Diplomarbeit

Analysis of the Impact of Income Inequality on CO_2 Emissions

zur Erlangung des akademischen Grades einer

Diplom-Ingenieurin

im Rahmen des Studiums

Statistik-Wirtschaftsmathematik

durch

Laura Granser, BSc.

Matrikelnummer: 01204392

unter der Anleitung von

Ao.Univ.Prof. Wolfgang Scherrer

ausgeführt am Institut für Stochastik und Wirtschaftsmathematik

der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Wien, am 12.1.2021

Unterschrift Verfasserin

Unterschrift Betreuer

Kurzfassung

Die vorliegende Diplomarbeit beschäftigt sich mit dem Einfluss von Einkommensungleichheit auf CO_2-Emissionen. Es wird untersucht, ob eine Senkung der Einkommensungleichheit zu einer Erhöhung der Pro Kopf-CO_2-Emissionen führen kann. Hierfür wird ein two-way error component Modell mittels eines grouped fixed effects (GFE) Schätzers von Bonhomme und Manresa (2015) geschätzt. Für eine zuvor festgelegte Anzahl an GFE Gruppen teilt der Schätzer jedes Land im Datensatz einer GFE Gruppe zu. Dadurch können gruppierte fixed effects geschätzt werden, welche zwischen den Gruppen variieren können. Dies ermöglicht, unterschiedliche Fortschritte bei der Entwicklung von sauberer Technologie sowie der Verbesserung von emissionsverursachenden Prozessen von Ländern im Modell zu berücksichtigen. Die Ergebnisse dieser Arbeit zeigen, dass eine Senkung der Einkommensungleichheit zu einer Erhöhung der Pro Kopf-CO_2-Emissionen führt, wenn das Pro Kopf-Einkommen unter einem bestimmten Schwellenwert liegt. Für Einkommen über diesem Schwellenwert ändert sich das Vorzeichen, jedoch ist dieser Schwellenwert so hoch, dass er nur von den reichsten Ländern erreicht wird. Daher kann gefolgert werden, dass für fast alle Länder außer jenen mit dem höchsten Pro Kopf-Einkommen eine Senkung der Einkommensungleichheit zu einer Erhöhung der CO_2-Emissionen führt. Länder, welche eine Senkung von Einkommensungleichheit sowie von CO_2-Emissionen erzielen wollen, sollten zusätzliche Maßnahmen treffen, um den entstehenden Zielkonflikt entgegenzusteuern.

Abstract

This thesis deals with the impact of income inequality on CO_2 emissions. It is investigated whether a reduction of income inequality can lead to an increase in per capita CO_2 emissions. A grouped fixed effects (GFE) estimator by Bonhomme and Manresa (2015) is being used to estimate this relationship in a two-way error component model. The estimator assigns for a previously set number of GFE groups every country in the sample to one of the GFE group. This allows the grouped fixed effects to vary between the GFE groups, whereby different stages of clean technology development and of improvements in emission creating processes between countries can be taken into account. The main findings are that the impact of income inequality on CO_2 emissions depends on the level of income. Reducing income inequality leads to an increase in CO_2 emissions below a certain threshold of income, while for income levels above this threshold the effect is reversed. However, this extremely high threshold value has been reached only by a few countries, implying that for almost all countries except ones with the highest per capita income levels reducing income inequality will lead to higher per capita CO_2 emissions. If countries want to cut down income inequality as well as per capita CO_2 emissions, they should consider taking additional measures to prevent a possible trade-off.

Danksagung

Hiermit möchte ich mich bei allen Personen für ihre Unterstützung im Laufe meines Studiums bedanken.

Zuallererst gilt mein Dank meiner Familie, welche mir immer zur Seite stand und mich während des Studiums maßgeblich unterstützt hat. Insbesondere möchte ich mich bei meinen Eltern bedanken, da sie mich dazu ermutigt haben, Mathematik zu studieren und mir in Momenten des Zweifels eine wichtige emotionale Stütze waren.

Für den emotionalen Rückhalt möchte ich mich natürlich auch bei Alex, Linda und Clara bedanken. Weiters möchte ich mich bei all meinen weiteren FreundInnen und StudienkollegInnen dafür bedanken, dass ich dank ihnen meine Studentenzeit in vollem Ausmaß genossen habe.

Zu guter Letzt bedanke ich mich auch bei Herrn Prof. Wolfgang Scherrer für die hervorragende Betreuung dieser Arbeit. Er hat mir mit seinen fachlichen Anmerkungen und Vorschlägen immer weitergeholfen, wenn ich im Entstehungsprozess dieser Arbeit einmal nicht weiter wusste.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 12.1.2021

Laura Granser

Contents

1 Introduction 1

2 Literature Review 3
- 2.1 Theory . . . 3
- 2.2 Empirical Findings . . . 6

3 Data Description 8
- 3.1 Carbon Dioxide Emissions . . . 9
- 3.2 GDP Per Capita . . . 10
- 3.3 Gini Index . . . 10
- 3.4 Polity Index . . . 12
- 3.5 Summary Statistics . . . 12

4 Model 13

5 Methodology 16
- 5.1 Ordinary Least Squares . . . 16
- 5.2 Fixed Effects . . . 17
  - 5.2.1 Consistency and Standard Errors . . . 17
  - 5.2.2 Fixed Effects Estimator for Balanced Panel . . . 18
  - 5.2.3 Fixed Effects Estimator for Unbalanced Panel . . . 18
- 5.3 Grouped Fixed Effects . . . 21
  - 5.3.1 Consistency . . . 21
  - 5.3.2 Asymptotic Distribution . . . 23
  - 5.3.3 GFE Estimator for Unbalanced Panel . . . 26
  - 5.3.4 Algorithms . . . 27

6 Results 29
- 6.1 Estimates . . . 29
- 6.2 GFE Groups . . . 33

7 Sensitivity Analysis 39

8 Conclusion 43

Appendix 45

List of Tables 51

List of Figures 52

List of Abbreviations 53

Bibliography 54

1 Introduction

In the last years, the climate crisis was one of the most debated topics. It is now clear, that the economic development of mankind has contributed to various kinds of environmental degradation including global warming.[1] There exist numerous studies about the impact of income and economic growth on the environment. Especially the relationship between economic growth and CO_2 emissions, which has mostly contributed to global warming (see IPCC (2014)), has been extensively investigated (e.g. Roberts and Grimes (1997), Martínez-Zarzoso and Bengochea (2004) and Zoboli et al. (2010)).

This thesis studies not only the nexus between income and CO_2 emissions, but also the effect of income distribution on CO_2 emissions, as for most countries not only the restoration of environmental damages, but also the reduction of inequality and poverty are major challenges. Several studies have already examined the impact of income inequality on environmental degradation. Theoretical studies like Scruggs (1998) and Boyce (2007) show that income inequality might influence environmental pollution through different transmission channels and various opposing effects, which will be further discussed in the following chapter. A full review of the theoretical arguments explaining the mechanisms, through which income inequality might impact environmental degradation, can be found in Berthe and Elie (2015). The empirical studies - such as Ravallion et al. (2000), Heerink et al. (2001), Magnani (2000), Borghesi (2006) and Grunewald et al. (2017) - have mixed outcomes depending on the chosen methods and data sets. They are discussed in section 2.2 in more detail.

This thesis contributes to the literature about the relationship between income inequality and CO_2 emissions by using a grouped fixed effects (GFE) estimator by Bonhomme and Manresa (2015), which can better handle time-varying unobserved heterogeneity in contrast to the fixed effects (FE) and ordinary least squares (OLS) estimator commonly used in the existing literature. The GFE estimator assigns for a previously set number of GFE groups every country in the sample to one of the GFE group. Therefore, different stages of

[1] See IPCC (2014) for further information.

clean technology development and of improvements in emission creating processes between countries can be taken into account as the grouped fixed effects can vary between the GFE groups. Nonetheless, the FE and OLS estimators are used to compare the estimated results from the grouped fixed effects estimator with them. Furthermore, the chosen panel data set is more extensive than most of the data sets used in existing studies, as it is now possible to include longer time series and improved measures for income inequality as well as CO_2, making the data better comparable among countries.
The thesis is most closely related to Grunewald et al. (2017), who introduced the GFE estimator to analyse the relationship between income inequality and CO_2 emissions in a two-way error component model. We mostly follow their approach, but additionally examine the composition of the estimated GFE groups in more detail to find a meaningful explanation for the characterization of group membership. Furthermore, the sensitivity analysis in this paper includes assessing the stability of the algorithm by Bonhomme and Manresa (2015).

The main findings resulting from the estimates by the GFE estimator are, that the impact of income inequality on CO_2 emissions depends on the income level. Increasing inequality has a negative effect on CO_2 emissions below a certain threshold of income, while for income levels above this threshold the effect is reversed. However, this extremely high threshold value is almost out of sample, implying that for low-, middle- and high-income countries reducing income inequality will lead to higher per capita CO_2 emissions. Moreover, the estimated results confirm the environmental Kuznet curve hypothesis.

The rest of this thesis is structured in the following way. Chapter 2 evaluates existing theoretical and empirical studies about the impact of income inequality on environmental degradation. The panel data set is outlined in chapter 3. Chapter 4 describes the model. The estimators are presented in chapter 5. Chapter 6 outlines the estimated results. Chapter 7 contains a sensitivity analysis. Lastly, chapter 8 completes this thesis with the conclusion.

2 Literature Review

2.1 Theory

Over the last decades several economists and researchers from various areas have examined how income inequality might affect environmental damage and developed theories to explain that relationship. The literature on this relationship has meanwhile become quite vast, therefore we will focus on the most important theories and hypotheses concerning this relationship.

The relationship between income inequality and environmental pressure is inextricably linked to the subject of economic development and economic welfare. One of the most famous hypotheses concerning the influence of economic development on environmental damage is the environmental Kuznet curve (EKC), which became known through the work of Grossman and Krueger (1991, 1995) and Shafik and Bandyopadhyay (1992).
The EKC shows that the influence of economic development on environmental damage is of a specific shape, namely an inverted-U:

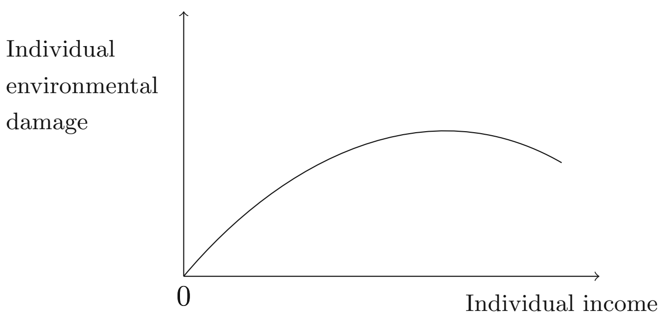

Figure 2.1: Environmental Kuznet curve

The graphic can be explained as follows: in the initial state the levels of both income and environmental damage are low. As income starts to rise, environmental damage grows as well. When income reaches a certain threshold, the direction of the relationship changes and environmental damage starts sinking.

If we assume that the EKC holds for household income, then this would mean: a decline of income inequality via a redistribution policy for example could have different outcomes concerning environmental pressure, as this relationship depends on the level of economic development.

Scruggs (1998) and Heerink et al. (2001) consider the EKC hypothesis to be the most realistic shape of the relationship between economic income and environmental pressure. Ravallion et al. (2000) receive similar results concerning this relationship. They find that there exists a trade off between lowering income inequality and reducing greenhouse gas emissions both across and within countries, which gets better for countries with a high average income.

Scruggs (1998) substantiates this hypothesis by using Inglehart's post-materialism theory (Inglehart, 1990), which claims that humans first want to satisfy their material needs and attain a certain level of prosperity before they change their preferences regarding the environment and start to take actions to limit environmental damages.
Heerink et al. (2001) argue that households start substituting polluting goods with environmental friendly goods after reaching a certain level of affluence. This would lead to a decrease of individual environmental pressure through the change of the individual household's behaviour and through producers making the manufacturing process of their goods more environmentally friendly.
However, in reality there exists a gap between developing environmental values and taking actions to limit environmental pressure. One possible explanation for this value-action gap might be that on one side individuals develop environmental concerns as they are getting more affluent, while on the other side they have more interest in reaching a higher standard of living, which would include more energy intensive transportation like flying and generally more frequent travelling.
This change of individual behaviour would imply that a decrease of economic inequality leads to a higher burden on the environment.

The emulation theory is closely linked to the above-mentioned points regarding individual economic behaviour. The theory was originally formulated by Veblen (1934), therefore also sometimes referred to as a Veblen effect.
It claims that in a very unequal society differences between the lifestyle and consumption behaviour of the affluent social groups compared to poorer social groups are bigger than in a more equal society and hence more conspicuous.

Members of a certain social group would try to emulate the consumption behaviour of the slightly wealthier social groups. This means that they would adopt a certain lifestyle and consume goods based on the social status they represent, rather than making their economic behaviour more environmentally friendly. This implies that higher economic inequality would lead to a more environmentally harmful lifestyle of all social groups.

Economic income inequality might influence environmental pressure not only through consumption patterns of individuals, but also through finding joint solutions to minimize environmental damages. As pointed out by Borghesi (2006), this might be more difficult for a society with unequal income distribution. Political agents like the government, lobbies, trade unions etc. generally have more difficulties to cooperate regarding social and environmental issues than in a more egalitarian society. Thus, a more unequal income distribution can hinder countries to implement effective environmental regulations.
However, for pollutants having a global impact like Greenhouse Gases international agreements and solutions are more relevant than national ones. Thus, this theory might be more applicable to inequality across countries than inequality within countries.

Also Boyce (2007) examined how economic inequality might influence environmental harm via the political channel. He makes two assumptions: firstly, economic income is generally correlated with political power; secondly, the wealthy part of a society tends to benefit the most from the economic activities generating environmental damages.
He explains the latter point by the argument that polluting companies are generally owned by the rich and goods and services with a higher emission rate are as well consumed by them. Damages to the environment arise in general locally. The wealthy and powerful do not suffer as much as the poor from those damages, because they can avoid those damages. They are able to move to more expensive neighborhoods with good environmental quality and to buy private substitutes for missing public environmental quality.
Thus it is not necessary for them to have environmental regulations implemented, because they will not benefit as much from those as from the benefits generated by their polluting companies and from consuming goods with high emission rates. Therefore, they will use their wealth and political power to hinder policymakers from formulating and implementing those costly environmental regulations.
Boyce (2007) argues that if those two assumptions are fulfilled, higher income inequality can be expected to lead to higher damages to the environment. This can be also applied to greenhouse gas emissions and their resulting environmental damages. In this case the damages are global, like rising sea levels, higher temperatures and more extreme weather phenomena. If we make the same assumptions for income distribution across countries,

then higher inequality across countries would cause higher greenhouse gas emissions and therefore more environmental damages.

Apart from theories being clear about whether income inequality influences environmental damages positively or negatively, some researchers like Ravallion et al. (2000) examined this relationship while staying vague about the sign of the direction. Nonetheless their results are in line with the theories by Heerink et al. (2001) & Scruggs (1998) mentioned above. Ravallion et al. (2000) assume that every individual causes emissions through their consumption behaviour, either directly by using goods or indirectly via the production process of goods. One can formulate for each individual an implicit demand function for carbon emissions, the derivative of this function with respect to income is called marginal propensity to emit (MPE).
Depending on the MPE lowering inequality could have different outcomes. If one assumes that poor people generally have a lower MPE than the more affluent ones, reducing inequality via a redistribution policy would lead to a decrease in emissions. However, if rich people have a lower MPE than the poor, such a redistribution policy would lead to rising emissions. It is difficult to say which case corresponds to reality. The consumption behaviour of affluent people include goods and activities with higher emission rates like cars and flying, while poor people would spend more income on clothes and food, which would correspond to the first case. On the contrary, the rich are able to use energy more efficiently than the poor, which would imply the latter case.

Following Berthe and Elie (2015), we can summarize, that economic inequality can influence environmental pressure via two main channels. The first channel describes the mechanism how income inequality affects environmental damage through consumption behaviour of households, while the second channel is characterized by social agreements and the implementation of environmental policies.
We have described a number of opposing effects. It is theoretically impossible to determine, which effect could dominate the others. Therefore, we will try to find the answer to this question empirically.

2.2 Empirical Findings

There exist a number of empirical studies about the nexus between income distribution and CO_2 emissions. Results by Ravallion et al. (2000) and Heerink et al. (2001) show that higher income inequality has a negative effect on CO_2 emissions, which would imply a trade off between lowering income inequality and reducing CO_2 emissions. Both studies use pooled OLS estimators to show the results. More recent results by Hübler (2017) are in line with this negative nexus. His empirical findings are achieved by using quantile regressions, however his FE estimations give no significant results concerning this relationship. Borghesi (2006) achieves similar results: his pooled OLS estimations give a negative coefficient for income inequality, however the FE estimations result in a statistically non-significant impact of income inequality on CO_2 emissions.

On the opposite, there are already studies, like Boyce (2007) and Magnani (2000) for example, providing possible explanations for a negative influence of income inequality on environmental degradation, implying therefore a negative influence on CO_2 emissions. Arguments by Boyce (2007) have been already mentioned in section 2.1. Magnani (2000) shows empirically that there exists a negative association between income inequality and environmental protection by doing OLS regressions with OECD data on research & development expenditure for pollution abatement.

Lastly, some studies come to the conclusion that the effect of income inequality on CO_2 emissions can change depending on other factors, such as the level of GDP per capita. For instance, Grunewald et al. (2017) empirically find evidence that for low- and middle-income countries income inequality has a negative effect on CO_2 emissions while for high-income countries the opposite is the case. They use a grouped fixed effects estimator, which is further presented in section 5.3, to estimate the effect of income inequality on CO_2 emissions for 158 nations between 1980 and 2008.
To sum it up, empirical studies investigating the relationship between income inequality and CO_2 emissions unfortunately give mixed results about the sign of this relationship depending on the chosen data set and the econometric methods applied.

3 Data Description

Our data set is a panel data set. Observations are collected for 140 countries over a time period of 25 years (1990 to 2014). The selected countries are:
Albania, Algeria, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Barbados, Belarus, Belgium, Belize, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Canada, Central African Republic, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, Arab Rep., El Salvador, Estonia, Eswatini, Ethiopia, Fiji, Finland, France, Gambia, The, Georgia, Germany, Ghana, Greece, Guatemala, Guinea, Guinea-Bissau, Guyana, Honduras, Hong Kong SAR, China, Hungary, Iceland, India, Indonesia, Iran, Islamic Rep., Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Rep., Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Micronesia, Fed. Sts., Moldova, Mongolia, Morocco, Mozambique, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, North Macedonia, Norway, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Rwanda, Senegal, Sierra Leone, Singapore, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, St. Vincent and the Grenadines, Sudan, Sweden, Switzerland, Tajikistan, Tanzania, Thailand, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Kingdom, United States, Uruguay, Venezuela, RB, Vietnam, West Bank and Gaza, Yemen, Rep., Zambia and Zimbabwe.

The selection of these countries is based on the available number of observations concerning GDP per capita, CO_2 emissions per capita and the Gini index. We denote a *complete* observation as one, where for a country i and year t the variables GDP per capita, CO_2 emissions per capita & the Gini index are not missing.
Therefore, all countries with 25 complete observations are selected. The other countries are selected in descending order of the number of complete observations. We select 140 countries and drop all countries with fewer complete observations than the selected ones.

This leads to a slight difference between our selection of countries and the one in Grunewald et al. (2017). Our data set contains the following countries, which are not included in the selection by Grunewald et al. (2017): Barbados, Burundi, Lebanon, Lesotho, Micronesia, Quatar, St.Vincent and the Grenadines, Sudan, Tonga, West Bank and Gaza & Zimbabwe. As countries were selected in descending order based on the number of complete observations, the following countries from the selection by Grunewald et al. (2017) are not included in our data set: Angola, Bangladesh, Benin, Chad, Comoros, Djibouti, Gabon, Haiti, St.Lucia & Suriname.

3.1 Carbon Dioxide Emissions

We choose Carbon Dioxide (CO_2) per capita emissions to measure environmental damage for the following reasons:

1. Among all the anthropogenic (i.e. man-made) Greenhouse Gases, CO_2 has contributed most to global warming in the last years according to the latest IPCC Synthesis report.[1]

2. Time series for CO_2 emissions are longer and more complete in comparison to all the other GHG (Greenhouse Gases).

The CO_2 emissions time series come from the US Oak Ridge National Laboratory (ORNL) at the Carbon Dioxide Information Analysis Center (CDIAC) by Boden et al. (2017). CO_2 per capita emissions are measured in metric CO_2 tons per capita and include emissions from burning of fossil fuels and cement production. Emissions for every year and country are calculated by taking the amount of burned fossil fuels and multiplying it with the average carbon content of each fuel type (oil, gas or coal). Emissions by cement production are measured by the average amount of CO_2, which is released during the procedure of cement production.

One disadvantage of the ORNL data set is the fact, that CO_2 emissions by deforestation, agriculture, livestock and land use change are not included. According to the latest IPCC Report (IPCC, 2014, page 45-46) those emissions have contributed since 1970 around one quarter of global CO_2 emissions. The available CO_2 emission time series could underestimate the real amount of CO_2 emissions, but probably only for countries with a strong

[1] See IPCC (2014) for further information.

agricultural sector for example. Nevertheless, the ORNL data set is the most complete data set compared to others concerning number of countries and years.

3.2 GDP Per Capita

GDP per capita is one of the *World Development Indicators* (2019) from the World Bank. GDP PPP denotes the gross domestic product converted to international dollars by using purchasing power parity and is measured in international 2011 Dollars. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United State, that means one can buy with one international dollar the same amount of goods which can be bought with one U.S. Dollar in the United States. The time series for GDP per capita is available for the years 1990 - 2018.

3.3 Gini Index

We choose the Gini Index to measure income inequality. The Gini index was developed by the Italian economist Gini and is one of the most common indicators to measure distribution inequality. It is defined as follows:

Following Gastwirth (1972), we define for a set of n ascending, ordered numbers $x_1, ..., x_n$ (e.g. income) the empirical **Lorenz** curve generated by the points $i/n, i = 0, ..., n$ by $L(0) := 0$ and $L(j/n) := s_j/s_n$ where $s_j := \sum_{i=1}^{j} x_i$.
The empirical Lorenz curve $L(p)$ is defined for all fractiles p in the interval $(0, 1)$ by linear interpolation of the points $L(p)$ and represents the fraction of the total variable measured (e.g. income). The Gini index G is defined as the ratio of the area between the empirical Lorenz curve and the 45° line to the area under the 45° line.[2]

Therefore, the Gini index can be a value between 1 and 0, where 0 stands for total equality and 1 for total inequality. For example, if the Lorenz curve is identical to the 45° line, then total equality is the case.

[2]See Gastwirth (1972) for further information.

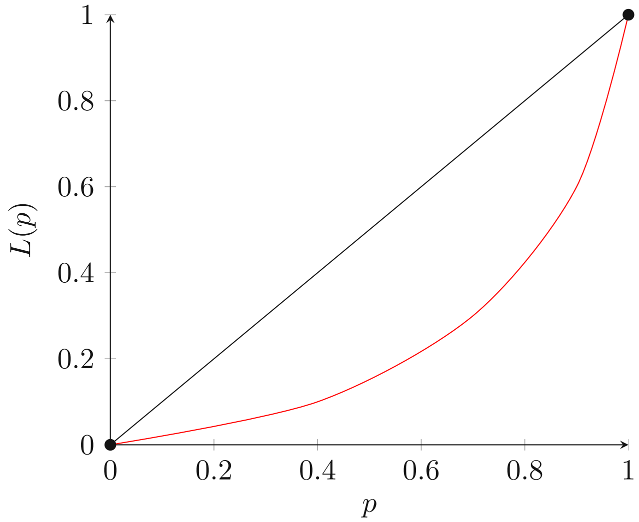

Figure 3.1: Empirical Lorenz curve

We use the Gini indices from the Standardized World Income Inequality (SWIID) Database with time series available from 1960 to 2018 and for up to 196 countries or territories. The SWIID routine, developed by Solt (2009), is based on the Luxembourg Income Study (LIS) Database and estimates relationships between Gini indices from the LIS and all the other Gini indices available for the same country-years. Gini indices for country-years that are not available in the LIS but from other sources, are estimated by using those estimated relationships.

The SWIID data set contains Gini indices for two different types of income, namely *before tax income* and *after tax income* (e.g. disposable income). As most of the SWIID sources use disposable income as welfare definition, the Gini indices based on disposable income are better comparable among countries and years, therefore we will use those Gini indices.

3.4 Polity Index

The Polity Index is part of the Integrated Network for Societal Conflict Research (INSCR) data set from the Center of Systemic Peace (Cole and Marshall, 2014).
The Polity Index measures the fragility of a state: The Index ranges between -10 and 10, where -10 stands for fully institutionalized autocracy and 10 stands for fully institutionalized democracy. Time series are available for the years 1800-2017 and for up to 167 countries.

We will use the Polity Index for our sensitivity analysis in chapter 7. Namely, the Polity Index can be interpreted as a measure of the quality of institutions, as a proxy variable for measuring environmental regulations e.g. the political channel through which income inequality might influence CO_2 emissions.

3.5 Summary Statistics

To conclude this chapter, the main variables used in the model are described in table 3.1, while table 3.2 displays a summary statistic of all variables used.

Table 3.1 Definition of variables

Variable	Definition	Unit	Number of observations
CO_2	CO_2 per capita	metric tons	3447
GINI	Gini Index	percentage scale	3224
GDP	GDP per capita	international 2011 Dollar	3440

Table 3.2 Summary statistic

Variable	Mean	St. Dev.	Min	Max
CO_2	4.48	6.42	0.04	70.92
GINI	38.18	8.39	17.30	62.60
GDP	14 252.96	16 225.05	354.28	129 349.90
Polity Index	4.41	6.03	−10.00	10.00

4 Model

Let us consider a data set where N individuals are repeatedly observed over a time period of length T. The set of the N individuals does not change over time. Such a data set is called a **panel** data set. A panel data set is called a **short** panel, if the number of observed subjects is relatively large compared to the length of the time period. If the contrary is true, i.e. if a small cross section of individuals is observed over a long time period, then the panel data set is referred to as a **long** panel. If the same time periods are available for all individuals, then the data set is called **balanced** panel data set. If this is not the case, i.e. when the number of time series observations is different across individuals, the panel is called **unbalanced**.
For reasons of simplicity the methods are first presented for a balanced panel data model as this makes the indexation and reading the equations considerably easier.

Consider the following model for individual $i, i = 1, ..., N$ and time $t, t = 1, ..., T$:

$$y_{it} = \alpha + \beta_1 x_{i,t}^1 + ... + \beta_K x_{i,t}^K + v_{it} \tag{4.1}$$

$$v_{it} = \theta_{it} + u_{it} \tag{4.2}$$

The term y_{it} denotes the dependent variable, α the intercept and $x_{i,t}^1, ..., x_{i,t}^K$ denote the K explanatory variables, also called **covariates**. The parameters of interest, which we want to estimate, are $\beta_1, ..., \beta_K$. If we define the vectors

$$X_{it} := \begin{pmatrix} x_{i,t}^1 \\ \vdots \\ x_{i,t}^K \end{pmatrix}, \quad \beta := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix} \tag{4.3}$$

then we can rewrite equation 4.1 in the following form:

$$y_{it} = \alpha + \beta' X_{it} + v_{it} \tag{4.4}$$

The error term $v_{it} = \theta_{it} + u_{it}$ is the sum of the **unobservable effects** θ_{it} and an **idiosyncratic error** u_{it}. We assume $u_{it} \sim IID(0, \sigma_u^2)$, meaning that u_{it} is independent and identically distributed (IID) with mean 0 and variance σ_u^2. Depending on the applied estimator, necessary assumptions concerning the unobservable effects θ_{it} and the idiosyncratic error u_{it} are described in the following chapter.

The precise model being used in this thesis is presented as follows: for country $i \in \{1, ..., 140\}$ and year $t \in \{1990, ..., 2014\}$:

$$\boxed{\begin{aligned} e_{it} &= \alpha + \beta_1 y_{it} + \beta_2 y_{it}^2 + \beta_3 g_{it} + \beta_4 g_{it} y_{it} + v_{it} && (4.5) \\ v_{it} &= \theta_{it} + u_{it} && (4.6) \end{aligned}}$$

The term e_{it} is the logarithm of CO_2, y_{it} the logarithm of GDP and g_{it} the logarithm of the GINI index. The error term v_{it} consists of the unobservable effects θ_{it} and the idiosyncratic error u_{it}.

The following methods, which will be presented and described in more detail in the following chapter, are being used to estimate the slope coefficients $\beta_1, ..., \beta_4$:

- Firstly, the coefficients are estimated by (pooled) OLS. This estimator requires no particular specification concerning the composition of the error term v_{it}. Therefore our precise model described by equations 4.5 and 4.6 stays the same.

- Secondly, we estimate the model by using the FE estimator. For the FE model we assume $\alpha = 0$ for the intercept α and the following structure for the error term:

$$v_{it} = \eta_i + \lambda_t + u_{it}, \tag{4.7}$$

  where v_{it} is the sum of the unobservable **individual effect** η_i, the unobservable **time effect** λ_t and an **idiosyncratic error** u_{it}. The term η_i accounts for unobserved individual heterogeneity, also called individual effect or fixed effects.[1]
  The time effect λ_t is individual-invariant and accounts for any time-specific effect not included in the regression. Such an effect could be for example a strike year leading to a disruption of production or an oil embargo resulting in higher oil prices and reduced oil consumption (Baltagi (2008)). Due to the construction of the error term, such a model is called a **two-way error component** model.
  The new structure of the error term means for our precise model described by equa-

[1]See Wooldridge (2010) for further information.

tions 4.5 and 4.6, that it can be now written as:

$$\boxed{e_{it} = \beta_1 y_{it} + \beta_2 y_{it}^2 + \beta_3 g_{it} + \beta_4 g_{it} y_{it} + \eta_i + \lambda_t + u_{it}} \tag{4.8}$$

The FE estimate can be received by first transforming equation 4.8 to eliminate the individual & time effects and then estimating the transformed equation by OLS.

- Lastly the model is being estimated by the GFE estimator. The GFE estimate is based on the assumption, that the unobserved (time) effects can vary between groups of countries, but have the same development over time for all countries within each group. If we define for a number of groups G, where each group is denoted by $g \in \{1, ..., G\}$, the group membership variables $g_i \in \{1, ..., G\}, i = 1, ..., N$, indicating to which group the country i is being assigned, then the unobserved effects can be modeled in the following way:

$$\theta_{it} = \lambda_{g_i t}$$

This specific structure of the unobserved effects results in our model to be specified to:

$$\boxed{e_{it} = \beta_1 y_{it} + \beta_2 y_{it}^2 + \beta_3 g_{it} + \beta_4 g_{it} y_{it} + \lambda_{g_i t} + u_{it}} \tag{4.9}$$

If we know the group membership for each country, then the GFE estimate can be received by estimating 4.9 by OLS.
However, in reality the group membership assignments are not known and must be estimated. It is outlined in section 5.3 of the following chapter how to estimate the group membership assignments. The algorithm, which is used to estimate the group membership assignments and the grouped fixed effects, is the algorithm 2 by Bonhomme and Manresa (2015).[2] We set the number of groups to 5 to make our results comparable to the ones by Grunewald et al. (2017). We use the resulting estimated grouping of the 140 countries into 5 groups by the GFE estimator to include an interaction term between group assignment of every country and year to estimate the grouped fixed effects.

[2]Presented on page 2 of the supplementary appendix available at https://www.dropbox.com/s/h2hk43owrl6rwh1/Bonhomme_Manresa_appendix.pdf?dl=0.

5 Methodology

Following Grunewald et al. (2017), we will focus on the grouped fixed effects estimator presented in section 5.3, therefore we will describe this estimator in more detail than the other two estimators. Nevertheless, the other two estimators will be briefly described, as we will use the estimated results to compare with the results from the grouped fixed effects estimator. This will be further explained in the following sections.

5.1 Ordinary Least Squares

The first method being described is OLS. For this estimator all observations are pooled across i and t, thus we receive a long regression with $N * T$ observations. Then this regression model is estimated by OLS.

The main assumption for consistency of the OLS estimator is the following:

Assumption 1 (Consistency)**.** *Let X be defined as on page 19, but here for a balanced panel. For all $t \in \{1, ..., T\}$ and $i \in \{1, ..., N\}$:*

$$\mathbb{E}[v_{it}|X] = 0 \tag{5.1}$$

The assumption 1 requires X_{it} and v_{it} to be uncorrelated. Due to the construction of the error term v_{it}, this implies that X_{it} is uncorrelated with the unobservable effects θ_{it} as well as with the idiosyncratic error u_{it}.
However, if in reality some of the unobserved effects θ are correlated with the regressors X, then regressing y on X will result in an omitted variable bias, as the error term v is as well correlated with X. The estimates of the coefficient vector β will be biased and inconsistent. Therefore, the OLS estimator is inconsistent if the true model is one, where some of the unobserved effects are correlated with the explanatory variables.

Since the panel, that we use, is a macro panel with a fixed sample, it is highly probable, that the unobserved effects are correlated with the explanatory variables.

The standard errors computed by the classical OLS formulas in addition are based on the following assumption:

$$\mathbb{E}[v_{it}v_{js}|X] = \sigma_v^2 \quad for \quad i = j, t = s \quad and \quad \mathbb{E}[v_{it}v_{js}|X] = 0 \quad else \tag{5.2}$$

This implies that the errors v_{it} are conditionally homoscedastic and uncorrelated. Nonetheless, OLS estimates of panel data regression models usually have largely biased standard errors, because the error terms are most likely auto-correlated for each individual and this auto-correlation is not being taken into account by the OLS estimator.

5.2 Fixed Effects

In contrast to the OLS estimator, the FE estimator allows the unobserved effects to be correlated with the explanatory variables.

5.2.1 Consistency and Standard Errors

Consistency of the FE estimator is based on the following main assumption:

Assumption 2 (Consistency)**.** *Strict exogeneity:*

$$\mathbb{E}[u_{it}|X_{it}, \eta_i, \lambda_t] = 0 \qquad \forall t \in \{1, ..., T\}, i \in \{1, ..., N\} \tag{5.3}$$

If consistency of the FE estimator is guaranteed, we can consistently measure the effect of X_{it} on y_{it} by controlling for the individual as well as for the time effect. Unobserved effects like the individual effects η_i are allowed to be correlated with the regressor X_{it}.

In addition, the computation of standard errors is based on the following assumption:

$$\mathbb{E}[u_{it}u_{js}|X, \eta, \lambda] = \sigma_v^2 \quad for \quad i = j, t = s \quad and \quad \mathbb{E}[u_{it}u_{js}|X, \eta, \lambda] = 0 \quad else \tag{5.4}$$

5.2.2 Fixed Effects Estimator for Balanced Panel

The idea behind the FE estimator is to transform the equations to eliminate the unobserved individual and time effect η_i, respectively λ_t. There are several transformations to do this. The most common transformation for balanced panels is the **within** transformation, originally formulated by Wallace and Hussain (1969).
This transformation can be obtained by the following procedure given by Baltagi (2008):

Let $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}, \bar{X}_i = T^{-1}\sum_{t=1}^{T} X_{it}, \bar{v}_i = T^{-1}\sum_{t=1}^{T} v_{it}$ be the average values over time for every individual $i, i = 1, ..., N$ and $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}, \bar{X}_t = N^{-1}\sum_{i=1}^{N} X_{it}, \bar{v}_t = N^{-1}\sum_{i=1}^{N} v_{it}$ the average values over individuals for $t, t = 1, ..., T$.
Let further $\overline{\overline{y}} := (TN)^{-1}\sum_{i}^{N}\sum_{t}^{T} y_{it}$ be defined as the average over time and individuals, which is the same definition for $\overline{\overline{X}}$ and $\overline{\overline{v}}$. We use the following transformation:

$$(y_{it} - \bar{y}_i - \bar{y}_t + \overline{\overline{y}}) = \beta(X_{it} - \bar{X}_i - \bar{X}_t + \overline{\overline{X}}) + (v_{it} - \bar{v}_i - \bar{v}_t + \overline{\overline{v}}) \tag{5.5}$$

The individual and time effects are eliminated through this transformation. If we define $\ddot{y} := (y_{it} - \bar{y}_i - \bar{y}_t + \overline{\overline{y}})$ and in the same way $\ddot{X}$ and $\ddot{v}$, we can reformulate the last equation as:

$$\ddot{y} = \beta\ddot{X} + \ddot{v} \tag{5.6}$$

Now we can estimate β by using OLS on the last equation. However, it is important to know, that we cannot use time invariant or individual invariant variables as the within transformation wipes out those variables making it impossible to estimate their effect.

5.2.3 Fixed Effects Estimator for Unbalanced Panel

However, as our data set is unbalanced, we cannot use the within transformation exactly as described above, as it was formulated for the case of a balanced panel. A within transformation for unbalanced panels is provided by Croissant, Millo, et al. (2019):

Let T_i denote the number of observations for individual i and $O := \sum_{i=1}^{N} T_i$ the total number of observations. We define the vectors y and v containing the response and the error, respectively. The vector β contains the coefficients. The matrix X contains the covariates, where the observations are ordered by individual first and second by period, which is the same way how the vectors v and y are ordered.

$$y := \begin{pmatrix} y_{1,1} \\ \vdots \\ y_{1,T_1} \\ y_{2,1} \\ \vdots \\ y_{2,T_2} \\ \vdots \\ y_{N,T_N} \end{pmatrix}, \quad v := \begin{pmatrix} v_{1,1} \\ \vdots \\ v_{1,T_1} \\ v_{2,1} \\ \vdots \\ v_{2,T_2} \\ \vdots \\ v_{N,T_N} \end{pmatrix}, \quad X := \begin{pmatrix} x^1_{1,1} & x^2_{1,1} & \dots & x^K_{1,1} \\ x^1_{1,2} & x^2_{1,2} & \dots & x^K_{1,2} \\ \vdots & & & \\ x^1_{1,T_1} & x^2_{1,T_1} & \dots & x^K_{1,T_1} \\ x^1_{2,1} & x^2_{2,1} & \dots & x^K_{2,1} \\ \vdots & & & \\ x^1_{2,T_2} & x^2_{2,T_2} & \dots & x^K_{2,T_2} \\ \vdots & & & \\ x^1_{N,1} & x^2_{N,1} & \dots & x^K_{N,1} \\ \vdots & & & \\ x^1_{N,T_N} & x^2_{N,T_N} & \dots & x^K_{N,T_N} \end{pmatrix}$$

We consider again equation 4.1 in matrix form:

$$y = \alpha j + X\beta + D_\eta \eta + D_\lambda \lambda + v \tag{5.7}$$

where j is a vector of ones of length O and D_η & D_λ are matrices of individual and time dummies, respectively. The vectors η and λ contain the individual effects and time effects. The matrices $D_\eta^T D_\eta$ and $D_\lambda^T D_\lambda$ are diagonal matrices containing the number of observations for each individual and time-series. If we pre-multiply a vector by $B_\eta := (D_\eta^T D_\eta)^{-1} D_\eta^T$ or by $B_\lambda := (D_\lambda^T D_\lambda)^{-1} D_\lambda^T$, we receive the individual and the time series means, respectively. The matrix $D_\lambda^T D_\eta$ is a $\max_{i\in\{1,\dots,N\}} T_i \times N$-matrix containing ones and zeros, indicating whether there is an observation for an individual in a certain year available.

The within transformation for an unbalanced panel can be received by applying twice the Frisch-Waugh theorem[1]: y, X and D_λ are regressed in a first stage on the matrix of individual dummies D_η, then the residual of y from this first regression is being regressed on the residuals from X and D_λ. This means that in the second stage, $W_\eta y$ is being regressed on the sum of $W_\eta X$ and $W_\eta D_\lambda$. We receive those residuals by pre-multiplying the variables with the projection matrix $W_\eta := I - D_\eta (D_\eta^T D_\eta)^{-1} D_\eta^T$. The matrix W_η projects onto the orthogonal complement of the column space of D_η, which is also the within transformation

[1] Also referred to as the Frisch-Waugh-Lovell theorem, named after Frisch and Waugh (1933) and Lovell (1963).

with respect to individuals.
The Frisch-Waugh theorem is applied a second time by regressing in a first stage $W_\eta y$ and $W_\eta X$ on $W_\eta D_\lambda$. The residuals from this regression in the first stage are calculated by pre-multiplying the variables by the projection matrix W, where W is defined as
$W := I - W_\eta D_\lambda (D_\lambda^T W_\eta D_\lambda)^- D_\lambda^T W_\eta$, where M^- stands for the generalized inverse of any matrix M. Lastly, the residuals of $W_\eta y$ are regressed on the residuals of $W_\eta X$ in the second stage.
Therefore for an unbalanced panel the within transformation with respect to individuals and time consists of pre-multiplying y and every column of X by the following matrix:

$$WW_\eta = (I - W_\eta D_\lambda (D_\lambda^T W_\eta D_\lambda)^- D_\lambda^T W_\eta) W_\eta = W_\eta - W_\eta D_\lambda (D_\lambda^T W_\eta D_\lambda)^- D_\lambda^T W_\eta$$

To sum it up, the two-ways error component fixed effects estimator can be obtained by the following procedure:

1. The individual within transformation is applied to X, y and D_λ,
2. $W_\eta X$ and $W_\eta y$ are regressed on $W_\eta D_\lambda$,
3. We obtain the residuals of $W_\eta X$ and $W_\eta y$ from those two regressions in step 2. The residuals of $W_\eta y$ are regressed on the residuals of $W_\eta X$.

In contrast to the OLS estimator, the FE estimator can deliver consistent estimates, if the individual effects as well as the time effects are correlated with the explanatory variables. However, one disadvantage of the FE model is the assumption that the time effect λ_t must be the same for all individuals. Therefore, the combined effect $\eta_i + \lambda_t$ would show the same development over time for every country. This implication could be problematic for the following reason: the development and deployment of more sustainable and more environmental friendly technology to reduce or mitigate environmental damages is not the same for every nation. The implementation of clean technologies is carried out at different rates for different parts of the world. As more environmental friendly technologies are normally quite expensive, developing nations tend to adopt those technologies more slowly than industrialized countries. Additionally, it seems more realistic, that certain shocks only affect specific regions and not all countries in the sample.

5.3 Grouped Fixed Effects

One alternative to the individual-invariant time effects estimated by the FE estimator is the GFE estimator by Bonhomme and Manresa (2015). The GFE estimator allows the time effects to vary between groups of individuals.

Let y_{it} be the response variable and X_{it} the vector of the covariates for all $i = 1, .., N$ and $t = 1, ..., T$ like we assumed earlier. We assume $\beta \in B$ to be the coefficient vector and B the set of all possible coefficient vectors. Let $\lambda_{gt} \in \Lambda$ for all $t = 1, ..., T$ and $g = 1, ..., G$ be the group specific time effect and the group membership variables g_i, which indicate to which group the individual i belongs for all $i = 1, ..., N$. We define $\gamma := (g_1, ..., g_N)$ as the set of all group membership variables g_i. In other words, $\gamma \in \Gamma_G$ describes a specific partition i.e. grouping of the N individuals into at most G groups, where Γ_G denotes the set of all partitions of $\{1, ..., N\}$ into at most G groups.

The grouped fixed estimator (GFE) is defined as the solution of the following minimization problem:

$$(\hat{\beta}, \hat{\lambda}, \hat{\gamma}) = \underset{(\beta,\lambda,\gamma)\in B\times\Lambda^{GT}\times\Gamma_G}{\arg\min} \sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \beta X_{it} - \lambda_{g_i t})^2 \tag{5.8}$$

where we minimize over all possible groupings $\gamma = \{g_1, ..., g_N\}$ of the N units into G groups, coefficient vectors β and group specific time effects λ.

5.3.1 Consistency

We assume from now on for this section and the following sections that we know the number of groups, $G = G^0$, where G^0 denotes the number of groups in the population (i.e. the true value for the number of groups). We consider the following equation:

$$y_{it} = \beta^0 X_{it} + \lambda^0_{g_i^0 t} + u_{it} \tag{5.9}$$

where $\beta^0, g_i^0 \in \{1, ..., G\}$ denote the true values of β and g_i.

We assume that $(\tilde{\beta}, \tilde{\lambda})$ is the infeasible version of the GFE estimator, where the group membership g_i is not being estimated but fixed to its population counterpart g_i^0 (i.e. the true value). We receive $(\tilde{\beta}, \tilde{\lambda})$ by using pooled regression of y_{it} on X_{it} with interaction terms of population group dummies and time dummies:

$$(\tilde{\beta}, \tilde{\lambda}) = \underset{(\beta,\lambda)\in B\times\Lambda^{GT}}{\arg\min} \sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \beta X_{it} - \lambda_{g_i^0 t})^2 \tag{5.10}$$

The following assumptions are needed to guarantee consistency of the grouped fixed effects estimator:

Assumption 3 (GFE). *There exists a constant $M > 0$ such that:*

1. *B and Λ are compact subsets of $\mathbb{R}^k$ and $\mathbb{R}$, respectively.*
2. $\mathbb{E}[\|X_{it}\|^2] \leq M$
3. $\mathbb{E}[u_{it}] = 0$ *and* $\mathbb{E}[u_{it}^4] \leq M$
4. $|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}[u_{it}u_{is}X_{it}'X_{is}]| \leq M$
5. $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[u_{it}u_{jt}]| \leq M$
6. $|\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}Cov[u_{it}u_{jt}, u_{is}u_{js}]| \leq M$
7. *We define $\bar{X}_{g\wedge\tilde{g},t} := \frac{\sum_i^N \mathbf{1}\{g_i^0=g\}\mathbf{1}\{g_i=\tilde{g}\}X_{it}}{\sum_i^N \mathbf{1}\{g_i^0=g\}\mathbf{1}\{g_i=\tilde{g}\}}$ as the mean of X_{it} in the intersection of groups $g_i^0 = g$, and $g_i = \tilde{g}$. For all partitions $\gamma = \{g_1, ..., g_N\} \in \Gamma_G$, we define $\hat{\rho}(\gamma)$ as the minimum eigenvalue of the following matrix:*

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \bar{X}_{g_i^0\wedge\tilde{g}_i,t})(X_{it} - \bar{X}_{g_i^0\wedge\tilde{g}_i,t})'. \tag{5.11}$$

Then we get: $\quad \text{plim}_{N,T\to\infty}\min_{\gamma\in\Gamma_G}\hat{\rho}(\gamma) = \rho > 0$

The first assumption 3.1. restricts the sets of all possible coefficient vectors and group specific time effects to be compact. Assumptions 3.2. and 3.3. restrict the covariates and idiosyncratic errors to be stationary. Assumptions 3.4. - 3.6. describe weak dependence conditions: 3.4. and 3.6. describe requirements on the time-series dependence of errors and covariates, while cross-sectional dependence of the errors is restricted through 3.5., which is automatically fulfilled if we assume the errors u_{it} to be *IID*. The last assumption 3.7. restricts the covariates X_{it} to have a sufficient variation across individuals and over time within the groups, which is automatically fulfilled for discrete non-invariant covariates.

The following results concerning consistency of the GFE estimator can be determined:

Theorem 1 (Consistency). *Let Assumptions 3.1 - 3.7 hold. Let $\hat{g}_i = \hat{g}_i(\hat{\beta}, \hat{\lambda})$ denote the GFE estimate for the group g_i^0 to which individual i belongs. We receive for $N, T \to \infty$:*

$$\hat{\beta} \xrightarrow{p} \beta^0, \tag{5.12}$$

$$\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\lambda}_{\hat{g}_i t} - \lambda^0_{g_i^0 t})^2 \xrightarrow{p} 0. \tag{5.13}$$

For the proof see Appendix A of Bonhomme and Manresa (2015).

5.3.2 Asymptotic Distribution

Let $(X_k, k \in \mathbb{Z})$ denotes a (not necessarily stationary) sequence of random variables. Following Bradley (2005), we define for $-\infty \leq J \leq L \leq \infty$ the σ-algebra $\mathcal{F}_J^L$ as the σ-algebra generated by $(X_k, J \leq k \leq L, k \in \mathbb{Z})$. We further define for $n \geq 1$ the coefficient

$$\alpha(n) := \sup_{j \in \mathbb{Z}} \sup_{A \in \mathcal{F}_{-\infty}^{j}, B \in \mathcal{F}_{j+n}^{\infty}} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)|$$

If $\alpha(n) \longrightarrow 0$ for $n \longrightarrow \infty$, then the random sequence is called **strongly mixing** with mixing coefficient α (or α-mixing).[2]

We assume the following assumptions that will be used in this section to describe the asymptotic properties of the GFE estimator:

Assumption 4. 1. *For all $g \in \{1, ..., G\}$:* $\quad \text{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{g_i = g\} =: \pi_g > 0.$

2. *For all $(g, \tilde{g}) \in \{1, ..., G\}^2$ with $g \neq \tilde{g}$:* $\quad \text{plim}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} (\lambda^0_{gt} - \lambda^0_{\tilde{g}t})^2 = c_{g,\tilde{g}} > 0.$

3. *There exist constants $a, d_1 > 0$ and a sequence $\alpha[t] \leq e^{-at^{d_1}}$ such that: for all $i \in \{1, ..., N\}$ and $(g, \tilde{g}) \in \{1, ..., G\}^2$ with $g \neq \tilde{g}$: $\{u_{it}\}_t$, $\{\lambda^0_{gt} - \lambda^0_{\tilde{g}t}\}_t$, and $\{(\lambda^0_{gt} - \lambda^0_{\tilde{g}t})u_{it})\}_t$ are strongly mixing processes with mixing coefficients $\alpha[t]$ and $\mathbb{E}[(\lambda^0_{gt} - \lambda^0_{\tilde{g}t})u_{it}] = 0$*

4. *There exist constants $b, d_2 > 0$ such that for all $i \in \{1, ..., N\}, t \in \{1, ..., T\}, m > 0$:*

[2] See Bradley (2005) for further information on mixing conditions.

$$\mathbb{P}[|u_{it}| > m] \leq e^{1-(m/b)^{d_2}}$$

5. *There exists a constant $M^* > 0$ such that for $N, T \to \infty$:*

$$\sup_{i \in \{1,\dots,N\}} \mathbb{P}\Big(\tfrac{1}{T}\textstyle\sum_{t=1}^{T} ||X_{it}|| \geq M^*\big) = o(T^{-\delta}) \quad \forall \delta > 0. \tag{5.14}$$

Assumptions 4.1. and 4.2. require that for each of the G population groups there exist enough observations and that the G population groups are well-separated, respectively. Conditions for the dependence and tail properties of the error u_{it} are imposed through assumptions 4.3. and 4.4. The last assumption 4.5. describes a condition on the distribution of the covariates X_{it}.

The following theorem states that under the assumptions 3 and 4 the infeasible least squares estimator from equation 5.8 and the GFE estimator are asymptotically equivalent:

Theorem 2 (Asymptotic Distribution)**.** *Let Assumptions 3 and 4 hold.*
For all $\delta > 0$ and $N, T \to \infty$:

$$\mathbb{P}(\sup_{i \in \{1,\dots,N\}} |\hat{g}_i - g_i^0| > 0) = o(1) + o(NT^{-\delta}) \tag{5.15}$$

$$\hat{\beta} = \tilde{\beta} + o_p(T^{-\delta}) \tag{5.16}$$

$$\hat{\lambda}_{gt} = \tilde{\lambda}_{gt} + o_p(T^{-\delta}) \qquad \forall g, t \tag{5.17}$$

For the proof see Appendix B of Bonhomme and Manresa (2015).

We need the following assumptions to be fulfilled to describe the asymptotic distribution of the least squares estimator $(\tilde{\beta}, \tilde{\alpha})$.

Assumption 5. *Let $\bar{X}_{gt}$ denote the mean of X_{it} in group $g_i^0 = g$.*

1. *For all $i, j \in \{1, ..., N\}, t \in \{1, ..., T\}$:* $\quad \mathbb{E}(X_{jt}u_{it}) = 0.$

2. *There exist positive definite matrices* Σ *and* Ω *such that:*

$$\Sigma := \operatorname*{plim}_{N,T\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it}-\bar{X}_{g_i^0 t})(X_{it}-\bar{X}_{g_i^0 t})' \tag{5.18}$$

$$\Omega := \lim_{N,T\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}[u_{it}u_{js}(X_{it}-\bar{X}_{g_i^0 t})(X_{js}-\bar{X}_{g_j^0 s})']. \tag{5.19}$$

3. *For* $N, T \to \infty$*:*

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}((X_{it}-\bar{X}_{g_i^0 t})u_{it} \xrightarrow{d} \mathcal{N}(0,\Omega) \tag{5.20}$$

4. *For all* $(g,t) \in \{1,...,G\}\times\{1,...,T\}$*:*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}\{g_i^0=g\}\mathbf{1}\{g_j^0=g\}u_{it}u_{jt}] =: \omega_{gt} > 0 \tag{5.21}$$

5. *For all* $(g,t) \in \{1,...,G\}\times\{1,...,T\}$ *and* $N, T \to \infty$*:*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{1}\{g_i^0=g\}u_{it} \xrightarrow{d} \mathcal{N}(0,\omega_{gt}) \tag{5.22}$$

The prior assumptions can be summed up as follows: assumptions 5.1. - 5.3. restrict the (infeasible) least squares estimator $\tilde{\beta}$ of the coefficient vector to have a standard asymptotic distribution. The least squares estimator $\tilde{\lambda}_{gt}$ of the grouped time effects is required to have a standard asymptotic distribution through assumptions 5.4. and 5.5.

Corollary 1 (Asymptotic Distribution)**.** *Let Assumptions 3, 4 and 5 hold, and let* $N, T \to \infty$ *such that, for some* $\nu > 0, N/T^\nu \to 0$*. Then we receive:*

$$\sqrt{NT}(\hat{\beta}-\beta^0) \xrightarrow{d} \mathcal{N}(0,\Sigma^{-1}\Omega\Sigma^{-1}) \tag{5.23}$$

$$\sqrt{N}(\hat{\lambda}_{gt}-\lambda_{gt}^0) \xrightarrow{d} \mathcal{N}(0,\frac{\omega_{gt}}{\pi_g^2}) \qquad \forall (g,t) \tag{5.24}$$

where π_g *is defined in assumption 4.1. and* Σ, Ω *and* ω_{gt} *are defined in assumption 5.2.*

For the proof see the supplementary appendix to Bonhomme and Manresa (2015).[3]

5.3.3 GFE Estimator for Unbalanced Panel

The GFE estimator can handle with unbalanced panel data sets. We define the indicator variable $d_{it} := 1$ for the case, when observations y_{it} and X_{it} are both available and $d_{it} := 0$ if one of the two observations is missing.

For a unbalanced panel with maximum time span T the GFE estimator is of the following form:

$$(\hat{\beta}, \hat{\lambda}, \hat{\gamma}) = \underset{(\beta,\lambda,\gamma)\in B\times\Lambda^{GT}\times\Gamma_G}{\arg\min} \sum_{i=1}^{N}\sum_{t=1}^{T} d_{it}(y_{it} - \beta X_{it} - \lambda_{g_i t})^2 \tag{5.25}$$

So the only difference to the GFE estimator for a balanced panel is the inclusion of the indicator variable d_{it} into the minimization problem.

[3]available at https://www.dropbox.com/s/h2hk43owrl6rwh1/Bonhomme_Manresa_appendix.pdf?dl=0

5.3.4 Algorithms

The algorithm 1 by Bonhomme and Manresa (2015) is an iterative algorithm, being used within the algorithm 2. In step 2 of algorithm 1, every individual i is being assigned to the group g_i, at which the objective function is being minimized. In step 3, an OLS regression, including an interaction term for the interaction between group membership and time dummies, provides the parameters β and λ. However, algorithm 1 has two drawbacks: firstly, it heavily depends on the choice of starting values and, secondly, it can happen that the algorithm calculates a so-called degenerate solution, which is a solution with less than G non-empty groups.[4]

Algorithm 1 Iterative

1: Let $(\beta^{(0)}, \lambda^{(0)}) \in B \times \Lambda^{GT}$ be some starting values. Set $s = 0$.
2: (Assignment) Calculate for all $i \in \{1, ..., N\}$:

$$g_i^{(s+1)} = \underset{g \in \{1,...,G\}}{\arg\min} \sum_{t=1}^{T} (y_{it} - \beta^{(s)} X_{it} - \lambda_{gt}^{(s)})^2$$

3: (Update) Calculate:

$$(\beta^{(s+1)}, \lambda^{(s+1)}) = \underset{(\beta,\lambda) \in B \times \Lambda^{GT}}{\arg\min} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - \beta X_{it} - \lambda_{g_i^{(s+1)} t})^2$$

4: Set $s = s + 1$ and got to Step 2. Continue until numerical convergence is fulfilled.

[4]The algorithm can easily be adapted to avoid degenerate solutions, see Hansen and Mladenovic (2001) for further information.

The second algorithm is based on the variable neighborhood search (VNS) heuristic proposed by Hansen and Mladenovic (2001). The algorithm is presented on page 2 in the supplementary appendix to Bonhomme and Manresa (2015).[5] Two search steps in algorithm 2 make this algorithm more efficient than the iterative algorithm 1, which has only one search step. The first search step is the algorithm 1 embedded into step 4 of this algorithm. The second search step (step 5) reassigns every individual to each group, updating the group membership of one individual, if the value of the objective function is being reduced by this reassignment. According to Hansen and Mladenovic (2001) the combination of those two search steps performs better than both of the search steps alone.
Another improvement in comparison to algorithm 1 is step 3, which consists of neighborhood jumps of increasing size, where the individuals are being reassigned to randomly chosen groups. This step makes it easier for the algorithm to leave local optima behind and reach a global optimum.

Before running the algorithm, the following parameters have to be set: the maximum neighborhood size $neigh_{max}$, the maximum number of iterations $iter_{max}$ and the number of starting values N_s.

Algorithm 2 Variable Neighborhood Search

1: Let $(\beta, \lambda) \in B \times \Lambda^{GT}$ be some starting values.
Perform one assignment step of Algorithm 1 and obtain an initial grouping γ_{init}.
Set $iter_{max}$ and $neigh_{max}$ to some desired values.
Set j=0.
Set $\gamma^* = \gamma_{init}$.
2: Set $n = 1$.
3: (Neighborhood jump) Relocate n randomly selected units to n randomly chosen groups, receiving a new grouping $\gamma^{'}$.
Perform one update step of Algorithm 1 and obtain new parameter values $(\beta^{'}, \lambda^{'})$.
4: Set $(\beta^{(0)}, \lambda^{(0)}) = (\beta^{'}, \lambda^{'})$ and apply Algorithm 1.
5: (Local Search) Starting from the grouping $\gamma = \{g_1, ..., g_N)$ obtained in step 4, systematically check all reassignments of units $i \in \{1, ..., N\}$ to groups $g \in \{1, ..., G\}$ (for $g \neq g_i$), updating g_i when the objective function decreases;
stop when no further re-assignment improves the objective function.
Let the resulting grouping be $\gamma^{''}$.
6: If the objective function using $\gamma^{''}$ improves relative to the one using γ^*, then set $\gamma^* = \gamma^{''}$ and go to step 2; otherwise, set $n = n + 1$ and go to step 7.
7: If $n \leq neigh_{max}$, then go to step 3; otherwise got to step 8.
8: Set $j = j + 1$. If $j > iter_{max}$, then Stop; otherwise go to step 2.

[5]available at https://www.dropbox.com/s/h2hk43owrl6rwh1/Bonhomme_Manresa_appendix.pdf?dl=0

6 Results

6.1 Estimates

Table 6.1 shows the main empirical results. The first column contains the estimated coefficients from the OLS estimator. The second and the third column contain the estimates from the FE and GFE estimator, respectively. For all 3 estimators the coefficient of income inequality is negative and statistically significant.

Table 6.1 Results

	Dependent variable:		
	$\log(CO_2)$		
	OLS	FE	GFE
log(GDP)	2.747***	1.536***	2.066***
	(0.293)	(0.268)	(0.117)
$\log(GDP)^2$	−0.166***	−0.140***	−0.112***
	(0.008)	(0.008)	(0.003)
log(GINI)	−4.304***	−4.612***	−3.359***
	(0.528)	(0.525)	(0.211)
log(GDP)*log(GINI)	0.378***	0.439***	0.293***
	(0.057)	(0.059)	(0.023)
Intercept	−7.143***		
	(2.229)		
Observations	3,154	3,154	3,154
R^2	0.871	0.305	0.981
Adjusted R^2	0.871	0.266	0.980
F Statistic	5,321.131*** (df = 4; 3149)	327.035*** (df = 4; 2986)	1,215.800*** (df = 129; 3025)

Note: Standard errors are given in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Our estimated results are slightly different to the ones in Grunewald et al. (2017), where the coefficients for income inequality & for the interaction term from the FE estimator are not statistically significant. Additionally, all coefficients for income inequality are lower than the ones estimated by Grunewald et al. (2017), which also applies for per capita GDP. Obviously, this is no surprise, as we have not chosen the exact same country selection as Grunewald et al. (2017).

Furthermore, the results show that the effect of income inequality on per capita CO_2 emissions depends on per capita GDP. This can be seen by the significant coefficient of the interaction term for all 3 estimators. For countries with a per capita GDP level below a certain threshold an increase in income inequality leads (ceteris paribus) to a decrease of per capita CO_2 emissions. If per capita income reaches this certain threshold, the negative effect of income inequality on CO_2 emissions is reversed as the coefficient of the interaction term counterbalances the coefficient of income inequality. This implies that for countries with income above the threshold of per capita GDP lowering inequality will reduce per capita CO_2 emissions.
The threshold level from the OLS model is 88 101 international dollars, which corresponds to high-income countries like Luxembourg (88 610 international dollars in 2005). The FE estimator provides a lower threshold level of 36 523 international dollars corresponding to the GDP per capita value of high-income countries like France (37 576 international dollars in 2014), Japan (37 337 international dollars in 2014) or New Zealand (34 608 international dollars in 2014). The threshold level from the GFE model are 95 241 international dollars. This threshold has been reached only by Quatar (112 532 international dollars) and by Luxembourg (97 864 international dollars in 2007).
While the threshold level from the FE estimator is reached by most high-income countries, it can be argued that the threshold levels from the OLS and GFE estimators might seem implausible as they both have been reached only by two countries from the sample. Those extremely high threshold values would imply that for most countries apart the richest ones reducing income inequality will result in higher CO_2 emissions.

Our estimated results are - similarly to the ones in Grunewald et al. (2017) - consistent with the hypothesis of the EKC. The estimated coefficient for the logarithm of GDP per capita is positive and statistically significant for all 3 estimators. The square of logarithm of GDP per capita has a negative coefficient, which is as well statistically significant for all 3 estimators. The estimated coefficients imply that increasing income per capita at an initial stage leads ceteris paribus to an increase of CO_2 emissions per capita until income reaches a certain threshold. Then further increasing income results in a decrease of CO_2 emissions. This relationship is consistent with the inverted-U relationship from the EKC hypothesis. The estimated turning point of the EKC depends for all 3 estimators on income inequality, as the coefficient of the interaction term is always statistically significant.

Figure 6.1: Scatter plots for the relationship between CO_2 emissions and the Gini Index as well as GDP per capita for the DACH region

Figure 6.1 shows scatter plots, where CO_2 emissions are plotted against each GDP per capita and the Gini Index for the countries Austria, Germany and Switzerland. The scatter plots, where CO_2 emissions are assigned to the y-axis and GDP per capita to the x-axis, are depicted in the left column of Figure 6.1, whereas the same is being done for the Gini Index in the right column. The EKC theory indicates a concave relationship, namely in the form of an inverted-U, between GDP per capita and CO_2 emissions. Interestingly this relationship is not being observed for any of the three countries, Germany and Switzerland rather show a linear downward trend for increasing per capita GDP values. The negative coefficient of income inequality from the estimated results would imply a linear downward trend of CO_2 emissions for higher values of the Gini Index, however this relationship is only being observed for Germany.

6.2 GFE Groups

We are not only interested in the estimated coefficients by the GFE estimator, but also in the estimated groups and the grouped fixed effects. Figure 6.2 shows the estimated grouped fixed effects. The GFE groups are named in descending order by the values of the grouped fixed effects. This means that the grouped fixed effects with the highest values belong to the GFE group A and the ones with the lowest values to the GFE group E. The estimated grouped fixed effects are all negative and appear to sink over time for each GFE group, except for group E.

Figure 6.2: Estimated grouped fixed effects for the GFE groups

Figure A.1 in the appendix shows the estimated time effects from the FE estimator. Comparing the estimated grouped fixed effects with the time effects by the FE estimator, we observe, that the time effects show a different behaviour over time than the grouped fixed effects. Firstly, all time effects are positive. Secondly, there is no downwards trend observable.
Additionally, it can be observed that the mean values of CO_2 emissions are ordered in a descending way from GFE group A to E, as can be seen in table 6.2. So the grouped fixed effects calculated by the GFE estimator are higher for GFE groups with higher mean values of CO_2 emissions.

Table 6.2 Summary statistics of main variables

GFE group	A	B	C	D	E
CO_2					
1Q	2.952	1.173	1.109	0.367	0.073
Mean	10.833	6.214	4.191	2.614	0.735
3Q	11.991	8.911	7.260	5.271	0.605
GDP					
1Q	3 485	3 426.1	5 463.2	2 725.9	1 463.4
Mean	14 006	15 559.1	15 939.8	14 006.9	7 236.8
3Q	13 738	20 361.3	24 049.5	24 592	5 803
GINI					
1Q	31.73	31.3	31.6	32.15	36.6
Mean	38.05	36.89	38.24	38.52	39.7
3Q	41.7	42.4	45.1	44.9	42.8

Table 6.3 and figure 6.3 show all 140 countries divided into the 5 GFE groups. The first and the second group contain most of the countries, group C even holds up to 45 countries. It is difficult to find a possible interpretation of the grouping of the 140 nations. Every group contains both high-income and low-income countries, thus it cannot be argued that countries are divided into those groups according to their per capita income levels. For example, on one hand group E contains low income countries like Burundi, Madagascar, Mali and Uganda, but on the other hand Switzerland, one of the countries with highest per capita income values within the data set, is being assigned to this group.

Table 6.3 Estimated GFE groups

	A	B	C	D	E
1	Azerbaijan	Armenia	Algeria	Albania	Burundi
2	Bosnia and Herzegovina	Australia	Argentina	Austria	Cameroon
3	China	Belarus	Barbados	Bangladesh	Lao PDR
4	Estonia	Bolivia	Belgium	Brazil	Madagascar
5	Kazakhstan	Bulgaria	Belize	Burkina Faso	Mali
6	Kyrgyz Republic	Canada	Botswana	Cambodia	Nepal
7	Lesotho	Czech Republic	Cabo Verde	Colombia	Paraguay
8	Mongolia	Guyana	Central African Republic	Costa Rica	Sri Lanka
9	Qatar	Honduras	Chile	Cote d'Ivoire	Sudan
10	Russian Federation	India	Denmark	Croatia	Switzerland
11	South Africa	Iran, Islamic Rep.	Dominican Republic	Cyprus	Tanzania
12	Trinidad and Tobago	Israel	Ecuador	El Salvador	Uganda
13	Turkmenistan	Jamaica	Egypt, Arab Rep.	Eswatini	Uruguay
14	Ukraine	Jordan	Finland	Ethiopia	West Bank and Gaza
15	Zimbabwe	Korea, Rep.	Gambia, The	Fiji	
16		Luxembourg	Georgia	France	
17		Malaysia	Germany	Ghana	
18		Micronesia, Fed. Sts.	Greece	Guatemala	
19		Moldova	Guinea	Guinea-Bissau	
20		Mozambique	Hungary	Hong Kong SAR, China	
21		North Macedonia	Indonesia	Iceland	
22		Papua New Guinea	Ireland	Italy	
23		Poland	Japan	Kenya	
24		Tajikistan	Lebanon	Latvia	
25		United States	Lithuania	Malta	
26		Venezuela, RB	Malawi	Mauritania	
27		Vietnam	Mexico	Mauritius	
28			Morocco	Niger	
29			Namibia	Nigeria	
30			Netherlands	Norway	
31			New Zealand	Pakistan	
32			Nicaragua	Panama	
33			Romania	Peru	
34			Senegal	Philippines	
35			Sierra Leone	Portugal	
36			Singapore	Rwanda	
37			Slovak Republic	Spain	
38			Slovenia	Sweden	
39			St. Vincent and the Grenadines	Zambia	
40			Thailand		
41			Tonga		
42			Tunisia		
43			Turkey		
44			United Kingdom		
45			Yemen, Rep.		

Looking at figure 6.3, the geographical proximity of countries does not seem to be relevant for any group except maybe group A. This group contains (with some exceptions) several countries in Central and East Asia like Azerbaijan, the Kyrgyz Republic, Kazakhstan, Mongolia, Russia and Turkmenistan.

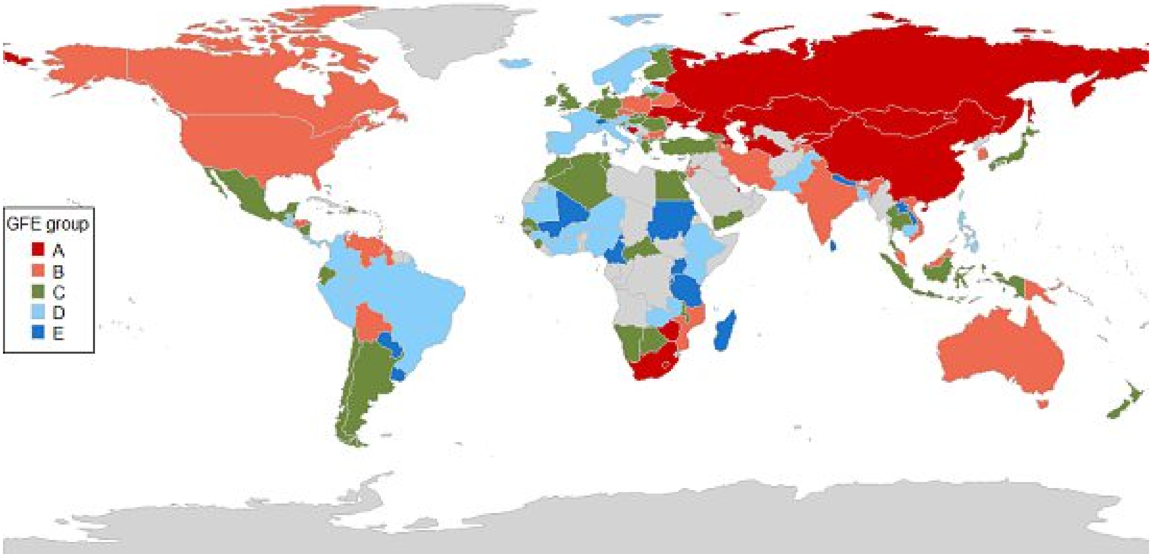

Figure 6.3: GFE groups

In comparison to the estimated grouped fixed effects and GFE groups in Grunewald et al. (2017), the following can be stated:

- The GFE groups cannot be meaningfully compared to the ones estimated in Grunewald et al. (2017), as the division of the countries into the GFE groups is quite different to the grouping in Grunewald et al. (2017). Except for group A, which can be most likely compared to the GFE group 3 in Grunewald et al. (2017). However our GFE group A contains in addition the countries Bosnia and Herzogovina, Estonia, Kyrgyz Republic, Lesotho, Quatar, Russia, South Africa, Trinidad and Tobago and Zimbabwe and excludes Moldova and Uzbekistan.

- The estimated grouped fixed effects in Grunewald et al. (2017) are located mostly near zero for all groups except group 3, where the grouped fixed effects sink over time below the value −2. The grouped fixed effects tend to decline over time only for the GFE groups 3 and 1. Our grouped fixed effects, depicted in Figure 6.2, are quite different, as they are all located in an interval between the values -4 and -7. Additionally, the grouped fixed effects for all GFE groups besides group E decrease over time, which is also different to the estimated GFE groups in Grunewald et al. (2017). Interestingly the grouped fixed effects for the GFE group A in our results and the GFE group 3 in Grunewald et al. (2017) have the same behavior, as the effects for both groups decrease relatively sharply over time. Therefore the GFE group A is the only one, which can be most likely meaningfully compared to a GFE group in Grunewald et al. (2017) concerning grouping and the behavior of the grouped fixed effects, namely the GFE group 3.

Another aspect concerning the characterization of GFE group membership, which we want to analyse, is emission intensity. Generally, emission intensity is measured as the level of GHG emissions per unit of GDP. [1] Therefore, emission intensity depends on energy intensity and the fuel mix being used to generate energy.
CO_2 emissions intensity denotes the level of CO_2 emissions per unit of GDP. We use the indicator "CO2 emissions (kg per 2017 PPP \$ of GDP)" from the World Bank Development Indicators to further analyze differences in CO_2 emissions intensity between countries in each GFE group. The CO_2 emissions intensity levels for each GFE group are depicted in the figures A.2, A.3, A.4, A.5 and A.6.

Table 6.4 Summary statistics of CO_2 emissions intensity

GFE group	A	B	C	D	E
1Q	0.481	0.238	0.175	0.109	0.058
Mean	0.763	0.351	0.222	0.146	0.077
3Q	0.937	0.416	0.267	0.180	0.096

Firstly, we observe for the GFE group A that the CO_2 emissions intensity levels are the highest among all GFE groups. The average CO_2 emissions intensity level for all countries in the GFE group A is 0.763 and almost all CO_2 emissions intensity levels are above 0.25, as can be seen in figure A.2. Furthermore, a clear downward trend can be identified for most of the countries within the group. Interestingly the group contains nations like Ukraine and Turkmenistan, that were able to decrease extremely high CO_2 emissions intensity levels in the nineties.
The following results can be observed for nations assigned to the GFE group B: the average is 0.351, most of the CO_2 emissions intensity levels are below 0.9. and a sinking trend is observable, as can be seen from figure A.3. This sinking trend is clearly evident for countries with quite high CO_2 emissions intensity values in the nineties like Belarus, Bulgaria,

[1] See Baumert et al. (2005, page 26) for further information.

Moldova and Poland.
The countries in the GFE group C have relatively low CO_2 emissions intensity values (average 0.222 and most of the CO_2 emissions intensity levels are below 0.6) and a clear downwards trend can be observed for nearly all countries within the group (see figure A.4), especially for Georgia, Romania and the Slovak Republic.
CO_2 emissions intensity values for the GFE group D are similar to the ones for GFE group C: the average is 0.146 and all of the CO_2 emissions intensity levels are below 0.4. However, there is only a slightly sinking trend observable for the GFE group D as the figure A.5 shows.
Lastly, countries in the GFE group E have the lowest CO_2 emissions intensity values (average 0.077 and below 0.15 for nearly all countries except West Bank and Gaza) and there is no trend recognizable for all countries within the GFE group. In fact, for some countries the is rather an upwards trend observable (see figure A.6).
The assignment of the countries to each GFE group can be characterized by differences between their CO_2 emissions intensity levels and their long time behaviour.

Additionally, we compare the endogenous variation for all countries within one GFE group and with the group-specific fixed effects to further assess the fit of the GFE model. We do this by first removing the exogenous variation explained by GDP per capita and the Gini Index. The figures A.7, A.8, A.9, A.10 and A.11 show the remaining variation of the countries within each GFE group. The black line in each figure represents the grouped fixed effects for each GFE group. Apart from a few outliers the values vary within an interval of approximately the same length for all GFE groups. The grouped fixed effects seem to approximate the endogenous variation of the countries in every GFE group quite well, as they are always in the middle of the interval.

7 Sensitivity Analysis

We employ several ways to test the robustness of our results. To begin with, we want to perform a sensitivity analysis by including another transmission channel, through which income inequality might influence CO_2 emissions, into the GFE model.
As already mentioned in section 2.1, income inequality might affect environmental pressure via two main transmission channels. The first channel is characterized by the way how income inequality affects CO_2 emissions via consumer behavior. The second channel describes how income inequality might influence CO_2 emissions through social agreements and the implementation of environmental policies. However, as it is very difficult to measure the amount and quality of environmental regulations, we will use the Polity Index, which has been presented in section 3.4, as a proxy variable.

Table 7.1 shows the result of including the Polity Index into the GFE model. More specifically, the Polity Index is being included as an additional covariate into the regression, but the number of GFE groups as well as the group membership assignments stay the same, as they are not being recalculated by the algorithm by Bonhomme and Manresa (2015).
Although the coefficient for the Polity Index is statistically significant, it is very small. The coefficients for the other regressors only change slightly without any change of sign or significance. Therefore, our main findings stay robust regarding the inclusion of the Polity Index into the GFE model.

Table 7.1 Sensitivity analysis - Polity Index

	Dependent variable:	
	$\log(CO_2)$	
	GFE	GFE with Polity Index
log(GDP)	2.066***	1.879***
	(0.117)	(0.123)
$\log(GDP)^2$	−0.112***	−0.109***
	(0.003)	(0.003)
log(GINI)	−3.359***	−3.647***
	(0.211)	(0.220)
log(GDP)*log(GINI)	0.293***	0.324***
	(0.023)	(0.024)
Polity Index		0.005***
		(0.001)
Observations	3,154	2,961
R^2	0.981	0.982
Adjusted R^2	0.980	0.981
F Statistic	1,215.800*** (df = 129; 3025)	1,175.714*** (df = 130; 2831)

Note: Standard errors are given in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Secondly we continue our sensitivity analysis by changing the number of GFE groups. As mentioned in chapter 4, we have set the number of GFE groups to 5, so we can compare our estimated results with the ones in Grunewald et al. (2017). We set the number of GFE groups to 4 and to 6 and let the algorithm 2 by Bonhomme and Manresa (2015) recalculate the group membership assignments as well as the grouped fixed effects.

Table 7.2 shows the resulting estimates. The first column shows the estimates from our main model, while the second and the third column show the estimates from the model with 4 GFE groups and the one with 6 GFE groups, respectively. Changing the number of GFE groups has no impact on our main findings. All estimated coefficients stay significant and no sign is reversed. Interestingly the coefficients from the model with 4 GFE groups are only slightly different to the ones from our main model, whereas in the model with 6 GFE groups bigger changes in the coefficients can be observed, especially for the coefficients belonging to the Gini Index and GDP per capita.

Table 7.2 Sensitivity analysis - number of GFE groups

	Dependent variable:		
	$\log(CO_2)$		
	5 GFE groups	4 GFE groups	6 GFE groups
log(GDP)	2.066***	2.828***	1.001***
	(0.117)	(0.130)	(0.112)
$\log(GDP)^2$	−0.112***	−0.147***	−0.106***
	(0.003)	(0.004)	(0.003)
log(GINI)	−3.359***	−3.236***	−5.310***
	(0.211)	(0.234)	(0.199)
log(GDP)*log(GINI)	0.293***	0.285***	0.539***
	(0.023)	(0.025)	(0.021)
Observations	3,154	3,154	3,154
R^2	0.981	0.976	0.984
Adjusted R^2	0.980	0.975	0.983
F Statistic	1,215.800*** (df = 129; 3025)	1,185.666*** (df = 104; 3050)	1,171.722*** (df = 154; 3000)

Note: Standard errors are given in parentheses.
*p<0.1; **p<0.05; ***p<0.01

We further continue our sensitivity analysis by assessing the impacts of changing the parameters for the algorithm 2 by Bonhomme and Manresa (2015). The algorithm depends on the maximum neighborhood size $neigh_{max}$, the maximum number of iterations $iter_{max}$ and the number of starting values N_s.[1] For the main model the 3 parameters of the algorithm were all set to 10 as previously mentioned in chapter 4. Now we set each of the three parameters to 20 and 50 to see whether an increase of this parameter leads to different estimates. Increasing ceteris paribus each of the parameters has absolutely no impact on the estimated coefficients. The algorithm always calculates the same composition of GFE groups, assigning the same countries to each GFE group, therefore we receive in each case the same resulting estimates. So our main findings stay robust with respect to changing the parameters of the algorithm.

[1]See the supplementary appendix to Bonhomme and Manresa (2015) for further information. Available at https://www.dropbox.com/s/h2hk43owrl6rwh1/Bonhomme_Manresa_appendix.pdf?dl=0.

8 Conclusion

This thesis reports on the potential impact of income inequality on CO_2 emissions. We use the GFE estimator by Bonhomme and Manresa (2015) which allows us to control for unobserved heterogeneity and for grouped fixed effects with different time patterns across the GFE groups, which makes this estimator arguably better suited for our model than other estimators, which are commonly used for panel data analysis. The main findings of this thesis are the following:

The estimated results indicate that the effect of income inequality on per capita CO_2 emissions depends on the income level. Increasing income inequality has a negative effect on CO_2 emissions below a certain threshold of income, while for income levels above this threshold the effect is reversed: higher income inequality leads to higher CO_2 emissions. However, the threshold value for the GFE model is so high that is has been reached only by two countries in the data set. This implies that for low-, middle- as well as for most high-income countries lowering income inequality leads ceteris paribus to higher per capita CO_2 emissions.
Furthermore, the results substantiate the well-known EKC hypothesis as per capita income has a positive effect on CO_2 emissions while the squared per capita income term negatively effects CO_2 emissions.
The conducted sensitivity analysis, which comprises changing the number of GFE groups and parameters of the GFE algorithm by Bonhomme and Manresa (2015) as well as including another possible transmission channel into the GFE model, does not change our main findings. They indicate that for countries aiming at reducing poverty, inequality and/or CO_2 emissions additional measures are needed to counterbalance the trade-off between environmental damage and economic development. This can be pursued by replacing carbon intensive energy consumption with renewable energy technologies, to name but one example.

However, existing theoretical and empirical studies analysing the impact of income inequality on environmental damage including those, that were discussed in the different sections of chapter 2, have contradictory outcomes depending on the chosen data set, estimators and assumptions. Therefore, our results should also be interpreted with caution. Nonetheless they are useful for giving a better insight into the mechanism of how income inequality and economic development in general influences environmental pollution. In addition, further research is necessary to fully understand this mechanism.

Appendix

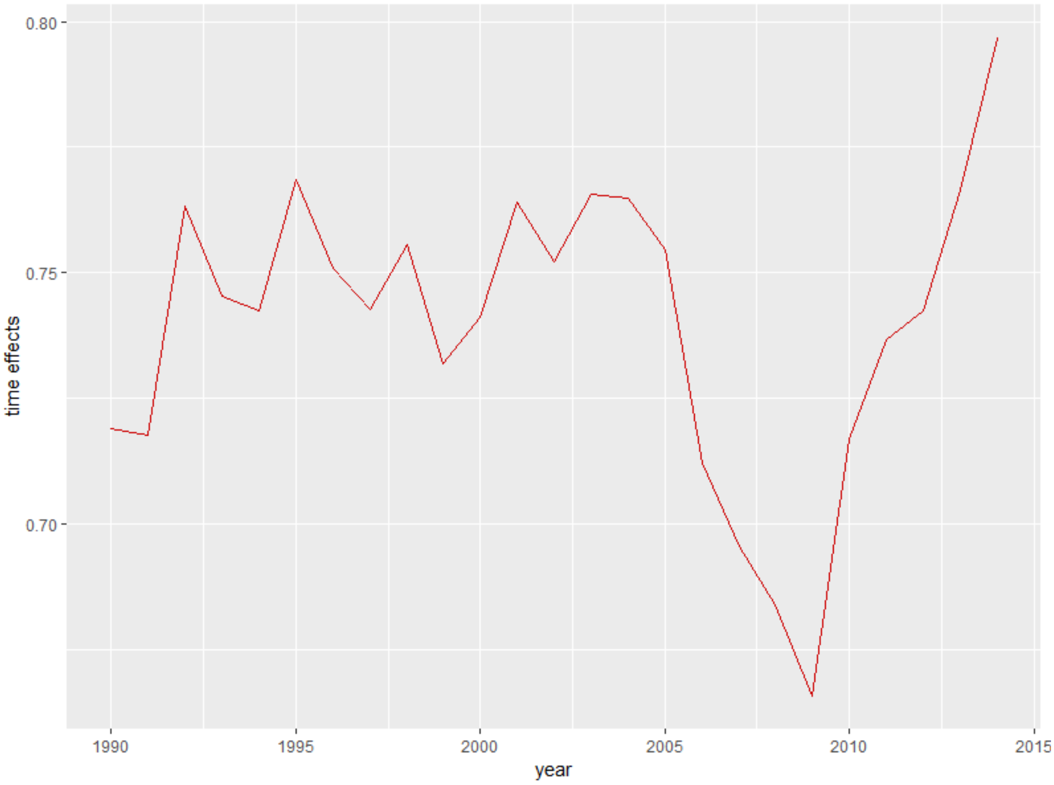

Figure A.1: Estimated time effects from the FE estimator

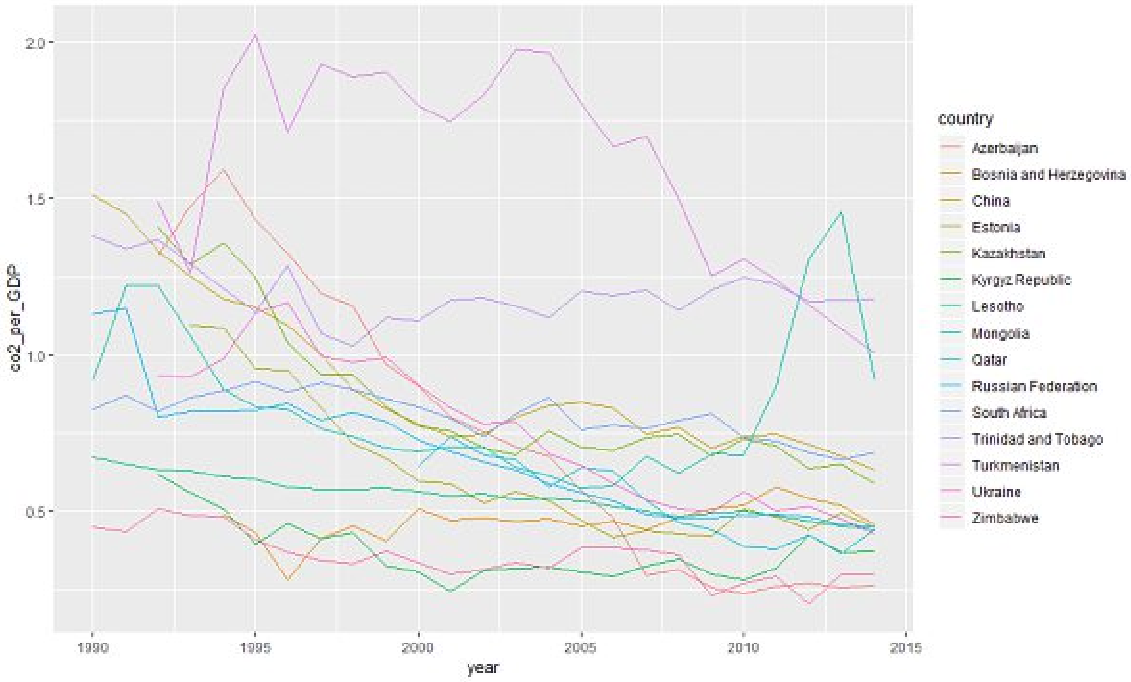

Figure A.2: Emission intensity values for all countries within the GFE group A

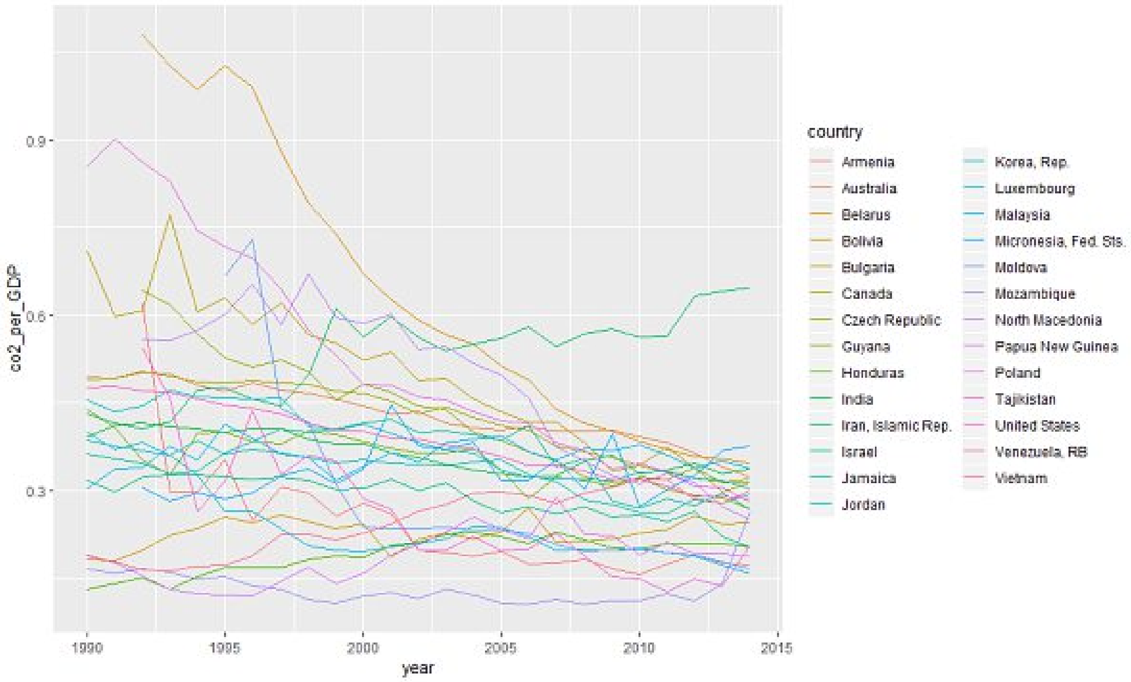

Figure A.3: Emission intensity values for all countries within the GFE group B

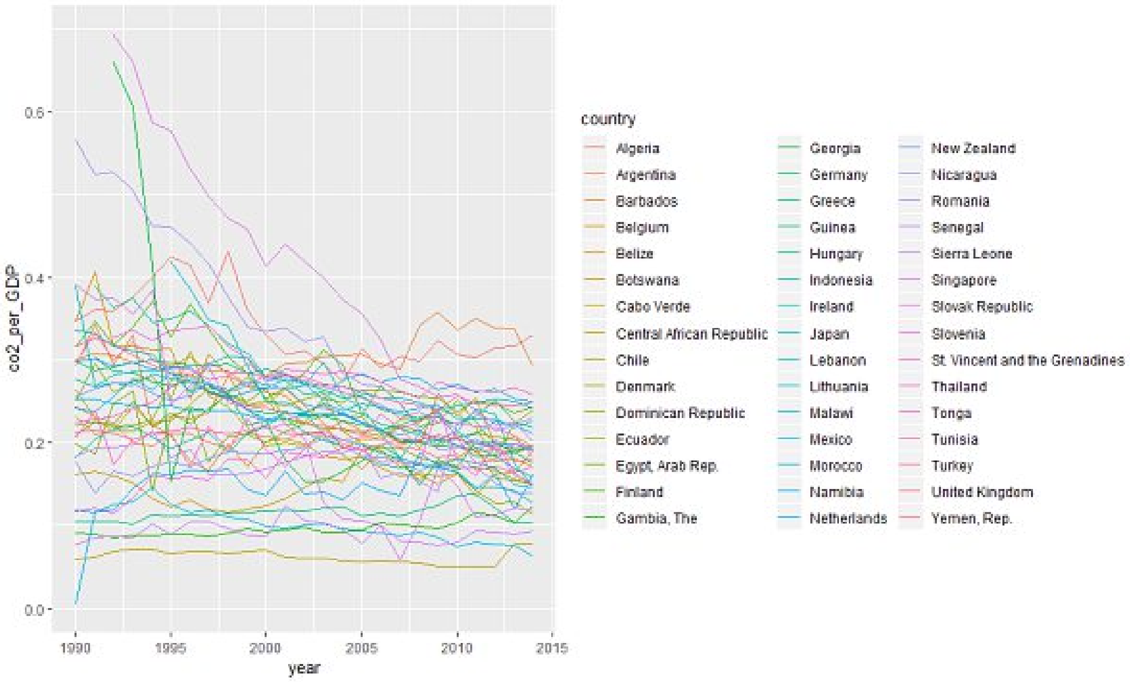

Figure A.4: Emission intensity values for all countries within the GFE group C

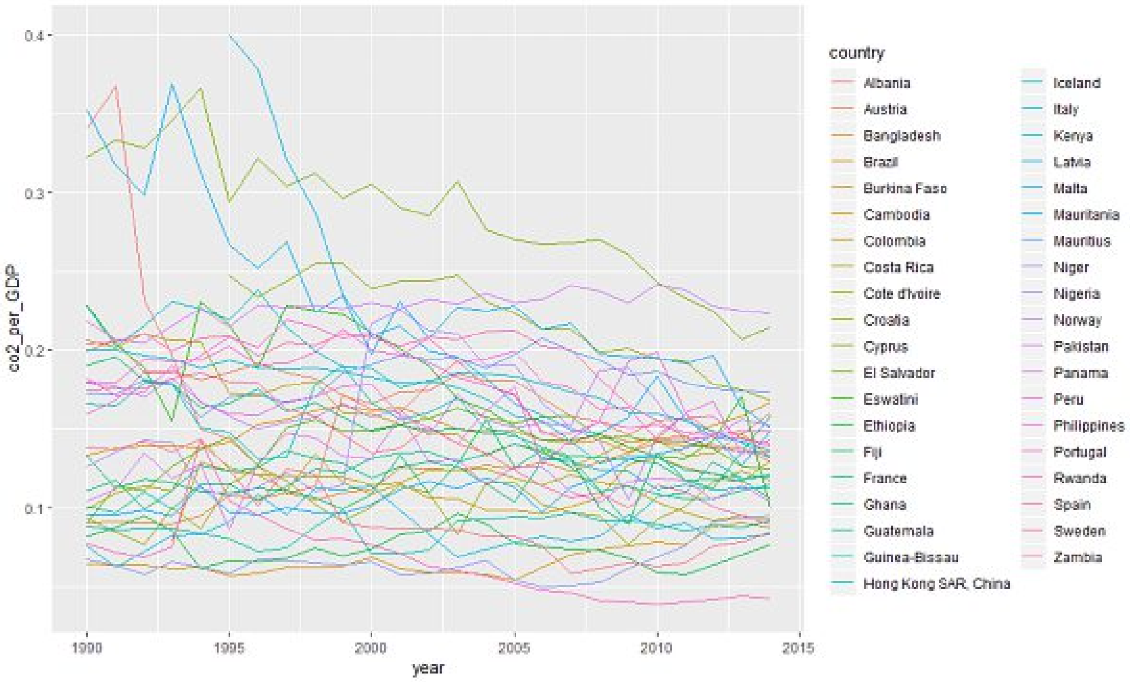

Figure A.5: Emission intensity values for all countries within the GFE group D

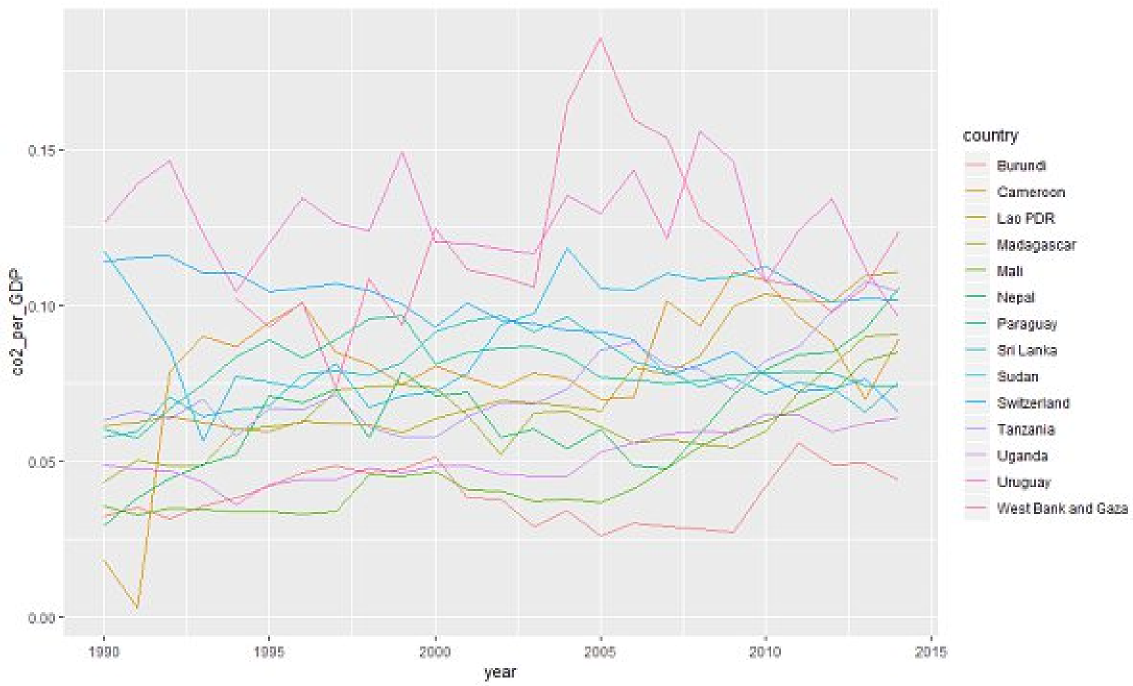

Figure A.6: Emission intensity values for all countries within the GFE group E

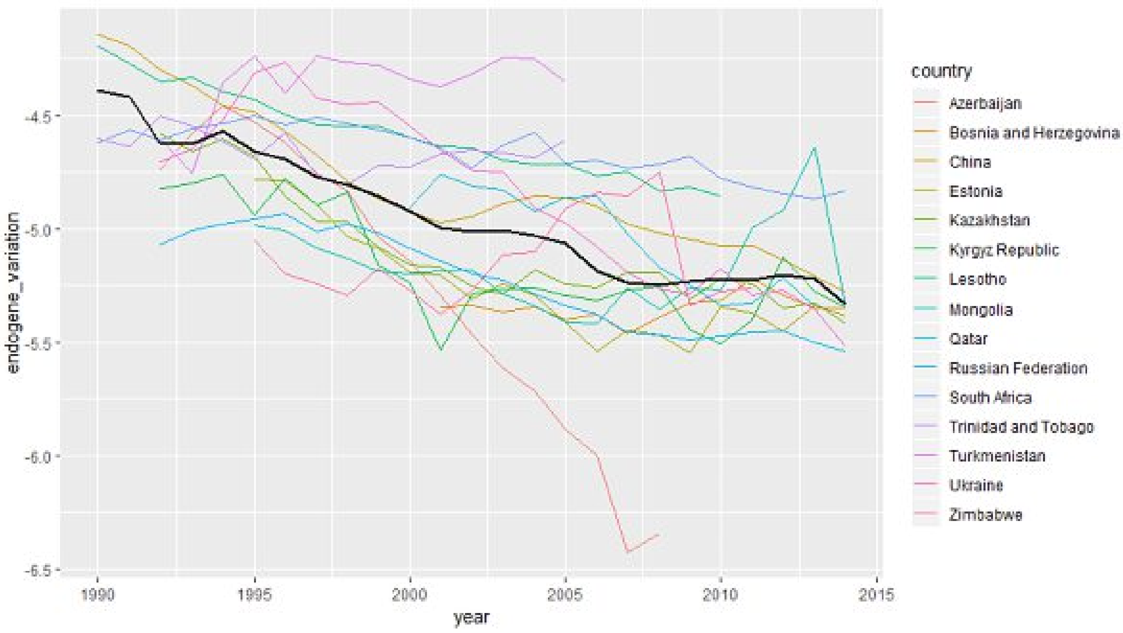

Figure A.7: Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group A

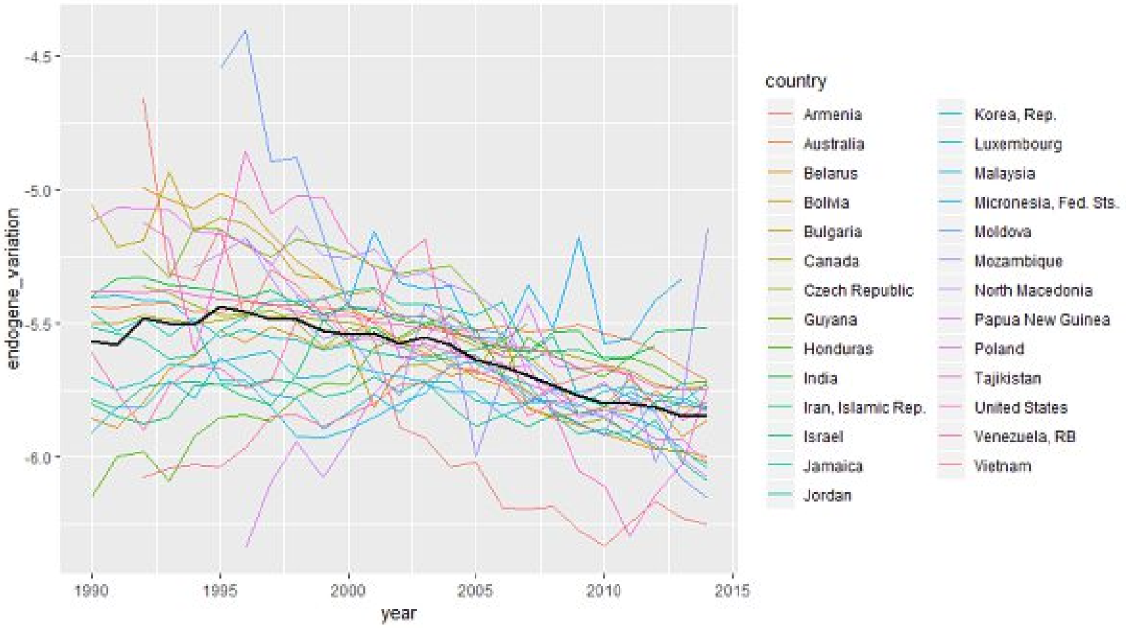

Figure A.8: Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group B

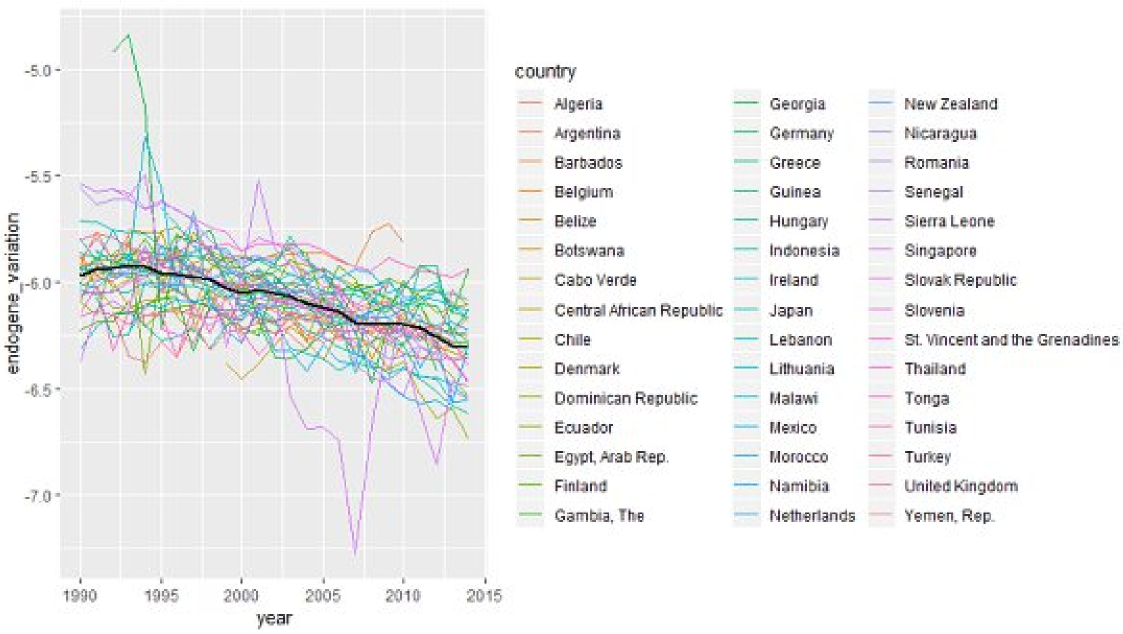

Figure A.9: Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group C

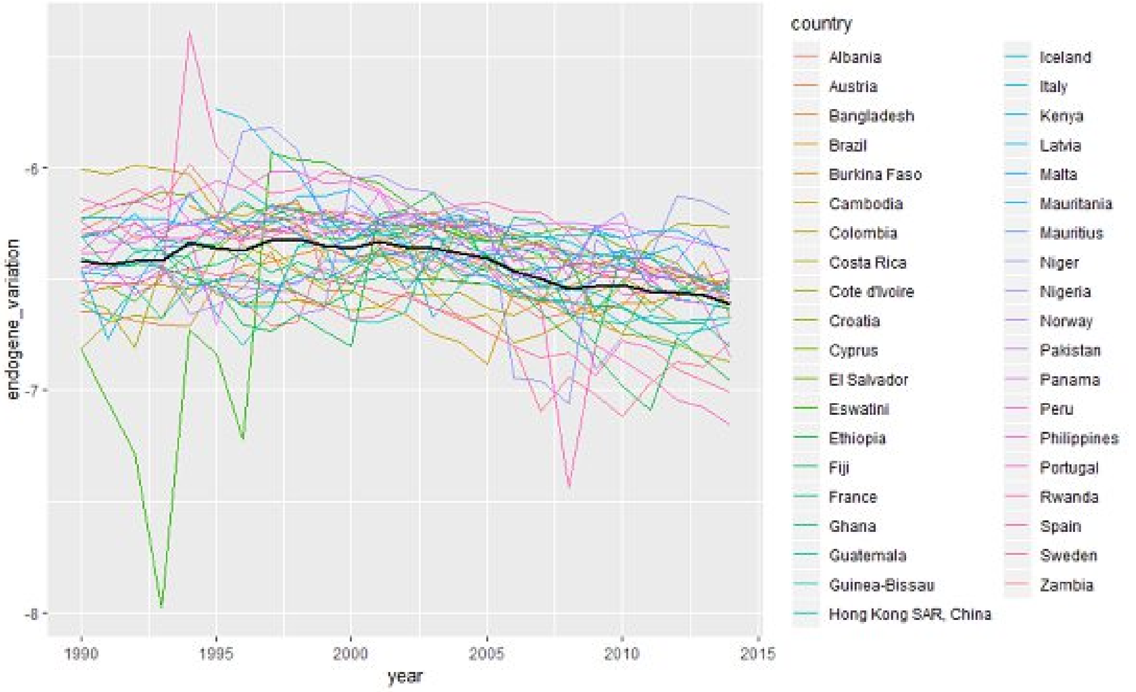

Figure A.10: Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group D

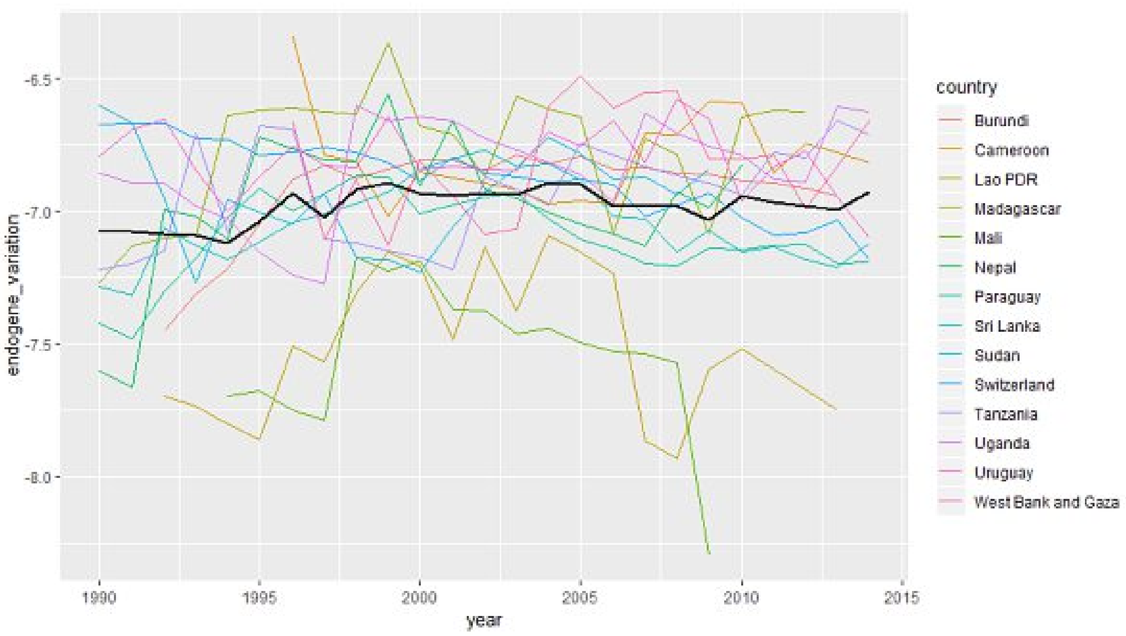

Figure A.11: Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group E

List of Tables

3.1 Definition of variables 12
3.2 Summary statistic 12

6.1 Results 29
6.2 Summary statistics of main variables 34
6.3 Estimated GFE groups 35
6.4 Summary statistics of CO_2 emissions intensity 37

7.1 Sensitivity analysis - Polity Index 40
7.2 Sensitivity analysis - number of GFE groups 41

List of Figures

2.1 Environmental Kuznet curve 3

3.1 Empirical Lorenz curve 11

6.1 Scatter plots for the relationship between CO_2 emissions and the Gini Index as well as GDP per capita for the DACH region 31
6.2 Estimated grouped fixed effects for the GFE groups 33
6.3 GFE groups 36

A.1 Estimated time effects from the FE estimator 45
A.2 Emission intensity values for all countries within the GFE group A 46
A.3 Emission intensity values for all countries within the GFE group B 46
A.4 Emission intensity values for all countries within the GFE group C 47
A.5 Emission intensity values for all countries within the GFE group D 47
A.6 Emission intensity values for all countries within the GFE group E 48
A.7 Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group A 48
A.8 Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group B 49
A.9 Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group C 49
A.10 Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group D 50
A.11 Comparison of endogenous variation with the estimated grouped fixed effects (black) for all countries within the GFE group E 50

List of Abbreviations

CDIAC Carbon Dioxide Information Analysis Center

EKC Environmental Kuznet curve

FE Fixed effects

GDP Gross domestic product

GFE Grouped fixed effects

GHG Grenhouse Gases

IID Independent and identically distributed

INSCR Integrated Network for Societal Conflict Research

IPCC Intergovermental Panel on Climate Change

MPE Marginal propensity to emit

OLS Ordinary least squares

ORNL Oak Ridge National Laboratory

PPP Purchasing power parity

SWIID Standardized World Income Inequality Database

VNS Variable neighborhood search

Bibliography

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

Baumert, Kevin, Tim Herzog, and Jonathan Pershing (2005). *Navigating the numbers: Greenhouse gas data and international climate policy.* World Resources Institute.

Berthe, Alexandre and Luc Elie (2015). „Mechanisms explaining the impact of economic inequality on environmental deterioration“. In: *Ecological economics* 116, pp. 191–200.

Boden, Tom A, Gregg Marland, and Robert J Andres (2017). *Global, regional, and national fossil-fuel CO2 emissions.* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tenn., USA. URL: https://cdiac.ess-dive.lbl.gov/trends/emis/tre_coun.html (visited on 2019-05-11).

Bonhomme, Stéphane and Elena Manresa (2015). „Grouped patterns of heterogeneity in panel data“. In: *Econometrica* 83.3, pp. 1147–1184.

Borghesi, Simone (2006). „Income inequality and the environmental Kuznets curve“. In: *Environment, inequality and collective action* 33.

Boyce, James (Jan. 2007). „Is Inequality Bad for the Environment?“ In: *Political Economy Research Institute, University of Massachusetts at Amherst, Working Papers* 15.

Bradley, Richard C. (2005). „Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions“. In: *Probability Surveys* 2, pp. 107–144.

Cole, Benjamin R and Monty G Marshall (2014). „Global report 2014 conflict, governance, and state fragility“. In: *Center for Systematic Peace.*

Croissant, Yves, Giovanni Millo, et al. (2019). *Panel Data Econometrics with R.* Wiley Online Library.

Frisch, Ragnar and Frederick Waugh (1933). „Partial Time Regressions as Compared with Individual Trends“. In: *Econometrica: Journal of the Econometric Society* 1, pp. 387–401.

Gastwirth, Joseph (Feb. 1972). „The Estimation of the Lorenz Curve and Gini Index“. In: *The Review of Economics and Statistics* 54.3, pp. 306–316.

Grossman, Gene and Alan Krueger (1991). *Environmental impacts of a North American free trade agreement*. NBER Working Papers 3914. National Bureau of Economic Research, Inc.

– (1995). „Economic Growth and the Environment“. In: *The Quarterly Journal of Economics* 110.2, pp. 353–377.

Grunewald, Nicole, Stephan Klasen, Inmaculada Martınez-Zarzoso, and Chris Muris (2017). „The trade-off between income inequality and carbon dioxide emissions“. In: *Ecological Economics* 142, pp. 249–256.

Hansen, Pierre and Nenad Mladenovic (2001). „J-MEANS: a new local search heuristic for minimum sum of squares clustering“. In: *Pattern Recognition* 34, pp. 405–413.

Heerink, Nico, Abay Mulatu, and Erwin Bulte (2001). „Income inequality and the environment: aggregation bias in environmental Kuznets curves“. In: *Ecological Economics* 38.3, pp. 359–367.

Hübler, Michael (2017). „The inequality-emissions nexus in the context of trade and development: a quantile regression approach“. In: *Ecological Economics* 134, pp. 174–185.

Inglehart, Ronald (1990). *Culture Shift in Advanced Industrial Society*. Princeton University Press.

IPCC (2014). *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ed. by Rajendra K Pachauri, Myles R Allen, Vicente R Barros, John Broome, Wolfgang Cramer, Renate Christ, John A Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta, et al. IPCC, Geneva, Switzerland.

Lovell, Michael (1963). „Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis“. In: *Journal of the American Statistical Association* 58.304, pp. 993–1010.

Magnani, Elisabetta (2000). „The Environmental Kuznets Curve, environmental protection policy and income distribution“. In: *Ecological economics* 32.3, pp. 431–443.

Martínez-Zarzoso, Inmaculada and Aurelia Bengochea (2004). „Pooled mean group estimation of an environmental Kuznets curve for CO2“. In: *Economics Letters* 82, pp. 121–126.

Ravallion, Martin, Mark Heil, and Jyotsna Jalan (2000). „Carbon emissions and income inequality“. In: *Oxford Economic Papers* 52.4, pp. 651–669.

Roberts, J. and Peter Grimes (1997). „Carbon Intensity and Economic Development, 1962-1991: A Brief Exploration of the Environmental Kuznets Curve". In: *World Development* 25, pp. 191–198.

Scruggs, Lyle A (1998). „Political and economic inequality and the environment". In: *Ecological Economics* 26.3, pp. 259–275. ISSN: 0921-8009.

Shafik, Nemat and Sushenjit Bandyopadhyay (1992). *Economic growth and environmental quality: time-series and cross-country evidence*. Vol. 904. World Bank Publications.

Solt, Frederick (2009). „Standardizing the world income inequality database". In: *Social Science Quarterly* 90.2, pp. 231–242.

Veblen, Thorstein (1934). „The theory of the leisure class: an economic study of institutions. 1899". In: *New York: The Modern Library*.

Wallace, T Dudley and Ashiq Hussain (1969). „The use of error components models in combining cross section with time series data". In: *Econometrica: Journal of the Econometric Society*, pp. 55–72.

Wooldridge, Jeffrey M (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Development Indicators (2019). World Bank. URL: http://datatopics.worldbank.org/world-development-indicators/ (visited on 2019-06-14).

Zoboli, Roberto, Antonio Musolesi, and Massimiliano Mazzanti (2010). „A panel data heterogeneous Bayesian estimation of environmental Kuznets curves for CO2 emissions". In: *Applied Economics* 42, pp. 2275–2287.