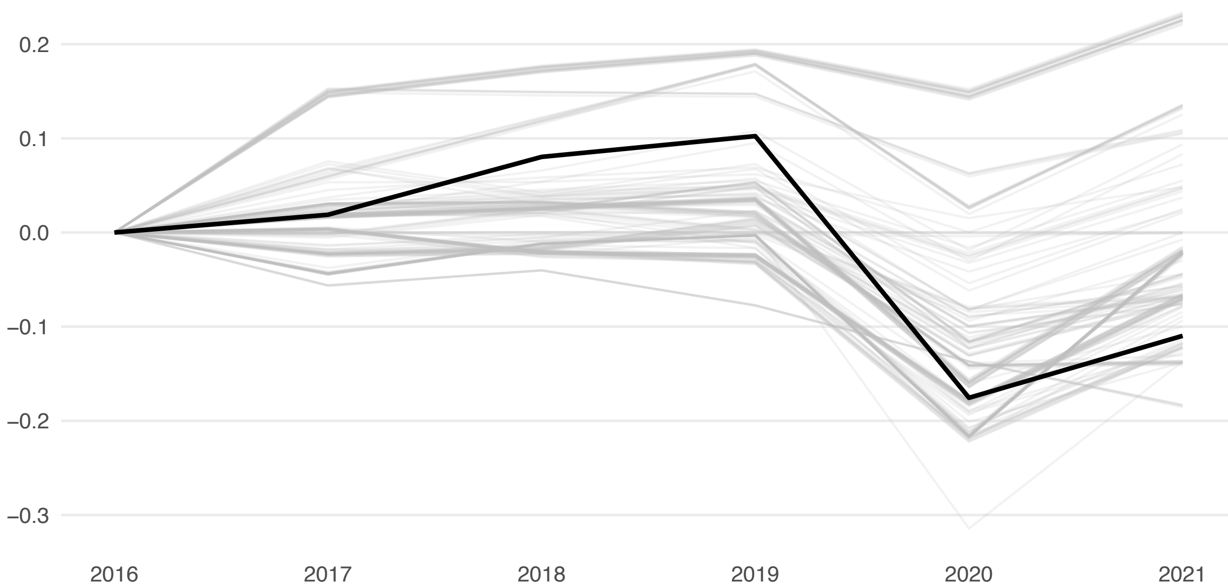
Luxembourg seems to have experienced a relatively large drop in 2020 relative to other regions, and emissions in 2021 stay consistently below pre-pandemic levels. Panel (b) shows

[3]We drop the United Kingdom, Norway, Romania, Sweden, Switzerland, Liechtenstein, and Lithuania due to missing data. Germany experienced methodological and quality breaks in EU-LFS data generation.

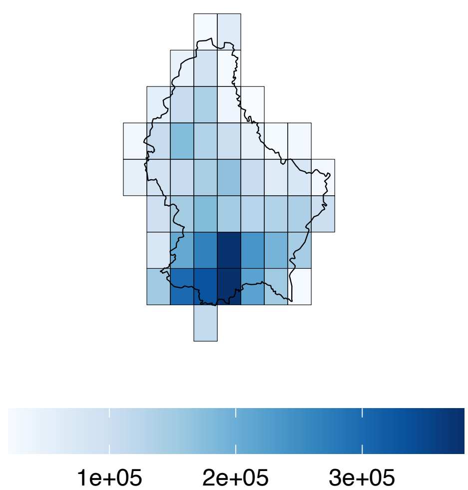
[4]Grid-cells that intersect with the NUTS 2 boundaries of Luxembourg are allocated according to their fraction that falls inside these boundaries.

**Figure 3.2:** Evolution of $CO_2$ emissions in Luxembourg over time and space

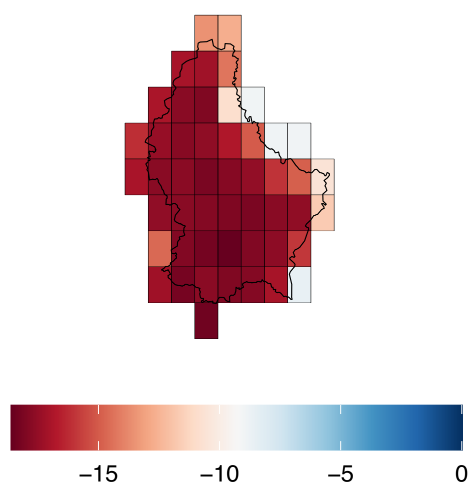
**(a)** Evolution of Log-$CO_2$ Emissions for Luxembourg and other NUTS 2 Regions




**(b)** Average Emissions 2016-2019




**(c)** %-Change 2020-2021 vs. 2016-2019




Note: Road transport $CO_2$ emissions (tons) are extracted from the EDGARv8.1 at 0.1x0.1 grid cells. (a) Shows the evolution of Log-$CO_2$ emissions, centered at zero in 2016. Luxembourg is indicated by the black line. (b) and (c) display spatial distributions of emissions for Luxembourg. (b) shows average emissions over the pre-treatment period, 2016-2019. (c) shows the %-change from average emissions over the post-treatment period (2020-2021) compared to the pre-treatment period.

the spatial distribution of average road transport emissions over the period 2016-2019, which constitutes our pre-treatment period. High emissions are indicated in dark blue and lower emissions in light blue. Emissions are concentrated around Luxembourg City and border regions with France. Panel (c) shows the percentage change of average post-treatment (2020-2021) emissions relative to average pre-treatment emissions. Emissions on average stayed below the pre-policy average in the entire country. The largest difference can be observed around Luxembourg City, while differences on the Eastern border of Luxembourg are less pronounced. The overall average emission reduction for the country for the post-treatment period relative to the pre-treatment period is around -17.5%. To extract the extent to which this reduction can be attributed to the free-public transport policy is the aim of our paper.

The reduction in $CO_2$ emissions shown in Figure 3.2 is directly related to a reduction in fuel consumption, indicating a shift in mobility patterns. This shift may be attributed to various factors. Our primary interest is the causal effect of the free public transport policy. To discern this causal effect, we need to account for potential variation caused by other confounding effects. These potential sources of variation in $CO_2$ emissions include COVID-19 related restrictions and reduced mobility, as well as an increase in the number people working from home and fewer commuting trips.

### 3.3.2 COVID-19 related variables

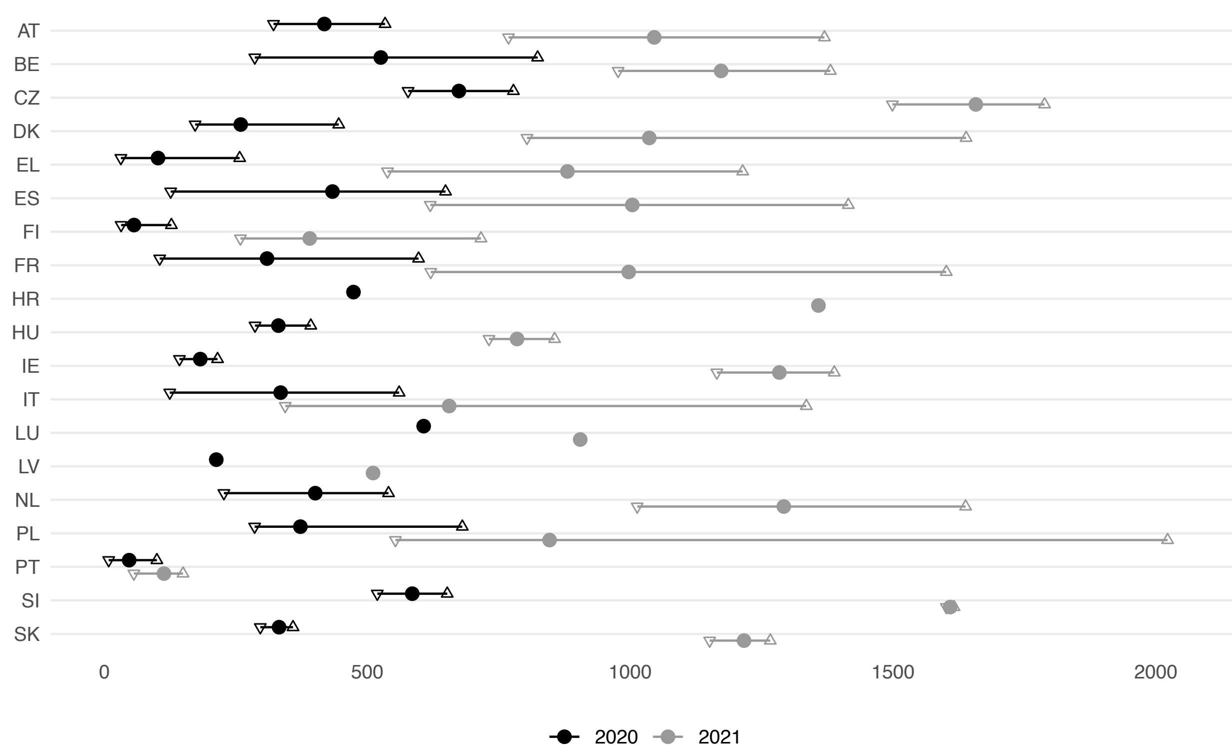
With the onset of the COVID-19 pandemic, many countries implemented lockdowns and travel restrictions to curtail the spread of the virus (Hale et al., 2021). Luxembourg was no exemption, with its government convening an extraordinary Government Council to respond to the pandemic on the 12th of March 2020. Subsequently, mobility restrictions aimed at containing the spread of the virus came into effect on the 13th of March, 2020 (Government of the Grand Duchy of Luxembourg, 2020).

The Our World in Data (OWID) COVID-19 Government policy stringency index, a composite index based on 9 response measures, illustrates that many countries, including Luxembourg, adopted similar measures during this period (Hale et al., 2021). These restrictions were often enforced at regional or local levels, triggered by the number of cases reported in specific areas. To capture the effect of the pandemic, we use data on confirmed COVID-19 cases as a proxy for various policy responses and reduced mobility.

This data is collected and reported by the COVID-19 European Regional Tracker at the NUTS 3 level (Naqvi, 2021). Information on the number of confirmed cases is taken from each country's official institutions responsible for providing COVID-19 related data. The regional data is then aggregated up to the country level and cross-checked against data from OWID, which provides confirmed COVID-19 cases at the country level (Mathieu et al., 2020). The data matches well for 2020 and 2021.

Data quality, however, deteriorates in 2022, because the number of countries regularly reporting cases decreases strongly in 2022. The COVID-19 European Regional Tracker reports cases for all regions that we consider in our study, except for Luxembourg. However, since the regional data is validated against the OWID data and matches well for our sample period, we resort to COVID-19 cases from OWID for Luxembourg. For our analysis, we aggregate the NUTS 3 level data in the COVID-19 European Regional Tracker to the NUTS 2 level.

**Figure 3.3:** Regional variation in COVID-19 cases for 2020 and 2021



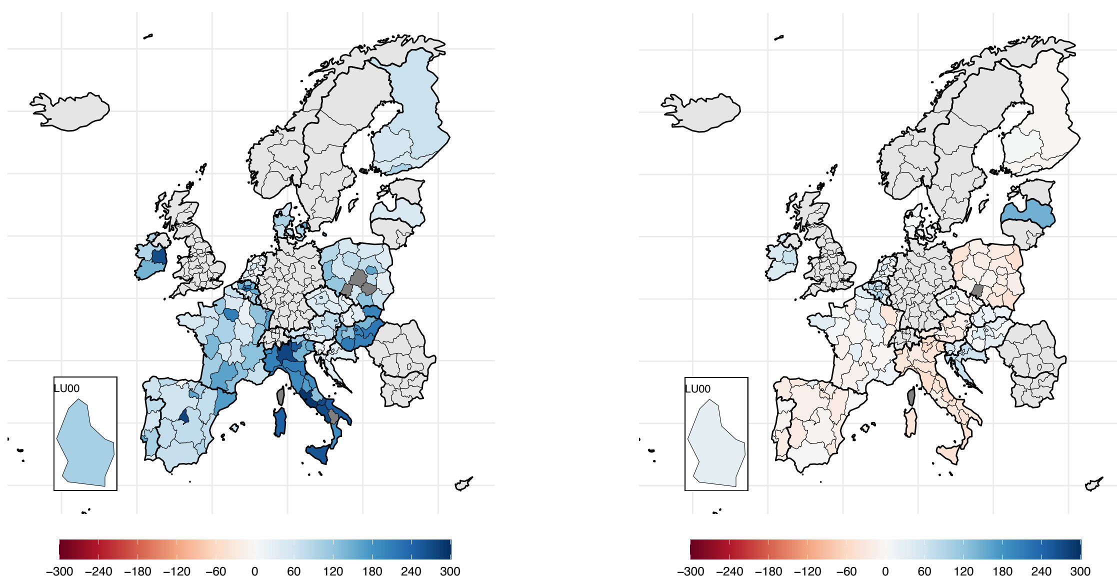

Note: The average daily confirmed COVID-19 cases and their spatial distribution across countries for 2020 and 2021. Data for Luxembourg is from Our Wold in Data (OWID), while data for NUTS 2 regions in other countries is taken from the COVID-19 European Regional Tracker (Naqvi, 2021).

Figure 3.3 shows the average regional variation in the number of confirmed daily COVID-19 cases per 10,000 persons for 2020 and 2021. Dots represent the mean of confirmed cases at the NUTS 0 level (i.e., country level), the downward-facing triangle represents the NUTS 2 region with the lowest and the upward-facing triangle the region with the highest number of confirmed cases per 10,000 persons within a country. The distance between these two points spans the spatial variation across NUTS 2 regions within a country. It is evident that this spatial variation is significant, which further motivates the choice to conduct our study at a regional level compared to the country level.

Overall, the number of cases per 10,000 persons as well as their spatial variation is smaller in 2020 compared to 2021. Countries with a larger population also tend to show a bigger variation in cases across their regions. Luxembourg does not show any regional variation because its NUTS 0 and NUTS 2 regional boundaries are identical. Average daily cases per 10,000 persons for Luxembourg in 2020 and 2021 are around 600 and 900, respectively. In 2020, this puts Luxembourg at the higher end of the spectrum of regional cases per 10,000 persons, while it puts it on the lower end in 2021. Compared to country averages, we find only few comparable units to Luxembourg. At the regional level, however, we find several regions with more cases in 2020 and fewer ones in 2021, further motivating our usage of regional data.

**Figure 3.4:** Change (%) in persons usually working from home for NUTS 2 regions

**(a)** 2019-2020 **(b)** 2020-2021




Note: Data is from a special extraction from the EU-LFS. Persons usually working from home with workplace at the NUTS 2 region shown in the figure and their location of residence in the associated country of the region.

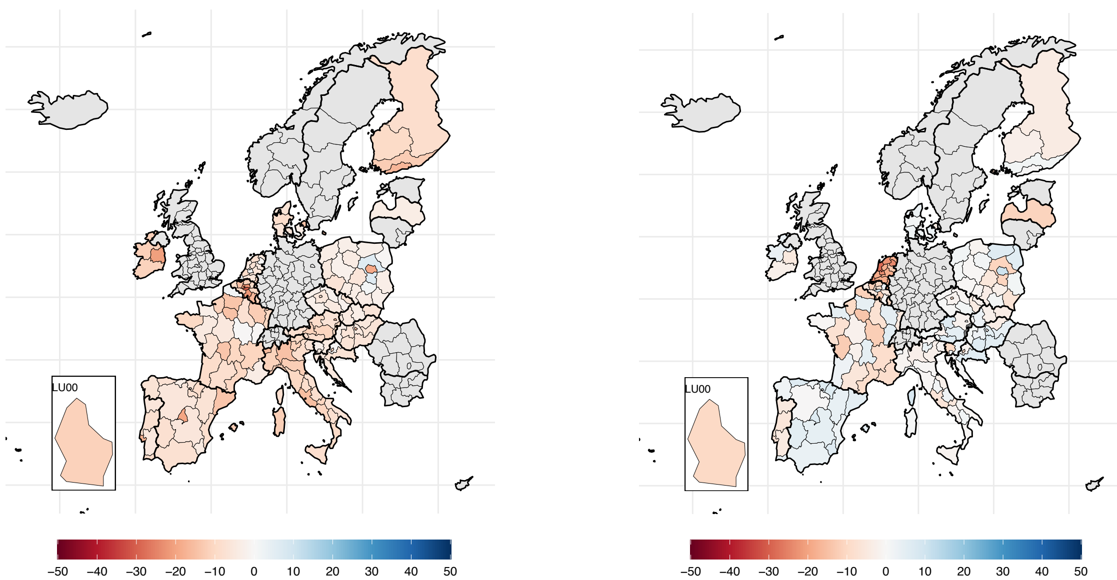
We use data on working from home and commuting inflow to further address changes in mobility behavior as a response to the pandemic. A person is classified as usually working from home when they were working at home half of the days that they worked in a reference period of four weeks preceding the end of the reference week in the EU-LFS survey. We focus on persons usually working at home with their workplace location in the associated NUTS 2 region and their location of residence within the same country.[5]

[5]Ideally, we would want to focus on persons working and living in the same NUTS 2 region. However, this is not available in the EU-LFS data structure.

However, this dataset does not capture commuting patterns across regions, which seems particularly important for Luxembourg, which traditionally experiences a large commuting inflow. To get a more complete picture of changes in mobility behavior with respect to work, we consider persons never working from home at a regional level. This category captures all persons commuting to work irrespective of their location of residence and thus incorporates commuting inflow from other regions and countries.

**Figure 3.5:** Change (%) of persons never working from home for NUTS 2 regions

**(a)** 2019-2020 **(b)** 2020-2021




Note: Data is from a special extraction from the EU-LFS. The figure shows yearly changes of persons never working at home for NUTS 2 regions which are the location of the workplace of these persons irrespective of their location of residence.

Figure 3.4 shows yearly changes of persons usually working from home for NUTS 2 regions. Figure 3.4a shows the change from 2019-2020, i.e., the immediate effect of the pandemic. Blue indicates an increase in working from home, whereas red indicates a decrease. As expected, almost all regions experienced an increase in people working from home. The figure zooms in on Luxembourg, which also experienced an increase, but notice that the change is not particularly strong relative to other regions, i.e., Luxembourg is not an outlier. In Luxembourg, the change of people usually working from home from 2019-2020 almost doubled at around +98%. Figure 3.4b shows the change from 2020-2021. The map now shows a more nuanced picture. Some regions experienced a decrease in working from home, while some experienced another increase. Luxembourg is among the latter group and experienced a change of around +28%.

Figure 3.5 shows yearly changes of persons never working at home for NUTS 2 regions.

Figure 3.5a shows percentage changes from 2020 to 2021. Overall, the map shows a decrease in persons never working from home, i.e. a decrease in commuters. This is to be expected since the pandemic caused an increase in working from home in most regions. Figure 3.5b shows percentage changes from 2020-2021 and shows a mixed picture. Some regions experienced a further decrease in persons never working from home, while others experienced an increase following the first year of the pandemic. Luxembourg experienced a decrease in 2019-2020 and 2020-2021 of $-12\%$ and $-10\%$, respectively. Again, Luxembourg does not appear to have experienced a particularly strong change relative to other countries.

## 3.4 Identification strategy

The inability to directly observe the potential outcomes of a specific unit both in the presence and in the absence of a policy event (treatment) complicates establishing causal relationships. In the case of Luxembourg, this translates to 'what would the $CO_2$ emissions from road transport have been if the free public transport policy had not been introduced?" To overcome this problem, it is necessary to design an appropriate identification strategy that constructs a credible comparison group to serve as a counterfactual for Luxembourg after the policy's introduction.

Given that Luxembourg differs significantly from other EU countries in observable characteristics such as $CO_2$ emissions per capita, GDP per capita, and motorization rates (refer to Section 3.2), we conduct our analysis at the NUTS 2 level. This approach is feasible because Luxembourg itself constitutes a NUTS 2 region, and it is likely that we can find more comparable units to construct the counterfactual for Luxembourg at the NUTS 2 regional level than at the country level. However, even at a NUTS 2 level, Luxembourg records the highest per capita $CO_2$ emissions from road transport. We therefore need an estimation strategy that can handle these complexities in our setting.

The canonical DID estimator calculates the difference in outcomes over time between treated and control units and relies on the parallel trends assumption. This assumption implies that, in the absence of treatment, the treated and control groups would have followed similar trends over time. By assuming parallel trends, the DID estimator controls for unobserved characteristics that remain constant over time, which might otherwise confound the results. Additionally, the DID method assumes that any time-varying shocks affecting the outcome are common to both treated and control groups, thereby isolating the treatment effect. However, the parallel trends assumption is often untestable, and in our specific setting, where Luxembourg already exhibits considerable differences in observable characteristics, we have reduced confidence that this assumption holds.

Some drawbacks of the DID method can be mitigated by the Synthetic Control (SC) method, which does not rely on the parallel trends assumption. Instead, the SC method creates a synthetic control unit as a weighted combination of units from the donor pool, ensuring that the pre-intervention outcomes of the synthetic unit closely match those of the treated unit. Importantly, not all units in the donor pool receive equal weights; higher weights are assigned to regions that are more similar to Luxembourg based on predictors of $CO_2$ emissions (Abadie, 2021).

The validity of the SC method depends on the trajectory of the outcome variable of the SC closely following that of the treated unit over a long pre-intervention period. This close alignment lends confidence that any deviations in outcome trends after the intervention can be attributed to the policy intervention. However, the substantial differences in predictors of $CO_2$ emissions between Luxembourg and other units, coupled with Luxembourg's status as the country and even the NUTS 2 region with the highest per capita emissions, challenge the applicability of this method in our context.

Therefore, we employ the recently proposed estimation procedure, the SDID approach introduced by Arkhangelsky et al. (2021). SDID combines the strengths of both DID and SC methods and circumvents the common drawbacks associated with traditional DID and SC methods. Specifically, it overcomes the challenge of estimating causal relationships when parallel trends are unlikely to hold in aggregate data for DID and eliminates the necessity for the treated unit to be within the convex hull of control units for SC. SDID essentially constructs a synthetic parallel trend for Luxembourg. Section 3.5 discusses the SDID estimation procedure in detail.

Identification is further complicated by the COVID-19 pandemic, which coincides with the policy's introduction. Since the pandemic was a global shock affecting all regions, its effects should not technically bias our analysis, as both the treated and control units were similarly exposed. However, regions adopted varying measures and policies to limit the spread of the virus, which could have differential impacts on mobility across regions. For instance, a higher number of COVID-19 cases may lead to shifts toward remote working, online education, and changes in consumer behavior. These policy responses, potentially influenced by the number of cases, could correlate with regional mobility restrictions. To account for these factors, we control for regional average daily COVID-19 cases across NUTS 2 regions.

Mobility patterns may have also shifted due to the pandemic. This is again only problematic insofar as regions experienced such shifts differently form one another. These changes include individuals who did not work from home prior to the pandemic but began and continued doing so after the COVID-19 outbreak. Consequently, mobility within countries (and within regions) and commuting patterns across borders might have changed. However, as discussed in detail in Section 3.3.2, Luxembourg did not experience particularly significant changes relative to other regions. This mitigates the associated threat to identification. It is nonetheless essential to control for these changes in the empirical analysis.

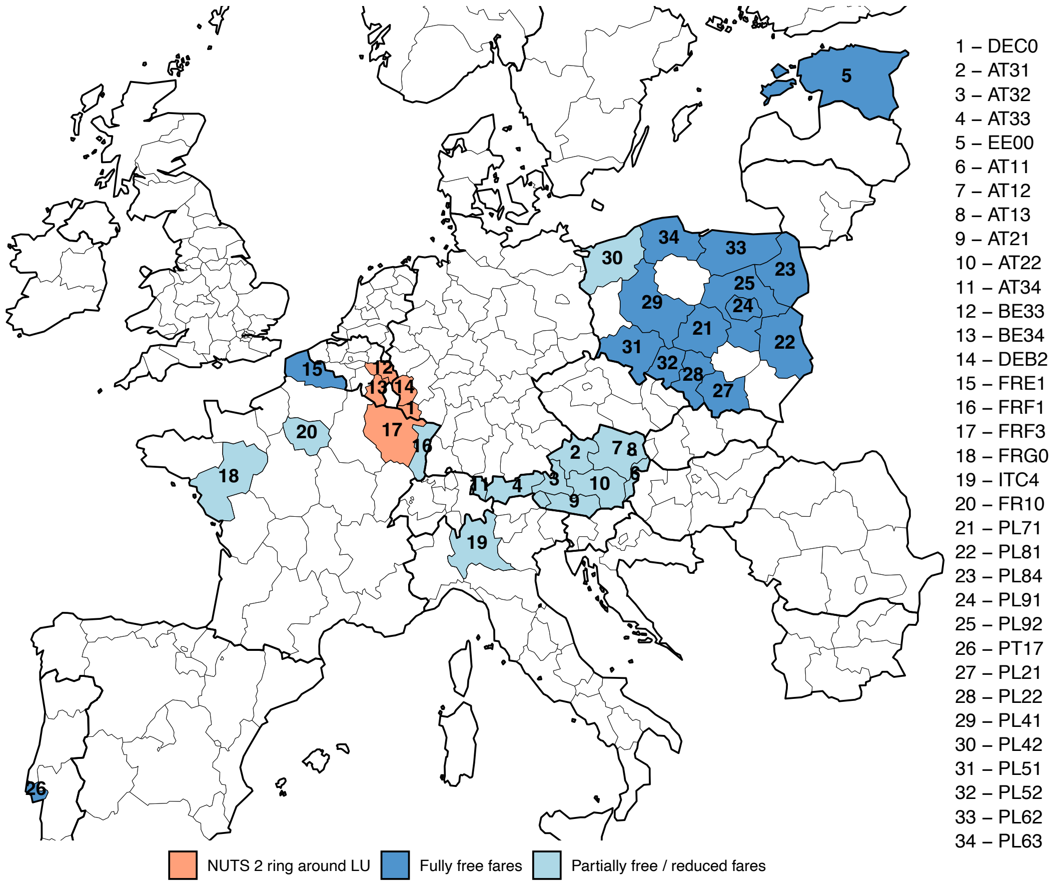
Finally, to avoid bad comparisons with already treated units, we excluded NUTS 2 regions that introduced free fares during our sample period. We drop the following regions before estimating our main results. Estonia (EE) introduced free public transport in Tallin in 2013 and further extended it in 2017. Given that Estonia is in itself a NUTS 2 region, we drop the whole country. Dunkirk and Calais in France introduced free public transport for all passengers in 2018 and 2020, respectively. Both are located within the same NUTS 2 region (FRE1) that we drop. We also drop Cascais in Portugal (PT17), which introduced free fares in 2020.

Several municipalities in Poland introduced some form of free public transport schemes during our sample period. Štraub et al. (2023) chart the spatial distribution of these policies in Poland, which covers over 90 free-fare programs since 2007. Polish municipalities that introduced free fares for everybody during our sample period cover 12 NUTS 2 regions which we drop (PL21, PL22, PL41, PL51, PL52, PL62, PL63, PL71, PL81, PL84, PL91, PL92). We also exclude the NUTS 2 regions surrounding Luxembourg to control for possible spillover effects. These regions include the Province of Luxembourg (BE34) and the Province of Liege (BE33) in Belgium, Trier (DEB2), and Saarland (DCE0) in Germany, and Lorraine in France (FRF3).

As a robustness check, we additionally drop regions that introduced free fares for specific groups (e.g., students, residents, elderly, etc.) or subsidized public transport during our sample period. These cases can distort the estimated effect if these policies significantly shifted the modal split in favor of public transport systems. Regions we drop in our robustness checks include the following. Attica in Greece (EL30), and Nantes (FRG0), Strasbourg (FRF1), and Paris (FR10) in France. These regions all introduced some form of free public transport for residents and/or students ("City Public Transport Information," 2024). Austria (AT) introduced a nationwide climate ticket for all public transport modes in 2021. This increased accessibility and significantly reduced prices for comparable tickets prior to the policy introduction.

The different regions that we drop in our main specification as well as in the robustness checks are shown in Figure 3.6. The figure zooms in on NUTS 2 regions in Europe to highlight potentially bad controls. NUTS 2 regions that introduced free fares for all passengers during our sample period are shown in darker blue. These are all the regions we drop in our specification to obtain our main results. Those that introduced free fares for specific groups

**Figure 3.6:** NUTS 2 regions - bad controls




Note: NUTS 2 regions that are potential bad control are highlighted. A NUTS 2 ring around Luxembourg in orange, regions that introduced free fares during our sample period in dark blue and regions that introduced reduced fares or partially free public transport in light blue.

only or introduced reduced fares are shown in lighter blue. These regions are additionally excluded from our sample in a robustness check. The NUTS 2 ring around Luxembourg is shown in orange and is dropped in all specifications.

## 3.5 Synthetic difference-in-differences (SDID)

We use the SDID methodology to estimate the impact of Luxembourg's free public transport policy on $CO_2$ emissions from road transport. The analysis covers a sample period from 2016 to 2021. As the policy is implemented in 2020, the analysis includes four years before the policy is introduced and two years after, which allows for a comparative analysis of the pre- and post-policy effects. Schenk (2023) shows that the SDID estimator performs remarkably well in short $T$ panels, is able to handle interactive fixed-effects that can influence the outcome, and provides conservative standard errors. Considering the few pre- and post-treatment periods in our sample, this reassures us that the applied methodology is consistent under our setting.

The SDID estimator aims to consistently estimate an ATT without relying on parallel pre-treatment trends between treated and not-treated units. In essence, SDID estimates the ATT, $\widehat{\tau}^{sdid}$, from a weighted two-way fixed-effects regression. Compared to SDID, DID approaches use an unweighted two-way fixed-effects regression, thus relying on parrallel pre-treatment trends in aggregate data. SC relaxes this requirement but uses only unit-specific weights and does not explicitly weigh time periods optimally. Contrary to SC method, SDID additionally allows for level differences between treatment and synthetic control units in estimating optimal weights. Following this rationale, Arkhangelsky et al. (2021) argue that SDID is more flexible compared to DID and SC methods.

The SDID-ATT is estimated by:

$$\left(\widehat{\tau}^{sdid}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \widehat{\omega}_i^{sdid} \widehat{\lambda}_t^{sdid} \right\}, \tag{3.1}$$

where the outcome of interest, $Y_{it}$, is observed for each unit $i$ at each time $t$, with $i = 1, ..., N$ and $t = 1, ..., T$. $W_{it}$, indicates treatment, with $W_{it} = 1$ if unit $i$ is treated at time $t$ and $W_{it} = 0$ else. $\mu$ is an intercept, $\alpha_i$ and $\beta_t$ are unit and time fixed-effects, respectively. $\widehat{\omega}_i^{sdid}$ and $\widehat{\lambda}_t^{sdid}$ are unit and time weights, respectively.

Unit weights are computed to align pre-treatments trends between treated and control units:

$$\left(\widehat{\omega}_0, \widehat{\omega}^{sdid}\right) = \underset{\omega_0 \in R, \omega \in \Omega}{\arg\min} \sum_{t=1}^{T_{pre}} \left( \omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 + \zeta^2 T_{pre} ||\omega||_2^2, \tag{3.2}$$

with $\Omega = \{\omega \in R_+^N,\ with\ \sum_{i=1}^{N_{co}} \omega_i = 1\ and\ \omega_i = 1/N_{tr}\ \forall\ i = N_{co}+1, ..., N\}$, where $||\omega||_2$ is the Euclidean norm and $R_+$ denotes the positive real line. $N_{co}$ and $N_{tr}$ are the number of untreated and treated units, respectively. Similarly, $T_{pre}$ is the number of pre-treatment periods. $\zeta$ is a regularization parameter to increase dispersion and ensure unique weights, it is defined in Arkhangelsky et al. (2021). Contrary to traditional synthetic control unit weights, these SDID weights do not aim to find comparable regions in absolute terms conditional on covariates, but the procedure rather assigns weights to align pre-treatment trends in the (adjusted) outcome.

Time weights are computed to align pre- and post-treatment periods for untreated units:

$$\left(\widehat{\lambda}_0, \widehat{\lambda}^{sdid}\right) = \underset{\lambda_0 \in R, \lambda \in \Lambda}{\arg\min} \sum_{i=1}^{N_{co}} \left( \lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right)^2 + \zeta^2 N_{co} ||\lambda||^2, \tag{3.3}$$

with $\Lambda = \{\lambda \in R_{+}^{T},\ with\ \sum_{t=1}^{T_{pre}} \lambda_t = 1\ and\ \lambda_t = 1/T_{post}\ \forall\ t = T_{pre}+1, ..., T\}$, where the regularization term ensures unique weights and is very small.

### 3.5.1 Handling covariates

We follow the procedure for handling covariates outlined in Arkhangelsky et al. (2021) and refined in Clarke et al. (2023). Handling covariates in this setting is treated as a pre-modeling approach, in which the outcome variable is adjusted by covariates before estimation. The procedure does not put any stationarity requirements on the covariates, i.e., they can be time-varying. This adjustment procedure contains two steps. In the first step, we estimate the coefficients of the covariates. To obtain estimates that are unconfounded by the treatment itself, we follow Kranz (2022) and exclude the treated unit in the estimation. We run the following model:

$$Y_{it}^{co} = \alpha_i + \gamma_t + X_{it}^{co}\beta + u_{it}, \tag{3.4}$$

where the super-script *co* indicates control units, $Y_{it}^{co}$ measures $CO_2$ emissions from road transport, $X_{it}^{co}$ collects covariates and may include daily COVID cases, the number of commuters, and the number of persons usually working from home, fuel prices, freight transportation, and GDP per capita. To capture differences between regions and time, we can include region-specific effects, $\alpha_i$, and time-specific effects, $\gamma_t$. In a second step, we adjust the outcome variable for the aforementioned effects for all units:

$$\widehat{Y}_{it}^{adj} = Y_{it} - X_{it}\hat{\beta}. \tag{3.5}$$

Finally, the SDID procedure is then applied to the adjusted outcome variable.

### 3.5.2 Placebo inference and event-study analysis

Arkhangelsky et al. (2021) show that the estimated ATT, $\widehat{\tau}^{sdid}$, is asymptotically normal. This means that conventional confidence intervals can be used to conduct asymptotically valid inference if the asymptotic variance, $\widehat{V}_\tau$, can be consistently estimated: $\tau \in \widehat{\tau}^{sdid} \pm z_{\alpha/2}\sqrt{\widehat{V}_\tau}$. Arkhangelsky et al. (2021) propose several estimators for the asymptotic variance (bootstrap, jackknife, placebo). But in cases where there is only one treated unit (i.e., $N_{tr} = 1$), only placebo estimates are well defined. The idea of this procedure is to replace the exposed unit with unexposed units, then randomly assign those units to a placebo treatment and compute a placebo ATT. This is repeated many times to obtain a vector of placebo ATTs. The variance of this vector can then be used to obtain an estimate for the asymptotic variance.

To evaluate the robustness of the results, we perform an event-study analysis, which enables us to study the dynamics of the policy effect and allow us to evaluate the credibility of pre-treatment parallel trends. We follow the discussion in Clarke et al. (2023) on how to compute these estimates manually. In principle, we want to estimate the differences in the outcome variable between treated and the non-treated synthetic control region for each time period $t$. This allows us to evaluate parallel pre-treatment trends by studying whether these differences changed over time prior to the policy adoption. Additionally, we can study the evolution of the treatment over each post-treatment period.

The difference at each time period $t$ is denoted as $d_t$ and given by:

$$d_t = (\bar{Y}_t^1 - \bar{Y}_t^0) - (\bar{Y}_{base}^1 - \bar{Y}_{base}^0), \tag{3.6}$$

where 1 indicates a treated unit and 0 the non-treated synthetic control unit. The first term in brackets calculates the difference in mean $CO_2$ emissions at time period $t$ for treated and control units. The second term in brackets captures the difference between the pre-treatment baseline means of these units. The baseline outcomes are weighted aggregates over pre-treatment periods rather than arbitrarily chosen time periods (as is usually done in DID applications). They are given by:

$$\bar{Y}_{base}^1 = \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} \bar{Y}_t^1 \qquad and \qquad \bar{Y}_{base}^0 = \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} \bar{Y}_t^0,$$

where the time weights, $\hat{\lambda}_t^{sdid}$, come from equation (3.3).

Confidence bands around the estimated $d_t$'s are generated with a placebo-based approach in the following sequence: (i) Exclude the treated unit (in our case Luxembourg) from the sample; (ii) Randomly assign treatment to a unit (from the remaining units, which are all controls units); (iii) Calculate the outcome adjusted for covariates following equations (3.4) and (3.5); (iv) Compute equation (3.6) and store the result; (v) Repeat 2-4 many times (e.g., 1,000 times); and (vi) Obtain the 5% quantile from the sample distribution of the stored results for each time period $t$.

## 3.6 Results

This section reports our main results as well as several robustness checks. We study several model specifications, which are outlined in section 3.6. section 3.7 tests the robustness of the main results. These checks include in-time placebo tests, specifications that exclude some of our controls, fuel-tourism effects, as well as results from a restricted sample. We find that our results are robust against these checks.

We provide results for three different model specifications. The first one does not adjust emissions for covariates; it is based on equation (3.1). The second specification adjusts the outcome variable for COVID-19 related covariates as described in Section 3.5.1. The auxiliary regression is given by:

$$\begin{aligned} log(CO_2/cap)_{it}^{co} =& \alpha_i + \gamma_t + \beta_1 asinh(cases)_{it}^{co} + \beta_2 asinh(nvrwfh)_{it}^{co} + \\ & \beta_3 asinh(wfh)_{it}^{co} + u_{it}, \end{aligned} \tag{3.7}$$

where the outcome variable is the log of road transport $CO_2$ emission per capita. It is regressed on the inverse hyperbolic sine (*asinh*) of COVID *cases*, on people usually working from home ($wfh$) with their work-place location in the associated NUTS 2 region, and on people never working from home ($nvrwfh$) with their work-place location in the associated NUTS 2 region. We use the inverse hyperbolic sine transformation on covariates that include zero-values because the natural logarithm of zero is undefined and the transformation approaches the natural log. This allows us to interpret the estimated coefficients as elasticities under certain assumptions.[6]

The third specification is our main specification and adjusts the outcome variable for additional covariates and is given by:

$$\begin{aligned} log(CO_2/cap)_{it}^{co} =& \alpha_i + \gamma_t + \beta_1 asinh(cases)_{it}^{co} + \beta_2 asinh(nvrwfh)_{it}^{co} + \\ & \beta_3 asinh(wfh)_{it}^{co} + \beta_4 log(gdp)_{it}^{co} + \beta_5 ei_{it}^{co} + \\ & \beta_6 diesel_{it}^{co} + \beta_7 petrol_{it}^{co} + \beta_8 log(frt)_{it}^{co} + u_{it}. \end{aligned} \tag{3.8}$$

The set of covariates that we consider in this specification additionally includes: the log of real GDP per capita, (*gdp*), energy intensity, (*ei*), measured as average $CO_2$ emissions of newly registered vehicles, (*diesel*) and (*petrol*) prices in real terms (adjusted with the harmonized index of consumer prices - HICP) to capture cross-unit variations in fuel prices, and the log of freight transport ($frt$), measured as tons of goods loaded in the region, to control for changes in freight transport. Estimation results for the auxiliary regressions based on Specifications (3.7) and (3.8) are shown in Table C.1 in Appendix D.1.

We provide estimates of the ATTs for the periods that the treatment is in effect, i.e., 2020-2021, as well as an event-study analysis over the period 2016-2021 in Figure 3.7 for the three different specifications. Estimates for the ATTs are shown in Figure 3.7a and the event-study estimates are shown in Figure 3.7b. Estimates are based on the following model specifications that differentiate in the way they adjust the outcome variable. 1) not adjusting for covariates - no covariates, 2) adjusting only for COVID-19 related covariates - adj COVID covariates, and 3) adjusting for the full set of covariates - adj all covariates. The latter specification produces our main results. The time weights for this variant are assigned to 2018 and 2019 with weights of 0.3348 and 0.6652, respectively. Figure 3.7b shows no

[6] As suggested by Bellemare and Wichman (2020), we multiply these covariates by a constant to generate average values greater than 10, which provides stable elasticities.

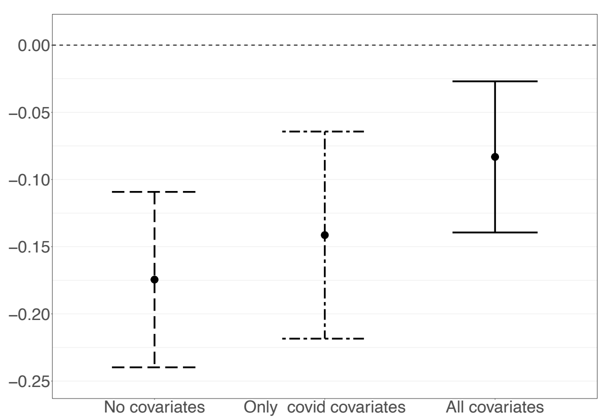
statistically significant violation of pre-treatment trends.

The estimated ATTs for the specification including all covariates indicate an effect at around $-0.083$, i.e., a 8.3% reduction in transport $CO_2$ emissions as a response to the free-public transport policy implemented in March 2020. This is less in magnitude compared to controlling only for COVID-19 related covariates, which yields an estimated ATT of around $-11.8\%$. The specification with no covariates provides the largest estimated ATT at almost $-15\%$. All estimates are statistically significant at the 5% significance level.

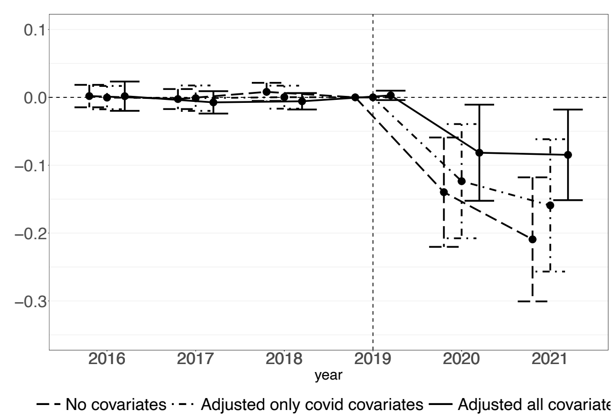
The event-study analysis shows no violation of parallel pre-treatment trends for all specifications. This also indicates that the tram extension in 2017 did not significantly alter Luxembourg's emissions trajectory compared to our synthetic control. Post-treatment effects show statistical significance in 2020 for all three specifications. In 2021, the confidence intervals based the specifications that adjusts the outcome variable for all covariates slightly cross the dashed zero-line at the 5% significance level.

**Figure 3.7:** ATTs and event study estimates

**(a)** ATTs since treatment in 2020




**(b)** Event study estimates for 2016-2021




Note: ATTs and event study estimates of the impact of free public transport on road emissions ($CO_2$) per capita in Luxembourg for different model specifications with 95% confidence bands based on placebo estimates.

The control units that contribute to the synthetic control together with their respective weights for the third specification are graphically shown in Figure C.1 in Appendix D.2. The regions with the largest weights come from Belgium, Denmark, Spain, Finland, Italy, Netherlands, and Portugal. In addition, Austria, Czechia, Greece, France, Hungary, Ireland, and Latvia receive weights. Belgium, Denmark, Finland, and the Netherlands are among the EU countries with the highest GDP per capita and thus most comparable to Luxembourg in this respect. While Italy is among EU countries with the highest motorization rate after Luxembourg. It is therefore quite reasonable that the regions contributing to the synthetic control are taken from these countries.

Figure C.2 in Appendix D.2 shows how well the SDID estimation aligns pre-treatment trends for Luxembourg and its synthetic control. Luxembourg is shown as a solid line and the weighted average across control regions according to the SDID unit weights as a dashed line. The figure also shows the average pre-treatment trend in the adjusted outcome variable over all regions and the unweighted average over regions that received a positive weight. Figure (a) shows the absolute level of trends, while Figure (b) standardizes the trends so that they are visually more easily comparable.[7]

The absolutes levels of the adjusted outcome differs markedly between Luxembourg and the different controls. This reinforces our argument that the SDID procedure is preferable over standard DID and SC methods. We can see from the standardized trends in part b of the figure that pre-treatment trends for Luxembourg and the average across all regions shows the biggest visual difference in trends. The unweighted average across regions that received a positive weight is a much better fit. The best fit seems to be between Luxembourg and the weighted average according to the SDID unit weights.

## 3.7 Robustness

The credibility of the SDID estimator depends on its ability to reproduce a counterfactual outcome for Luxembourg in the absence of the free public transport policy. In this section, we conduct standard robustness tests commonly used for synthetic controls, an *in-time placebo test*, where the policy is backdated to a fictitious date, as well as *leave-one-out placebo tests* to assess the sensitivity of the synthetic control to the composition of the donor pool (Abadie, 2021). Additionally, we examine the robustness of our results to different model specifications. Finally, we apply the SDID method to $CO_2$ from energy use in the building sector to assess whether there was an effect attributable to COVID-19. Finally, we extend our post-treatment period by one year and discuss the results of this analysis.

**In-time placebo**: We perform an in-time placebo (also referred to as back-dating test) as suggested by Abadie (2021). Here, we assign the free public transport policy to 2019, the year before its actual introduction. Since the treatment is artificially assigned to a prior date we should not observe a significant post-placebo treatment effect. Figure C.1 in Appendix C.3 shows these results. The solid black line represents our main specification with all covariates, and the dot-dash line represents the specification without covariates.[8]

[7]Standardization is performed by subtracting the mean and dividing by the standard deviation within each group.

[8]We do not estimate the specification adjusted only for COVID-19 covariates since the policy is back-dated before the pandemic.

The confidence bands at the 5% significance level encompass the zero line, indicating no significant treatment effect in 2019.

**Leave-one-out placebo**: We use a donor pool of 137 NUTS 2 regions in our analysis. To assess the sensitivity of our results, we conduct a leave-one-out robustness check by iteratively excluding one region at a time, re-estimating the SDID model, and obtaining a distribution of ATT estimates. The resulting distribution is presented in Figure C.2a in Appendix C.3. The estimated ATTs from this exercise range from -0.085 to 0.081, with our main estimate of -0.083 positioned near the center of the distribution. These estimates are not statistically different from our main result, indicating that our findings are robust to the exclusion of individual regions from the donor pool.

Next, we extend this robustness check by iteratively excluding one country at a time. Since the 137 NUTS 2 regions in the donor pool in our sample span 18 countries, this approach removes multiple regions at once. The resulting ATT estimates, plotted in Figure C.2b in Appendix C.3, range from -0.0962 to -0.0793. The results exhibit greater sensitivity compared to the leave-one-NUTS 2 region at a time analysis, as dropping an entire country removes a substantial number of regions simultaneously. Nonetheless, our main estimate of -0.083 remains centrally located within the distribution. The most pronounced deviations occur when we exclude Italy (-0.0962) and the Netherlands (-0.0902). Dropping Italy removes 21 NUTS 2 regions, while excluding the Netherlands removes 12 NUTS 2 regions. Notably, both Italian and Dutch regions receive high weights in our main specification (see Appendix C.1). The fact that our estimates shift in a direction that strengthens our main result suggests that, if anything, our primary findings are conservative.

**Restricted sample:** We also conduct our analysis on a more restricted donor sample by excluding regions that introduced any form of public transport subsidy affecting specific segments of the population, as described in Section 3.3. We further exlcude Torrevieja in Spain, Livigno in Italy, Attica in Greece, and Nantes, Strasbourg, and Paris in France, all of which introduced some form of free public transport for residents and/or students ("City Public Transport Information," 2024). We also exclude all Austrian regions due to the nationwide climate ticket introduced in 2021, which increased accessibility and significantly reduced prices for comparable tickets. These results are reported in Figure C.3 in Appendix C.3. Part (a) of the figure shows the estimated ATTs of our three specifications. The specification that includes all covariate adjustments estimates the ATT at $-0.06$, statistically identical to our main results. Part (b) of the figure shows the corresponding event-studies. Again, the trajectories and confidence bands are visually indistinguishable from the ones based on the larger sample.

**Alternative specifications:** To evaluate the robustness of our main results in Figure 3.7, we explored sensitivity across alternative model specifications. Given that our measures for people working from home and those commuting to work likely capture similar dynamics[9] to a certain degree, we test the sensitivity of our results by excluding one or the other from our specifications. Additionally, Table C.1 shows that the coefficient for $log(frt)$ (log of freight transport) is statistically insignificant. Consequently, we estimate the following specifications, each excluding different combinations of these covariates: a model excluding controls for freight transport (Spec 1), a model omitting controls for working from home (Spec 2), a model excluding both freight transport and working from home (Spec 3), a model excluding the commuting variable, $nvrwfh$ (Spec 4), and a model excluding both the commuting variable and freight transport (Spec 5). The results of these sensitivity analyses are displayed in Figure C.4 and Table C.1 in Appendix C.3. All five alternative specifications yield estimates similar to our main specification.

**Fuel tourism:** Luxembourg's lower fuel prices compared to neighbouring regions can attract fuel tourism, which can then lead to increased fuel consumption and higher emissions. This effect would be unrelated to the free public transport policy and confound our estimates. We already control for absolute fuel prices in our main specification, which should capture this effect to some degree. Arguably fuel tourism is more adequately accounted for by fuel prices of Luxembourg relative to its neighbours. Figure C.5 in Appendix C.3 compares both absolute and relative fuel prices between Luxembourg and its neighbouring regions. Throughout our sample period, Luxembourg's absolute fuel prices are consistently lower than those of its neighbours, resulting in relative prices below one. To test the robustness of our estimates, we re-estimate our main specification incorporating relative fuel prices, calculated as the fuel price of a NUTS 2 region relative to the mean of its neighbours that are not part of the same country. The estimated ATT is $-0.0839$ and is statistically indistinguishable from our main result ($-0.0832$). Similarly, the event-study estimates align closely with our main results. We attribute this consistency to several factors. First, absolute fuel prices may partly reflect the effects of relative prices. Second, the relative fuel price in Luxembourg remained below one throughout the sample period, maintaining an incentive for fuel tourism. Third (and arguably most importantly), the estimated ATT is based on a comparison between weighted averages of the pre-and post-treatment periods. As shown in Table C.2 in Appendix C.3, there is no significant difference between these weighted averages for diesel and petrol prices in Luxembourg relative to its neighbors.

**Energy for buildings**: While the pandemic undeniably led to temporary reductions in mobility and emissions globally, our SDID approach inherently accounts for this, as the donor pool is composed of NUTS 2 regions also affected by COVID-19. If COVID-19 were

[9] They show a moderate correlation of around 0.6.

the primary driver of the observed emission reductions, we would expect similar declines in road emissions in synthetic Luxembourg as well, yet this is not what we find. Moreover, we conducted a placebo test using $CO_2$ from energy use in buildings, which should also have been affected by pandemic-related shifts in energy demand (e.g., increased residential electricity use due to lockdowns and remote work). If COVID-19 were driving a broad reduction/increase in emissions, we would expect to see an effect in this sector as well. However, we find a null effect, suggesting that the reduction in road $CO_2$ is not merely a byproduct of the pandemic. This strengthens the argument that the free public transport policy itself played a causal role in reducing $CO_2$ rather than reflecting a general COVID-19 induced effect. The results of this analysis is reported in Figure C.6 in Appendix C.3.

**Extending the post-treatment period to 2022**: We extend the post-treatment period by one year to 2022, with the results presented in Figure C.7 in Appendix C.3. When including this additional year, we observe a further reduction in road transport $CO_2$ emissions, leading to an increase in the estimated ATT for the post-treatment period. This result is expected, as Luxembourg's free public transport policy is an ongoing intervention rather than a short-term measure, allowing individuals to gradually adjust their travel behavior over time. However, we do not adopt this as our main specification for three reasons. First, the 2022 EDGAR data are still estimates and subject to revision. Second, COVID-19 case reporting in 2022 was inconsistent, with most regions ceasing to report cases after September, making it difficult to control for pandemic-related effects. Finally, Luxembourg made additional investments in public transport infrastructure in 2022. This makes it more challenging to isolate the effect of the free public transport policy from the broader improvements in public transit accessibility.

## 3.8 Discussion

We now discuss the estimated effect size of Luxembourg's free public transport policy. We attribute the estimated ATT of $-8.3\%$ to a modal shift from private motorized transport to public transport and ask whether this estimated effect size is reasonable.

Consider the following back-of-the-envelop calculation. Following figures from the European Commission and Directorate-General for Mobility and Transport (2021), we assume a modal split between private vehicles and public transport of 82 and 18%, respectively. We further assume that the observed reduction in $CO_2$ emissions results from a modal shift from private vehicles to public transport. An 8.3% reduction in $CO_2$ emissions from road transport then implies a corresponding decrease in private vehicle usage by approximately 6.8%.

This decrease is derived from the fact that private vehicles represent 82% of the modal split and thus contribute the majority of emissions reductions (calculated as 82% of the 8.3% reduction). To maintain the overall transport capacity, public transport usage must increase by approximately 38%, calculated by dividing the reduction in private vehicle usage (6.8%) by the initial share of public transport (18%).

To assess the credibility of this effect size, we utilize data on the average daily number of people using trams on weekdays from the OECD (2023). In February 2020, this average tram usage was at around 31,000 persons. This increased to around 36,000 in February 2021 and to around 53,000 in February 2022. This amounts to an increase of around 16% and 47% from 2020-2021 and 2021-2022, respectively. These numbers align with our estimates, suggesting that our effect size is reasonable.

Additionally, we can relate these results to the LUXmobile survey, conducted by the Luxembourg City Council (Luxmobile, 2020). This survey reports that the free public transport policy has led to an average increase in public transport usage of around 34% and a 38% increase among residents in 2022, further adding credibility to our estimate. While the descriptive analysis does not direclty validate the causal estimates, the observed figures are consistent with our estimated effect size, lending further credibility to our findings.

Further, we calculate the associated marginal abatement cost of carbon for the policy as the government expenditure per ton of $CO_2$ abated. A simple calculation takes the foregone revenue from ticket sales of around 41 Mio. Euros and compares it to the tons of $CO_2$ emissions abated according to our estimates. The latter are calculated as the counterfactual post-treatment emissions for Luxembourg: $\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} CO2_t^{tr} / (1 - \hat{\tau})$, where $tr$ indicates the treated unit. With this back-of-the-envelop calculation, we estimate a marginal abatement cost of EUR 114 per ton of carbon. This is, of course, a crude estimate and does not capture the full costs nor the additional non-$CO_2$-benefits of the policy. As Hahn et al. (2024) argue, such calculations overlook the benefits to inframarginal individuals—those who do not alter their behavior in response to the policy—thereby potentially underestimating the policy's overall effectiveness. They suggest a more comprehensive approach, the Marginal Value of Public Funds (MVPF) framework, which captures these benefits and provides a more accurate assessment of the policy's impact. We leave such detailed calculations to future research.

Finally, we attempt to reconcile our findings with other studies, particularly those examining the German 9-Euro ticket. This short-term policy was implemented in Germany for three months, from June to August 2022. Its temporary nature may explain why the 9-Euro ticket did not result in a substantial shift away from car usage (Liebensteiner et al., 2024). In contrast, Luxembourg's free public transport policy was introduced as a long-term measure with no specified end date, potentially allowing for more enduring impacts on travel behavior and emissions, as documented in our study. Further, recent work by researchers at the Mercator Institute of Global Commons and Climate in Germany examines the impact of the German 49-Euro ticket, a policy introduced in May 2023 that remains in effect today. Their findings also document a significant shift in traffic from road to rail (Koch et al., 2024). Together, these results emphasize that long-term measures are essential for systematically and meaningfully changing individual behavior and reducing emissions.

## 3.9 Conclusion

We estimate the causal impact of Luxembourg's 2020 free public transport policy on road transport emissions and find a 8.3% reduction in $CO_2$ emissions. Our analysis remains robust across various models that consider the effects of COVID-19, fuel prices, and commuting patterns. It is further validated through placebo tests, sample restrictions, and fuel tourism analyses. Our findings hold high policy relevance, particularly for policymakers in urbanized, affluent areas with robust public transport networks like Luxembourg. Demonstrating the policy's effectiveness in reducing $CO_2$ emissions, our study highlights the potential of integrating free public transport into comprehensive sustainable transport and urban planning initiatives to meet climate targets and foster a sustainable future.

# Chapter 4

# The Development of Austrian Greenhouse Gas Emissions since 2021

Greenhouse gas emissions in Austria in 2023 were 14% below 1990 levels, matching those last observed in 1970. Particularly strong decreases occurred in 2022 and 2023, with emissions falling by 5.8% and 6.4%, respectively. The buildings sector in 2023 was over 50% below its 1990 baseline. It experienced a 20% drop in 2023, with 0.7 percentage points explained by a milder winter and the remainder driven by an increased share in renewables. Two-thirds of this uptake can be traced to high energy prices since 2021. Emissions in remaining sectors declined by 4.9% in 2023, with weak economic performance contributing 0.86 percentage points and the majority attributed to a higher share of renewables, around 60% of which can be explained by rising energy prices since 2021. A hypothetical scenario, assuming average economic conditions and winter temperatures, indicates that emissions would have been lower than the ones observed in 2021 and 2022 but slightly higher in 2023.

## Summary

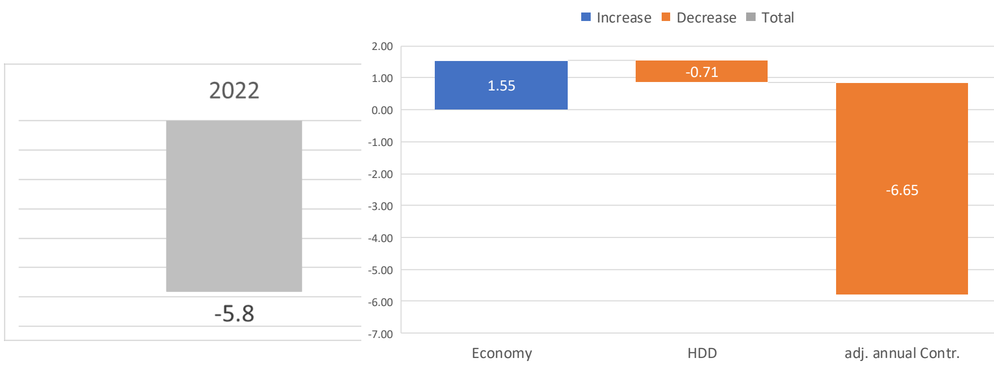
This report analyzes greenhouse gas emissions (GHG) in Austria since 2021. Greenhouse gas emissions have decreased by approximately 14% compared to 1990, with particularly strong reductions achieved in the past two years. In the buildings sector, emissions have recently been more than 50% lower (equivalent to a reduction of 7 million tons of $CO_2$-eq), while emissions in other sectors are now 22% below their 2005 peak (17.6 million tons of $CO_2$-eq lower) and, for the first time since 1990, significantly below 1990 levels by 6% (3.8 million tons of $CO_2$-eq). Emissions have now fallen to levels last seen in 1970.

The following sections present developments in (a) the buildings sector and (b) all other sectors. In the buildings sector, emissions in 2023 decreased by 1.5 million tons of CO2 equivalents compared to the previous year, corresponding to a 20.2% reduction in emissions from the sector. Of this reduction, an estimated 0.7 percentage points can be attributed to a milder winter (measured using population-weighted heating degree days observed regionally across Austria). The remaining reduction is structural, with 17 percentage points attributed to the transition in heating systems, i.e., an increased share of renewable energy sources. Rising energy prices since late 2020 account for approximately 67% of this increase in renewables, serving as a key driver of the heating system transition. In 2022, emissions decreased by nearly 17% compared to the previous year. This was a significantly warmer year than 2021, with 7 percentage points of the reduction attributed to the milder winter and 6 percentage points to the higher share of renewables. Other broad factors likely reflect year-specific issues and initiatives, such as those prompted by the intensified energy crisis following Russia's invasion of Ukraine, including temperature reductions in public buildings and reductions in gas demand, particularly in private households.

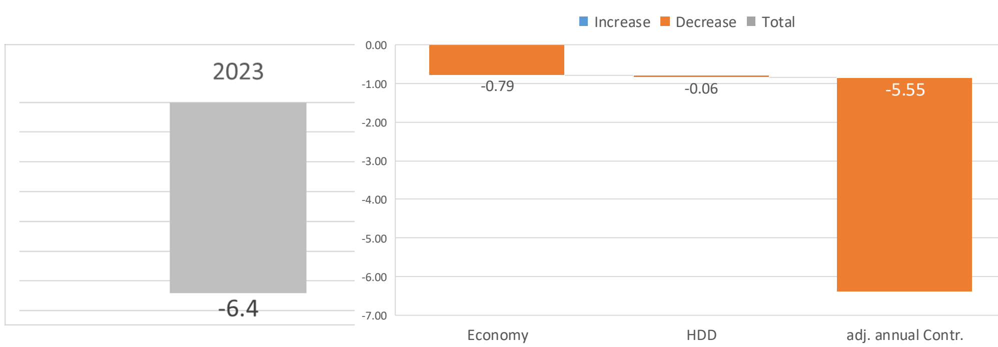
In the remaining sectors (i.e., non-buildings sectors), the energy crisis had significant impacts on emissions. In the first crisis year, 2022, emissions decreased by 4.5% compared to the previous year. Economic activity (GDP) increased significantly in 2022 compared to the 10-year trend, contributing an estimated 1.73 percentage points to emissions growth. The observed emissions reduction was supported by the increased share of renewables, which accounted for a reduction of 1.04 percentage points. Most of the emissions reduction is attributed to significant changes in energy markets and energy policy triggered by Russia's invasion of Ukraine, including sharply increased fossil fuel prices (which in the short term led primarily to an overall reduction in energy demand) and changes in fuel price structures relative to neighboring countries, which reduced fuels exported in fuel tanks, i.e. fuels tanked in Austria but driven abroad (and thus reduced transport emissions recorded for Austria).

In the following year, 2023, emissions in non-buildings sectors decreased by approximately 3 million tons of CO2 equivalents (-4.9%) compared to the previous year. Weak economic performance relative to the long-term trend contributed 0.86 percentage points to the reduction, while the increased use of renewable energy sources could explain a nearly 6 percentage point reduction over the long term. However, the absence (or diminished presence) of the exceptional factors observed during the 2022 energy crisis led to a slight rebound in emissions relative to that year, resulting in a net observed reduction of "only" 4.9%. The 2023 increase in the share of renewables (+15%) can be attributed to rising energy prices since 2020, explaining approximately 60% of the increase. Additionally, a range of other measures, including energy-saving initiatives during the energy crisis, building renovation programs, and initiatives in public transport and cycling infrastructure, were fundamental to emissions reductions.

Regarding the total emissions for the two years, as shown in Figure Z.1, strong economic performance in 2022 (panel a) caused emissions to increase by 1.55% relative to the previous year, while the comparatively warm winter reduced emissions by 0.71%. After adjusting for deviations in economic activity and heating degree days from long-term averages, the annual adjusted contribution shows a 6.65% emissions reduction, higher than the net reduction of 5.8%. In 2023 (panel b), the total observed emissions reduction (6.4%) was attributable to both weak economic performance (0.79 percentage points) and the even warmer winter (0.06 percentage points). After adjustment for those two factors, the annual contribution amounts to an emissions reduction of 5.55%.



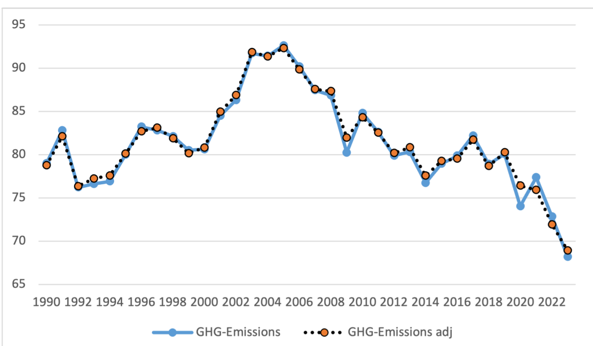

**(a)** Change in GHG Emissions Compared to the Previous Year [%] - Adjusted for Economic and Heating Degree Day (HDD) Deviations



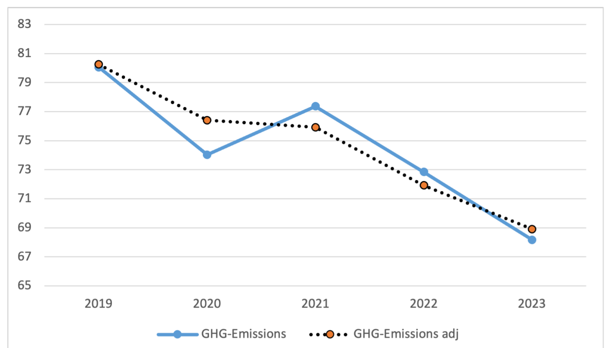

**(b)** Change in GHG Emissions Compared to the Previous Year [%] - Adjusted for Economic and Heating Degree Day (HDD) Deviations

**Figure Z.1:** Adjustments to GHG emission changes compared to the previous year for economic and heating degree day (HDD) deviations from long-term averages for 2022 and 2023.

An alternative perspective estimates what GHG emissions would have been in a year with average economic activity and heating degree days (and thus winter temperatures) at their long-term levels. This approach calculates hypothetical emissions under such assumptions and compares them to actual emissions. Figure Z.2 shows the results of this analysis for 1990–2023 (panel a), with a specific focus on recent years (panel b). It reveals that the sharp economic downturn in 2020 significantly contributed to emissions reductions (to a much greater extent than the slightly warmer winter), meaning the adjusted emissions in that year would have been significantly higher. In contrast, both above-average economic growth and a cold winter in 2021 led to higher emissions than the adjusted (hypothetical) levels for a year with average conditions. In 2022, strong economic activity pushed emissions above their adjusted hypothetical level, whereas in 2023, both weak economic activity and a warm winter contributed to emissions reductions, resulting in adjusted emissions being higher than the observed levels.




**(a)** GHG Emissions Austria 1990-2023 Adjusted for Economic and Heating Degree Day Deviations




**(b)** GHG Emissions Austria 2019-2023 Adjusted for Economic and Heating Degree Day Deviations

**Figure Z.2:** Hypothetical emission levels assuming average economic development and winter temperatures (sum of heating degree days, population-weighted).

## 4.1 Introduction

This memo examines the development of greenhouse gas (GHG) emissions in Austria in the years 2023 and 2022 compared to the respective previous year. Data for 2023 are sourced from the UBA Nowcast and the preliminary energy balance for Austria. Emissions decreased significantly in 2023 and 2022 relative to the respective prior year. Figure 4.1 shows the trend of emissions from 1990 to 2023 across all sectors. Emissions in 2023 compared to 2022 decreased by 6.4%, and in 2022 compared to 2021, by 5.8%. We aim to determine the factors responsible for these reductions. We illustrate our method using the emission changes in 2023 and then apply it to those in 2022. Here, we differentiate between emissions from the buildings sector and all other emissions. The latter includes greenhouse gases from the energy, industry, transport, waste management, F-gases, and agriculture sectors. GHG emissions from the buildings sector decreased by 20% in 2023 compared to 2022 and by 17% in 2022 compared to 2021. In the remaining sectors, emissions decreased by 4.9% (2023 vs. 2022) and 4.4% (2022 vs. 2021).

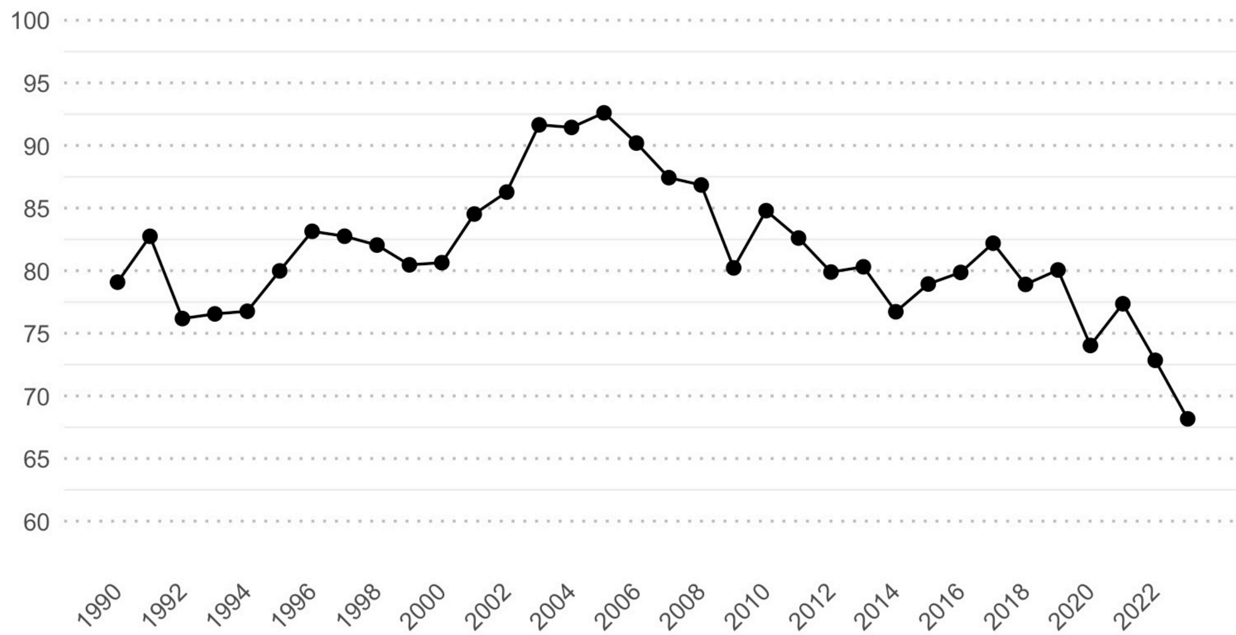



**Figure 4.1:** Austria's GHG emissions, in million tons $CO_2$-eq. (1990–2023)

We are particularly interested in estimating the effect of structural changes and separating it from other effects, such as the impacts of mild winters or changes in overall economic activity (e.g., recession). In the buildings sector, emissions are influenced by reduced heating demand in mild winters. Structural effects may arise from the share of renewable energy sources, which, in turn, can increase due to higher fossil fuel prices. Emissions in the remaining sectors are linked to the level of economic activity and the share of renewable energy, which is also affected by price effects.

To analyze these factors, we use econometric time-series methods to explain greenhouse gases through various explanatory factors. For our estimations, we use annual data from 1990 to 2022. The estimated relationships based on this data are applied to the changes in 2023 to determine the explanatory power of the different factors for the emissions reduction. The interpretation of such models always assumes that one explanatory variable changes while all others remain constant. Therefore, we proceed in two steps for both emission areas.

In the buildings sector, we first estimate the relationship between emissions, heating degree days, and the share of renewables. We use the share instead of absolute values to allow for changes in absolute consumption. For emissions from the remaining sectors, we proceed similarly and first estimate the relationship between emissions, economic growth, and the share of renewables. In a second step, we estimate the effect of energy prices on the share of renewables for both areas.

We discuss the results in the following sections. A formal representation of the estimated models is provided in Appendix D.1, and detailed results and model quality tests are presented in Appendix D.2.

## 4.2 Emissions from the Buildings Sector

This analysis focuses on emissions covered by the Climate Protection Act (Klimaschutzgesetz, KSG), meaning emissions outside the EU Emissions Trading System (ETS) that are attributable to the buildings sector. Figure 4.2 shows the emissions trend in the buildings sector from 1990-2023. Buildings sector emissions decreased by over 50% between 1990 and 2023. The largest reductions were achieved in 2022 and 2023. In 2022, emissions decreased by nearly 17% compared to 2021 (1.5 million tons of $CO_2$-eq). According to the UBA Nowcast, emissions in 2023 compared to 2022 fell again by 1.5 million tons of $CO_2$-eq, amounting to a reduction of approximately 20%.

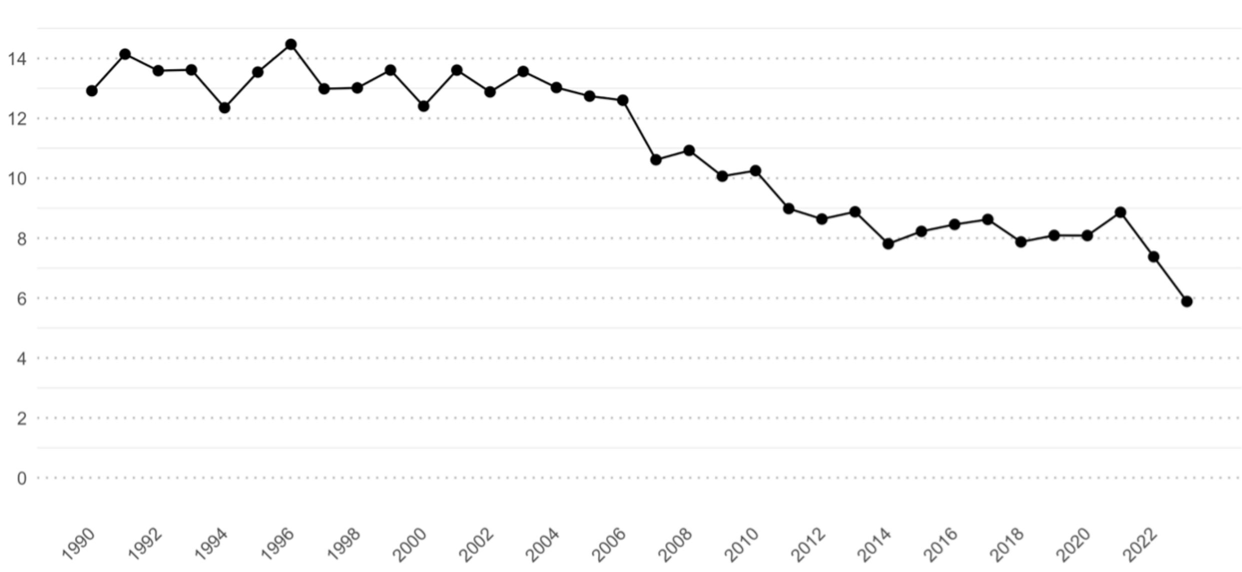



**Figure 4.2:** Austria's GHG Emissions in the Buildings Sector, in million tons of $CO_2$-eq. (1990–2023)

The substantial reduction in emissions can be explained by the development of several key indicators related to buildings emissions. These are summarized in Table 4.1. Emissions from heating and hot water in the buildings sector are primarily attributed to the consumption of fossil fuels. For example, in 2023, according to the preliminary energy balance of Statistics Austria (2024a), oil and gas consumption significantly decreased. Oil and gas consumption dropped by around 21%, and coal by approximately 16%.

District heating and electricity are recorded with zero emissions in the buildings sector since they are attributed to the energy sector. Consequently, these energy sources do not contribute to emissions in the buildings sector and are considered entirely non-fossil (i.e., renewable). By this definition, renewable energy consumption in 2023 increased slightly by approximately 1.5%. Although the absolute increase in renewable energy consumption was minor, the share of renewables increased significantly, rising by 4.9 percentage points from 71% in 2022 to 76% in 2023, representing a total increase of 6.9%. The sharp decline in fossil fuel consumption can be attributed to several factors.

In 2023, Austria experienced an unusually mild winter. As a result, heating demand likely decreased due to warmer temperatures during the winter months. Heating demand can be measured using heating degree days (HDD), defined as the sum of daily differences between indoor temperature (20°C) and outdoor mean temperature on days where the outdoor mean temperature falls below 12°C. Heating degree days in 2023 decreased by approximately 1% compared to the already mild previous year.

**Table 4.1:** Changes 2021–2023

| Year | GHG (%) | HDD (%) | Energy Price (%) | Oil (%) | Gas (%) | Coal (%) | Renewables (%) | Renewables Share (pp) |
|---|---|---|---|---|---|---|---|---|
| 2021 | 9.6 | 9.8 | 7.7 | 8.4 | 10.2 | -15.3 | 11.5 | 0.4 |
| 2022 | -16.8 | -10.2 | 36.8 | -13.9 | -18.3 | -26.6 | -9.8 | 1.7 |
| 2023 | -20.2 | -1.0 | 16.5 | -21.4 | -20.7 | -15.6 | 1.5 | 4.9 |

*Note:* GHG from Buildings sector. Rounded to one decimal digit. Changes compared to the previous year. Changes in the share of renewables are given in percentage points (pp); all other changes are percentages.

Another reason for the reduction may be lower consumption due to price increases. According to the energy price index (base year 1986) from Statistics Austria (2024b), energy prices for space heating and hot water increased by approximately 16% in 2023 relative to 2022. Price effects can also have a delayed impact. Energy prices increased by nearly 37% in 2022 relative to 2021, which likely affected heating behavior or led to a transition to renewable energy effective from the following year (2023) onwards.

### 4.2.1 Heating Degree Days and Share of Renewables

We focus on differences (annual changes) in the variables. Changes in greenhouse gas emissions are explained by changes in heating degree days and the share of renewables. The results of our regression indicate an elasticity of emissions to heating degree days of 0.69. This can be interpreted as follows: A 1% reduction in heating degree days leads to a 0.69% reduction in GHG emissions on average, keeping other factors constant. For the share of renewables, we estimate that a one-percentage-point increase in the share of renewables reduces GHG emissions by 3.5%.

GHG emissions in the buildings sector decreased by 20.2% in 2023 compared to the previous year. Heating degree days decreased by 0.98%, and the share of renewables increased by 6.9% or 4.9 percentage points. According to our estimates, the reduction in heating degree days accounts for a 0.68 percentage point reduction in GHG emissions (equivalent to 3.4% of the total emissions reduction). The increase in the share of renewables accounts for a reduction of approximately 17.1 percentage points. Our results are summarized in Table 4.2.

**Table 4.2:** Explanation of GHG Reduction in the Building Sector 2023 Compared to the Previous Year

| | |
|---|---|
| Reduction in GHG Emissions Building Sectors | 20.2% |
| Contributing Factors: | |
| Heating Degree Days | 0.68 |
| Share of Renewables | 17.1 |

*Note:* GHG emissions in the buildings sector decreased by 20.2% in 2023 compared to 2022. Heating degree days contributed to a reduction of 0.68 percentage points, and the increase in the share of renewables accounted for 17.1 percentage points.

### 4.2.2 Price Effect

For the relationship between the share of renewables and energy prices, we estimate an error correction model. This model assumes a long-term equilibrium relationship between the two variables. If such a relationship exists, it allows for more efficient estimation while also capturing dynamic relationships. Statistical tests in Table D.3 in Appendix D.2 suggest the presence of such a long-term relationship. Figure 4.3 shows the estimated dynamic effect of a change in energy prices over time based on the error correction model. The x-axis represents time periods in years, and the y-axis represents the estimated effect of a price change. The solid black line indicates the estimated effects, and the statistical confidence intervals are shown in grey around this line. Whenever this grey area lies above the zero line, the estimated effects are statistically significant.

We observe that a price increase in the same period (time = 0) has a small and statistically insignificant effect. However, the effect one period into the future is much stronger and statistically highly significant. The effect diminishes in subsequent periods and eventually trends toward zero.

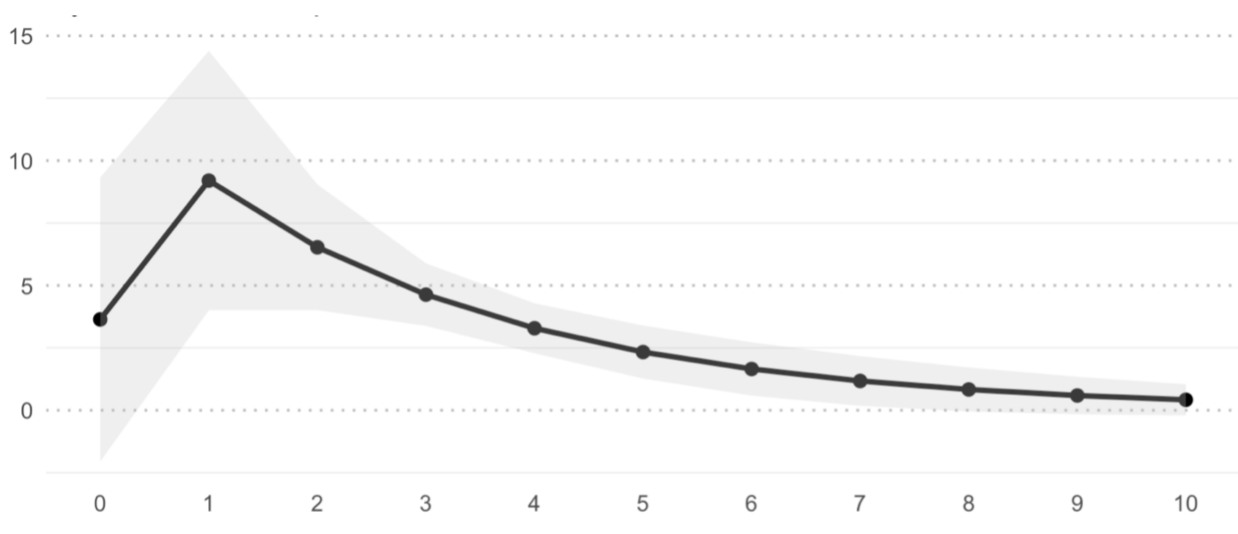



**Figure 4.3:** Dynamic Multipliers for Energy Price Index

**Table 4.3:** Explanation of Renewable Energy Increase in 2023 Due to Energy Price Index Rise

| Period | Contribution to Increase (%) |
|---|---|
| Since January 1, 2022 | 59.6 |
| Since January 1, 2021 | 67.1 |

*Note:* The share of renewables increased by 4.9 percentage points (+6.9%) in 2023 compared to the previous year. This increase can be attributed to the rise in the energy price index since January 2022 by 59.6%. Considering price increases since January 2021, the contribution is 67.1%.

To interpret these estimated effects, it is useful to consider the following scenario. Energy prices increased by approximately 16% in 2023 compared to 2022. This leads to a contemporaneous increase in the share of renewables by 0.6 percentage points, explaining 9% of the increase in renewables in 2023. However, this effect is statistically insignificant.

Energy prices increased by approximately 37% in 2022 compared to 2021. Considering that this price increase also affects the share of renewables in 2023, we can already explain 60% of the increase in 2023. Energy prices rose by approximately 7% in 2021 compared to 2020. Including this effect for 2023, price increases since January 1, 2021, explain about 67% of the increase in the share of renewables. These effects are summarized in Table 4.3.

### 4.2.3 Hypothetical Emissions in a Year with an Average Winter

In 2023, Austria experienced a mild winter compared to the long-term average (2014–2023). We now ask: What would emissions have been in 2023 with an average winter? To answer this, we calculate the mean heating degree days (HDD) over the past 10 years (2014–2023), resulting in an average of 3,792 HDD. In 2023, there were 3,602 HDD. Emissions in 2023 with an average winter would therefore have been slightly higher. Using the previously estimated elasticity (around 0.7), we calculate an emissions increase of approximately 3.5% (+0.2 million tons). Similarly, 2022 also recorded a warm winter compared to the long-term average, with 3,638 HDD. For 2022, we calculate an emissions increase of approximately 2.8% (+0.21 million tons) with an average winter. These calculations are summarized in Table 4.4.

**Table 4.4:** Hypothetical Emissions with an Average Cold Winter

| Year | HDD | Difference to 10-Year Avg. | Hypothetical Emissions Increase (%) | Hypothetical Emissions Increase (Mio. T.) |
|---|---|---|---|---|
| 2023 | 3602 | -189.2 | 3.5 | 0.20 |
| 2022 | 3638 | -153.5 | 2.8 | 0.21 |

*Note:* The average heating degree days (HDD) between 2014–2023 were 3,792. The "Difference to 10-Year Avg." represents the deviation of actual HDD from this average. The hypothetical emissions increase is calculated using the estimated HDD elasticity of 0.7.

## 4.3 Emissions from Other Sectors

Emissions from non-buildings sectors include energy, industry, transport, waste management, F-gases, and agriculture. These emissions are not limited to those covered by the Climate Protection Act (KSG) but also include emissions under the European Emission Trading System (EU ETS), such as those in the industrial sector. The trend in these emissions is shown in Figure 4.4. A significant decline is observed in the last two years, with emissions decreasing by 3.2 million tons (-4.4%) in 2023 and 3 million tons (-4.9%) in 2022.

Emissions from the non-buildings sector are closely linked to the level of economic activity (GDP growth). Following the onset of the COVID-19 pandemic in 2020 and significant decline in economic activity in that year, economic activity increased again. In 2021 and 2022, real GDP (adjusted for prices, in national currency) grew by 4.2% and 4.8% respectively compared to the previous year, while in 2023, GDP declined by 0.8%.

We also consider the share of renewables in non-buildings sectors. Since emissions in these sectors are partially covered by the EU Emissions Trading System (ETS), the share of renewables here also includes district heating and electricity. In 2023, 46.5% of district heating and nearly 90% of electricity were derived from renewable energy sources. The share of renewables increased by 3.77 percentage points in 2023, corresponding to a 15.2% increase. The development of these variables is listed in Table 4.5.

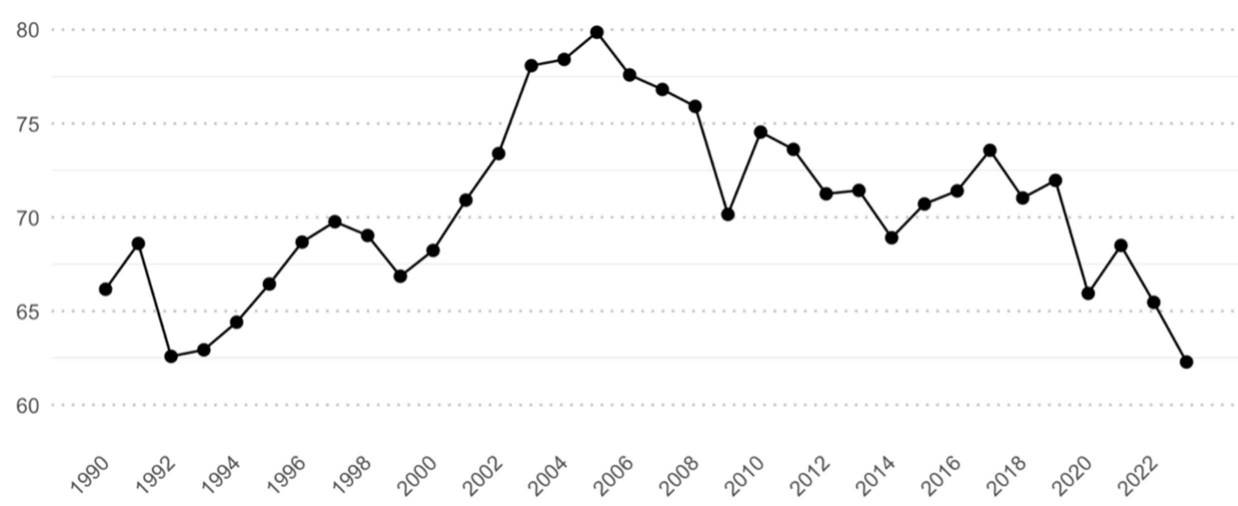



**Figure 4.4:** Austria's GHG Emissions in Non-Building Sectors, in million tons of $CO_2$-eq (1990–2023)

**Table 4.5:** Changes 2021–2023

| Year | GHG (%) | GDP (%) | Energy Price (%) | Oil (%) | Gas (%) | Coal (%) | Renewables (%) | Renewables Share (pp) |
|---|---|---|---|---|---|---|---|---|
| 2021 | 3.9 | 4.2 | 7.7 | 5.2 | 2.1 | 1.9 | -0.1 | -1.08 |
| 2022 | -4.4 | 4.8 | 36.8 | -0.6 | -10.0 | -13.3 | 1.3 | 0.66 |
| 2023 | -4.9 | -0.8 | 16.5 | 0.5 | -18.1 | -5.8 | 11.1 | 3.77 |

*Note:* Rounded to one decimal digit. Changes compared to the previous year. Changes in the share of renewables are given in percentage points (pp); all other changes are percentages. GDP values are adjusted for prices in domestic currency units.

### 4.3.1 GDP and Share of Renewables

We explain GHG emissions in non-buildings sectors through the development of GDP (adjusted for prices, in domestic currency units) and the share of renewable energy sources (including district heating and electricity). GHG emissions decreased by 4.9% in 2023 compared to the previous year. GDP declined by 0.8%. Relative to the average economic growth of the past 10 years (+1.04%), GDP growth in 2023 was 1.84 percentage points below the long-term average. We estimate a GDP elasticity for GHG of 0.46. Thus, the lower GDP growth relative to the long-term average explains an emissions reduction of approximately 0.86% or 0.53 million tons of $CO_2$-eq.

The share of renewables increased by 3.77 percentage points (15.2%) in 2023. We estimate an elasticity of approximately -1.58%, indicating that the increased share of renewables reduces emissions by an estimated 5.95 percentage points. The total GHG emissions did not decline as much, suggesting that other factors had a net emissions-increasing effect. Following the energy crisis in 2022, these factors could include compensatory effects, such as the reversal of extreme energy demand reductions and a return toward pre-crisis levels. The results are summarized in Table 4.6.

**Table 4.6:** Explanation of GHG Reduction in Non-Building Sectors, 2023 Compared to the Previous Year

| | |
|---|---|
| Reduction in GHG Emissions Non-Building Sectors | 4.9% |
| Contributing Factors: | |
| Deviation of GDP Change from Long-Term Avg. | 0.86 |
| Share of Renewables | 5.95 |

*Note:* GHG emissions in the non-building sector decreased by 4.9% in 2023 compared to 2022. A reduction of 0.86 percentage points is attributable to the deviation of GDP growth from the long-term average, and 5.95 percentage points are attributable to the increase in the share of renewables. Other factors not listed here had a net emissions-increasing effect, likely reflecting compensatory effects after the emissions reductions in the 2022 energy crisis, such as the reversal of extreme energy demand reductions and a return toward pre-crisis levels.

### 4.3.2 Price Effect

As in the analysis of the buildings sector, we examine the effect of energy prices on the share of renewables. Cointegration tests in Table D.3 suggest a cointegrating relationship between the two variables. We therefore estimate an error correction model to capture the dynamic effect of prices. The regression results are listed in Table D.1 in Appendix D.2. Figure 4.5 shows this dynamic effect. We observe no statistically significant contemporaneous effect. However, price changes in period 0 have a statistically significant effect in the following period. In subsequent periods, this effect steadily decreases and tends toward zero with increasing time lag.

The share of renewables increased by 15.2% in 2023 compared to the previous year. Energy prices increased by 16.5% in 2023, 36.8% in 2022, and 7.7% in 2021 relative to the previous year. Considering price changes since 2022, they explain approximately 55% of this increase, while price changes since 2021 explain approximately 61%. These estimates are summarized in Table 4.7.

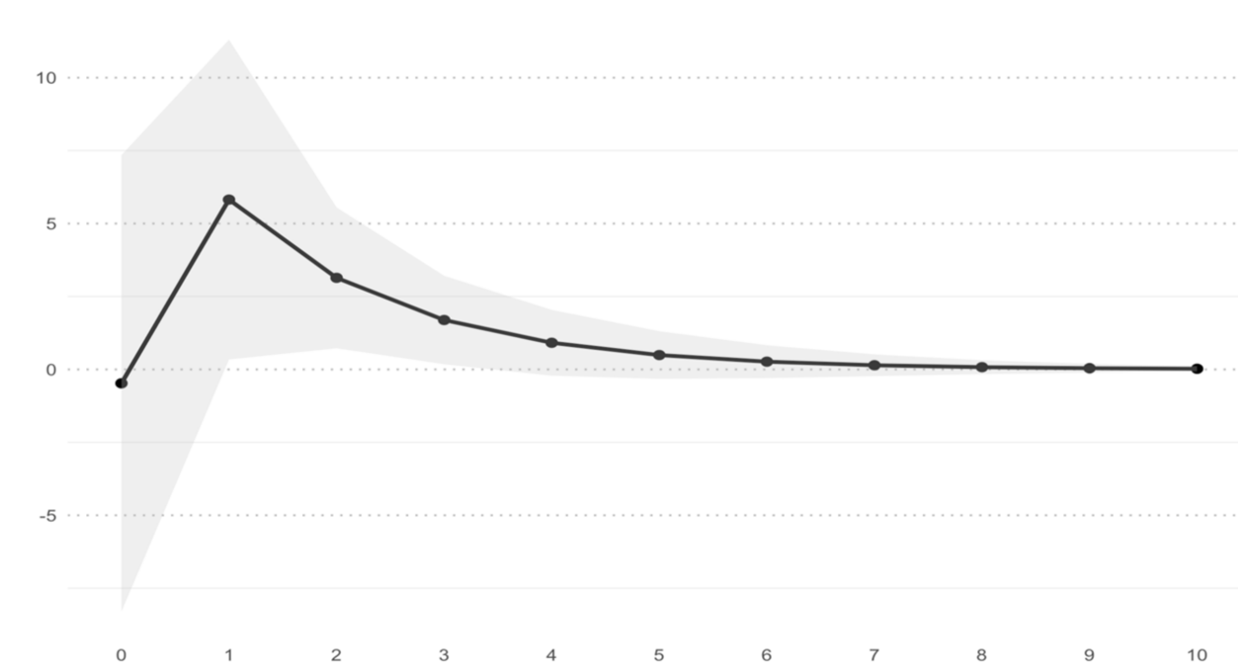



**Figure 4.5:** Dynamic Multipliers of the Energy Price Index

### 4.3.3 Hypothetical Emissions with Average GDP Growth

In this section, we calculate hypothetical emissions for a year with average GDP growth. The average economic growth over the past 10 years is +1.04%. In 2023, GDP growth was

**Table 4.7:** Explanation of Renewable Energy Increase in 2023 Due to Energy Price Index Rise

| | |
|---|---|
| Since January 1, 2022 | 54.7% |
| Since January 1, 2021 | 61.0% |

*Note:* The share of renewables increased by 3.8 percentage points (+15%) in 2023 compared to the previous year. This increase can be attributed to the rise in the energy price index since early 2022 by 54.7%. Considering price increases since early 2021, the contribution is 61.0%.

-0.8% compared to the previous year. Thus, GDP growth in 2023 was 1.84 percentage points below the long-term average. Consequently, hypothetical emissions in 2023 would have been higher with average GDP growth. Considering the previously estimated elasticity of GDP growth (0.46), we calculate a hypothetical emissions increase of approximately 0.86% or 0.53 million tons for 2023.

In 2022, GDP growth was 4.8%, significantly above the long-term average. In this year, hypothetical emissions would have been lower. We calculate lower emissions of approximately 1.73% or 1.13 million tons. These calculations are summarized in Table 4.8.

**Table 4.8:** Hypothetical Emissions with Average GDP Growth

| **Year** | **GDP Growth** (%) | **Difference to Average** (% points) | **Hypothetical Emissions Change** (%) | **Hypothetical Emissions Change** (Mio. T.) |
|---|---|---|---|---|
| 2023 | -0.8 | -1.84 | 0.86 | 0.53 |
| 2022 | 4.8 | 3.77 | -1.73 | -1.13 |

*Note:* The average GDP growth between 2014–2023 is 1.04%. "Diff. to Avg." represents the deviation of actual GDP growth from this average. The emissions change is calculated using the estimated elasticity of 0.46.

# Bibliography

2. Abgabenänderungsgesetz, Federal Law Gazette No. 645 (1977).

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425.

Adler, M. W., & van Ommeren, J. N. (2016a). Does public transit reduce car travel externalities? quasi-natural experiments' evidence from transit strikes. *Journal of Urban Economics*, *92*, 106–119.

Adler, M. W., & van Ommeren, J. N. (2016b). Does public transit reduce car travel externalities? quasi-natural experiments' evidence from transit strikes. *Journal of Urban Economics*, *92*, 106–119. https://doi.org/https://doi.org/10.1016/j.jue.2016.01.001

Administration des transports publics. (2024). Régime Général des Transports Routiers (RGTR) - Acteurs [Accessed: March 27, 2024]. https://transports.public.lu/fr/acteurs/r-z/rgtr.html

Ahn, S. C., & Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, *81*(3), 1203–1227.

Anderl, M., Bartel, A., Geiger, K., Gugele, B., Gössl, M., Haider, S., Heinfellner, H., Heller, C., Köther, T., Krutzler, T., Kuschel, V., Lampert, C., Neier, H., Pazdernik, K., Perl, D., Poupa, S., Prutsch, A., Purzner, M., amd Wolfgang Schieder, E. R., ... Zechmeister, A. (2021). *Klimaschutzbericht 2021*. https://www.umweltbundesamt.at/fileadmin/site/publikationen/rep0776.pdf

Anderl, M., Gössl, M., Haider, S., Heinfellner, H., Krutzler, T., Lampert, C., Pazdernik, K., Purzner, M., Poupa, S., Perl, D., Schieder, W., Schmid, C., Stranner, G., Schodl, B., Storch, A., Wiesenberger, H., Weiss, P., & Zechmeister, A. (2021). *GHG Projections and Assessment of Policies and Measures in Austria. Reporting under Regulation (EU) 2018/1999 15 March 2021*. https://www.umweltbundesamt.at/fileadmin/site/publikationen/rep0766.pdf

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, *104*(9), 2763–2796.

Andersson, J. J. (2019a). Carbon taxes and co2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy*, *11*(4), 1–30.

Andersson, J. J. (2019b). Carbon taxes and CO2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy*, *11*(4), 1–30. https://doi.org/10.1257/pol.20170144

Änderung des Bundesgesetzes über Verkehrsbeschränkungen zur Sicherung der Treibstoffversorgung, Federal Law Gazette No. 5 (1974).

Andrés, L., & Padilla, E. (2018). Driving factors of GHG emissions in the EU transport activity. *Transport Policy*, *61*, 60–74.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Auffhammer, M., & Kellogg, R. (2011). Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, *101*(6), 2687–2722.

Aydin, E., & Kürschner Rauck, K. (2023). Public transport subsidization and air pollution: Evidence from the 9-euro-ticket in germany. *Available at SSRN 4505249*.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, *77*(4), 1229–1279.

Bai, J., & Ng, S. (2004). A panic on unit roots and cointegration. *Econometrica*, *72*, 1127–1177.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Baltagi, B., Feng, Q., & Kao, C. (2016). Estimation of heterogeneous panels with structural breaks. *Journal of Econometrics*, *191*(1), 176–195.

Baltagi, B., Feng, Q., & Kao, C. (2019). Structural changes in heterogeneous panels with endogenous regressors. *Journal of Applied Econometrics*, *34*(6), 883–892.

Banerjee, A., Marcellino, M., & Osbat, C. (2004). Some cautions on the use of panel methods for integrated series of macroeconomic data. *The Econometrics Journal*, *7*(2), 322–340.

Bauernschuster, S., Hener, T., & Rainer, H. (2017). When labor disputes bring cities to a standstill: The impact of public transit strikes on traffic, accidents, air pollution, and health. *American Economic Journal: Economic Policy*, *9*(1), 1–37.

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50–61. https://doi.org/https://doi.org/10.1111/obes.12325

Berger, W. J., Kräutler, C., Sammer, G., Schopf, J. M., Schützhofer, B., & Snizek, S. (2022). *Ein neuer ansatz für höchstzulässige geschwindigkeiten im straßenverkehr in öster-*

*reich aus synergetischer, nachhaltiger sicht* (Vol. 25). Österreichische Forschungsgesellschaft Straße – Schiene – Verkehr.

Bersvendsen, T., & Ditzen, J. (2021). Testing for slope heterogeneity in stata. *The Stata Journal*, *21*(1), 51–80.

Bigi, F., Schwemmle, N., & Viti, F. (2023). Evaluating the impact of free public transport using agent-based modeling: The case-study of luxembourg.

BKA. (2020). Aus Verantwortung für Österreich. Regierungsprogramm 2020-2024 der Österreichischen Bundesregierung. https://www.bundeskanzleramt.gv.at/bundeskanzleramt/die-bundesregierung/regierungsdokumente.html

BMDV. (2019). *Verkehr in Zahlen - Archiv*. https://www.bmvi.de/SharedDocs/DE/Artikel/G/verkehr-in-zahlen-archiv.html

BMK. (2021). *Austria's 2030 mobility master plan*. https://www.bmk.gv.at/en/topics/mobility/mobilitymasterplan2030.html

Borsati, M., Albalate, D., & Gragera, A. (2023). Free rides to cleaner air? examining the impact of massive public transport fare discounts on air quality. *SSRN: https://ssrn.com/abstract=4603867*.

Botta, E., & Koźluk, T. (2014). Measuring environmental policy stringency in oecd countris: A composite index approach. *OECD Economics Department Working Papers*, *1177*, 1–48.

Breitung, J., & Pesaran, M. H. (2008). Unit roots and cointegration in panels. In L. Matyas & P. Sevestre (Eds.), *The econometrics of panel data. fundamentals and recent developments in theory and practice*. Springer.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, *17*, 334–355.

Brunel, C., & Levinson, A. (2016). Measuring the stringency of environmental regulations. *Review of Environmental Economics and Policy*, *10*(1). https://www.journals.uchicago.edu/doi/abs/10.1093/reep/rev019?journalCode=reep

Bull, O., Muñoz, J. C., & Silva, H. E. (2021). The impact of fare-free public transport on travel behavior: Evidence from a randomized controlled trial. *Regional Science and Urban Economics*, *86*, 103616. https://doi.org/doi.org/10.1016/j.regsciurbeco.2020.103616

Cao, S., & Zhou, Q. (2022). Common correlated effects estimation for dynamic heterogeneous panels with non-stationary multi-factor error structures. *Econometrics*, *10*(3).

Carbone, J. C., Rivers, N., Yamazaki, A., & Yonezawa, H. (2020). Comparing applied general equilibrium and econometric estimates of the effect of an environmental policy shock.

*Journal of the Association of Environmental and Resource Economists*, *7*(4). https://doi.org/10.1086/708734

Cats, O., Susilo, Y. O., & Reimal, T. (2017). The prospects of fare-free public transport: Evidence from tallinn. *Transportation*, *44*(5), 1083–1104. https://doi.org/10.1007/s11116-016-9695-5

Cavaliere, G. (2005). Unit root tests under time-varying variances. *Econometric Reviews*, *23*(3), 259–292.

Chen, Y., & Whalley, A. (2012). Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal: Economic Policy*, *4*(1), 58–97.

Chudik, A., Mohaddes, K., Pesaran, M. H., & Raissi, M. (2016). Long-run effects in large heterogeneous panel data models with cross-sectionally correlated errors. In G. González-Rivera, R. C. Hill, & T. H. Lee (Eds.), *Essays in honor of man ullah (advances in econometrics, vol. 36)*. Emerald Group Publishing Limited.

Chudik, A., & Pesaran, M. H. (2015a). Large panel data models with cross-sectional dependence. a survey. In B. H. Baltagi (Ed.), *The oxford handbook of panel data*. Oxford University Press.

Chudik, A., & Pesaran, M. H. (2015b). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, *188*(2), 393–420.

City public transport information [Accessed: 2024-04-26]. (2024). https://freepublictransport.info/city/

Clarke, D., Pailañir, D., Athey, S., & Imbens, G. (2023). Synthetic difference in differences estimation. https://arxiv.org/abs/2301.11859

Coakley, J., Fuertes, A.-M., & Smith, R. (2002). *A principal components approach to cross-section dependence in panels* (10th International Conference on Panel Data, Berlin, July 5-6, 2002 No. B5-3). International Conferences on Panel Data.

Copernicus Climate Change Service. (2024). Global climate highlights 2024. https://climate.copernicus.eu/sites/default/files/custom-uploads/GCH-2024/GCH2024-PDF-1.pdf

Corrocher, N., & Mancusi, M. L. (2021). International collaborations in green energy technologies: What is the role of distance in environmental policy stringency? *Energy Policy*, *156*, 112470. https://doi.org/https://doi.org/10.1016/j.enpol.2021.112470

Creutzig, F., Roy, J., Lamb, W. F., Azevedo, I. M. L., Bruine de Bruin, W., Dalkmann, H., Edelenbosch, O. Y., Geels, F. W., Grubler, A., Hepburn, C., Hertwich, E. G., Khosla, R., Mattauch, L., Minx, J. C., Ramakrishnan, A., Rao, N. D., Steinberger, J. K., Tavoni, M., Ürge-Vorsatz, D., & Weber, E. U. (2018). Towards demand-side solutions for mitigating climate change. *Nature Climate Change*, *8*(4), 260–263. https://doi.org/10.1038/s41558-018-0121-1

Crippa, M., Guizzardi, D., Solazzo, E., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Banja, M., Oliver, J., Grassi, G., Rossi, S., & Vignati, E. (2022). Ghg emissions of all world countries - 2022 report. *EUR 30831 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-41547-3, doi:10.2760/173513, JRC126363*.

Davis, L. W. (2008). The effect of driving restrictions on air quality in mexico city. *Journal of Political Economy*, *116*(1), 38–81.

Davis, L. W. (2017). Saturday driving restrictions fail to improve air quality in mexico city. *Scientific reports*, *7*(1), 41652.

Département de la mobilité et des transports. (2020). Train - mobility - transport portal - luxembourg [Accessed on April 4, 2024]. https://transports.public.lu/fr/mobilite/transport-en-commun/train.html

Département de la mobilité et des transports. (2024). Tram - mobilité [Accessed: March 27, 2024]. https://transports.public.lu/fr/mobilite/transport-en-commun/tram.html

Dietz, T., & Rosa, E. (1997). Effects of population and affluence on CO emissions. *PNAS*, *94*(1), 175–179.

Dippold, M., Rexeis, M., & Hausberger, S. (Eds.). (2012). *Nemo - a universal and flexible model for assessment of emissions on road networks*.

Ditzen, J. (2016a). XTCD2: Stata module to test for weak cross sectional dependence. Statistical Software Components, Boston College Department of Economics.

Ditzen, J. (2016b). XTDCCE2: Stata module to estimate heterogeneous coefficient models using common correlated effects in a dynamic panel. Statistical Software Components, Boston College Department of Economics.

Ditzen, J. (2018). Estimating dynamic common-correlated effects in stata. *The Stata Journal*, *18*(3), 585–617.

Ditzen, J. (2021). Estimating long-run effects and the exponent of cross-sectional dependence: An update to xtdcce2. *Stata Journal*, *21*(3), 687–707.

Ditzen, J., & Reese, S. (2023). Xtnumfac: A battery of estimators for the number of common factors in time series and panel-data models. *The Stata Journal*, *23*(2), 438–454.

Dong, K., Hochman, G., Zhang, Y., Sun, R., Li, H., & Liao, H. (2018). Co2 emissions, economic and population growth, and renewable energy: Empirical evidence across regions. *Energy Economics*, *75*, 180–192.

Doornik, J., & Hendry, D. (1997). *Modelling dynamic systems using pcfiml 9.0 for windows*. International Thomson Business Press.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics*, *80*(4), 549–560.

Dugan, A., Mayer, J., Thaller, A., Bachner, G., & Steininger, K. W. (2022). Developing policy packages for low-carbon passenger transport: A mixed methods analysis of trade-offs and synergies. *Ecological Economics*, *193*, 107304.

EAA. (2019). *Dokumentation der Emissionsfaktoren*. https://www.umweltbundesamt.at/fileadmin/site/themen/mobilit%C3%A4t/daten/ekz_doku_verkehrsmittel.pdf

EC. (2021). *Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 establishing the framework for achieving climate neutrality and amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law')*. http://data.europa.eu/eli/reg/2021/1119/oj

Edgerton, D., & Shukur, G. (1999). Testing autocorrelation in a system perspective testing autocorrelation. *Econometric Review*, *18*(4), 343–386.

EEA. (2021). *Annual European Union greenhouse gas inventory 1990–2019 and inventory report 2021. Submission to the UNFCCC Secretariat*. f454ae4c825646b2a15497b17a71dbf0

EEA. (2024). *Average CO2 emissions per km from new passenger cars*. https://www.eea.europa.eu/en/datahub/datahubitem-view/5d252092-d328-40d8-bca2-c0734bd6143b

Ehlrich, P., & Holdren, J. (1971). Impact of population growth. *Science*, *171*(3977), 1212–1217.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64*(4), 813–836.

European Commission. (2022). Nuts - nomenclature of territorial units for statistics [Accessed: February 22, 2022]. https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts

European Commission. (2024a). Effort sharing 2021-2030: Targets and flexibilities [Accessed: January 29, 2025]. https://climate.ec.europa.eu/eu-action/effort-sharing-member-states-emission-targets/effort-sharing-2021-2030-targets-and-flexibilities_en

European Commission. (2024b). *Weekly Oil Bulletin*. https://energy.ec.europa.eu/data-and-analysis/weekly-oil-bulletin_en#price-developments

European Commission and Directorate-General for Mobility and Transport. (2021). *Eu transport in figures – statistical pocketbook 2021*. Publications Office. https://doi.org/doi/10.2832/27610

European Environment Agency (EEA). (2024). Greenhouse gases data viewer [Accessed: [Insert Date]]. https://www.eea.europa.eu/en/analysis/maps-and-charts/greenhouse-gases-viewer-data-viewers

Eurostat. (2024). *Regional statistics*. https://ec.europa.eu/eurostat/web/regions/database

Federal Transit Administration. (2010). Public transportation's role in responding to climate change [Accessed: December 28, 2023]. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/PublicTransportationsRoleInRespondingToClimateChange2010.pdf

Forchini, G., Jiang, B., & Peng, B. (2015, July). *Consistent Estimation in Large Heterogeneous Panels with Multifactor Structure and Endogeneity* (School of Economics Discussion Papers No. 0315). School of Economics, University of Surrey.

Gagliardini, P., Ossola, E., & Scaillet, O. (2019). A diagnostic criterion for approximate factor structure. *Journal of Econometrics*, *212*(2), 503–521.

Galeotti, M., Salini, S., & Verdolini, E. (2020). Measuring environmental policy stringency: Approaches, validity, and impact on environmental innovation and energy efficiency. *Energy Policy*, *136*, 1–14.

Gallego, F., Montero, J.-P., & Salas, C. (2013). The effect of transport policies on car use: Evidence from latin american cities. *Journal of Public Economics*, *107*, 47–62.

Gendron-Carrier, N., Gonzalez-Navarro, M., Polloni, S., & Turner, M. A. (2022). Subways and urban air pollution. *American economic journal: Applied economics*, *14*(1), 164–196.

Georgatzi, V. V., Stamboulis, Y., & Vetsikas, A. (2020). Examining the determinants of CO2 emissions caused by the transport sector: Empirical evidence from 12 european countries. *Economic Analysis and Policy*, *65*(1), 11–20.

Gerlagh, R., van den Bijgaart, I., Nijland, H., & Michielsen, T. (2018). Fiscal policy and $CO_2$ emissions of new passenger cars in the eu. *Environmental and Resource Economics*, *69*(1), 103–134. https://doi.org/10.1007/s10640-016-0067-6

Geschwindigkeitsbeschränkungs-Verordnung, Federal Law Gazette No. 562 (1973).

Gibson, M., & Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, *89*, 62–73. https://doi.org/https://doi.org/10.1016/j.jue.2015.06.005

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, *46*, 1293–1301.

Gohl, N., & Schrauth, P. (2024). Ticket to paradise?: The effect of a public transport subsidy on air quality. *Journal of Urban Economics*.

Gonçalves, S., & Perron, B. (2014). Bootstrapping factor-augmented regression models [Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions]. *Journal of Econometrics*, *182*(1), 156–173.

González, R. M., Marrero, G. A., Rodríguez-López, J., & Marrero, Á. S. (2019). Analyzing CO2 emissions from passenger cars in europe: A dynamic panel data approach. *Energy Policy*, *129*, 1271–1281.

Government of the Grand Duchy of Luxembourg. (2020, March). Measures taken by the government council on 12 march 2020 in response to the coronavirus [Accessed: 2024-08-07].

Greene, W. H. (2018). *Econometric analysis* (8th). Pearson.

Guo, M., Chen, S., Zhang, J., & Meng, J. (2022). Environment kuznets curve in transport sector's carbon emission: Evidence from china. *Journal of Cleaner Production*, *371*, 133504.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *The Econometrics Journal*, *3*(2), 148–161.

Hahn, R. W., Hendren, N., Metcalfe, R. D., & Sprung-Keyser, B. (2024, July). *A welfare analysis of policies impacting climate change* (Working Paper No. 32728). National Bureau of Economic Research. https://doi.org/10.3386/w32728

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, *5*(4), 529–538. https://doi.org/10.1038/s41562-021-01079-8

Hanck, C. (2012). Do panel cointegration tests produce "mixed signals"? *Annals of Economics and Statistics*, (107/108), 299–310.

Hanck, C. (2013). An intersection test for panel unit roots. *Econometric Reviews*, *32*(2), 183–203.

Harding, M., & Lamarche, C. (2011). Least squares estimation of a panel data model with multifactor error structure and endogenous covariates. *Economics Letters*, *111*(3), 197–199.

Hashmi, R., & Alam, K. (2019). Dynamic relationship among environmental regulation, innovation, co2 emissions, population, and economic growth in oecd countries: A panel investigation. *Journal of Cleaner Production*, *231*, 1100–1109.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*(6), 1251–1271.

Hille, E., & Möbius, P. (2019). Environmental policy, innovation, and productivity growth: Controlling the effects of regulation and endogeneity. *Environmental and Resource Economics*, *73*(4), 1315–1355.

Hintermann, B., Schoeman, B., Götschi, T., Axhausen, K., Molloy, J., & Tchervenkov, C. (2022). *Pigovian transport pricing in practice* (VfS Annual Conference 2022 (Basel): Big Data in Economics). Verein für Socialpolitik / German Economic Association. https://EconPapers.repec.org/RePEc:zbw:vfsc22:264014

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, *115*(1), 53–74.

Immissionsschutzgesetz – Luft, Federal Law Gazette I No. 115 (1997).

International Transport Forum. (2020). *Decarbonising transport in europe: The way forward* (tech. rep.) (Accessed: December 28, 2023). International Transport Forum at the OECD. https://www.itf-oecd.org/sites/default/files/docs/decarbonising-transport-europe-way-forward.pdf

IPCC. (2018). *Global warming of 1.5°c. an ipcc special report on the impacts of global warming of 1.5°c above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P. Shukla, A. Pirani, W. Moufouma-Okia, C. Peéan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T. Waterfield, Eds.).

IPCC. (2023). *Sections* (H. L. Core Writing Team & J. Romero, Eds.). https://doi.org/10.59327/IPCC/AR6-9789291691647

Jochem, P., Doll, C., & Fichtner, W. (2016). External costs of electric vehicles. *Transportation Research Part D: Transport and Environment*, *42*, 60–76. https://doi.org/https://doi.org/10.1016/j.trd.2015.09.022

Juodis, A., & Reese, S. (2021). The incidental parameters problem in testing for remaining cross-section correlation. *Journal of Business and Economic Statistics.*, *40*(3), 1191–1203.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, *90*(1), 1–44.

Kao, C., & Chiang, M.-H. (2001). On the estimation and inference of a cointegrated regression in panel data. In B. H. Baltagi, T. B. Fomby, & R. Carter Hill (Eds.), *Nonstationary panels, panel cointegration, and dynamic panels* (pp. 179–222, Vol. 15). Emerald Group Publishing Limited.

Kapetanios, G., Pesaran, M. H., & Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, *160*(2), 326–3488.

Karabiyik, H., Palm, F. C., & Urbain, J.-P. (2019). Econometric analysis of panel data models with multifactor error structures. *Annual Review of Economics*, *11*(1), 495–522.

Kilian, L., & Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

Koch, N., Illenseer, N., Amberg, M., & Treichel-Grass, K. (2024). Effect of the deutschland-ticket on mobility and emissions [Ariadne compact].

Koch, N., Naumann, L., Pretis, F., Ritter, N., & Schwarz, M. (2022). Attributing agnostically detected large reductions in road CO2 emissions to policy mixes. *Nature Energy*, *7*(9), 844–853. https://doi.org/10.1038/s41560-022-01095-6

Kraftfahrzeugsteuergesetz, Federal Law Gazette No. 110 (1952).

Kraftfahrzeugsteuergesetz, Federal Law Gazette No. 449 (1992).

Kranz, S. (2022). *Synthetic difference-in-differences with time- varying covariates. technical report*. https://github.com/skranz/xsynthdid/blob/main/paper/synthdid%20with%20covariates.pdf

Kripfganz, S., & Sarafidis, V. (2021). Instrumental-variable estimation of large-t panel-data models with common factors. *The Stata Journal*, *21*(3), 659–686.

Kuss, P., & Nicholas, K. A. (2022). A dozen effective interventions to reduce car use in european cities: Lessons learned from a meta-analysis and transition management. *Case Studies on Transport Policy*, *10*(3), 1494–1513.

Lalive, R., Luechinger, S., & Schmutzler, A. (2018). Does expanding regional train service reduce air pollution? *Journal of Environmental Economics and Management*, *92*, 744–764.

Le Ministre du Développement durable et des Infrastructures. (2017). Loi du 14 juillet 2017 sur l'égalité de traitement entre hommes et femmes [Accessed: 2024-05-15]. https://legilux.public.lu/eli/etat/leg/rmin/2017/07/14/a663/jo

Lee, C., Wu, J.-L., & Yang, L. (2016). A simple panel unit-root test with smooth breaks in the presence of a multifactor error structure. *Oxford Bulletin of Economics and Statistics*, *78*(3), 365–393.

Levin, A., Lin, C.-F., & Chu, C.-S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, *108*(1), 1–24.

Li, S., Liu, Y., Purevjav, A.-O., & Yang, L. (2019). Does subway expansion improve air quality? *Journal of Environmental Economics and Management*, *96*, 213–235.

Liebensteiner, M., Losert, J., Necker, S., Neumeier, F., Paetzold, J., & Wichert, S. (2024). Almost fare free: Impact of a cheap public transport ticket on mobility patterns and infrastructure quality.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Luxmobile. (2020). Online survey - notre plan de mobilité pour demain. https://www.vdl.lu/en/getting-around/notre-plan-de-mobilite-pour-demain/online-survey

Luxtoday. (2022, August). Luxembourg to have 4 tram lines by 2035 [Accessed: March 27, 2024]. https://luxtoday.lu/en/luxembourg-en/luxembourg-to-have-4-tram-lines-by-2035

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, *61*(S1), 631–652.

Mark, N. C., & Sul, D. (2003). Cointegration vector estimation by panel dols and long-run money demand*. *Oxford Bulletin of Economics and Statistics*, *65*(5), 655–680.

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2020). Coronavirus pandemic (covid-19) [https://ourworldindata.org/coronavirus]. *Our World in Data*.

Mineralölsteuergesetz, Federal Law Gazette No. 140 (1949).

Mineralölsteuergesetz, Federal Law Gazette II No. 630 (1994).

Ministère du Développement Durable et des Infrastructures. (2018). Modu 2.0 - Stratégie pour une mobilité durable [Accessed: March 27, 2024]. https://transports.public.lu/fr/planifier/strategie/modu2.html

Ministère du Développement Durable et des Infrastructures. (2020). Bus - mobilité [Accessed: March 27, 2024]. https://transports.public.lu/fr/mobilite/transport-en-commun/bus.html

Ministry of Mobility and Public Works. (2020). Faq: Free public transport [Accessed: February 29, 2024]. https://transports.public.lu/dam-assets/publications/mobilite/transports-collectifs/20200212-mobilite-gratuite-en-texte-faq.pdf

Naqvi, A. (2021). Covid-19 european regional tracker. *Scientific data*, *8*(1), 181.

Neal, T. (2015). Estimating heterogeneous coefficients in panel data models with endogenous regressors and common factors. *Working paper*, (2009), 1–29.

Neves, S. A., Marques, A. C., & Patrício, M. (2020). Determinants of CO2 emissions in european union countries: Does environmental regulation reduce environmental pollution? *Economic Analysis and Policy*, *68*, 114–125.

Ng, S. (2008). A simple test for nonstationarity in mixed panels. *Journal of Business & Economic Statistics*, *26*(1), 113–127.

Norkutė, M., Sarafidis, V., Yamagata, T., & Cui, G. (2021). Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure. *Journal of Econometrics*, *220*(2), 416–446.

Normverbrauchsabgabegesetz, Federal Law Gazette No. 695 (1991).

Odyssee-Mure. (2021). *Odyssee Database*. https://www.indicators.odyssee-mure.eu/energy-efficiency-database.html

OECD. (2023). *Tram usage has increased steadily*. https://doi.org/doi.org/10.1787/9409d9b6-en

OENB. (2022). *Nationaler Verbraucherpreisindex*. https://www.oenb.at/isaweb/report.do;jsessionid=1D06BAE627C2C24182856C6667D340A2?report=6.4

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, *92*(4), 1004–1016.

Opoku, E. E. O., Dogah, K. E., & Aluko, O. A. (2022). The contribution of human development towards environmental sustainability. *Energy Economics*, *106*, 105782.

Ostermeijer, F., Koster, H. R., & van Ommeren, J. (2019). Residential parking costs and car ownership: Implications for parking policy and automated vehicles. *Regional Science and Urban Economics*, *77*, 276–288. https://doi.org/https://doi.org/10.1016/j.regsciurbeco.2019.05.005

Pablo-Romero, M. P., Cruz, L., & Barata, E. (2017). Testing the transport energy-environmental kuznets curve hypothesis in the EU27 countries. *Energy Economics*, *62*, 257–269.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, *61*(S1), 653–670.

Pedroni, P. (2001). Fully modified ols for heterogeneous cointegrated panels. In B. H. Baltagi, T. B. Fomby, & R. Carter Hill (Eds.), *Nonstationary panels, panel cointegration, and dynamic panels* (pp. 93–130, Vol. 15). Emerald Group Publishing Limited.

Pedroni, P. (2004). Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the ppp hypothesis. *Economic Theory*, *20*, 597–625.

Persyn, D., & Westerlund, J. (2008). Error-correction–based cointegration tests for panel data. *Stata Journal*, *8*(2), 232–241.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*(4), 967–1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265–312.

Pesaran, M. H. (2012). On the interpretation of panel unit root tests. *Economics Letters*, *116*(3), 545–546.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, *6-10*(34), 1089–1117.

Pesaran, M. H., L., V. S., & Yamagata, T. (2013). Panel unit root tests in the presence of a multifactor error structure. *Journal of Econometrics*, *175*(2), 94–115.

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

Pesaran, M. H., Shin, Y., & Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, *94*(446), 621–634.

Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, *68*(1), 79–113.

Pesaran, M. H., & Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, *142*(1), 50–93.

Phillips, P. C. B., & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, *67*(5), 1057–1111.

Phillips, P. C., & Moon, H. R. (2000). Nonstationary panel data analysis: An overview of some recent developments. *Econometric Reviews*, *19*(3), 263–286.

Pretis, F. (2022a). Does a carbon tax reduce CO2 emissions? evidence from british columbia. *Environmental and Resource Economics*, *83*(1), 115–144. https://doi.org/10.1007/s10640-022-00679-w

Pretis, F. (2022b). Does a carbon tax reduce co2 emissions? evidence from british columbia. *Environmental and Resource Economics*, *83*(1), 115–144.

Probst, M., & Sauter, C. (2015, January). *CO2 Emissions and Greenhouse Gas Policy Stringency - An Empirical Assessment* (IRENE Working Papers No. 15-03). IRENE Institute of Economic Research. https://ideas.repec.org/p/irn/wpaper/15-03.html

Rafiq, S., Salim, R., & Nielsen, I. (2016). Urbanization, openness, emissions, and energy intensity: A study of increasingly urbanized emerging economies. *Energy Economics*, *56*, 20–28.

Research Luxembourg. (2021, March). Luxembourg celebrates 1 year of free public transport [Accessed: March 26, 2024]. https://www.researchluxembourg.org/en/luxembourg-celebrates-1-year-of-free-public-transport/

Robinson, S., Yúnez-Naude, A., Hinojosa-Ojeda, R., Lewis, J. D., & Devarajan, S. (1999). From stylized to applied models:: Building multisector cge models for policy analysis.

*The North American Journal of Economics and Finance*, *10*(1), 5–38. https://doi.org/https://doi.org/10.1016/S1062-9408(99)00014-5

Sangiacomo, M. (2014). XTCIPS: Stata module to compute Pesaran Panel Unit Root Test in the Presence of Cross-section Dependence. Statistical Software Components, Boston College Department of Economics.

Santos, G., Behrendt, H., Maconi, L., Shirvani, T., & Teytelboym, A. (2010). Part i: Externalities and economic policies in road transport [Road Transport Externalities, Economic Policies And Other Instruments For Sustainable Road Transport]. *Research in Transportation Economics*, *28*(1), 2–45. https://doi.org/https://doi.org/10.1016/j.retrec.2009.11.002

Sarafidis, V., & Wansbeek, T. (2012). Cross-Sectional Dependence in Panel Data Analysis. *Econometric Reviews*, *31*(5), 483–531.

Sarmiento, L., Wägner, N., & Zaklan, A. (2023). The air quality and well-being effects of low emission zones. *Journal of Public Economics*, *227*, 105014.

Schenk, T. (2023). *Time-weighted difference-in-differences: Accounting for common factors in short t panels*. https://papers.tinbergen.nl/23004.pdf

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1–48.

Song, M. K. (2013). Essays on large panel data analysis: Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications. *Dissertation, Department of Economics, Columbia University*. https://doi.org/10.7916/D8MS3S40

Statistics Austria. (2024a). *Energy balances* [Accessed: 2024-08-22]. Retrieved August 22, 2024, from https://www.statistik.at/en/statistics/energy-and-environment/energy/energy-balances

Statistics Austria. (2024b). *Energy prices* [Accessed: 2024-08-22]. Retrieved August 22, 2024, from https://www.statistik.at/en/statistics/energy-and-environment/energy/energy-prices-taxes

Statistik Austria. (2021). *Jahresdurchschnittsbevölkerung 1870-2020*. https://tinyurl.com/48fayh7u

Stechemesser, A., Koch, N., Mark, E., Dilger, E., Klösel, P., Menicacci, L., Nachtigall, D., Pretis, F., Ritter, N., Schwarz, M., Vossen, H., & Wenzel, A. (2024). Climate policies that achieved major emission reductions: Global evidence from two decades. *Science*, *385*(6711), 884–892. https://doi.org/10.1126/science.adl6547

Steg, L., & Gifford, R. (2005). Sustainable transportation and quality of life [Sustainability and the Interaction Between External Effects of Transport (Part Special Issue, pp. 23-99)]. *Journal of Transport Geography*, *13*(1), 59–69. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2004.11.003

Steininger, K. W., Posch, A., Thaller, A., Romirer, C., Maier, R., Fleiß, E., Jäger, G., Nabernegg, S., Plakolb, S., Heinfellner, H., Thaller, O., & Schulev-Steindl, E. (2024). Knowledge integration in policy development and assessment for sustainable passenger transport transformation. *Transportation Research Part D: Transport and Environment*, *128*, 104081. https://doi.org/https://doi.org/10.1016/j.trd.2024.104081

Steininger, K. W., Riahi, K., Stagl, S., Kromp-Kolb, H., Kirchengast, G., Leschka, N., Williges, K., Scheifinger, K., Henner, D. N., Reisinger, M., & Wolf, A. (2024). Nationaler Energie- und Klimaplan (NEKP) für Österreich - Wissenschaftliche Bewertung der in der Konsultation 2023 vorgeschlagenen Maßnahmen [[Steininger, Karl W., Riahi, K., Stagl, S., Kromp-Kolb, H., Kirchengast, G., Leschka, N., Williges, K., Scheifinger, K., Henner, D. N., Reisinger, M., Wolf, A.]].

Štraub, D., Kebłowski, W., & Maciejewska, M. (2023). From Bełchatów to Żory: Charting Poland's geography of fare-free public transport programmes. *Journal of Transport Geography*, *111*, 103650. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2023.103650

Trapani, L. (2021). Inferential theory for heterogeneity and cointegration in large panels. *Journal of Econometrics*, *220*(2), 474–503.

U.S. Energy Information Administration. (2022a). *Crude oil prices: Brent - europe*. https://fred.stlouisfed.org/series/ACOILBRENTEU

U.S. Energy Information Administration. (2022b). *Spot crude oil price: West texas intermediate (wti)*. https://fred.stlouisfed.org/series/WTISPLC

Wagner, M. (2015). The environmental kuznets curve, cointegration and nonlinearity. *Journal of Applied Econometrics*, *30*(6), 948–967.

Wagner, M., & Reichold, K. (2023). Panel cointegrating polynomial regressions: Group-mean fully modified ols estimation and inference. *Econometric Reviews*, *42*(4), 358–392.

Wang, K., Yan, M., Wang, Y., & Chang, C.-P. (2020). The impact of environmental policy stringency on air quality. *Atmospheric Environment*, *231*, 117522.

Wang, W.-Z., Liu, L.-C., Liao, H., & Wei, Y.-M. (2021). Impacts of urbanization on carbon emissions: An empirical analysis from OECD countries. *Energy Policy*, *151*, 112171.

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, *69*(6), 709–880.

Westerlund, J. (2014). Heteroscedasticity robust panel unit root tests. *Journal of Business Economic Statistics*, *32*(1), 112–135.

Westerlund, J., & Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks*. *Oxford Bulletin of Economics and Statistics*, *70*(5), 665–704.

Westerlund, J., Petrova, Y., & Norkute, M. (2019). Cce in fixed-t panels. *Journal of Applied Econometrics*, *34*(5), 746–761.

Westerlund, J., & Urbain, J.-P. (2015). Cross-sectional averages versus principal components. *Journal of Econometrics*, *185*(2), 372–377.

Winkler, L., Pearce, D., Nelson, J., & Babacan, O. (2023). The effect of sustainable mobility transition policies on cumulative urban transport emissions and energy demand. *Nature Communications*, *14*(1), 2357. https://doi.org/10.1038/s41467-023-37728-x

WKO. (2021). *Bip und Wirtschaftswachstum*. https://tinyurl.com/yc58htfj

Wolff, H. (2014). Keep your clunker in the suburb: Low-emission zones and adoption of green vehicles. *The Economic Journal*, *124*(578), 481–512.

Xu, B., & Lin, B. (2016). Differences in regional emissions in china's transport sector: Determinants and reduction strategies. *Energy*, *95*(1), 459–470.

Yirong, Q. (2022). Does environmental policy stringency reduce co2 emissions? evidence from high-polluted economies. *Journal of Cleaner Production*, *341*, 130648. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.130648

Zhang, N., Yu, K., & Chen, Z. (2017). How does urbanization affect carbon dioxide emissions? a cross-country panel data analysis. *Energy Policy*, *107*, 678–687.

Zheng, L., Yuan, L., Khan, Z., Badeeb, R. A., & Zhang, L. (2023). How G-7 countries are paving the way for net-zero emissions through energy efficient ecosystem? *Energy Economics*, *117*, 106428.

# Appendix A

# Appendix for Chapter 1

## A.1 Macropanel Econometrics

### A.1.1 Spurious Regression Case

The pooled OLS estimator for equation (1.1) is given by

$$\hat{\beta} = \left( \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} x_{it}' \right) \left( \sum_{i=1}^{N} \sum_{t=1}^{T} x_{it} x_{it}' \right)^{-1}, \tag{A.1}$$

which pools information across individuals $i$. Applying the sequential limit theory from Phillips and Moon (1999) with $T, N \to \infty_{seq}$[1], and noting that under the functional panel limit theorem the random walks converge to Brownian motions, we get:

$$\hat{\beta} \Rightarrow \left( \frac{1}{N} \sum_{i=1}^{N} \int B_{y_i} B_{x_i}' \right) \left( \frac{1}{N} \sum_{i=1}^{N} \int B_{x_i} B_{x_i}' \right)^{-1} \quad as\ T \to \infty\ for\ fixed\ N,$$

$$\xrightarrow{p}\ E\left( \int B_y B_x' \right) \left[ E\left( \int B_x B_x' \right) \right]^{-1} = \Omega_{yx} \Omega_{xx}^{-1} = \beta \quad as\ N \to \infty,$$

where $B_i$ is a Brownian motion and integrals are taken over the interval $[0, 1]$. The first line shows the limit distribution of $\hat{\beta}$ as $T \to \infty$ and the second line shows the convergence in probability as $N \to \infty$ by applying the law of large numbers. $\Omega_{yx}$ and $\Omega_{xx}$ denote long-run covariance matrices. We can see that $\hat{\beta}$ is a consistent estimator for $\beta$.

[1] In panel analysis a distinction is made between sequential asymptotics, where first $T \to \infty$ with $N$ fixed and then $N \to \infty$, and joint asymptotics in which both $T$ and $N$ go to infinity jointly, typically with restrictions on their relative rates of converges, e.g., $T/N \to 0$.

### A.1.2 Cointegration Case

The pooled least squares estimator of model (1.2) with a homogeneous cointegrating relation is given by:

$$\hat{\beta} = \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}x'_{it} \left(\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x'_{it}\right)^{-1} = \beta + \sum_{i=1}^{N}\sum_{t=1}^{T} e_{it}x'_{it} \left(\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x'_{it}\right)^{-1}. \tag{A.2}$$

The limit distribution of the rescaled and centered estimator is then:

$$\begin{aligned}\sqrt{N}(\hat{\beta}-\beta) \Rightarrow & \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\int B_{e_i}B'_{x_i} + \Delta_{ex}\right)\right) \times \\ & \left(\frac{1}{N}\sum_{i=1}^{N}\left(\int B_{x_i}B'_{x_i}\right)\right)^{-1}, \quad as\ T \to \infty\ for\ fixed\ N,\end{aligned}$$

where Phillips and Moon (1999) show that

$$E\left(\int B_{e_i}B'_{x_i} + \Delta_{ex}\right) = \frac{1}{2}\Delta_{ex} \neq 0 \quad as\ N \to \infty.$$

The bias $\Delta_{ex}$ arises from endogeneity of the regressor, which can be induced by the temporal correlation between $e_{it}$ and $u_{it}$ and direct correlation between $x_{it}, e_{it}$. This second-order bias is persistent and the estimator is inconsistent when $\Delta_{ex} \neq 0$. Note that the asymptotic covariance matrix is given by $\Omega$, which can be decomposed as: $\Omega = \sum_{j=-\infty}^{\infty} E(w_{ij}w'_{i0}) = \Sigma + \Gamma + \Gamma'$, with contemporaneous covariance $\Sigma = E(w_{i0}w'_{i0})$ and sum of autocovariances $\Gamma = \sum_{j=1}^{\infty} E(w_{ij}w'_{i0})$. Then define $\Delta = \Sigma + \Gamma$.

### A.1.3 Panel Unit-Root Test

Formally, the CIPS test is based on the regression

$$\Delta y_{it} = \alpha_i + \rho_i^* y_{i,t-1} + \theta_0 \bar{y}_{t-1} + \sum_{j=0}^{p} \theta_{j+1}\Delta\bar{y}_{t-j} + \sum_{k=1}^{p} c_k \Delta y_{i,t-k} + u_{it}, \tag{A.3}$$

where $\bar{y}_t = \frac{1}{N}\sum y_{it}$ is the time $t$ cross-sectional average and $\Delta\bar{y}_{t-1}$ its first difference. The lagged cross-sectional average and its first-differences account for the cross-sectional dependence, which is assumed to be described by a factor structure. The degree of augmentation can be chosen by an information criterion or by sequential testing (Baltagi, 2008).

The CIPS test statistic is given by

$$CIPS = \frac{1}{N}\sum_{i=1}^{N} CADF_i, \tag{A.4}$$

where $CADF_i$ is the cross-sectionally augmented Dickey-Fuller (CADF) test for each unit $i$ in the panel. The CIPS statistic has a nonstandard joint asymptotic limit and critical values are provided in Pesaran (2007) for various choices of $N$ and $T$.

### A.1.4 Westerlund Panel Cointegration Test

The proposed test statistics from Westerlund (2007) are based on the regression

$$\Delta y_{it} = \delta_i' d_t + \alpha_i y_{it-1} + \lambda_i' x_{it} + \sum_{j=1}^{p_i} \alpha_{ij}\Delta y_{it-j} + \sum_{j=0}^{p_i} \gamma_{ij}\Delta x_{it-j} + u_{it}, \tag{A.5}$$

where $d_t = (1, t)'$ denotes the deterministic components with $\delta_i = (\delta_{1i}, \delta_{2i})'$ being the associated vector of parameters. The least square estimate of $\alpha_i$ can then be used to provide valid tests to determine the presence of cointegration.

The $G_\tau$ test statistic is given by:

$$G_\tau = \frac{1}{N}\sum_{i=1}^{N} \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)}, \tag{A.6}$$

where $SE(\hat{\alpha}_i)$ is the conventional standard error of $\hat{\alpha}_i$. Bootstrapped versions of these statistics perform well under cross-sectional dependence.

### A.1.5 Static CCE

We assume that $\beta_i'$ from equation (1.5) follows a random coefficient model and we are interested in estimating the cross-sectional mean, $\beta$, of these individual-specific slope coefficients. Pesaran (2006) proposes two versions of estimators: The mean group and the pooled CCE estimator. The mean-group estimator is given by a simple average of the individual CCE estimators $\hat{\beta}_i$ of $\beta_i$

$$\hat{\beta}_{cce} = N^{-1}\sum_{i=1}^{N} \hat{\beta}_i, \tag{A.7}$$

with

$$\hat{\beta}_i = (X_i'\bar{M}_w X_i)^{-1} X_i'\bar{M}_w y_i, \tag{A.8}$$

where $X_i = (x_{i1}, x_{i2}, \ldots, x_{iT})'$, $y_i = (y_{i1}, y_{i2}, \ldots, y_{iT})'$ and $\bar{M}_w = I_T - \bar{H}_w(\bar{H}_w'\bar{H}_w)^{-1}\bar{H}_w'$ with $I_T$ being an identity matrix and $\bar{H}_w$ consists of matrices of observations on the observed common effects $d_t$ and weighted cross-sectional averages $\bar{z}_{wt} = \sum_{j=1}^{N} w_j z_{jt}$ with $z_{it} = \binom{y_{it}}{x_{it}}$ and weights $\{w_j\}$. The weights satisfy following conditions: $w_i = O(\frac{1}{N})$, $\sum_{i=1}^{N} w_i = 1$ and $\sum_{i=1}^{N} |w_i| < K$. They are determined such that the asymptotic variance of the estimators of interest are minimized subject to these conditions. The weights are not unique and do not affect asymptotic distributions of the estimator. When $N$ is reasonably large, equal weights $w_i = 1/N$ may be used.

The pooled CCE estimator assumes homogeneous slope coefficients and is given by:

$$\hat{\beta}_P = (\theta_i X_i' \bar{M}_w X_i)^{-1} \theta_i X_i' \bar{M}_w y_i, \tag{A.9}$$

where $\theta_i$ are the pooling weights and usually set to equal $1/N$.

### A.1.6 Dynamic CCE - CS-DL

Consider the following ARDL(1,0) model:

$$y_{it} = \varphi_i y_{i,t-1} + \beta_i' x_{it} + \lambda_i' f_t + u_{it}, \qquad i = 1, 2, \ldots, N; \; t = 1, 2, \ldots T, \tag{A.10}$$

with

$$x_{it} = \delta_i' f_t + v_{it}, \tag{A.11}$$

where the observed common effects and deterministic components are omitted for simplification.

Multiplying (A.10) by $(1 - \varphi_i L)^{-1}$ yields:

$$\begin{aligned} y_{it} &= (1 - \varphi_i L)^{-1} \beta_i' x_{it} + (1 - \varphi_i L)^{-1} \lambda_i' f_t + (1 - \varphi_i L)^{-1} u_{it} \\ &= \theta_i x_{it} - \alpha_i'(L) \Delta x_{it} + \lambda_i' \bar{f}_{it} + \tilde{u}_{it}, \end{aligned} \tag{A.12}$$

where $\theta_i = \frac{\beta_i}{1-\varphi_i}$, $\Delta x_{it} = x_{it} - x_{i,t-1}$, $\alpha_i(L) = \sum_{l=0}^{\infty} \varphi_i^{l+1}(1-\varphi_i)^{-1}\beta_i L^l$,$\bar{f}_{it} = (1-\varphi_i L)^{-1} f_t$ and $\tilde{u}_{it} = (1 - \varphi_i L)^{-1} u_{it}$. This distributed lag specifications does not include lagged values of $y_{it}$ therefore CCE estimation procedure can be applied directly. Augmentation by the cross-section averages takes care of the effects of unobserved common factors and level regression of $y_{it}$ on $x_{it}$ is estimated by augmenting the individual regressions by differences of unit specific regressors $x_{it}$ and their lags. The goal is to estimate the mean long-run coefficients

$\theta = E(\theta_i)$. The CS-DL mean group estimator of the long-run coefficients is defined by

$$\hat{\theta}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i, \tag{A.13}$$

with

$$\hat{\theta}_i = (X_i' M_{qi} X_i)^{-1} X_i' M_{qi} y_i, \tag{A.14}$$

where $M_{qi} = I_{T-p} - Q_{wi}(Q_{wi}'Q_{wi})^{+}Q_{wi}'$ being a projection matrix with $Q_{wi} = (\bar{Z}_w, \Delta\bar{X}_{wp}, \Delta X_{ip})$ where $\bar{Z}_w = (\bar{z}_{w,p+1}, \bar{z}_{w,p+2}, \ldots, \bar{z}_{w,T})$, $\Delta\bar{X}_{wp} = \sum_{i=1}^{N} w_i \Delta X_{ip}$ and

$$\Delta X_{ip} = \begin{pmatrix} \Delta x_{i,p+1}' & \Delta x_{i,p}' & \ldots & \Delta x_{i2}' \\ \Delta x_{i,p+2}' & \Delta x_{i,p+2}' & \ldots & \Delta x_{i3}' \\ \vdots & \vdots & & \vdots \\ \Delta x_{iT}' & \Delta x_{i,T-1}' & \ldots & \Delta x_{i,T-p+1}' \end{pmatrix},$$

$p = p(T)$ is a chosen non-decreasing truncation lag function such that $0 \leq p < T$, and $A^{+}$ is the Moore-Penrose pseudoinverse of matrix $A$. The cross-section averages are defined by $\bar{z}_{wt} = (y_{wt}, \bar{x}_{wt}') = \sum_{i=1}^{N} w_i z_{it}$ with weights $w_i$ satisfying the granularity conditions $||w|| = O(N^{-\frac{1}{2}})$ and $\frac{w_i}{||w||} = o(N^{-\frac{1}{2}})$ uniformly for $i$ and the normalization condition $\sum_{i=1}^{N} w_i = 1$, where $w = (w_1, w_2, \ldots, w_N)'$ is an $N \times 1$ vector of weights.

The pooled version of the estimator is given by:

$$\hat{\theta}_P = (\omega_i X_i' M_{qi} X_i)^{-1} \omega_i X_i' M_{qi} y_i. \tag{A.15}$$

### A.1.7 CCE with IV

Assuming endogenous regressors in the sense that the regressors are correlated with the error term ($E(x_{it}, u_{it}) \neq 0$) would lead to inconsistent estimation in the case of static as well as dynamic CCE. Therefore, IV/2SLS may be included into the estimation procedure as explained in Neal (2015). Consider model (A.10), which can be rewritten as

$$A_{0i}\tau_{it} = A_{1i}\tau_{i,t-1} + C_i f_t + e_{it}, \tag{A.16}$$

where $\tau_{it} = (y_{it}, x_{it})$, $C_i = (\lambda_i, \delta_i)$ are the factor loadings, $e_{it} = (u_{it}, v_{it})$ is the error process, $A_{0i} = \begin{pmatrix} 1 & -\beta_i \\ 0 & I_k \end{pmatrix}$ and $A_{1i} = \begin{pmatrix} \varphi_i & 0 \\ 0 & 0 \end{pmatrix}$.

$A_{0i}$ is invertible and reduces model (A.16) to

$$\tau_{it} = A_{0i}^{-1} A_{1i} \tau_{i,t-1} + A_{0i}^{-1} C_i f_t + A_{0i}^{-1} e_{it}. \tag{A.17}$$

Furthermore, in the case of static CCE $A_{1i} = 0$ (because $\varphi_i = 0$) yielding

$$\tau_{it} = A_{0i}^{-1} C_i f_t + A_{0i}^{-1} e_{it}. \tag{A.18}$$

A set of $J$ instruments can be defined

$$Z_{iw} = \begin{pmatrix} z_{1i,pT+1} & \cdots & z_{Ji,pT+1} \\ z_{1i,pT+2} & \cdots & z_{Ji,pT+2} \\ \vdots & & \vdots \\ z_{1i,pT+T} & \cdots & z_{Ji,pT+T} \end{pmatrix} \tag{A.19}$$

which should be exogenous ($E(Z_{it} u_{it}) = 0$), linearly independent ($rank(Z_{iw}' Z_{iw}) = J$), sufficiently correlated with the regressors to contain full rank ($rank(Z_{iw}' \Xi_i) = K+1$), and satisfy the order condition ($J \geq K+1$) for the complete identification of coefficients.

In general, the set of instruments $Z_{iw}$ can be populated with any exogenous regressors, the cross section averages, and lags of the endogenous regressors and/or dependent variable. The set of first-stage fitted values is defined as

$$\hat{\Xi}_i = Z_i \zeta^{-1} Z_i' \Xi_i \tag{A.20}$$

where $\zeta$ s a positive semi-definite weight matrix and

$$\Xi_i = \begin{pmatrix} y_{i,pT} & x_{i,pT+1} \\ y_{i,pT+1} & x_{i,pT+2} \\ \vdots & \vdots \\ y_{i,T-1} & x_{i,T} \end{pmatrix}. \tag{A.21}$$

CCE or dynamic CCE can then be estimated with 2SLS by[2]

$$\hat{\pi}_{2SLS} = (\Xi_i' Z_i (Z_i' Z_i)^{-1} Z_i' \bar{M}_q Z_i (Z_i' Z_i)^{-1} Z_i' \Xi_i)^{-1} \Xi_i' Z_i (Z_i' Z_i)^{-1} Z_i' \bar{M}_q y_i \tag{A.22}$$

[2]The original formula in Neal (2015) for Eq. 36 contains typos and is corrected here.

where $\bar{M}_q = I_{T-pT} - \bar{Q}_w(\bar{Q}_w'\bar{Q}_w)^{-1}\bar{Q}_w$ is the projection matrix with

$$\bar{Q}_w = \begin{pmatrix} 1 & \bar{\tau}_{w,1} \\ 1 & \bar{\tau}_{w,2} \\ \vdots & \vdots \\ 1 & \bar{\tau}_{w,T} \end{pmatrix} \tag{A.23}$$

in the case of a static panel and

$$\bar{Q}_w = \begin{pmatrix} 1 & \bar{\tau}_{w,pT+1} & \bar{\tau}_{w,pT} & \dots & \bar{\tau}_{w,1} \\ 1 & \bar{\tau}_{w,pT+2} & \bar{\tau}_{w,pT+1} & \dots & \bar{\tau}_{w,2} \\ \vdots & \vdots & & \vdots & \\ 1 & \bar{\tau}_{w,T} & \bar{\tau}_{w,T-1} & \dots & \bar{\tau}_{w,T-pT} \end{pmatrix} \tag{A.24}$$

in case of a dynamic panel. $\bar{\tau}_{wt} = (\bar{y}_{wt}, \bar{x}_{wt}) = \sum_{i=1}^{N} w_i \tau_{wt}$ with the weights of Chudik and Pesaran (2015b).

## A.2 Guides

### A.2.1 Step-by-Step Guide

1. Test all variables relevant for estimation for weak cross-sectional dependence. Apply the test proposed by Pesaran (2015).

   – If the null of weak CSD cannot be rejected, proceed with first-generation panel unit root and cointegration tests.

     i If the variables are stationary, proceed with standard panel data models; see, e.g., Baltagi (2008).

     ii If variables are nonstationary and integrated of order one, proceed with a first-generation cointegration test. If there is sufficient evidence in favor of cointegration, appropriate cointegrating regression models should be applied. If there is no cointegrating relation, standard panel data models may be applied to first differences of the variables.

   – If the null of weak CSD is rejected, proceed with 2.

2. Second-generation unit-root test: Apply, e.g., the CIPS-test proposed by Pesaran (2007).

   – If the test indicates that all variables are stationary, proceed with CSD-robust panel models. These include variants of the panel common-correlated effects model (PCCE) proposed by Pesaran (2006). Alternatively, a principal-components model proposed by Bai (2009) may be considered. Use instrumental-variable extensions of these estimators when endogeneity of the regressors is suspected.

   – If there is sufficient evidence for nonstationary series that are integrated of order one, $I(1)$, proceed with 3.

3. Second-generation cointegration test: Apply a CSD-robust test, e.g., the one proposed by Westerlund (2007).

   – If there is a lack of sufficient evidence for cointegration and one is interested the estimation of long-run relations, static CCE can still produce consistent estimates when the defactored CCE residuals are stationary. Dynamic versions of CCE can also produce consistent estimates of long-run relations. The CS-DL model appears particularly robust in such settings. Alternatively (or in addition), transform the

$I(1)$ variables using first differences and apply static and/or dynamic CSD-robust panel models.

- If there is sufficient evidence in favor of cointegration, one can proceed with a static CCE model as well as dynamic specifications including error-correction variants. The latter are the preferred option as they can distinguish between long-run and short-run effects as well as the speed of adjustment towards the long-run equilibrium following shocks.
- Use instrumental-variable extensions of these estimators when endogeneity of the regressors is suspected. Bias-corrected versions of the dynamic estimators as well as bootstrapped standard errors and confidence intervals may be considered.

4. Test the residuals post estimation:

   i. Check residuals for CSD. Residuals from a CCE model should be tested with the weighted-CSD test proposed by Juodis and Reese (2021).

   ii. Test residuals for a unit root with an appropriate (first-generation) unit-root test, e.g., IPS (Im, Pesaran, and Shin, 2003).

   iii. Test for remaining unobserved common factors, e.g., Gagliardini et al. (2019).

   When a version of the CCE estimator is applied, consistency requires that the number of unobserved common factors does not exceed the number of regressors. This assumption should be tested.

   CCE estimators are available in pooled and mean-group version. Test for parameter homogeneity when necessary. This is more important in dynamic relative to static specifications.

### A.2.2 STATA Packages and Literature

**Table A.1:** STATA Packages for Testing and CCE-type Estimation under CSD

| Package | Description | References |
|---|---|---|
| **Pre-Estimation** | | |
| xtcd2 | Tests for weak cross-sectional dependence | Pesaran (2015); Ditzen (2016a) |
| xtcips | Second-generation panel unit-root test | Pesaran (2007); Sangiacomo (2014) |
| xtwest | Second-generation cointegration test | Westerlund (2007); Persyn and Westerlund (2008) |
| xthst | Tests for heterogeneous slopes | Pesaran and Yamagata (2008); Bersvendsen and Ditzen (2021) |
| xtnumfac | Dynamic specification of common-correlated effects models require that the number of unobserved common factors is lower than the number of cross-sectional averages for consistent estimation. | Chudik and Pesaran (2015b); Ahn and Horenstein (2013); Onatski (2010); Ditzen and Reese (2023) |
| **Estimation** | | |
| xtdcce2 | Implements pooled and mean-group common-correlated effects models. The package can implement the static as well as the dynamic versions of this model. The dynamic versions include the CS-DL and CS-ECM models. Additionally, the models can be augmented by IV specifications | Pesaran (2006); Kapetanios et al. (2011); Chudik and Pesaran (2015b); Chudik et al. (2016); Ditzen (2016b) |
| xtivdfreg | Combines CCE and PC approach to accommodate endogenous regressors | Norkutė et al. (2021); Kripfganz and Sarafidis (2021) |
| **Post-Estimation** | | |
| xtcd2 | Checks residuals for remaining cross-sectional dependence. For fixed-effects of common-correlated effects estimators the weighted version of the test should be applied. | Juodis and Reese (2021); Ditzen (2016a) |
| xtnumfac | Estimates the number of remaining unobserved common factors. When applied post-estimation, the GOL criterion is the relevant statistic. | Gagliardini et al. (2019); Ditzen and Reese (2023) |

**Table A.2:** Overview of the Related Econometric Literature

| **Testing** | |
|---|---|
| Description | References |
| DF-type unit-root tests under CSD - CIPS. | Pesaran (2007) |
| PC-based unit root and cointegration tests under CSD. | Bai and Ng (2004) |
| ADF-type tests for panels with heteroscedastic errors. | Cavaliere (2005) and Westerlund (2014) |
| CIPS test for multiple common factors and structural breaks. | Lee et al. (2016) and Pesaran et al. (2013) |
| Unit root panel nulls considering heterogeneous stationarity. | Pesaran (2012) |
| Error-correction based cointegration tests with CSD. | Westerlund (2007) |
| Residual-based cointegration tests accommodating CSD, heteroscedasticity, serial correlation, and structural breaks. | Westerlund and Edgerton (2008) |
| Cointegration and unit-root tests with mixed $I(1)$ and $I(0)$ series. | Trapani (2021) |
| **CCE Estimation** | |
| Description | References |
| Static CCE model with stationary factors. | Pesaran (2006) |
| Static CCE model with nonstationary factors | Kapetanios et al. (2011) |
| Dynamic ARDL-type CCE model (CS-ARDL). | Chudik and Pesaran (2015b) |
| Dynamic distributed-lag type CCE model (CS-DL). | Chudik et al. (2016) |
| Error-correction type dynamic CCE model (CS-ECM). | Ditzen (2021) |
| Endogenous regressors in homogeneous panels. | Harding and Lamarche (2011) |
| Endogeneity in heterogeneous panels. | Forchini et al. (2015) |
| Endogenous regressors in dynamic heterogeneous panels with lagged instruments. | Neal (2015) |
| CCE with structural breaks in two time regimes. | Baltagi et al. (2016) |
| CCE with endogenous regressors and unknown common structural breaks. | Baltagi et al. (2019) |
| **Alternative Estimation** | |
| Description | References |
| Principal Components (PC) method for handling CSD. | Bai (2009) and Coakley et al. (2002) |
| Comparison of CCE and PC methods under multifactor error structures. | Karabiyik et al. (2019) and Westerlund and Urbain (2015) |
| Performance analysis of CCE vs. iterative PC under various conditions. | Sarafidis and Wansbeek (2012) |
| PC-based IV estimators for large dynamic panels with multifactor error structure. | Norkutė et al. (2021) |
| Mean-group fully-modified OLS for polynomial cointegration under CSD. | Wagner and Reichold (2023) |
| PC with hetreogeneous coefficients in an dynamic setting. | Song (2013) |

# Appendix B

# Appendix for Chapter 2

## B.1 Descriptives

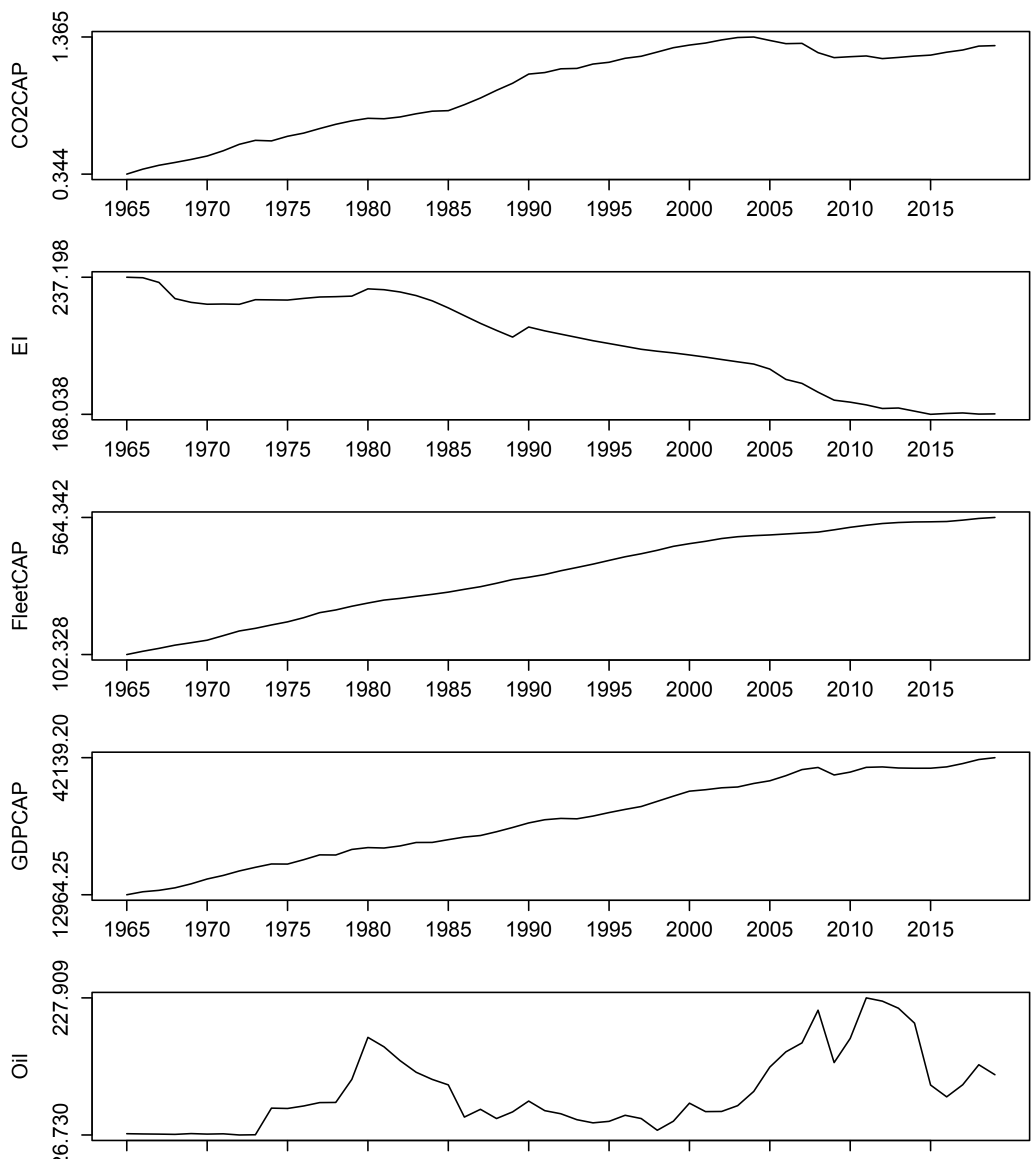
**Table B.1:** Description of Variables

| Variable | Description |
|---|---|
| CO2/CAP | CO2 emissions from combustion engine passenger cars (diesel, petrol, hybrids, and plug-in hybrids) in 1000t divided by the average population in a given year in 1000 persons. |
| EI | Energy intensity measured by grams CO2 emitted per 100km |
| Fleet/CAP | Total number of passenger cars with combustion engines (including hybrids and plug-in hybrids) divided by the average population in a given year in 1000 persons. |
| GDP/CAP | Real gross domestic product in 2015 prices divided by the average population in a given year. |
| Oil | Real international oil prices in 2015 prices (WTI up to 1986, BRENT thereafter). |

**Table B.2:** Summary Statistics of Variables

| Variable | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|
| CO2/CAP | 55 | 0.981 | 0.321 | 0.344 | 0.728 | 1.250 | 1.365 |
| EI | 55 | 204.994 | 22.368 | 168.038 | 188.258 | 225.730 | 237.198 |
| Fleet/CAP | 55 | 372.211 | 145.580 | 102.328 | 258.820 | 506.959 | 564.342 |
| GDP/CAP | 55 | 28,972.920 | 8,957.980 | 12,964.250 | 22,032.390 | 37,761.610 | 42,139.200 |
| Oil | 55 | 89.790 | 56.120 | 26.730 | 49.994 | 122.567 | 227.909 |

**Figure B.1:** Time Series in Levels, 1965-2019



## B.2 Policy Stringencies

**Table B.1:** Score Assignment to Policies

| Fuel Tax | SFC Tax | Insurance Tax | Score |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| $0 < c_t < 0.0996$ | $0 < c_t < 0.1$ | $0 < c_t < 82.42$ | 1 |
| $0.0996 \leq c_t < 0.1878$ | $0.1 \leq c_t < 0.11$ | $82.42 \leq c_t < 129.96$ | 2 |
| $0.1878 \leq c_t < 0.2759$ | $0.11 \leq c_t < 0.12$ | $129.96 \leq c_t < 177.50$ | 3 |
| $0.2759 \leq c_t < 0.3642$ | $0.12 \leq c_t < 0.13$ | $177.50 \leq c_t < 225.04$ | 4 |
| $0.3642 \leq c_t < 0.4524$ | $0.13 \leq c_t < 0.14$ | $225.04 \leq c_t < 272.58$ | 5 |
| $c_t \geq 0.4524$ | $c_t \geq 0.14$ | $c_t \geq 272.58$ | 6 |

Note: Score assignments are shown for policies which can be attributed a cost in monetary terms. Scores range from 0 (not in effect) to 6 (most stringent implementation in the observed period). The first three columns show the intervals that are matched with policy costs $c_t$ for a given year and the last column shows the associated scores with these intervals. A policy that is not in effect is assigned a score of 0 and a policy that has cost higher or equal than the 90th percentile of the distribution of the costs of a policy over the years is assigned a score of 6. The remaining scores are assigned according to the bins with width $w$. The policy with cost $c_t$ in a given year is then assigned the score according to the associated bin.

**Table B.2:** Policy costs and scores

| Year | Ins Tax | | SFC Tax | | Fuel Tax | | Speed Limit | | Car-Free Days | | IG-L | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. |
| 1950 | NA | 0 | NA | 0 | 0.0114 | 1 | NA | 0 | NA | 0 | NA | 0 |
| 1952 | 34.88 | 1 | NA | 0 | 0.0114 | 1 | NA | 0 | NA | 0 | NA | 0 |
| 1960 | 34.88 | 1 | NA | 0 | 0.0202 | 1 | NA | 0 | NA | 0 | NA | 0 |
| 1961 | 34.88 | 1 | NA | 0 | 0.1140 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1966 | 34.88 | 1 | NA | 0 | 0.1284 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1971 | 34.88 | 1 | NA | 0 | 0.1455 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1973 | 34.88 | 1 | NA | 0 | 0.1455 | 2 | 0.25 | 1.5 | NA | 0 | NA | 0 |
| 1974 | 34.88 | 1 | NA | 0 | 0.1455 | 2 | 0.25 | 1.5 | 0.0833 | 0.5 | NA | 0 |
| 1975 | 34.88 | 1 | NA | 0 | 0.1455 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1977 | 104.65 | 2 | NA | 0 | 0.1455 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1978 | 104.65 | 2 | 0.12 | 4 | 0.1455 | 2 | NA | 0 | NA | 0 | NA | 0 |
| 1980 | 104.65 | 2 | 0.12 | 4 | 0.2186 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1981 | 104.65 | 2 | 0.12 | 4 | 0.2333 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1984 | 156.97 | 3 | 0.12 | 4 | 0.2333 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1985 | 156.97 | 3 | 0.12 | 4 | 0.2262 | 3 | NA | 0 | NA | 0 | NA | 0 |

Continued on next page

**Table B.2 – continued from previous page**

| Year | Ins Tax | | SFC Tax | | Fuel Tax | | Speed Limit | | Car-Free Days | | IG-L | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. | Cost | Scr. |
| 1987 | 156.97 | 3 | 0.12 | 4 | 0.2338 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1992 | 156.97 | 3 | 0.09 | 1 | 0.2668 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1993 | 180.04 | 4 | 0.09 | 1 | 0.2668 | 3 | NA | 0 | NA | 0 | NA | 0 |
| 1994 | 180.04 | 4 | 0.09 | 1 | 0.2906 | 4 | NA | 0 | NA | 0 | NA | 0 |
| 1995 | 180.04 | 4 | 0.09 | 1 | 0.3641 | 4 | NA | 0 | NA | 0 | NA | 0 |
| 1997 | 180.04 | 4 | 0.09 | 1 | 0.3641 | 4 | NA | 0 | NA | 0 | 1 | 6 |
| 2000 | 272.58 | 6 | 0.09 | 1 | 0.3641 | 4 | NA | 0 | NA | 0 | 1 | 6 |
| 2004 | 272.58 | 6 | 0.09 | 1 | 0.3769 | 5 | NA | 0 | NA | 0 | 1 | 6 |
| 2005 | 272.58 | 6 | 0.09 | 1 | 0.3752 | 5 | NA | 0 | NA | 0 | 1 | 6 |
| 2007 | 272.58 | 6 | 0.09 | 1 | 0.4089 | 5 | NA | 0 | NA | 0 | 1 | 6 |
| 2010 | 272.58 | 6 | 0.11 | 3 | 0.4089 | 5 | NA | 0 | NA | 0 | 1 | 6 |
| 2011 | 272.58 | 6 | 0.11 | 3 | 0.4524 | 6 | NA | 0 | NA | 0 | 1 | 6 |
| 2013 | 272.58 | 6 | 0.12 | 4 | 0.4524 | 6 | NA | 0 | NA | 0 | 1 | 6 |
| 2014 | 310.06 | 6 | 0.14 | 6 | 0.4524 | 6 | NA | 0 | NA | 0 | 1 | 6 |
| 2016 | 310.06 | 6 | 0.15 | 6 | 0.4524 | 6 | NA | 0 | NA | 0 | 1 | 6 |

Note: Only years in which policy changes took place are shown. Costs for Ins Tax and Fuel Tax are in EUR and for SFC Tax in percent of the price of the average new vehicle. Costs for qualitative measures (Speed Limit, Car-Free Days, IG-L) are indicated by a dummy variable. This dummy takes a value of 1 if the policy was in effect throughout an entire year and is otherwise weighted according to the fraction of a year that it was in force. Only the IG-L legislation is in full effect since its introduction in 1997 and receives a cost of 1 and score of 6. Speed limits were in effect for one quarter in 1973 and 1974, respectively. Car-free days were in effect for only one month in 1973.

## B.3 Model Adequacy Tests

**Table B.1:** Elliott, Rothenberg, and Stock (1996) Unit Root Test (DF-GLS)

| | Levels | | Differenced | |
|---|---|---|---|---|
| Variable | trend | constant | trend | constant |
| EI | $-1.850$ | 0.700 | $-3.700^{***}$ | $-3.430^{***}$ |
| Fleet/CAP | $-0.640$ | $-0.020$ | $-3.420^{**}$ | $-2.090^{**}$ |
| CO2/CAP | $-0.460$ | 0.280 | $-3.210^{**}$ | $-2.480^{**}$ |
| GDP/CAP | $-0.630$ | 0.680 | $-5.110^{***}$ | $-3.890^{***}$ |
| Oil | $-2.190$ | $-1.240$ | $-5.480^{***}$ | $-5.270^{***}$ |
| Comp | $-2.470$ | 0.660 | $-5.710^{***}$ | $-5.440^{***}$ |
| Invest | $-2.070$ | 0.160 | $-5.310^{***}$ | $-5.170^{***}$ |
| Use | $-1.870$ | $-0.030$ | $-5.800^{***}$ | $-5.660^{***}$ |

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$; Null hypothesis: unit root.

**Table B.2:** Johansen trace test, with 2 lags and linear trend

| cointegrating vectors r | test | p-value |
|---|---|---|
| r <= 2 | 3.34 | 0.8264 |
| r <= 1 | 10.99 | 0.8707 |
| r = 0 | 27.91 | 0.6311 |

Note: Null hypothesis: number of cointegrating vectors is r.

**Table B.3:** VAR Order Selection Criteria

| | Eq. (2.1) | Eq. (2.2) | Eq. (2.3) | Eq. (2.4) |
|---|---|---|---|---|
| AIC | 2 | 1 | 3 | 1 |
| HQ | 1 | 1 | 1 | 1 |
| SC | 1 | 1 | 1 | 1 |
| FPE | 2 | 1 | 1 | 1 |

Note: The VAR order selected by the respective information criteria are shown for model specifications based on the respective Eq. numbers.

**Table B.4:** Edgerton and Shukur (1999) test for residual autocorrelation

| | Eq. (2.1) | Eq. (2.2) | Eq. (2.3) | Eq. (2.4) |
|---|---|---|---|---|
| Order | P-Value | | | |
| 1 | 0.0745 | 0.3333 | 0.7749 | 0.1360 |
| 2 | 0.1050 | 0.2070 | 0.4168 | 0.2769 |
| 3 | 0.3672 | 0.3828 | 0.2952 | 0.0963 |
| 4 | 0.5257 | 0.8202 | 0.7327 | 0.3823 |
| 5 | 0.4486 | 0.4537 | 0.3783 | 0.3853 |

Note: P-values are shown for model specifications based on the respective Eq. numbers. Null hypothesis: no residual autocorrelation.

**Table B.5:** Doornik and Hendry (1997) multivariate LM-test for ARCH effects in residuals

| | Eq. (2.1) | Eq. (2.2) | Eq. (2.3) | Eq. (2.4) |
|---|---|---|---|---|
| Order | P-Value | | | |
| 1 | 0.9189 | 0.9504 | 0.8354 | 0.8436 |
| 2 | 0.6279 | 0.4348 | 0.5714 | 0.559 |
| 3 | 0.5087 | 0.4098 | 1.0000 | 1.0000 |
| 4 | 0.6808 | 0.9999 | 1.0000 | 1.0000 |
| 5 | 0.6169 | 1.0000 | 1.0000 | 1.0000 |

Note: P-values are shown for model specifications based on the respective Eq. numbers. Null hypothesis: no ARCH in residuals.

# Appendix C

# Appendix for Chapter 3

## C.1 TWFE Regression Output

**Table C.1:** TWFE regression for specification projected with all covariates and only adjusted for COVID-related controls

| | (1) | | (2) | |
|---|---|---|---|---|
| | Coef. | SE | Coef. | SE |
| asinh(cases) | $-0.0300^{***}$ | (0.0032) | $-0.0181^{***}$ | (0.0060) |
| asinh(nvrwfh) | 0.0412 | (0.0303) | $0.1881^{***}$ | (0.0553) |
| asinh(wfh) | $-0.0446^{***}$ | (0.0071) | $-0.0443^{***}$ | (0.0106) |
| log(gdp) | $0.2649^{***}$ | (0.0776) | | |
| ei | $0.0036^{***}$ | (0.0004) | | |
| diesel | $-0.6570^{***}$ | (0.0747) | | |
| petrol | 0.0631 | (0.1274) | | |
| log(frt) | 0.0001 | (0.0081) | | |
| Obs | 822 | | 822 | |
| N | 137 | | 137 | |

Note: Dependent variable is log of CO2 per captia, $log(co2)$, standard errors are in parantheses and clustered at the regional level. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

## C.2 Unit Weights and Pre-Treatment Trends

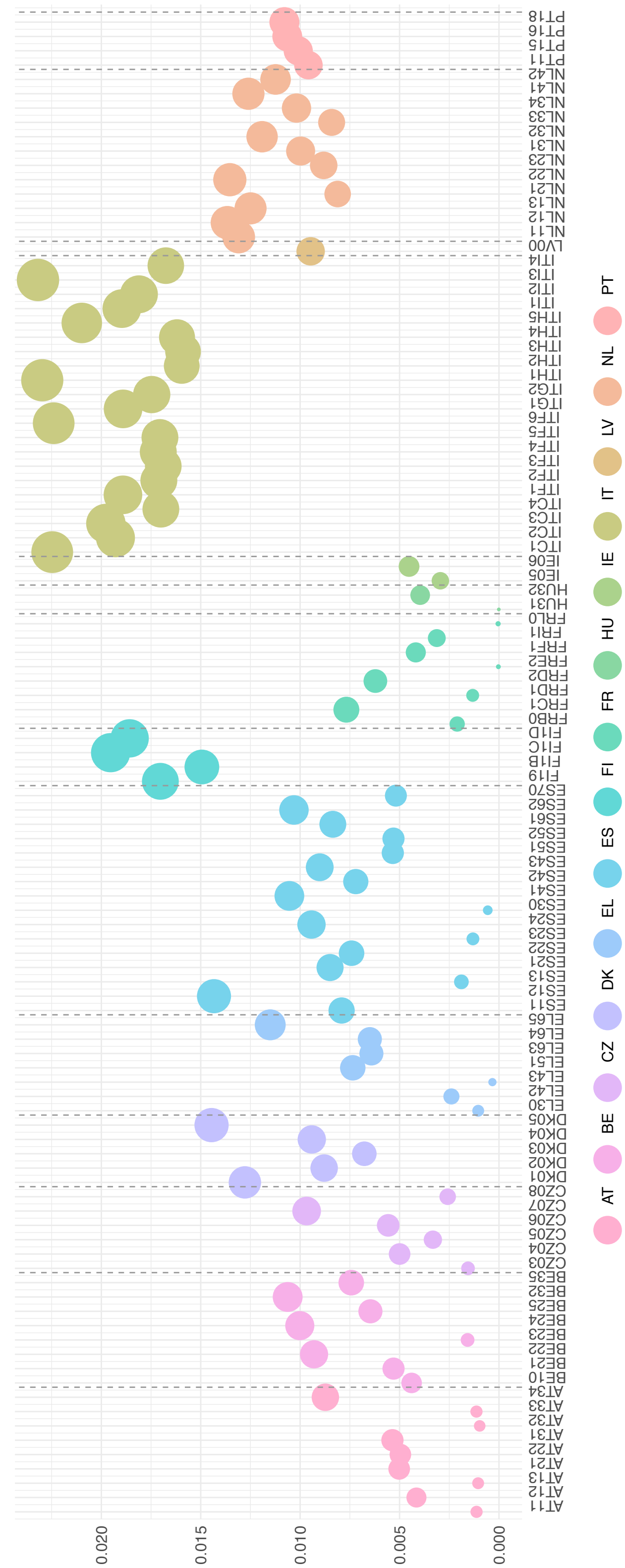
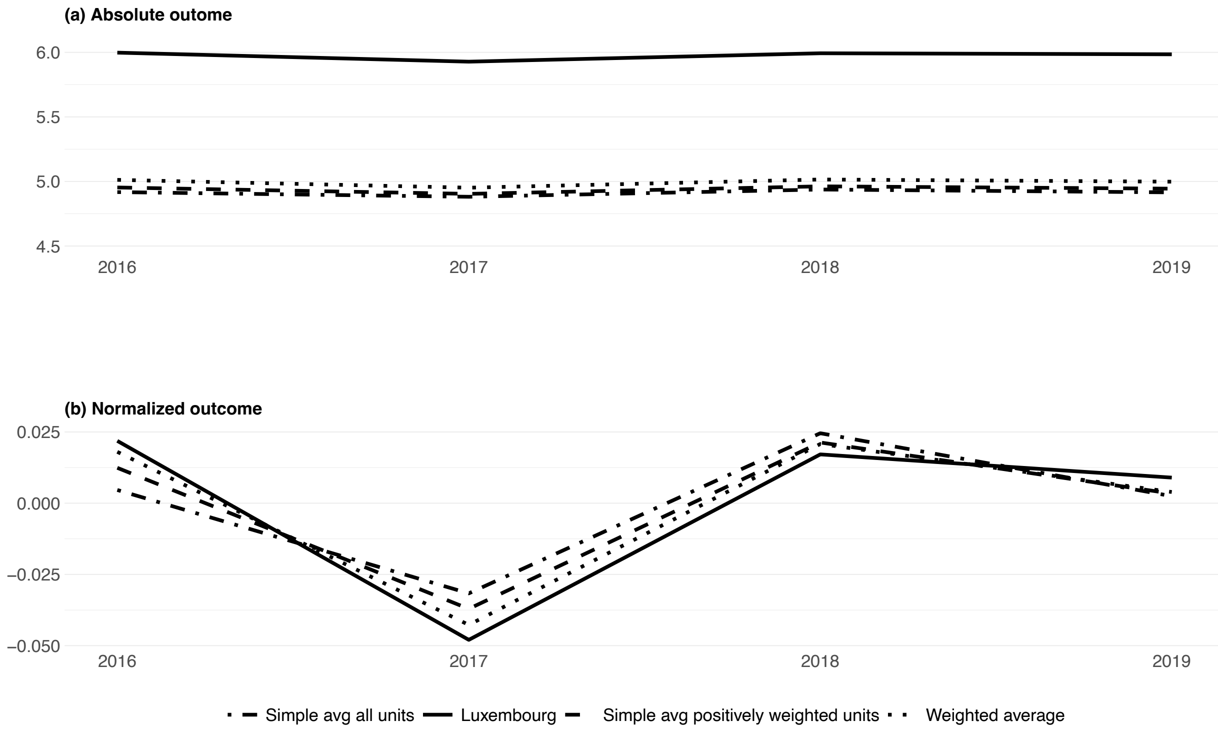



**Figure C.1:** Unit weights - all covariates

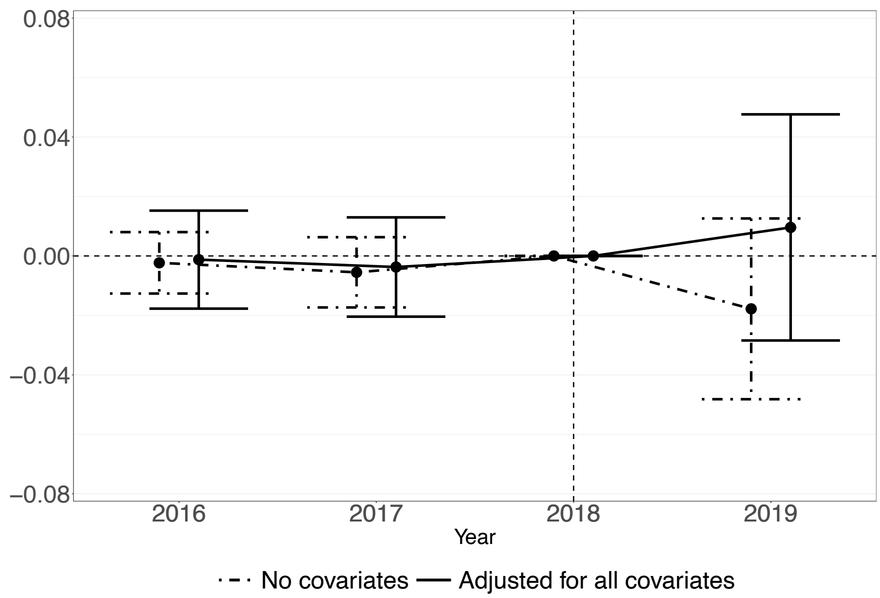
**Figure C.2:** Pre-treatment trends of the adjusted log $CO_2$ per capita emissions




Note: *Luxembourg* is the pre-treatment time series trend for Luxembourg (treated unit). *Simple avg all units* is the pre-treatment average trend of all units in the donor pool. *Simple avg positively weighted units* is the pre-treatment average trend of the units in the donor pool that received positive weights. *Weighted average* is the pre-treatment weighted average of the units that received positive weights.

## C.3 Robustness Tests

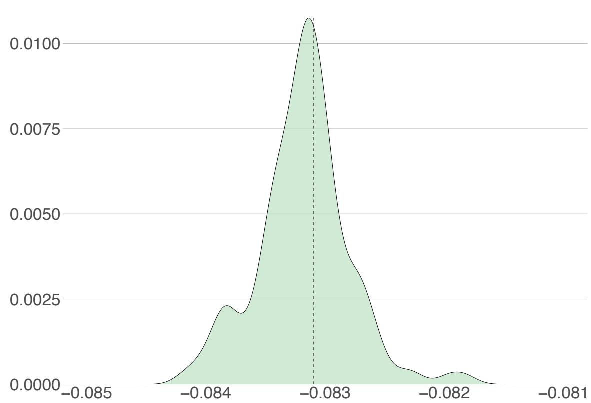
**Figure C.1:** In-time placebo test



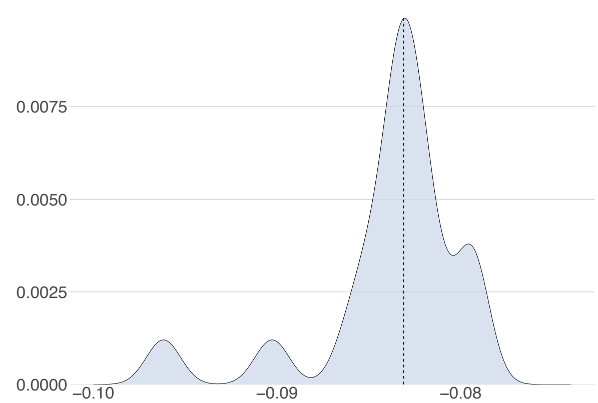

Note: Results are re-estimated by back dating the policy to 2019, prior to the actual policy implementation.

**Figure C.2:** Distribution of ATT: leave one out analysis

**(a)** Dropping a regions from the donor pool

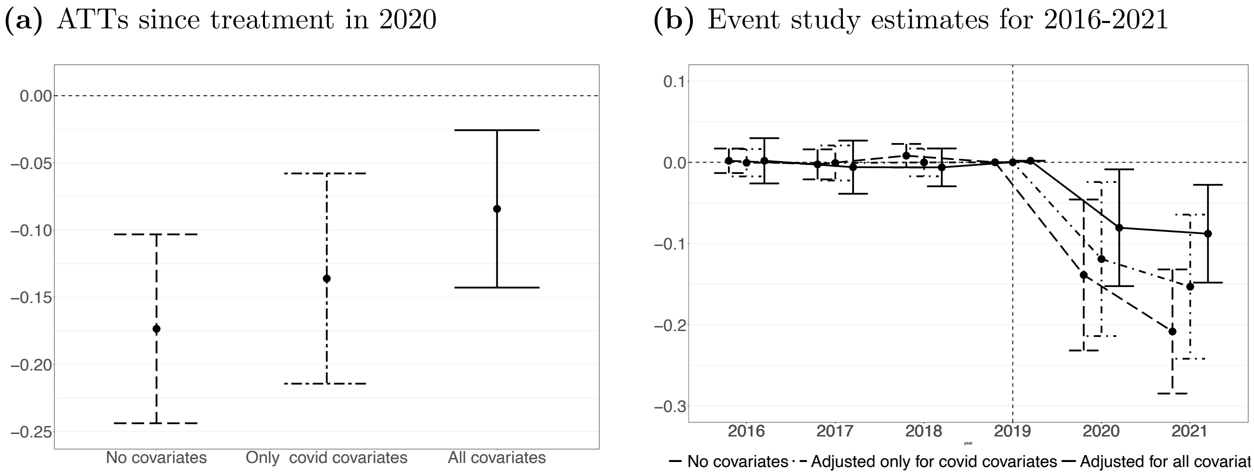



**(b)** Dropping a country from the donor pool



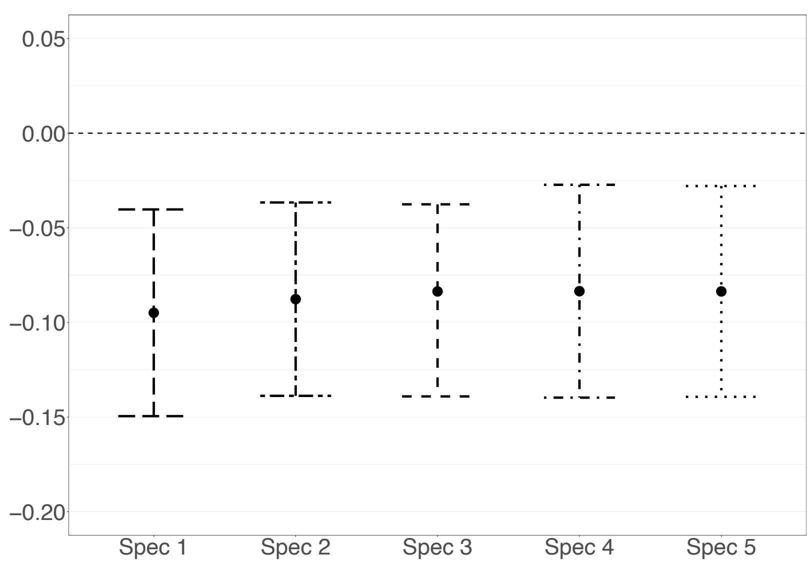

Note: Panel (a) presents the distribution of ATT estimates obtained by iteratively excluding one region at a time and re-estimating the SDID model. Panel (b) displays the distribution of ATT estimates obtained by iteratively excluding one country at a time and re-estimating the SDID model.

**Figure C.3:** ATTs and event study estimates restricted sample




Note: ATTs and event study estimates of the estimated impact of free public transport on road emissions $CO_2$ per capita in Luxembourg using the restricted sample for different model specifications with 95% confidence bands based on placebo estimates.

**Figure C.4:** ATTs across different model specifications




Note: Spec 1 excludes controls for freight transport; Spec 2 excludes controls for working from home; Spec 3 excludes controls for both freight and working from home, Spec 4 excludes controls for commuting (never working from home); Spec 5 excludes controls for both freight and commuting.

**Table C.1:** Sensitivity analysis across different model specifications

| | (1) lco2cap | (2) lco2cap | (3) lco2cap | (4) lco2cap | (5) lco2cap |
|---|---|---|---|---|---|
| acases | -0.0300*** | -0.0296*** | -0.0294*** | -0.0306*** | -0.0306*** |
| | (0.00315) | (0.00305) | (0.00303) | (0.00327) | (0.00320) |
| lgdp | 0.265*** | 0.268** | 0.270** | 0.259*** | 0.259*** |
| | (0.0774) | (0.0946) | (0.0942) | (0.0770) | (0.0768) |
| anever_all | 0.0412 | 0.125*** | 0.125*** | | |
| | (0.0301) | (0.0354) | (0.0353) | | |
| ausual | -0.0446*** | | | -0.0479*** | -0.0479*** |
| | (0.00715) | | | (0.00662) | (0.00662) |
| ei | 0.00359*** | 0.00304*** | 0.00304*** | 0.00366*** | 0.00366*** |
| | (0.000383) | (0.000390) | (0.000389) | (0.000380) | (0.000380) |
| diesel_real | -0.657*** | -0.650*** | -0.653*** | -0.670*** | -0.671*** |
| | (0.0738) | (0.0765) | (0.0759) | (0.0736) | (0.0726) |
| super_real | 0.0631 | -0.0115 | -0.0101 | 0.0635 | 0.0639 |
| | (0.127) | (0.129) | (0.129) | (0.127) | (0.126) |
| lload | | 0.00314 | | 0.000612 | |
| | | (0.00741) | | (0.00780) | |
| Observations | 822 | 822 | 822 | 822 | 822 |

Note: Size Standard errors in parentheses. Dependent variable is log(co2cap).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

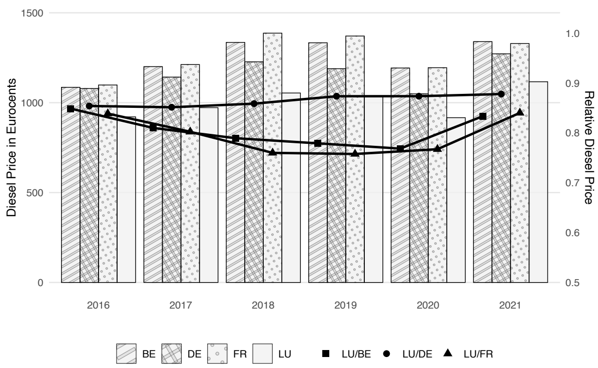
**Table C.2:** Pre- and post-treatment averages of relative fuel prices for Luxembourg

| | Diesel | | Petrol | |
|---|---|---|---|---|
| | Pre-Avg | Post-Avg | Pre-Avg | Post-Avg |
| BE | 0.7825 | 0.8028 | 0.8869 | 0.8814 |
| DE | 0.8684 | 0.8759 | 0.8493 | 0.8368 |
| FR | 0.7585 | 0.8056 | 0.8001 | 0.8182 |

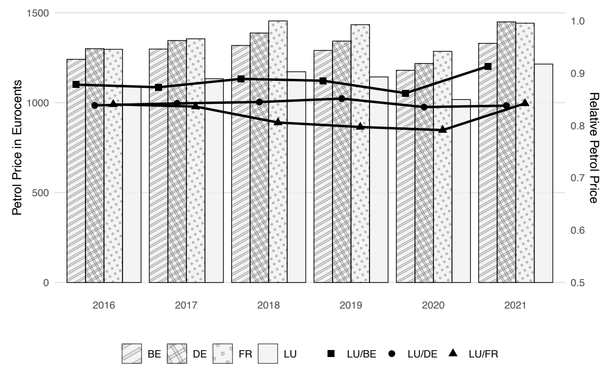
Note: Relative fuel prices of LU with respect to its neighboring countries. Pre-Avg are relative fuel prices based on time-weighted pre-treatment fuel prices, where time weights are taken from the SDiD main specification. Post-Avg are relative fuel prices based on post-treatment fuel prices.

**Figure C.5:** Absolute and relative fuel prices for LU and neighbouring countries

**(a)** Diesel



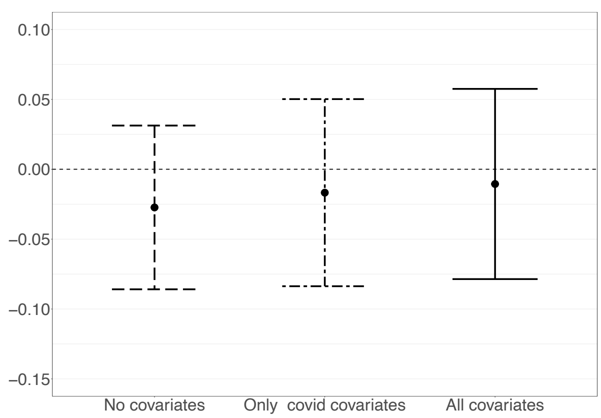

**(b)** Petrol



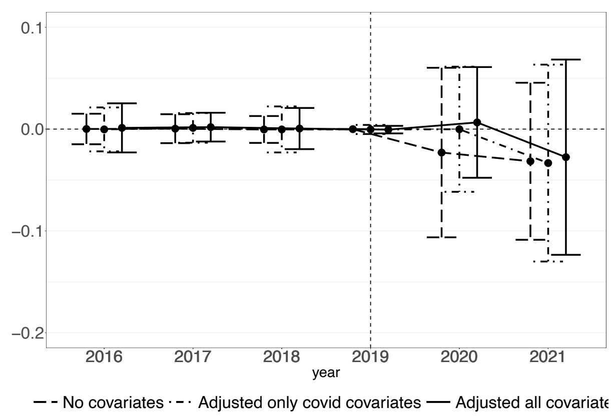

Note: Bars show fuel prices in Eurocents per 1,000 litres adjusted for inflation (HICP). Lines indicate fuel prices of Luxembourg relative to its neighbouring countries over time.

**Figure C.6:** Energy use in the building sector

**(a)** ATT since treatment in 2020




**(b)** Event study estimates for 2016-2021

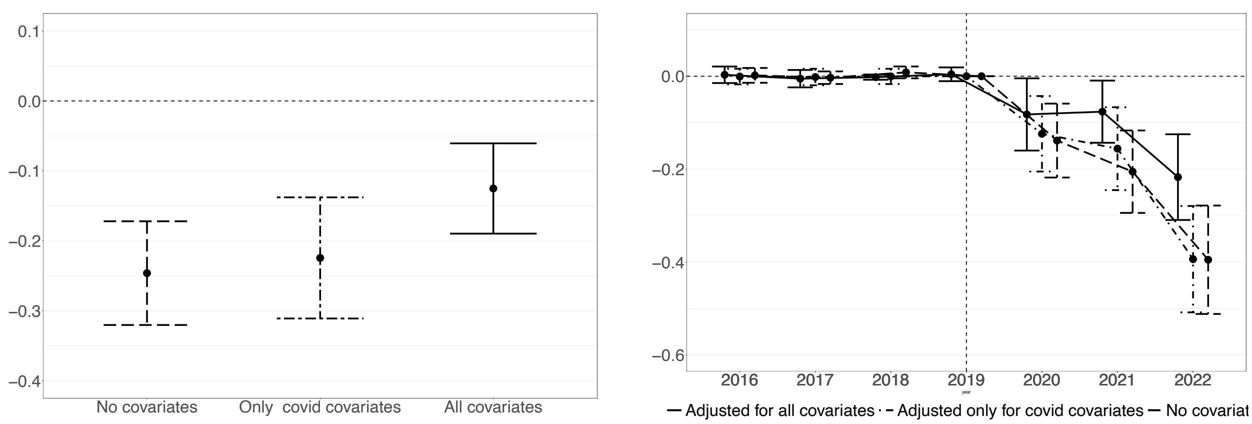



Note: ATTs and event study estimate of the estimated impact of COVID-19 on $CO_2$ emissions from the Energy use in the building sector. The specification with all covariates includes controls for the log of GDP per capita, commuting, working from home, COVID-19 cases, and additionally employment rate.

**Figure C.7:** Extending the post-treatment period to 2022

**(a)** ATT since treatment in 2020

**(b)** Event study estimates for 2016-2022




Note: The sample used for the extended post-treatment period has 129 regions (instead of 137). We lose 8 regions, 6 from Greece (EL42, EL43, EL51, EL63, and EL64), and 2 Italian regions (ITC4, and ITF2) due to missing data in covariates).

# Appendix D

# Appendix for Chapter 4

## D.1 Time Series Models

The OLS model is specified and estimated in first differences:

$$\Delta Y_t = \alpha + \beta \Delta X_{j,t} + \epsilon_t,$$

where $Y$ represents the dependent variable, which may refer to GHG emissions, while $X$ denotes the set of explanatory variables ($j = 1, \ldots, k$), which may include Heating Degree Days (HDD) or GDP. The autoregressive distributed lag model of the first order, i.e., ARDL(1,1), is estimated in levels and includes one lag of both the dependent and independent variables:

$$Y_t = \alpha + \lambda Y_{t-1} + \beta X_t + \gamma X_{t-1} + \epsilon_t.$$

If the dependent and independent variables exhibit a long-term equilibrium relationship, the ARDL model can be written as an error correction model (ECM). This transformation allows for more efficient estimation and distinguishes between the adjustment mechanism to deviations from equilibrium and short-term effects. The ECM can be written as (e.g., Greene (2018)):

$$\Delta Y_t = \alpha + \beta \Delta X_t - \phi(Y_{t-1} - \theta X_{t-1}) + \epsilon_t,$$

where:

$$\phi = 1 - \lambda \quad \text{and} \quad \theta = -\frac{\beta + \gamma}{1 - \lambda}.$$

The term $\phi(Y_{t-1} - \theta X_{t-1})$ is the error correction term and indicates the speed of adjustment back to equilibrium after a deviation.

The ARDL model can also be expressed as an infinite distributed lag model:

$$Y_t = \sum_{i=0}^{\infty} \rho^i X_{t-i} + \epsilon_t.$$

This form allows us to trace the dynamic effect of changes in the price index over multiple periods. These multipliers are then given by:

$$M_k = \sum_{i=0}^{k} \rho^i.$$

The multipliers are based on a level-log model. A 1% change in the price index leads to a change in the share of renewables by $M_k$ units. For 2023, the price index increased by 16.5%. The contemporaneous effect on the share of renewables is:

$$\Delta R = M_0 \cdot 16.5\%.$$

For the buildings sector, we estimate an increase in the share of renewables by 0.4 percentage points. The overall increase in the share of renewables in 2023 is 6.7 percentage points. Thus, contemporaneous price effects explain about 6% of the change. Effects of past price changes can be calculated similarly and lead to the results presented in Table 4.3.

## D.2 Detailed Results and Model Adequacy

### Regression Results

Table D.1 shows the regression results based on the OLS model and the ARDL model in its equivalent error correction form (ECM).

**Table D.1:** Regression Results for OLS and ARDL / ECM

| | Buildings | | | Other Sectors | | |
|---|---|---|---|---|---|---|
| Variable | OLS | ARDL | ECM | OLS | ARDL | ECM |
| Δlog(HDD) | $0.6958^{***}$ | | | | | |
| | (0.0993) | | | | | |
| Δlog(GDP) | | | | $0.4599^{*}$ | | |
| | | | | (0.2061) | | |
| ΔShare Renew | $-0.0351^{***}$ | | | $-0.0158^{***}$ | | |
| | (0.0045) | | | (0.0029) | | |
| Lag(Share Renew) | | $0.7095^{***}$ | | | $0.5393^{***}$ | |
| | | (0.0787) | | | (0.1385) | |
| ECT | | | $-0.2905^{***}$ | | | $-0.4607^{**}$ |
| | | | (0.0734) | | | (0.1316) |
| log(Pidx) | | 3.6409 | | | −0.4757 | |
| | | (2.9012) | | | (3.9891) | |
| Lag(log(Pidx)) | | 6.6179 | | | 6.0775 | |
| | | (4.243) | | | (4.638) | |
| Δlog(Pidx) | | | 3.6409 | | | −0.4757 |
| | | | (2.8165) | | | (3.8582) |
| Intercept | 0.0074 | $-32.5610^{***}$ | $-32.5610^{***}$ | −0.0051 | $-17.7431^{*}$ | $-17.7431^{**}$ |
| | (0.0043) | (8.5601) | (8.3550) | (0.0053) | (6.4420) | (5.1138) |
| $T$ | 32 | 33 | 33 | 32 | 33 | 33 |
| $R^2$ | 0.907 | 0.987 | 0.405 | 0.543 | 0.8282 | 0.2989 |
| Adjusted $R^2$ | 0.901 | 0.9857 | 0.364 | 0.513 | 0.8098 | 0.2505 |

*Note:* $^{*}p < 0.1$. $^{**}p < 0.05$. $^{***}p < 0.01$. OLS is estimated in first differences; ARDL is an Autoregressive Distributed Lag model with a single lag. ECM is the error correction form of the ARDL model, where ECT represents the error correction term.

### Model Diagnostics

Table D.2 presents tests for model fit. This includes results of the Breusch-Godfrey LM test (Breusch, 1978; Godfrey, 1978) for autocorrelation and the Ljung-Box ARCH test (Ljung & Box, 1978) for heteroscedasticity, each for up to 1, 2, 3, and 4 lags. The null hypotheses are no autocorrelation and no heteroscedasticity, respectively. Autocorrelation tests for the OLS model reject the null for up to 1–3 lags. Therefore, robust standard errors are estimated for the OLS model.

**Table D.2:** Tests for Autocorrelation and Heteroscedasticity

| 2*Sector | | *OLS* | | *ARDL* | |
|---|---|---|---|---|---|
| | | *B-G LM* | *ARCH* | *B-G LM* | *ARCH* |
| 4*Buildings | Lag 1 | 0.0246 | 0.4769 | 0.2993 | 0.9850 |
| | Lag 2 | 0.0659 | 0.4566 | 0.5780 | 0.8242 |
| | Lag 3 | 0.0588 | 0.5292 | 0.7480 | 0.6181 |
| | Lag 4 | 0.1085 | 0.6426 | 0.8686 | 0.3990 |
| 4*Other Sectors | Lag 1 | 0.0221 | 0.2768 | 0.7724 | 0.9850 |
| | Lag 2 | 0.0413 | 0.3827 | 0.8355 | 0.3474 |
| | Lag 3 | 0.0823 | 0.5661 | 0.5751 | 0.4776 |
| | Lag 4 | 0.1086 | 0.6955 | 0.3424 | 0.2653 |

*Note:* Breusch-Godfrey Serial Correlation LM Test (B-G LM) with H0: no serial correlation. ARCH Test for heteroscedasticity (ARCH) with H0: no heteroscedasticity. The p-values are provided for each test.

## Cointegration Tests

Table D.3 shows cointegration tests, including the Bounds F-Test and Bounds T-Test (Pesaran et al., 2001), based directly on the ARDL model.

**Table D.3:** Cointegration tests

| Sector | Bounds T-Test | Bounds F-Test |
|---|---|---|
| Buildings | $p = 0.023$ | $p = 0.024$ |
| Other Sectors | $p = 0.048$ | $p = 0.061$ |

*Note:* The Bounds T-Test and the Bounds F-Test are based on the ARDL model. The Bounds F-Test does not reject the absence of cointegration at the 5% level. The null hypothesis for both the Bounds T-Test and the Bounds F-Test is that no cointegration exists.

## Dynamic Multipliers

Table D.4 presents the dynamic multipliers for up to 10 periods and their associated standard errors, as illustrated in Figures 4.3 and 4.5 for the buildings sector and other sectors, respectively.

**Table D.4:** Dynamic Multipliers (s-Period Lags)

| **Period** | **Buildings** | | **Other Sectors** | |
|---|---|---|---|---|
| | Delay-Multip. | Std.-Error | Delay-Multip. | Std.-Error |
| 0 | 3.64 | 2.90 | -0.48 | 3.99 |
| 1 | 9.20 | 2.65 | 5.82 | 2.80 |
| 2 | 6.53 | 1.28 | 3.14 | 1.23 |
| 3 | 4.63 | 0.64 | 1.69 | 0.77 |
| 4 | 3.29 | 0.51 | 0.91 | 0.57 |
| 5 | 2.33 | 0.54 | 0.49 | 0.42 |
| 6 | 1.65 | 0.54 | 0.27 | 0.29 |
| 7 | 1.17 | 0.51 | 0.14 | 0.19 |
| 8 | 0.83 | 0.45 | 0.08 | 0.12 |
| 9 | 0.59 | 0.38 | 0.04 | 0.08 |
| 10 | 0.42 | 0.32 | 0.02 | 0.05 |

*Note:* Dynamic multipliers for s-period lags and the corresponding standard errors for two sectors.