The Effect of Climate Change on Gross Domestic Product per Capita

Master's Thesis

Ann-Katrine Kjærsgaard Nielsen

Aalborg University
Mathematics-Economics

AALBORG UNIVERSITY
STUDENT REPORT

Mathematics-Economics
Aalborg University
http://www.aau.dk

Title:
The Effect of Climate Change on Gross Domestic Product per Capita

Theme:
Climate Econometrics

Project Period:
Spring Semester 2019

Project Group:
5.218b

Participants:
Ann-Katrine Kjærsgaard Nielsen

Supervisor:
Eduardo Vera Valdés

Copies: 3

Page Count: 84

Date of Completion:
7th June, 2019

Abstract:

The aim of this paper is to examine the relationship between global CO_2 emissions, atmospheric CO_2 concentration, temperature anomalies, and the first principal component of GDP pr capita for different economies. The time series are modelled using time trends, and ARIMA and ARFIMA models. They predict that the temperature has increased 1.5°C and 2°C compared to pre-industrial levels in 2058 and 2083, respectively. The relationships between the time series are examined using regression with ARMA errors. Continuing to increase global CO_2 emissions will cause the atmospheric CO_2 concentration and temperature to rise at a higher rate than previously. The models predict that temperature rises will cause economic growth, indicating that the models are too simple to capture the relationship between temperature changes and GDP per capita, and that they should include more information about the causes and consequences of climate change.

Contents

Resumé

Der er stor enighed blandt verdens forskere om, at menneskeskabte klimaforandringer påvirker kloden. Derfor blev Parisaftalen, som vil styrke det globale svar mod klimaforandringer, vedtaget i 2015 af FN. Ét af målene i Parisaftalen er at sørge for, at temperaturen ikke stiger mere end 2°C sammenlignet med før den industrielle revolution og at forsøge at begrænse temperaturstigningen til 1.5°C.

Da klimaforandringer vil påvirke lande forskelligt, er det relevant at undersøge, hvordan forskellige lande med forskellige økonomier vil blive påvirket af klimaforandringer. Denne rapport undersøger derfor relationenerne mellem globale CO_2 udledninger, atmosfærisk CO_2 koncentration, temperaturafvigelser samt Bruttonationalprodukt (BNP) per indbygger i forskellige økonomier.

I stedet for at undersøge BNP per indbygger for alle lande deles de op i fire grupper baseret på Verdensbankens lande grupper: Lav-indkomst, lavere-middel-indkomst, øvre-middel-indkomst og høj-indkomst økonomier. For at reducere dimensionaliteten i hver gruppe anvendes det første principale komponent af landende i hver gruppe.

Ved anvendelse af tidsrækkeanalyse modelleres de fire tidsrækker. Baseret på temperaturstigninger siden år 1850 forudsiges det, at en temperaturstigning på 1.5°C og 2°C sammenlignet med før den industrielle revolution vil blive opnået i henholdsvis år 2058 og 2083. Både de globale CO_2 udledninger og den atmosfæriske CO_2 koncentration vil fortsætte med at stige. Det første principale komponent af BNP per indbygger for alle fire grupper vil fortsat stige, og forskellen mellem grupperne vil blive større, hvilket vil øge den globale ulighed.

For at undersøge en tidsrækkes påvirkning af en anden tidsrække anvendes regressionsmodeller med ARMA fejl. Desuden konstrueres forskellige stier af forudsigelser for globale CO_2 udledninger for at undersøge effekten af at ændre niveauet af CO_2 udledninger.

Ved at regressere atmosfærisk CO_2 koncentration på globale CO_2 udledninger forudsiges det, at den atmosfæriske CO_2 koncentration vil stige endnu hurtigere end tidligere. Det samme resultat fås ved at regressere temperaturafvigelser på atmosfærisk CO_2 koncentration: Temperaturen vil stige hurtigere end hidtil, og dermed er det sandsynligt, at temperaturen er steget med 1.5°C og 2°C før henholdvis år 2058 og 2083.

Ved at regressere det første principale komponent af BNP per indbygger for de fire grupper på temperaturafvigelser forudsiges det, at en reduktion af globale CO_2 udledninger vil reducere den økonomiske vækst i alle typer økonomier. Dette giver mening, da økonomisk vækst ofte hænger sammen med øget produktion, som ofte hænger sammen med større CO_2 udledninger. Dermed fanger modellen det positive forhold mellem CO_2 udledninger og økonomisk vækst men ikke det forventede negative forhold mellem temperaturstigninger og økonomisk vækst. Dette indikerer, at regressionsmodellen med ARMA fejl er for simpel til at fange det sande forhold mellem klimaforandringer og BNP per indbygger.

Modellerne forudsiger også at forskellen mellem BNP per indbygger i de fire grupper vil stige. Da modellen dog ikke tager højde for, at klimaforandringer påvirker lande forskelligt, og at nogle lande bliver ubeboelige i fremtiden grundet slem varme og tørke, bliver forskellen mellem forskellige økonomier muligvis endnu større, end modellen forudsiger.

Da det nuværende niveau af globale CO_2 udledninger påvirker de globale temperaturer og forårsager klimaforandringer, vil den tidligere og nuværende økonomiske vækst højst sandsynligt ikke være bæredygtig i fremtiden. Men da de nuværende økonomiske modeller ikke tager højde for udgifterne fra klimaforandringer eller indtægterne fra klimavenlige forretningsmuligheder, er der nogle som argumenterer for, at det er muligt at opnå den samme eller endda en større økonomisk vækst, der kommer fra klimavenlige tiltag sammenlignet med den økonomiske vækst, der kommer fra at fortsætte som hidtil.

Preface

This paper has been compiled by Ann-Katrine Kjærsgaard Nielsen in the period from 1st February, 2019 to 7th June, 2019 on the basis of the fourth semester of the Master's programme in Mathematics-Economics at Aalborg University. The target group of this paper is mainly students at university level, lecturers, advisers, or others who may be interested. The Harvard method is used for references and abbreviations are introduced in parentheses.

I thank my supervisor Eduardo Vera Valdés for his devoted help.

Chapter 1

Introduction

This chapter is based on [NASA, 2019a], [NASA, 2019b], [NASA, 2019d], [United Nations, 2015], and [United Nations, 2019].

Small changes in the amount of energy received from the sun due to small variations in the Earth's orbit has caused the Earth's climate to change throughout history. However, most climate scientists agree that there is more than 95 percent probability that human activity, such as human-produced greenhouse gasses, has affected the climate and is causing the Earth to warm roughly ten times faster than the average rate of ice-age-recovery warming, and at a rate that is unprecedented over decades to millennia. This is due to certain gasses in the atmosphere, called greenhouse gasses, blocking heat from escaping. Such greenhouse gases include carbon dioxide (CO_2), methane, and nitrous oxide (N_2O). The evidence for greenhouse gases affecting the climate has been found in several places, such as tree rings, coral reefs, layers of sedimentary rocks, and ice cores drawn from Greenland, Antarctica, and tropical mountain glaciers.

Increased atmospheric greenhouse gas concentration has caused the planet's average surface temperature to rise. This rise in temperature has had wide spread consequences in the form of shrinking ice sheets, decreased snow cover, sea level rise, declining Arctic sea ice, and more extreme weather events, such as record high temperature events, severe droughts, and intense rainfall events.

It is estimated that further climate changes will affect regions differently, with temperature increases benefiting some regions and harming others: The effect will be determined by, among other things, regions' ability to mitigate or adopt to change. However, it is estimated that net annual costs will increase over time as global temperatures increase.

The Paris Agreement

Therefore, at COP 21 in Paris, on 12 December 2015, the Parties to the United Nations Framework Convention on Climate Change (UNFCCC) agreed on the global response to climate change in the form of the Paris Agreement. With the Paris Agreement the Parties seek to "accelerate and intensify the actions and investments needed for a sustainable low carbon future." [United Nations, 2019]. The Paris Agreement became effective on 4 November 2016, 30 days after the so-called "double threshold", which is ratification by 55 countries that account for at least 55% of global greenhouse gas emissions, had been met. Since then, more countries have ratified and continue to ratify the Agreement, reaching a total of 185 Parties as of today.

The Paris Agreement has three aims in order to strengthen the global response to climate change:

- To hold the temperature increase "well below" 2°C above pre-industrial levels, that is, the temperature before the industrial revolution, and to pursue efforts to limit the temperature increase to 1.5°C above pre-industrial levels,
- To increase the ability of countries to deal with the impacts of climate change,
- To make finance flows consistent with a low greenhouse gas emission.

The Paris Agreement recognises that limiting the temperature increase to 1.5°C or well below 2°C above pre-industrial levels will reduce the risks and impacts of climate change significantly. However, as it is estimated that temperature increases will benefit some regions and harm others, it is relevant to examine how countries with different economies will respond to climate changes, even when the temperature increase is limited.

Problem Statement

This paper examines the effect of global CO_2 emissions and the consequential global warming on the development in Gross Domestic Product (GDP) per capita. First it examines global CO_2 emissions and atmospheric CO_2 concentration after the industrial revolution. Then, using the HadCRUT4 data set of marine and land temperatures, it examines global temperature changes and makes predictions of future global temperature changes, more specifically predictions of when global temperatures have increased 1.5°C and 2°C compared to pre-industrial levels. The paper then examines GDP per capita for a variety of countries and estimates their future growth in GDP per capita at the predicted times of global temperature increases of 1.5°C and 2°C. In order to reduce the dimensionality of the data, Principal Component Analysis (PCA) is used, and the first principal component of the GDP per capita for similar countries is found.

Finally, the paper examines the relationship between global CO_2 emissions, atmospheric CO_2 concentration, global temperature changes, and development in GDP per capita, and compares the results of this analysis with the results from the individual analyses above.

Chapter 2

Exploratory Data Analysis

This chapter is based on [NASA, 2019a], [Ritchie and Roser, 2019], [Wikipedia, 2019f], and [Morice et al., 2012].

2.1 CO_2

CO_2 is a minor but very important component of the atmosphere, released both naturally, for example through volcano eruptions, and through human activities, such as deforestation and burning fossil fuels. CO_2 emissions are often a part of the discussion about global warming, since evidence suggests that increased atmospheric greenhouse gas concentration has caused the planets average surface temperature to rise.[1]

Figure 2.1 shows the global CO_2 emissions from 1751 to 2016 and global average atmospheric concentration of CO_2 from 1600 to 2016.[2] The figure illustrates how the amount of global CO_2 emissions has increased explosively since approximately 1960. This explosive growth comes mainly from the development of developing countries such as China and India, but also from developed countries such as the United States of America.[3]

The figure also illustrates how the level of atmospheric CO_2 concentration has been stable at approximately 275 Parts Per Million (PPM) until the late 18th century, which corresponds to the time before the industrial revolution. The atmospheric CO_2 concentration begins to increase during and after the industrial revolution,

[1] [NASA, 2019a]
[2] Data from [Ritchie and Roser, 2019]
[3] [Ritchie and Roser, 2019]

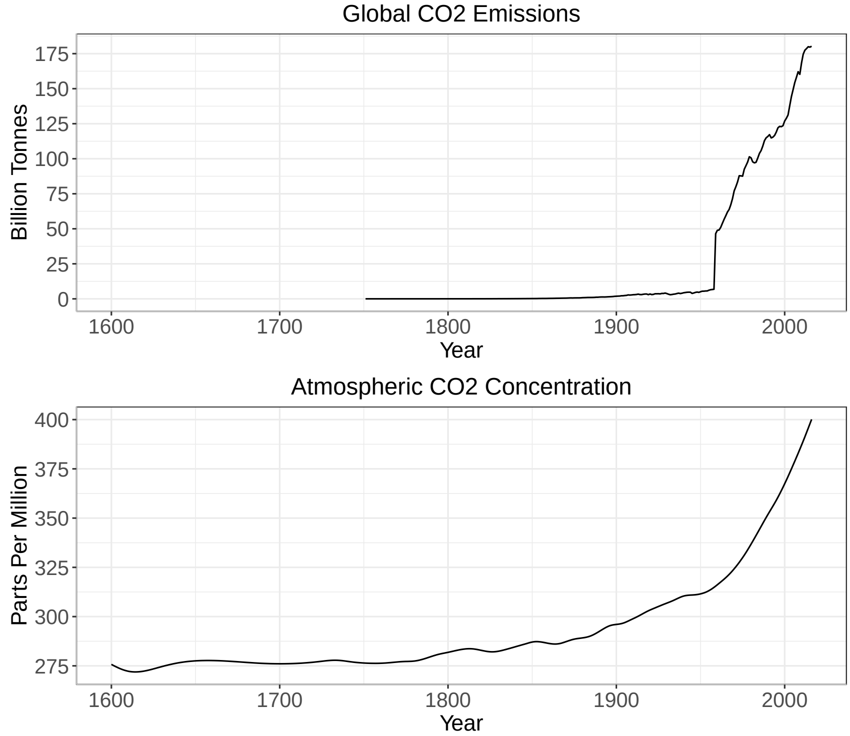


Figure 2.1: Global CO_2 emissions measured in billion tonnes from 1751 to 2016 and global average long-term atmospheric concentration of CO_2 measured in Parts Per Million (PPM) from 1600 to 2016.

which is the period from approximately 1760 to 1840 where the industry transitioned to new manufacturing processes.[4]

The atmospheric CO_2 concentration has changed naturally through time: During ice ages atmospheric CO_2 concentration has been approximately 200 PPM and between ice ages it has been approximately 280 PPM. The atmospheric CO_2 concentration is now higher than it has been in the past 400,000 years, and it has passed 400 PPM for the first time ever recorded.[5]

Even though global CO_2 emissions have been stabilised during the last few years, the atmospheric CO_2 concentration continues to increase, and it would take a substantial decrease in global CO_2 emissions for the atmospheric CO_2 concentration to stabilise.[6] This is due to CO_2 accumulating in the atmosphere based on residence time, which is the time required for emitted CO_2 to be removed from the atmosphere through natural processes in Earth's carbon cycle.

[4] [Wikipedia, 2019f]
[5] [NASA, 2019c]
[6] [Ritchie and Roser, 2019]

From Figure 2.1 it seems that the development in global CO_2 emissions first began around 1960 and not during the industrial revolution. However, this is due to the explosive increase in the amount of global CO_2 emissions in the last 60 years. Figure 2.2 shows the global CO_2 emissions from 1750 to 1900, and, thus, makes it possible to take a closer look at the development in the CO_2 emissions during and after the industrial revolution. The figure shows how the amount of global CO_2

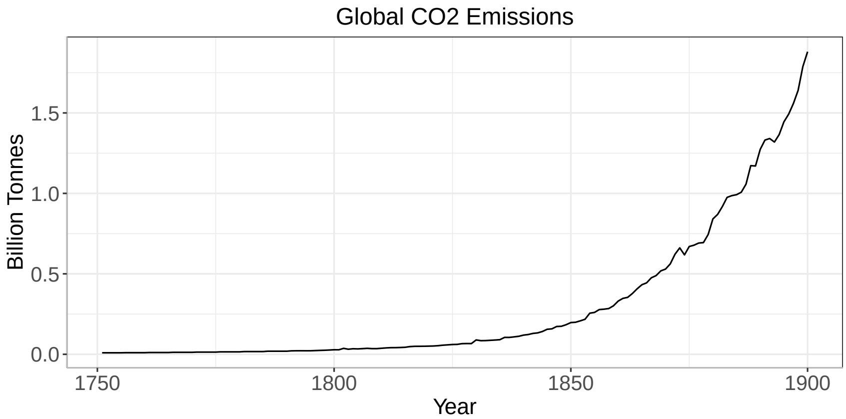


Figure 2.2: Global CO_2 emissions measured in billion tonnes from 1750 to 1900.

emissions before the industrial revolution was stable and close to zero; however, after the industrial revolution, the amount of global CO_2 emissions begins to increase rapidly, leading to an increase in the global atmospheric CO_2 concentration.

2.2 Temperature

As mentioned in Chapter 1, increased atmospheric greenhouse gas concentration has caused the planet's average surface temperature to rise. In order to examine the rise in global average temperature, the HadCRUT4 data set[7] is used, which contains monthly surface temperature anomalies globally from 1850 to 2018 relative to a 1961-1990 reference period in 5° by 5° grids. The surface temperature anomalies are a combination of land temperature anomalies compiled by the Climatic Research Unit of the University of East Anglia and marina temperature anomalies compiled by the Met Office Hadley Centre.

Two months of the HadCRUT4 data set is seen in Figure 2.3, which shows the temperature anomalies for January 1850 and January 2018, that is, so-called slices of the data. The figure shows the 5° by 5° grids in which temperature anomalies have been measured. The white squares illustrate the grids in which the temperature

[7] [Morice et al., 2012]

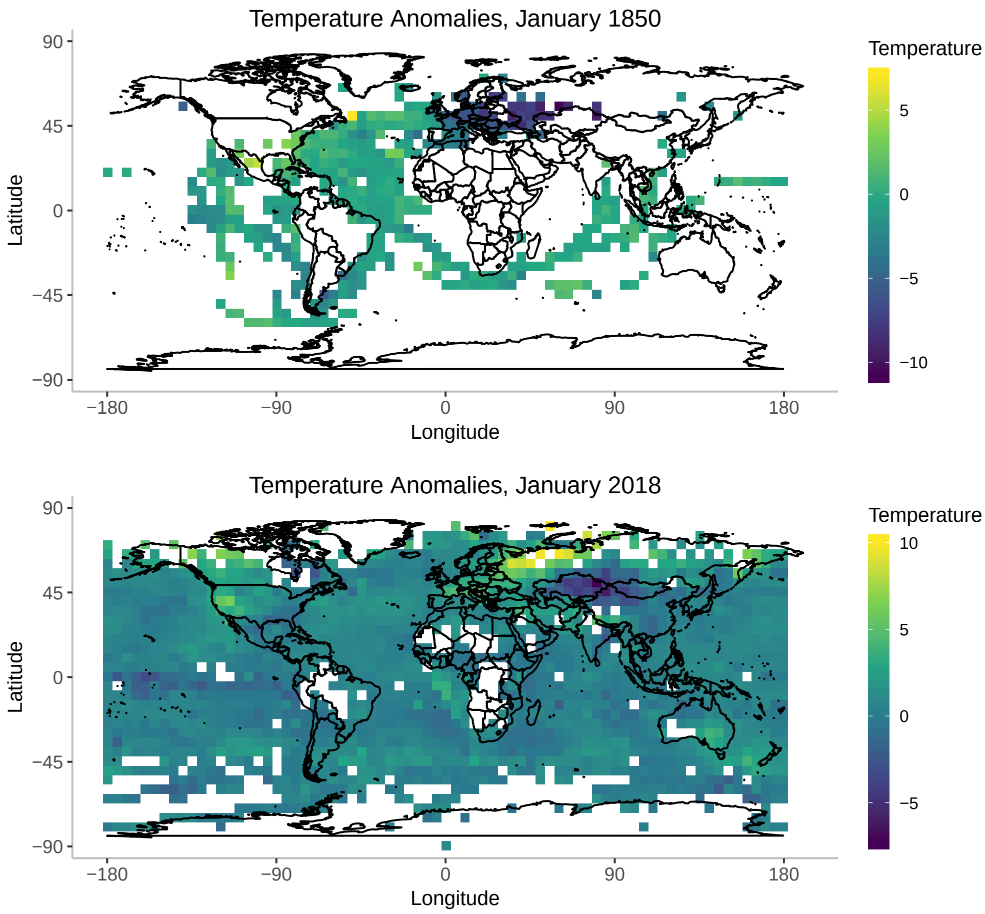


Figure 2.3: Heatmaps of temperature anomalies in January 1850 and January 2018 relative to a 1961-1990 reference period. The white squares illustrate the grids in which the temperature measurement was not available at the chosen time.

measurement was not available in the period illustrated in the figure. The grids that are not white in January 1850, that is, the grids in which the temperature has been measured, are mostly in Europe and in the Atlantic ocean. In January 2018 there are substantially less white grids, which are mostly in the most northern part of the northern hemisphere, the most southern part of the southern hemisphere, and in Africa.

In January 1850 there seems to be temperature anomalies from approximately -11°C to 7°C and in January 2018 there seems to be temperature anomalies from approximately -7°C to 10°C. The largest temperature anomalies in January 1850 seem to be in Europe and in January 2018 they seem to be in the north-western part of Russia and in the area around the border between Russia, Kazakhstan, and China.

From the HadCRUT4 data set there can be made illustrations similar to those in Figure 2.3 for each month from January 1850 to December 2018; however, it is difficult to gain an overview of the development in the temperature from 1850 to 2018 just from examining each slice of the data. Therefore, a time series has been created from the data set by taking the medians of regional time series computed for 100 ensemble member realisations.[8] There are several uncertainties in the data that needs to be taken into account: Measurement, sampling, and coverage uncertainties in the HadCRUT4 data, which are described in the HadCRUT4 paper[9], and bias from the 100 ensemble member realisations, which is accounted for by integrating across the distributions they describe.[10]

Figure 2.4 shows the decadally smoothed time series of global temperature anomalies relative to a 1961-1990 reference period. The figure visualises how the temperature anomalies are relative to a 1961-1990 reference period, since the temperature anomalies in this period are centred around and close to zero. It is also seen that there tends to be mostly negative temperature anomalies before this reference period and only positive temperature anomalies after the reference period, indicating that the global average surface temperature has increased from 1850 to 2018. In fact, a temperature anomaly of approximately -0.25°C in 1850 and a temperature anomaly of almost 0.75°C in 2018 means that there has been an increase in the temperature of approximately 1°C from the industrial revolution to today, where most of the increase has happened from 1975 to 2018.

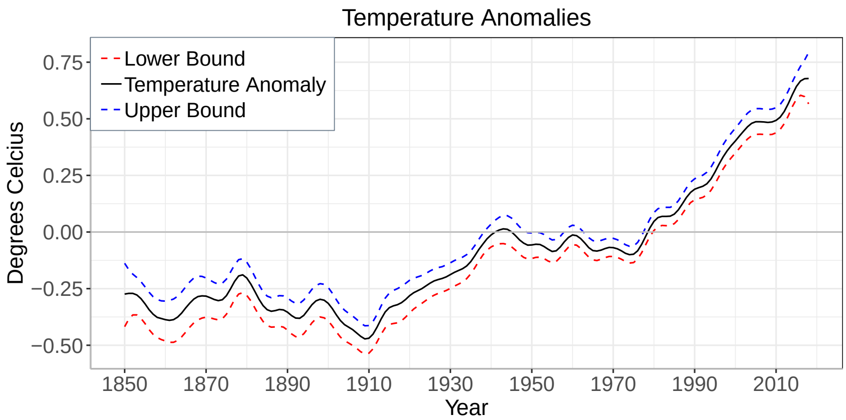


Figure 2.4: Temperature anomalies (black solid line) from 1850 to 2018 relative to a 1961-1990 reference period. The red and blue dashed lines represent the lower and upper bounds of the 95% confidence interval of the combined effects of measurement, sampling, coverage, and bias uncertainties.

[8] [Morice et al., 2012]
[9] [Morice et al., 2012]
[10] [Morice et al., 2012]

There seems to be more uncertainty in the beginning of the time series compared to the end. This could be due to the large amount of missing data in the beginning of the data set, meaning less coverage, compared to the most resent measurements as illustrated in Figure 2.3.

2.3 GDP per Capita

Instead of examining each country's GDP per capita individually, the countries are divided into four groups: Low-income, lower-middle-income, upper-middle-income, and high-income economies. These groups are chosen based on the World Bank Country Groups[11]. See Appendix A for a list of the countries and the group they belong to.[12] The division of the countries into the four groups is illustrated in Figure 2.5. The figure illustrates how most of the countries in the low-income

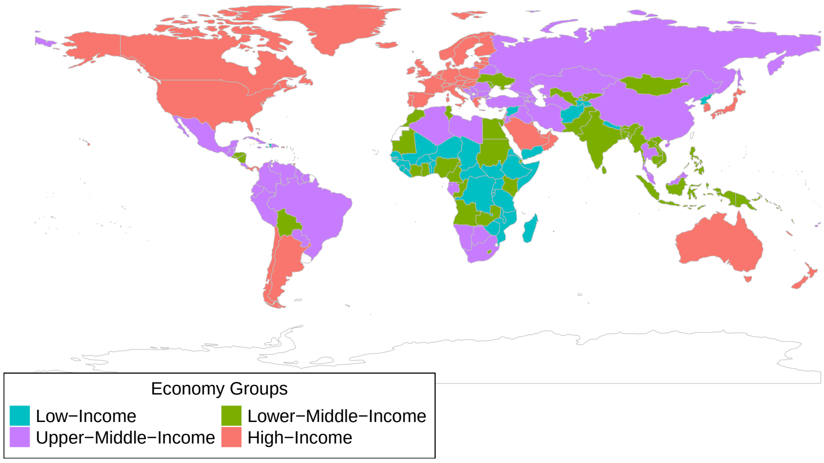


Figure 2.5: The division of countries into the low-income, lower-middle-income, upper-middle-income, and high-income economy groups. A list of the countries in each group can be found in Appendix A.

economy group (blue countries) are located in central Africa, many of the countries in the lower-middle-income economy group (green countries) are located close to the equator, many of the countries in the upper-middle-income economy group (purple countries) are located in South America and Asia, and most of the countries in the high-income economy group (red countries) are located in North America, Europe, and Oceania.

[11] [The World Bank, 2019b]

[12] Data from [The World Bank, 2019a].

Approximately 50% of the countries' GDP per capita data contain missing data. Countries for which the GDP per capita data contains more than 20% missing values are excluded from the analysis; in Appendix A these countries' NA value is coloured red. Approximately 43% of the countries are removed due to missing values. The GDP per capita for the remaining countries containing missing values are imputed.[13] The results can be seen in Figure 2.6, which shows the GDP per capita for the remaining imputed data divided into the four groups.

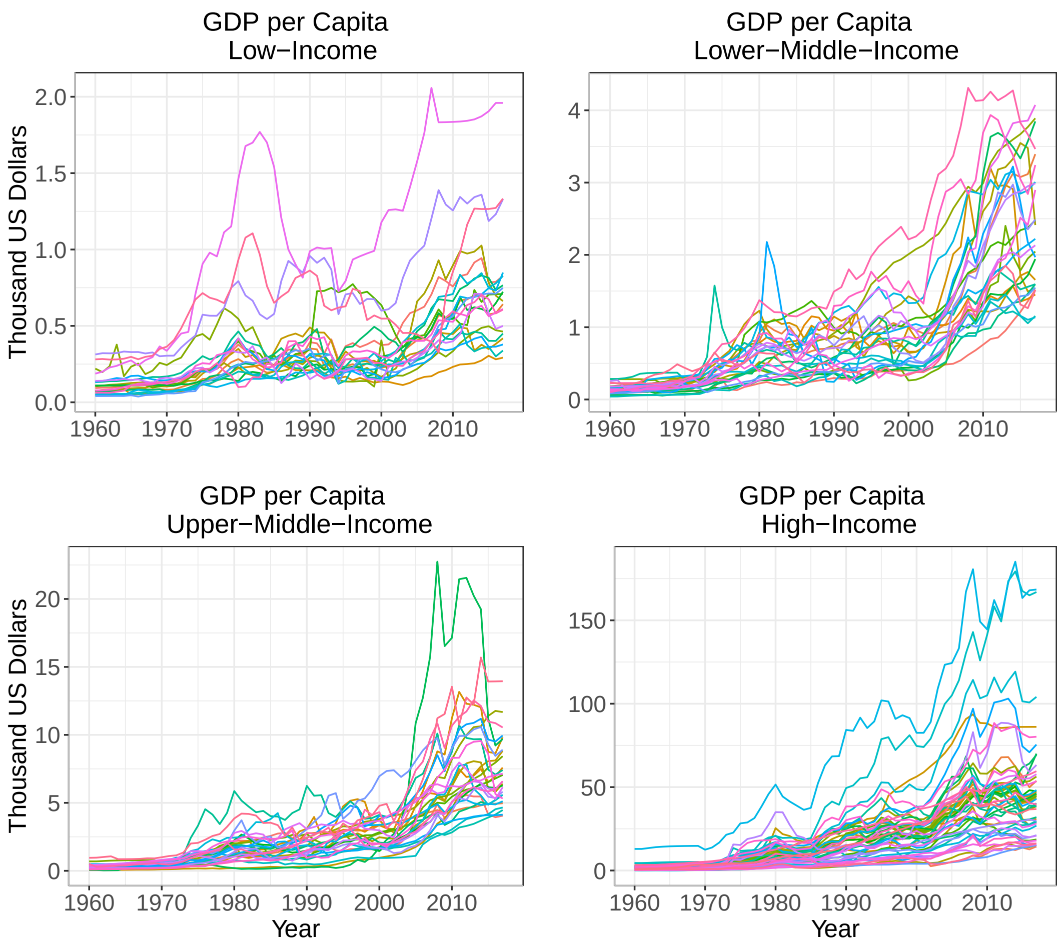


Figure 2.6: GDP per capita for low-income, lower-middle-income, upper-middle-income, and high-income economies measured in thousand US dollars from 1960 to 2017. The division of countries into these categories can be found in Appendix A.

[13] The "imputeTS" package in R is used for imputation. The missing values for all countries but one are imputed by a weighted moving average with exponential weighting and window size equal to 10. The missing values for the United States Virgin Islands are imputed by a weighted moving average with linear weighting and window size equal to 20. [Moritz, 2018, pp. 4-5, 8-9]

The figure shows how the division of the countries into the four groups is based on their level of GDP per capita, since the GDP per capita in 2018 for low-income, lower-middle-income, upper-middle-income, and high-income economies is between approximately 250 and 1,250 US dollars, 1,000 and 4,000 US dollars, 4,000 and 14,000 US dollars, and 15,000 and 170,000 US dollars, respectively. In all of the four groups, drops in the GDP per capita can be seen around 1980, which is at the time of the second major oil crisis, and around 2008, which is at the time of the latest financial crisis.[14] Apart from these two crises, many of the countries' GDP per capita in all four groups seem to follow approximately the same growth trends.

The country with the highest GDP per capita in the low-income economy group, represented by the pink line, is Syria. The Syrian economy grew in the 1970's after General Hafiz al-Assad took power.[15] The up-turn in the 1990's came after the institution of a series of economic reforms. The data for Syria contained missing values after 2007, partly because of Syria's civil war, and has therefore been imputed. The country with the lowest GDP per capita, represented by the yellow line, is Burundi, which is a landlocked, resource-poor country with an underdeveloped manufacturing sector.[16]

The country represented by the highest green line in the upper-middle-income economy group is Equatorial Guinea. The sudden growth in the GDP per capita came mainly from oil, since Equatorial Guinea has become one of sub-Saharan Africa's largest oil producers.[17] It has the highest GDP per capita in Africa, however, the wealth is very uneven, since only few people have benefited from the oil production.

The two countries with the highest GDP per capita in the high-income economy group are not surprisingly Monaco and Lichtenstein.

In the lower-middle-income economy group, there are no countries that are extremely noticeable at the top or bottom as in the other economy groups. Many of the countries seem to have a sudden increase from around 2005 to approximately 2015.

That many of the countries in all four groups seem to follow approximately the same growth trends is illustrated in Figure 2.7, which shows the GDP per capita growth in the four groups. Some countries experience extreme growth in some years, making it difficult to see the general trends in the data and to compare the behaviour in the four groups. Therefore, Figure 2.7 is recreated where the countries

[14] [Wikipedia, 2019a], [Wikipedia, 2019e]
[15] [Wikipedia, 2019c]
[16] [Wikipedia, 2019b]
[17] [Wikipedia, 2019d]

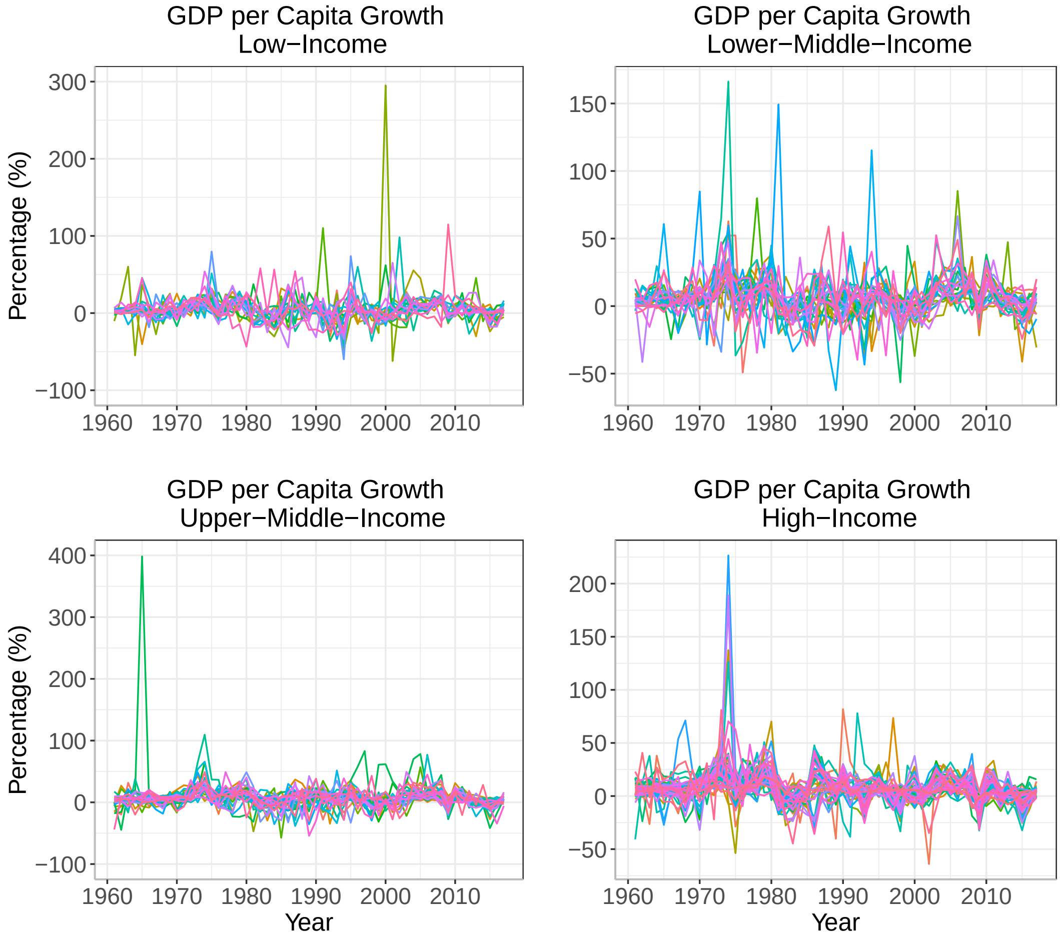


Figure 2.7: Growth in GDP per capita for low-income, lower-middle-income, upper-middle-income, and high-income economies measured in percentage from 1961 to 2017.

in each group having extreme fluctuations are excluded.

The green line in the low-income economy group showing the extreme growth in 2000 represents the GDP per capita growth for the Democratic Republic of the Congo. The turquoise line in the lower-middle-income economy group showing the extreme growth in 1974 represents the GDP per capita growth for Kiribati and the blue line showing the extreme growth in 1981 represents the GDP per capita growth for Nigeria. The green line in the upper-middle-income economy group showing the extreme growth in 1965 represents the GDP per capita growth for Equatorial Guinea. Five countries in the high-income economy group experience extreme growth in 1974; Kuwait, Brunei, Qatar, Saudi Arabia, and Oman. These countries are some of the largest oil producers in the world, and experienced large growths in their economies in 1974 due to the 1973 oil crisis, where the Organisation of Arab Petroleum Exporting Countries proclaimed an oil embargo, causing the oil prices to increase.[18]

Figure 2.8 shows the result of excluding these countries from Figure 2.7. Note that

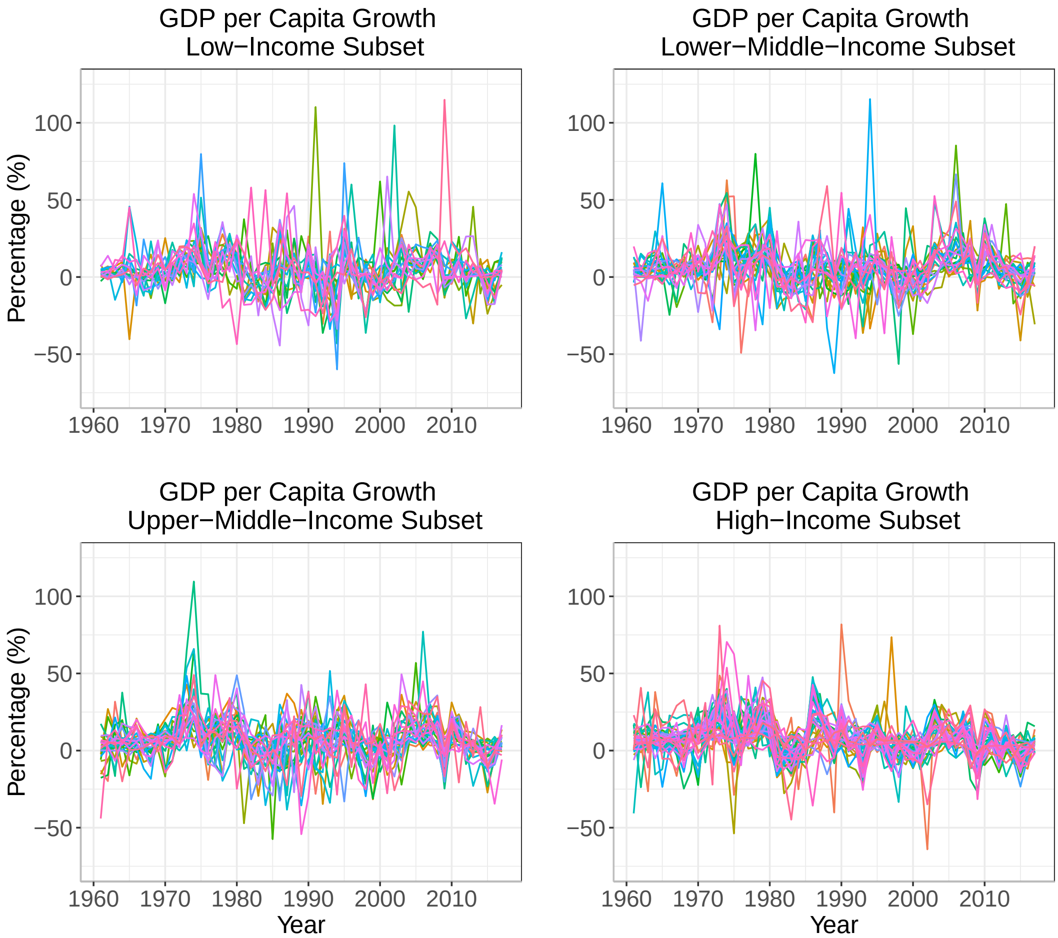


Figure 2.8: Growth in GDP per capita for low-income, lower-middle-income, upper-middle-income, and high-income economies measured in percentage from 1961 to 2017 excluding the Democratic Republic of the Congo, Nigeria, Kiribati, Equatorial Guinea, Kuwait, Brunei, Qatar, Saudi Arabia, and Oman.

all four plots in Figure 2.8 now have the same y-axis. The figure illustrates that most of the countries seem to have the same growth trends despite their economy group: However, the growths of the countries in the high-income economy group seem to be more alike than the growths of the countries in the other groups, that is, there seems to be less variance in the high-income economy group.

[18] [Wikipedia, 2019a]

2.3.1 Principal Component Analysis

Modelling and making predictions for the GDP per capita for all countries in the four groups and examining the relationship between the countries' GDP per capita and temperature changes is a lengthy process. Therefore, PCA is used on the countries' GDP per capita in each group. PCA is an unsupervised learning method and a way of presenting high-dimensional data in a lower-dimensional space while preserving as much variability in the data as possible.[19] Using PCA will not only reduce the dimensionality of the analysis, but also make it easier to compare the four groups and illustrate their differences.

In order to determine how many principal components to use of the data in each

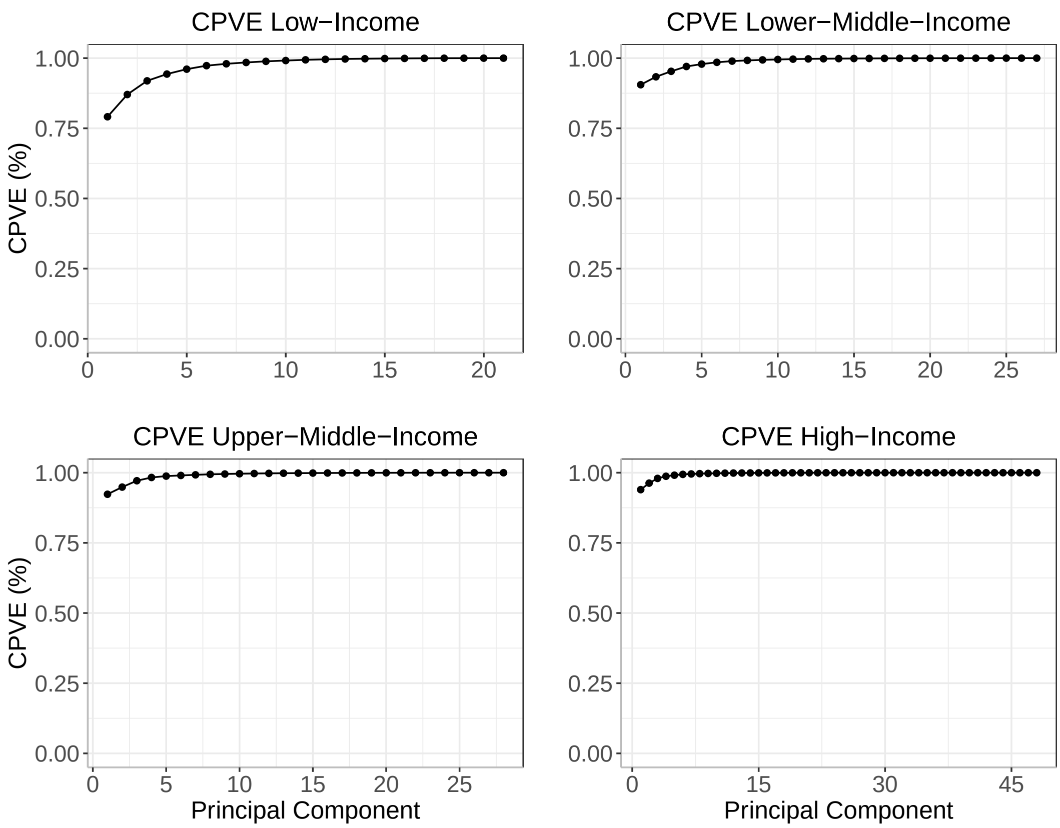


Figure 2.9: Cumulative Percentage of variance explained for low-income, lower-middle-income, upper-middle-income, and high-income economies. The x-axis indicates the number of principal components for each group, which corresponds to the number of countries in each group.

[19]See Appendix B for a short introduction to PCA.

group the Cumulative Percentage of Variance Explained (CPVE) is examined.[20] The CPVE for the four groups is shown in Figure 2.9. The figure shows that the first principal component of the data in each group explains more than 75% of the variance in the data in that group. More specifically the variance explained by the first principal component of the low-income, lower-middle-income, upper-middle-income, and high-income economies is 79.12%, 90.53%, 92.33%, and 93.97%, respectively. The figure also shows that using additional principal components only increases the CPVE at a slow rate at the expense of increasing the dimension. Therefore, only the first principal component of each group is used in the analysis in this paper.

The loading vector for the first principal component of each group can be seen in Appendix C along with the means and standard deviations used for standardising the data.

In Table C.1 it is seen that the weights in the loading vector for the low-income group vary between 0.100 and 0.242, but that most of the weights are very similar, meaning that the countries are weighted almost equally. This is also the case in the loading vectors for the lower-middle-income, upper-middle-income, and high-income groups seen in Tables C.2, C.3, and C.4, respectively, where the weights vary between 0.180 and 0.200, 0.176 and 0.195, and 0.130 and 0.148. Thus, in practice the first principal components almost reduce to simple averages.

The tables in Appendix C also illustrate the difference between the GDP per capita in the four groups. The average of the means in the low-income, lower-middle-

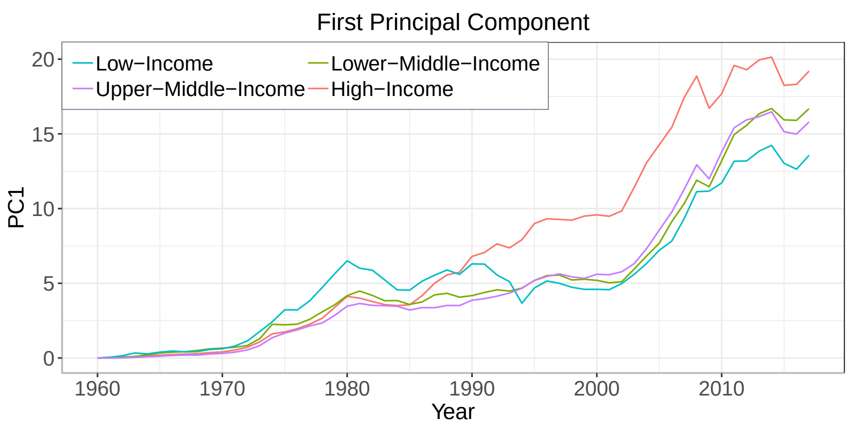


Figure 2.10: The first principal component of GDP per capita for low-income, lower-middle-income, upper-middle-income, and high-income economies from 1960 to 2017 shifted to have initial value equal to zero.

[20]See Appendix B.

income, upper-middle-income, and high-income groups are approximately 366, 871, $2,677$, and $19,344$ thousand US dollars, respectively.

Figure 2.10 shows the first principal component for the four groups, shifted such that each principal component has initial value equal to zero. The figure shows that the developments in the GDP per capita for the lower-middle-income economies (green solid line) and upper-middle-income economies (purple solid line) are almost identical, and that, of the four groups, the GDPs per capita for the high-income economies (red solid line) have increased the most and the GDPs per capita for the low-income economies (blue solid line) have increased the least.

During the 1970s the GDP per capita for low-income economies increased at a higher rate than the GDP per capita for the other groups; however, after approximately 1980 the increase stopped for all groups and the GDP per capita for the low-income economies decreased until approximately 2000.

The GDP per capita for high-income economies increased at a higher rate than the GDP per capita for the other groups from approximately 1985 to approximately 1995.

From approximately year 2000 the GDP per capita began to increase for the countries in all four groups at approximately the same rate. However, the larger increase in GDP per capita for the high-income economies from approximately 1985 to 1995 means that the countries in the other three groups are not able to catch up to the high-income economies with regards to development since 1960.

Chapter 3

Linear Time Series Models

This chapter is based on [Fan and Yao, 2017, pp. 33-36].

Often one of the main goals of modelling time series is to use the data of the past to forecast the future. Being able to do this requires that the underlying process of the data has time-invariance properties. These properties characterise stationarity.

Definition 3.1 (Weak Stationarity) *A time series $\{x_t\}$ is said to be weakly stationary if $\mathbb{E}\left[x_t^2\right] < \infty$ and, for any integer k, neither $\mathbb{E}[x_t]$ nor $Cov(x_t, x_{t+k})$ depend on t.*

The reason for the stationarity in the definition above being called weak, is due to it being a weak notion of stationarity: For example, $\{x_t\}$ being weakly stationary does not imply that $\{x_t^2\}$ is weakly stationary. A stronger definition is therefore stated below.

Definition 3.2 (Strong Stationarity) *A time series $\{x_t\}$ is said to be strongly stationary if the k-dimensional distribution of $(x_1, \ldots, x_k)$ is the same as that of $(x_{t+1}, \ldots, x_{t+k})$ for any $k \geq 1$ and t.*

Strong stationarity implies weak stationarity in the cases where $\mathbb{E}\left[x_t^2\right] < \infty$. Furthermore, strong stationarity of the time series $\{x_t\}$ implies strong stationarity of $\{g(x_t)\}$ for any function g.[1]

Strong stationarity is, however, often too strong of an assumption for real-world data, and therefore, weak stationarity is often used for determining if a model is stationary.

[1][Fan and Yao, 2017, pp. 35-36]

3.1 The ARMA Model

This section is based on [Fan and Yao, 2017, pp. 36, 41, 46, 62, 64] and [Shumway and Stoffer, 2011, pp. 88, 92-95].

One of the most frequently used time series models is the stationary AutoRegressive Moving Average (ARMA) model.

Definition 3.3 (ARMA(p,q)) *An ARMA model with the order (p,q), written ARMA(p,q), is the stationary solution to*

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \tag{3.1}$$

where p and q are non-negative integers, $\varepsilon_t \sim WN\left(0,\sigma^2\right)$, $\phi_1, \cdots, \phi_p$, $\theta_1, \cdots, \theta_q$ are parameters, and $\phi_p, \theta_q \neq 0$.[2]

The ARMA(p,q) model is created by combining an AutoRegressive (AR) model of order (p), written AR(p), given by

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t,$$

and a Moving Average (MA) model of order (q), written MA(q), given by

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}.$$

Definition 3.4 (AR and MA Polynomials) *The AR(p) and MA(q) polynomials are defined as*

$$\begin{aligned} \phi(z) &= 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p, \quad && \phi_p \neq 0, \\ \theta(z) &= 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q, \quad && \theta_q \neq 0, \end{aligned}$$

respectively, where z is a complex number.

Using Definition 3.4, the ARMA(p,q) model in Equation (3.1) can be written as

$$\phi(B) x_t = \theta(B) \varepsilon_t, \tag{3.2}$$

where B is the backshift operator defined as

$$B^k x_t = x_{t-k} \quad \text{for } k \in \mathbb{Z}.$$

It is assumed that the two equations $\phi(z) = 0$ and $\theta(z) = 0$ do not have common roots, since the common factors from Equation (3.2) then can be cancelled out, which could mean that (p,q) is not the genuine order of the model.

[2] $\varepsilon_t \sim WN\left(0,\sigma^2\right)$ means that $\{\varepsilon_t\}$ is a white noise process, which is characterised by $\rho(k) = \text{Corr}(\varepsilon_t, \varepsilon_{t+k}) = 0$ for any $k \neq 0$. That is, a white noise process is a sequence of uncorrelated random variables with the same mean and variance.

Properties of ARMA Models

Definition 3.5 (Causality) *An ARMA(p,q) model is said to be causal, if the time series $\{x_t\}$ can be written as a one-sided linear process:*

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(B)\varepsilon_t, \tag{3.3}$$

where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, $\sum_{j=0}^{\infty} |\psi_j| < \infty$, and where $\psi_0 = 1$.

Thus, causality implies that the ARMA(p,q) model does not depend on the future, and that it can be written as a model that only depends on past innovations to infinity, that is, as an MA(∞) model.

From Definition 3.5 it follows that an ARMA(p,q) model is causal if and only if the roots of $\phi(z)$ lie outside the unit circle; that is, $\phi(z) = 0$ if and only if $|z| > 1$.[3] The coefficients of the linear process given in Equation (3.3) can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

Definition 3.6 (Invertibility) *An ARMA(p,q) model is said to be invertible, if the time series $\{x_t\}$ can be written as*

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = \pi(B)x_t, \tag{3.4}$$

where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$, $\sum_{j=0}^{\infty} |\pi_j| < \infty$, and where $\pi_0 = 1$.

Thus, invertibility implies that an ARMA(p,q) model can be written as a model that only depends on past values to infinity, that is, as an AR(∞) model. Invertibility is a way of choosing an MA(q) model: Since only the time series can be observed and not the innovations, the model is chosen such that it can be written as an AR(∞) model, mimicking the causality property of AR models.[4]

From Definition 3.6 it follows that an ARMA(p,q) model is invertible if and only if the roots of $\theta(z)$ lie outside the unit circle; that is, $\theta(z) = 0$ if and only if $|z| > 1$.[5]

[3] [Shumway and Stoffer, 2011, p. 95]
[4] [Shumway and Stoffer, 2011, p. 92]
[5] [Shumway and Stoffer, 2011, p. 95]

The coefficients of the linear process given in Equation (3.4) can be determined by solving

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1.$$

Parameter Estimation

There are several methods for estimating the parameters of an ARMA(p, q) model. One method is Gaussian maximum likelihood estimation, which I will briefly introduce.

When estimating the model parameters it is assumed that the model order (p, q) is known and that the process has mean zero. The idea in Gaussian maximum likelihood estimation is to assume that $\{x_t\}$ is Gaussian meaning that $\phi(B)x_t = \theta(B)\varepsilon_t$, where ε_t is i.i.d. Gaussian, that is, $\varepsilon_t \sim N(0, \sigma^2)$ and independent, and then choose ϕ_i, θ_j, and σ^2 that maximise the likelihood function

$$\mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = f_{\boldsymbol{\phi},\boldsymbol{\theta},\sigma^2}(x_1, \ldots, x_n),$$

where $f_{\boldsymbol{\phi},\boldsymbol{\theta},\sigma^2}$ is the simultaneous Gaussian probability density function for the given ARMA model, and where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)$.

That is, assume that $x_1, \ldots, x_n$ are observations from a Gaussian ARMA(p, q) process with mean zero. The value of the likelihood function in the parameters $\boldsymbol{\phi} \in \mathbb{R}^p$, $\boldsymbol{\theta} \in \mathbb{R}^q$, and $\sigma^2 \in \mathbb{R}_+$ is defined as the density of $\mathbf{x} = (x_1, \ldots, x_n)^T$ under the Gaussian model with these parameters:

$$\mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = \frac{1}{(2\pi)^{n/2}|\Sigma_n|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T \Sigma_n^{-1} \mathbf{x}\right),$$

where $|\Sigma_n|$ denotes the determinant of Σ_n, which is the variance-covariance matrix of $\mathbf{x}$ with the given parameter values. The log-likelihood function is the logarithm of the likelihood function, yielding

$$\begin{aligned} l(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) &= \log(\mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2)) \\ &= -\frac{1}{2}\log\left(|\Sigma_n|\right) - \frac{1}{2}\boldsymbol{x}^T \Sigma_n^{-1} \boldsymbol{x}, \end{aligned}$$

where the constant term is omitted. The maximum likelihood estimator of $(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2)$ maximises this quantity.

3.2 The ARIMA Model

This section is based on [Shumway and Stoffer, 2011, p. 141, 277, 279-280] and [Pfaff, 2008, pp. 53-54, 57].

Some time series do not meet the conditions required to be classified as ARMA models, but instead need to be differenced in order to become stationary processes. Therefore, the class of ARMA models is broadened to include differencing, leading to the Integrated ARMA (ARIMA) model.

Definition 3.7 (ARIMA(p, d, q)) *A process $\{x_t\}$ is said to be an ARIMA model with order (p, d, q), written ARIMA(p, d, q), if*

$$\nabla^d x_t = (1 - B)^d x_t$$

is ARMA(p, q), where p, d, and q are non-negative integers. In general, the ARIMA(p, d, q) model is written as

$$\phi(B)(1 - B)^d x_t = \theta(B)\varepsilon_t.$$

Note that an ARMA(p, q) process is also an ARIMA$(p, 0, q)$ process. If $d \geq 1$ then the time series $\{x_t\}$ is not stationary, and to obtain a stationary process the time series needs to be differenced d times. It is said that the time series $\{x_t\}$ is integrated of order d, and is written $\{x_t\} \sim \mathrm{I}(d)$.

When a time series does not meet the conditions required to be classified as an ARMA model, it can be difficult to determine whether the process is non-stationary due to the corresponding AR polynomial having a root on the unit circle, that is, a unit root, and therefore needs to be differenced, or due to the process having a time trend and therefore needs to be detrended.[6] In the first case, the time series is I(1) and in the second case the time series is I(0).

Unit root testing can be used in order to determine if the non-stationarity of the time series is due to a time trend or a root on the unit circle.

3.2.1 Unit Root Testing

As the name suggests, the aim of unit root testing is to test if the AR polynomial has a root on the unit circle. However, this is not as simple as it appears.

Let $\{x_t\}$ be an AR(1) process, that is, $x_t = \phi_1 x_{t-1} + \varepsilon_t$. It would be simple to test for the following hypothesis:

$$\mathcal{H}_0 : \phi_1 = 0, \quad \mathcal{H}_a : \phi_1 \neq 0,$$

[6] [Pfaff, 2008, p. 53]

since then a simple t-test could be used:

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1}{\text{sd}\left(\hat{\phi}_1\right)} \sim t(n-1).$$

However, it should be tested if the time series $\{x_t\}$ has a unit root and, therefore, the following hypothesis needs to be tested:

$$\mathcal{H}_0 : \phi_1 = 1, \quad \mathcal{H}_a : -1 < \phi_1 < 1.$$

That is, the null hypothesis is that the process has a unit root and the alternative hypothesis is that the process is stationary and causal. This hypothesis causes issues for the t-test, since $\phi_1 = 1$ means that the process is a random walk and therefore non-stationary, meaning that the standard deviation of ϕ_1 does not converge to a constant, but instead goes to infinity. This means that $t_{\hat{\phi}_1}$ no longer follows a t-distribution, but instead goes to a non-standard distribution. In order to limit this paper no further details about the theory behind this is presented.

The solution to this problem is the Dickey-Fuller (DF) test.

In the DF-τ test, the following model is considered:

$$x_t = \mu + \tau t + \phi_1 x_{t-1} + \varepsilon_t,$$

which, under the assumption $\mathcal{H}_0 : \phi_1 = 1$, defines a random walk model with drift and a linear time trend. The model is transformed to

$$\nabla x_t = \mu + \tau t + \gamma x_{t-1} + \varepsilon_t, \tag{3.5}$$

where the assumption now is $\mathcal{H}_0 : \gamma = 0$.

Other versions of the DF test are the DF-μ test, where $\tau = 0$ in Equation (3.5), and the DF-0 test, where $\mu = \tau = 0$ in Equation (3.5).

Since the DF tests are not regular statistical tests in the way that the test statistics go to non-standard distributions, regular critical values can not be used. Instead critical values from special simulated tables found for example in [Fuller, 1996] should be used.

The DF tests mentioned above are based on an AR(1) model, which can sometimes prove adequate for the time series at hand. However, at other times the regression in Equation (3.5) is too simple.

The solution to this problem is to use the Augmented Dickey-Fuller (ADF) test, which uses higher orders of AR processes:

Every AR(p) model

$$x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + \varepsilon_t$$

can be rewritten as

$$\nabla x_t = \gamma x_{t-1} + \sum_{j=1}^{p-1} \tilde{\gamma}_j \nabla x_{t-j} + \varepsilon_t, \tag{3.6}$$

where there is a bijection between the parameters $\phi_1, \ldots, \phi_p$ and $\gamma, \tilde{\gamma}_1, \ldots, \tilde{\gamma}_{p-1}$. Similarly to the DF test, if $\gamma = 0$ in Equation (3.6) then the AR polynomial has a unit root.

There are several ways of choosing the augmentation, p, in Equation (3.6); as a function of the sample size, using information criteria such as AIC or BIC, or such that there is no longer any autocorrelation in the residuals.

3.3 Long Memory

This section is based on [Ruppert, 2004, pp. 270-272], [Nielsen and Frederiksen, 2005, pp. 406, 409, 411-412], [Sibbertsen, 2004, p. 476], and [Shumway and Stoffer, 2011, pp. 181-182, 267-270].

The conventional ARMA(p,q) model is often referred to as a short-memory model because the coefficients in the MA(∞) representation

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

obtained by solving

$$\phi(z)\psi(z) = \theta(z)$$

are dominated by exponential decay. This implies that the ACF of the short memory process goes to zero, $\rho(k) \to 0$, exponentially fast as $k \to \infty$ and that $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

If the average of $x_1, x_2, \ldots, x_n$ is written as $\bar{x}_n = \dfrac{x_1 + \cdots + x_n}{n}$ then

$$\mathrm{Var}(\bar{x}_n) = \frac{1}{n} \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma(h).$$

Assuming $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ means that

$$\sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma(h) \to \sum_{h=-\infty}^{\infty} \gamma(h) \quad \text{for} \quad n \to \infty.$$

This means that

$$n\mathrm{Var}(\bar{x}_n) \to \sum_{h=-\infty}^{\infty} \gamma(h)$$

or

$$\mathrm{Var}(\bar{x}_n) \to \frac{1}{n}\left(\sum_{h=-\infty}^{\infty} \gamma(h)\right) + o\left(\frac{1}{n}\right).$$

Thus, for a short memory time series $\mathrm{Var}(\bar{x}_n)$ goes to zero with the usual speed $\frac{\sigma^2}{n}$ when the sample size increases, meaning that $\bar{x}_n$ is a good estimate of the mean of the time series. When this is not the case, the advice is often to difference the time series until it fluctuates around a well-defined mean value and its ACF decays fairly rapidly to zero. However, differencing can be too severe a modification in the sense that the non-stationary model might represent an overdifferencing of the original process.

These stationary time series that do not fulfil $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ and, therefore, can not be modelled by ARMA models, instead have $\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$. This means that the calculations of $\mathrm{Var}(\bar{x}_n)$ do not work and no matter how large the sample size is, a good estimate can not be made of the mean of the time series.

Furthermore, seen from a prediction point of view, short memory of a time series implies that after a certain number of lags, the past values can be disregarded from the prediction without serious consequences. However, when the autocorrelations do not go to zero fast enough no past values can be disregarded, since their significance will be too large, and the prediction will have to include all past information to infinity in order to be a good prediction.

3.3.1 The ARFIMA Model

Therefore, the concept of differencing in the ARIMA model is expanded to allowing d in Definition 3.7 to be any real number, leading to the Fractionally Integrated ARMA (ARFIMA) model.

Definition 3.8 (ARFIMA(p, d, q)) *A time series $\{x_t\}$ is said to be an ARFIMA model with order (p, d, q), written ARFIMA(p, d, q), if*

$$\nabla^d x_t = (1 - B)^d x_t \tag{3.7}$$

is ARMA(p, q), where p and q are non-negative integers and $d \in (-0.5, 0.5)$.

If $d \in (0.5, 1)$, the time series is differenced once, and then $d \in (-0.5, 0)$. When d is negative, it is said that the time series is antipersistent.

The term $(1-B)^d$ in Equation (3.7) is to be interpreted the following way: For d a positive integer the well-known binomial formula is given by

$$(x+y)^d = \sum_{j=0}^{d} \binom{d}{j} x^j y^{d-j}.$$

This formula can be written as

$$(x+y)^d = \sum_{j=0}^{\infty} \binom{d}{j} x^j y^{d-j} \tag{3.8}$$

if it is implicit that $\binom{d}{j} = 0$ for d integer and $j > d$.

When d is not an integer, there is a generalised binomial formula, which shows the series expansion for powers that are not necessarily positive integers:

$$(x+y)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j y^{\alpha-j}, \tag{3.9}$$

where it is assumed that $|x| < |y|$. If α is a positive integer then Equation (3.9) is interpreted as Equation (3.8). Letting $x = -B$, $y = 1$, and $\alpha = d$ yields

$$(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-B)^j 1^{d-j} = 1 + \sum_{j=1}^{\infty} (-1)^j \binom{d}{j} B^j. \tag{3.10}$$

If α is not an integer then $\binom{\alpha}{j}$ is defined using the gamma function as follows: The gamma function is formally defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad for \quad x \neq 0, -1, -2, \ldots.$$

It follows that $\Gamma(1) = 1$ and that the following recursion applying to all x: $\Gamma(x+1) = x\Gamma(x)$. If x is a positive integer, this means that $\Gamma(x+1) = x!$.

A simple closed expression for the gamma function does not exist, however, normally it is tabulated in the area $0 < x < 1$, and then the function value for all other positive numbers can be found using the recursion formula. For negative non-integer x, the function values can be found using the reverse of the recursion formula: $\Gamma(x) = \frac{1}{x}\Gamma(x+1)$.

The gamma function can now be used to define the generalised binomial coefficients: For positive integers $d > j$, the binomial coefficient is given by

$$\binom{d}{j} = \frac{d!}{j!(d-j)!}.$$

Using the gamma function to rewrite this yields

$$\binom{d}{j} = \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)}. \tag{3.11}$$

If j is a positive integer and $|d| < 1$, then using the reverse of the recursion function on the term $\Gamma(d-j+1)$ in the denominator in Equation (3.11) and the recursion formula on $\Gamma(j-d)$ yields

$$\Gamma(d-j+1) = (-1)^j \frac{\Gamma(d+1)\Gamma(-d)}{\Gamma(j-d)},$$

and substituting this in Equation (3.11) yields

$$\binom{d}{j} = (-1)^j \frac{\Gamma(j-d)}{\Gamma(-d)j!}.$$

Substituting this in Equation (3.10) yields

$$(1-B)^d = 1 + \sum_{j=1}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)j!} B^j. \tag{3.12}$$

It can be shown[7] that

$$(1-B)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(d)j!} B^j. \tag{3.13}$$

The ARFIMA$(0,d,0)$ model is given by

$$(1-B)^d x_t = \varepsilon_t. \tag{3.14}$$

Substituting $(1-B)^d$ with Equation (3.12) yields

$$\varepsilon_t = (1-B)^d x_t = x_t + \sum_{j=1}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)j!} B^j x_t = x_t + \sum_{j=1}^{\infty} \pi_j x_{t-j},$$

where $\pi_j = \dfrac{\Gamma(j-d)}{\Gamma(-d)j!}$. Alternatively, the model can be written as an MA(∞) model by using Equation (3.13):

$$x_t = (1-B)^{-d} \varepsilon_t = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(d)j!} B^j \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

where $\psi_j = \dfrac{\Gamma(j+d)}{\Gamma(d)j!}$.

[7] [Shumway and Stoffer, 2011, p. 269]

For the ARFIMA$(0, d, 0)$ process in Equation (3.14), where ε_t is a white noise process and $|d| < 0.5$, it can be shown[8] that $\{x_t\}$ is stationary and invertible, and that its ACF is

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \sim h^{2d-1},$$

for large h, meaning that the rate of decay is hyperbolic.[9] It follows then for $d \in (0, 0.5)$ that

$$\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$$

Estimation of d

There are several methods for estimating the fractional differencing parameter d.

One parametric method is Gaussian maximum likelihood in the time domain. This method is similar to the estimation method described in Section 3.1. Assuming that the time series has zero mean and that the innovations are Gaussian the log-likelihood function is given by

$$l(d, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = -\frac{1}{2}\log(|\Sigma_n|) - \frac{1}{2}\mathbf{x}^T\Sigma_n^{-1}\mathbf{x},$$

where Σ_n is now also a complicated function of d.

Other estimation methods exist, and among them is the frequently used Geweke-Porter-Hudak (GPH) Log-Periodogram Regression (LPR), often referred to as the GPH estimator. Contrary to the Gaussian maximum likelihood method, the GPH estimator works in the frequency domain, using an approximation to the spectral denisty.

Definition 3.9 (Spectral Density) *If a time series $\{x_t\}$ has an autocovariance γ satisfying $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, then the spectral density for $\{x_t\}$ is defined as*

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h)\exp(-2\pi i\omega h),$$

for $-\infty < \omega < \infty$.

[8] [Shumway and Stoffer, 2011, p. 26]
[9] [Shumway and Stoffer, 2011, p. 269]

Properties of the Spectral Density

Some properties of the spectral density are:

1. $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| \exp(-2\pi i\omega h) < \infty$, due to $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ and $|\exp(i\theta)| = |\cos(\theta) + i\sin(\theta)| = \left(\cos^2(\theta) + \sin^2(\theta)\right)^{1/2} = 1$.
2. f is periodic with period 1, due to $\exp(-2\pi i\omega h)$ being a periodic function of ω with period 1. This means that the domain of f can be constrained to be $-1/2 \leq \omega \leq 1/2$.
3. f is an even function, that is, $f(-\omega) = f(\omega)$.
4. $f(\omega) \geq 0$.
5. $\gamma(h) = \int\limits_{-1/2}^{1/2} \exp(2\pi i\omega h) f(\omega) d\omega$.

Note from item 5 that $\gamma(0) = \mathrm{Var}(x_t) = \int\limits_{-1/2}^{1/2} f(\omega) d\omega$, meaning that the total variance is the spectral density integrated over all frequencies. Note also that the autocovariance function $\gamma(h)$ and the spectral density $f(\omega)$ contain the same information; the autocovariance function contains the information in the form of lags and the spectral density contains the information in the form of cycles.

Furthermore, note that $f(0) = \sum\limits_{h=-\infty}^{\infty} \gamma(h)$. This means that for short memory processes, where $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, the spectral density goes to a constant as the frequency, ω, goes to zero. However, for long memory processes, where $\sum\limits_{h=-\infty}^{\infty} |\gamma(h)| = \infty$, the spectral density goes to infinity when $0 < d < 0.5$ and to zero when $-0.5 < d < 0$. These scenarios are illustrated in Figure 3.1. The GPH estimator uses these properties to determine if the time series has long memory. In the GPH estimator a bandwidth $\omega_{\min}$ to $\omega_{\max}$ is chosen, where the time series is examined. However, instead of examining the spectral density, the method uses the periodogram, which is an estimate of the spectral density. The method then takes the logarithm of the periodogram, which is illustrated in Figure 3.2.

Thus, instead of examining whether the spectral density goes to zero, a constant, or infinity in order to determine if the time series has long memory, the GPH estimator examines the slope of the line resulting from taking the logarithm of the periodogram in the chosen bandwidth. Moreover, it turns out that the slope of the logarithm of the periodogram is the estimate for the fractional differencing

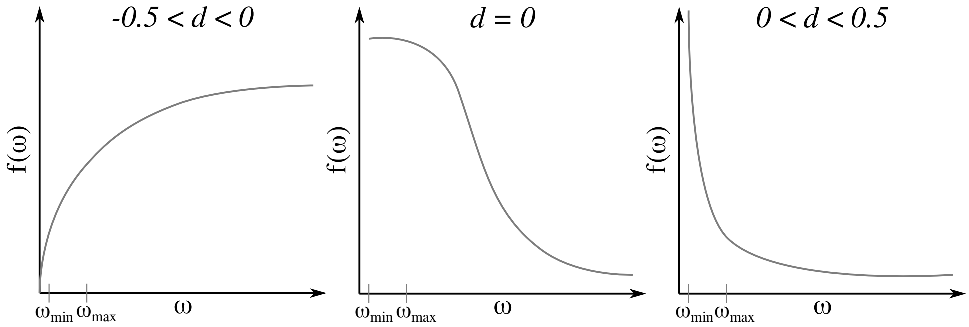


Figure 3.1: Illustrations of the spectral density for an anti persistent time series with $-0.5 < d < 0$, a stationary time series with $d = 0$, and a long memory time series with $0 < d < 0.5$.

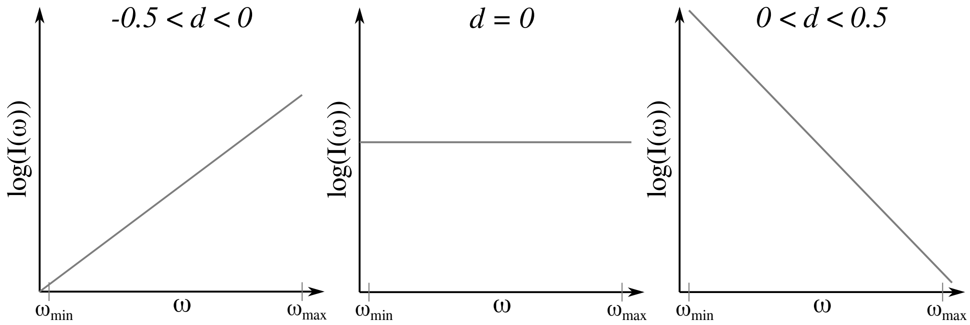


Figure 3.2: Illustrations of the logarithm of the spectral density for an anti persistent time series with $-0.5 < d < 0$, a stationary time series with $d = 0$, and a long memory time series with $0 < d < 0.5$ in the bandwidth ω_{min} to ω_{max}.

parameter, d, multiplied by -2.

In this paper the GPH estimator is used to estimate the fractional differencing parameter d. This is due to the disadvantages of the Gaussian maximum likelihood method in the time domain: The method requires the covariance matrix, Σ_n, of the process which is an extensive computation, and parametric time domain procedures can sometimes have severe bias. This bias of parametric time domain methods is often alleviated when larger sample sizes are considered, however, the sample sizes in this paper are not very large. Another disadvantage of parametric methods is that the number of parameters in the model should be known; thus, the orders of the AR and MA polynomials need to be assumed. One disadvantage of the GPH estimator is that the periodogram is used to estimate the spectral density of the process: The periodogram is not a consistent estimator which causes difficulties for the estimator. However, the GPH estimator has been showed to outperform even correctly specified parametric time domain methods, and it shows a much lower bias. Furthermore, it does not need to make assumptions about the model order, making it more robust.[10]

3.4 Regression with ARMA Errors

This section is based on [Ruppert and Matteson, 2015, pp. 367-373, 377-380].

The focus of this chapter so far has been on modelling time series only using the time series itself. However, the purpose of time series analysis is often to examine the relationship between several time series. This is usually done using multiple linear regression; however, multiple linear regression assumes that the innovations, ε, are mutually independent, which is often not the case.[11] Therefore, another method is to use linear regression with ARMA errors. As the name suggests, the method involves a simple multiple linear regression of one time series on other time series and then assuming that the errors are correlated; therefore, an ARMA model is fitted to the residuals of the regression model:

Assume that the relationship between a dependent time series $\{x_t\}$ and b exogenous time series $\{y_{t,1}\}$, ..., $\{y_{t,b}\}$ is to be examined. Then the linear regression model with ARMA errors is given by

$$x_t = \beta_1 y_{t,1} + \cdots + \beta_b y_{t,b} + \varepsilon_t,$$

where ε_t is an ARMA process; that is,

$$\phi(B)\varepsilon_t = \theta(B)u_t,$$

where $u_t \sim WN(0, \sigma^2)$. As in regular multiple linear regression, the coefficients $\beta_1, \ldots, \beta_b$ signify the effect of changing the exogenous time series, $y_{t,1}, \ldots, y_{t,b}$, on the dependent time series, x_t.

The procedure of fitting a regression model with ARMA errors is to first use multiple linear regression on the dependent time series and the exogenous time series, and then fit an ARMA model to the residuals from the multiple linear regression.

In order to make predictions, the procedure is as follows: Make predictions for the ARMA model fitted to the residuals from the multiple linear regression. Make predictions for the exogenous time series and add them to the predictions from the ARMA model.

[10] [Sibbertsen, 2004, p. 476] and [Nielsen and Frederiksen, 2005, p. 406]

[11] [Ruppert and Matteson, 2015, p. 367]

Chapter 4

Modelling the Data

In order to determine which model best fits a time series, there are some steps to consider.

- Plot and examine the time series and its ACF and PACF. Try to determine if the time series looks stationary and look for trends, seasonal components, level shifts, extreme values, increasing variance, et cetera. For time series that do not seem to be stationary, it can be difficult to determine if the time series has a unit root and should be differenced, or a time trend which should be removed: Use a unit root test, such as the ADF. The ACF should also indicate if the time series has long memory.
- Apply relevant transformations to the data such that the residuals look stationary.
- After having stationarised the data, use the ACF and PACF to determine preliminary values of p and q.
- Estimate the parameters of the preliminary chosen models.
- Use diagnostic tests to confirm that the residuals of the models behave as Gaussian white noise.
- Choose p and q.

4.1 Global CO_2 Emissions

The global CO_2 emission time series in Figure 2.1 does not look stationary. Until approximately 1960, the time series appears almost linear, however, after 1960 there appears to be an increasing trend. Taking the logarithm of the time series yields the result in Figure 4.1. Taking the logarithm of the global CO_2 emissions seems

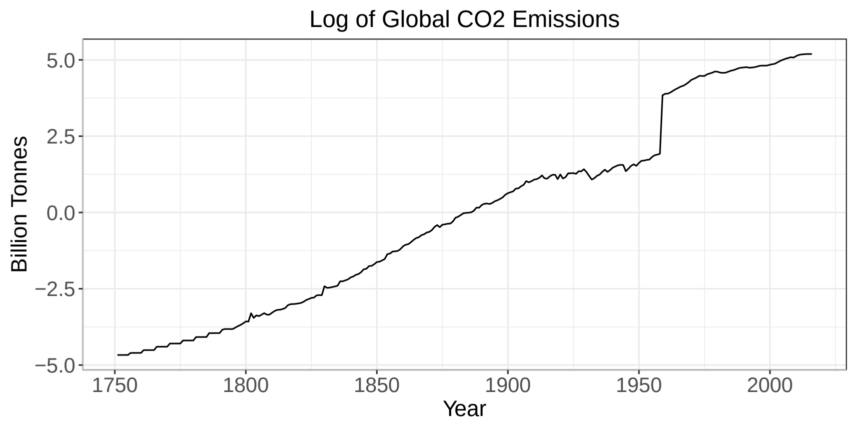


Figure 4.1: The logarithm of global CO_2 emissions measured in billion tonnes from 1750 to 2016.

to have changed the exponential time trend to a linear time trend: However, the sudden jump in approximately 1960 is still present in the data.

One way of removing this linear time trend with a jump at 1958 and to make the time series stationary is by fitting the following model to the data

$$x_t = \beta_0 + \beta_1 \mathbb{1}_{(t>1958)} + \beta_2 t + \beta_3 \mathbb{1}_{(t>1958)} t.$$

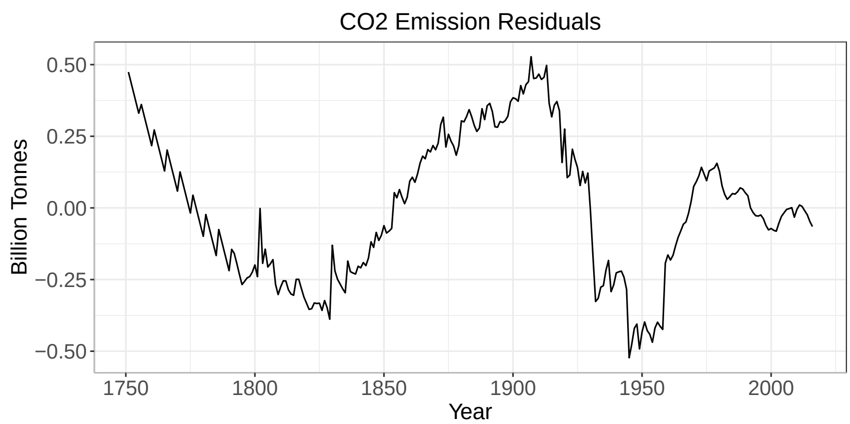


Figure 4.2: The residuals from the model in Equation (4.1), where x_t is the logarithm of the global CO_2 emissions at time t.

Fitting this model to the data yields the following significant[1] parameter estimates:

$$x_t = -5.181 + 4.715 \cdot \mathbb{1}_{(t>1958)} + 0.036t - 0.015 \cdot \mathbb{1}_{(t>1958)}t. \tag{4.1}$$

The residuals from the model can be seen in Figure 4.2. The figure shows that the model has removed the time trend successfully: However, the time series still does not look completely stationary. This is confirmed by the ACF shown in Figure 4.3.

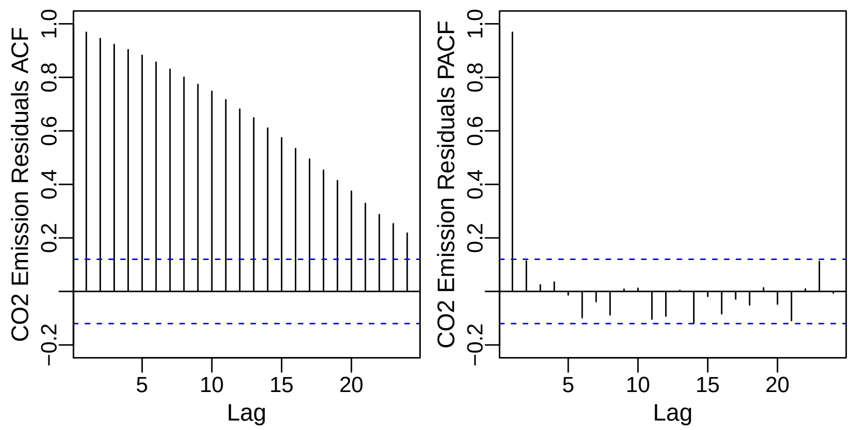


Figure 4.3: ACF and PACF for the residual time series in Figure 4.2.

When performing an ADF test on the residuals, the null hypothesis is not rejected, meaning that it can not be rejected that the residual time series has a unit root. The differenced residual time series, which now looks stationary, is shown in Figure 4.4. This is confirmed by the ACF and PACF shown in Figure 4.5. The ACF and

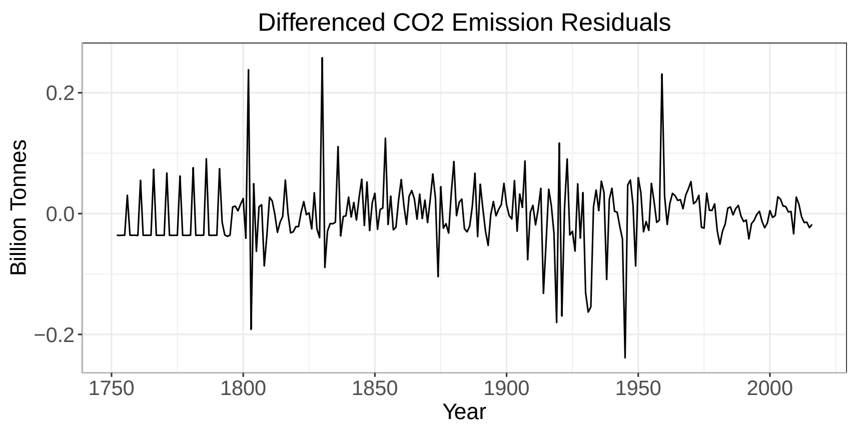


Figure 4.4: The differenced residual time series from 1751 to 2016.

[1] Using a significance level of 0.05.

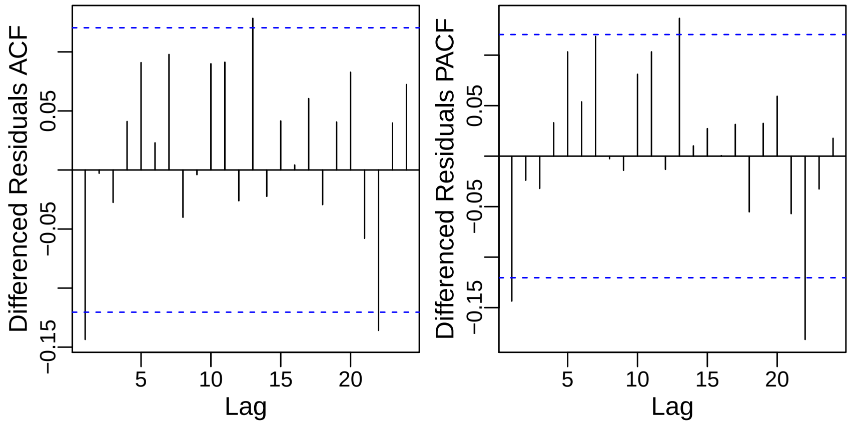


Figure 4.5: ACF and PACF for the differenced residual time series in Figure 4.4.

PACF could indicate an AR(1) model or an MA(1) model since it could be argued that the ACF is tailing off and the PACF cuts off at lag 1, or it could be argued that the PACF is tailing off and the ACF cuts off at lag 1. Since these two models are only preliminary choices, similar models are also fitted, and the best model is chosen based on BIC and that the residuals behave as white noise. This results in the best model for the differenced residuals from the global CO_2 emission model in Equation (4.1) being an $\mathrm{MA}(1)$ model with the following significant parameter:

$$x_t = \varepsilon_t - 0.149\varepsilon_{t-1}.$$

In conclusion, the logarithm of the global CO_2 emissions where a time trend and drift have been removed using the model in Equation (4.1), follows an $\mathrm{ARIMA}(0,1,1)$ model.

4.2 Atmospheric CO_2 Concentration

The atmospheric CO_2 concentration time series in Figure 2.1 does not look stationary. There seems to be either a time trend in the time series, which should be removed, or a unit root, meaning that the time series should be differenced.

The following model

$$x_t = \beta_0 + \beta_1 \mathbb{1}_{(t>1958)} + \beta_2 \mathbb{1}_{(t>1958)} t + \beta_3 t^2 + \beta_4 \mathbb{1}_{(t>1958)} t^2$$

with the significant parameters

$$x_t = 277.3 + 416.4 \cdot \mathbb{1}_{(t>1958)} - 4.399 \cdot \mathbb{1}_{(t>1958)} t + 0.0009t^2 + 0.012 \cdot \mathbb{1}_{(t>1958)} t^2 \quad (4.2)$$

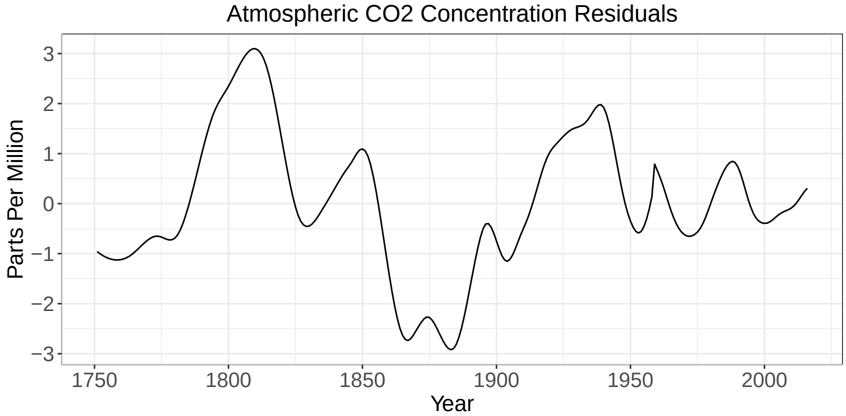


Figure 4.6: The residuals from the model in Equation (4.2), where x_t is the atmospheric CO_2 concentration at time t.

yields the residuals in Figure 4.6. It is clear that the model in Equation (4.2) has removed the time trend, however, the ACF and PACF of the residuals seen in Figure 4.7 still indicate some autocorrelation in the time series. This is confirmed by an

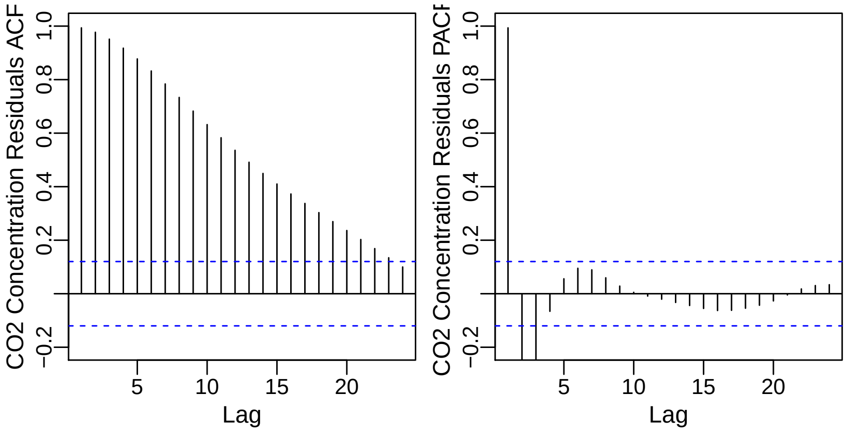


Figure 4.7: ACF and PACF for the residual time series in Figure 4.6.

estimated fractional integration parameter, d, of approximately 0.93: However, as a d parameter estimate of 0.93 makes the differenced time series antipersistent, it is chosen to fractionally difference the atmospheric CO_2 concentration residuals with $d = 0.49$. This yields the time series in Figure 4.8. The ACF and PACF of the fractionally differenced time series, seen in Figure 4.9, could indicate several different models, since it could be argued that both the ACF and PACF are tailing off, indicating an ARMA model. After fitting several different ARMA models to the time series, the best model for the fractionally differenced atmospheric CO_2 concentration residuals is found to be an $\text{ARMA}(2, 2)$ model with the following

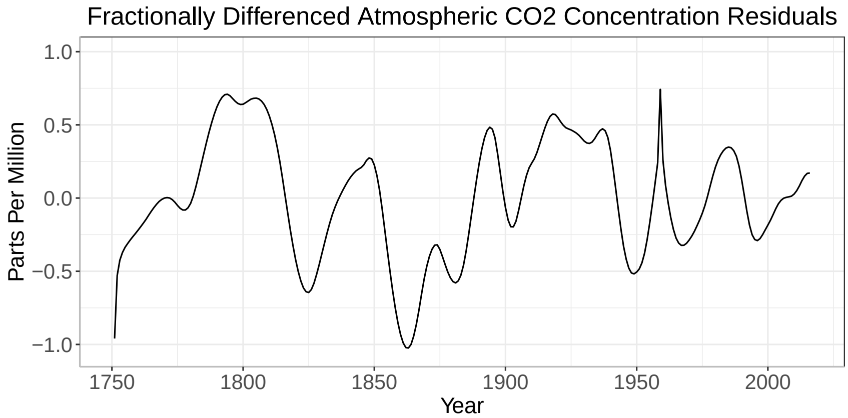


Figure 4.8: The fractionally differenced atmospheric CO_2 concentration residual time series from 1751 to 2016, where $d = 0.49$.

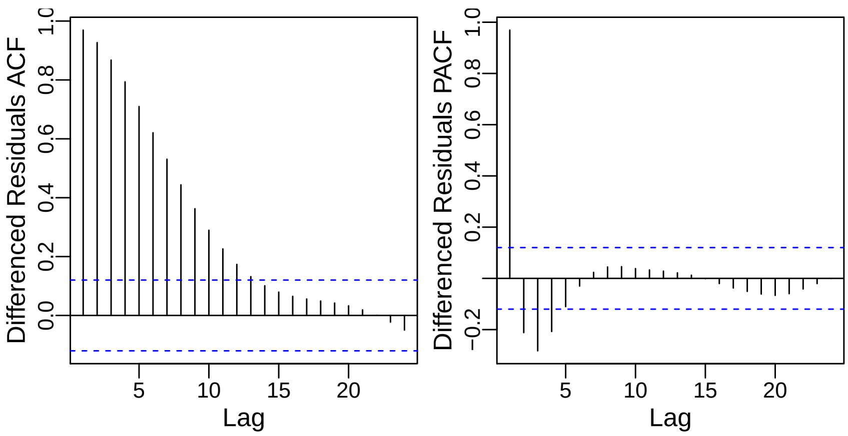


Figure 4.9: ACF and PACF for the fractionally differenced residual time series in Figure 4.8.

significant parameters:

$$x_t = 1.822x_{t-1} - 0.847x_{t-2} + \varepsilon_t - 0.473\varepsilon_{t-1} + 0.181\varepsilon_{t-2}.$$

This means that the best model for the atmospheric CO_2 concentration residuals is an ARFIMA$(2, 0.49, 2)$ model.

4.3 Temperature Anomalies

The temperature time series in Figure 2.4 seems to have an upwards-going trend, that should be removed. In order to remove the time trend, the following model is fitted to the data

$$x_t = \beta_0 + \beta_1 t + \beta_2 t^2.$$

Fitting this model to the temperature time series yields the following significant parameter estimates:

$$x_t = -0.266 - 0.004t + 0.0001t^2. \tag{4.3}$$

The residuals from the model is shown in Figure 4.10. The figure shows that the

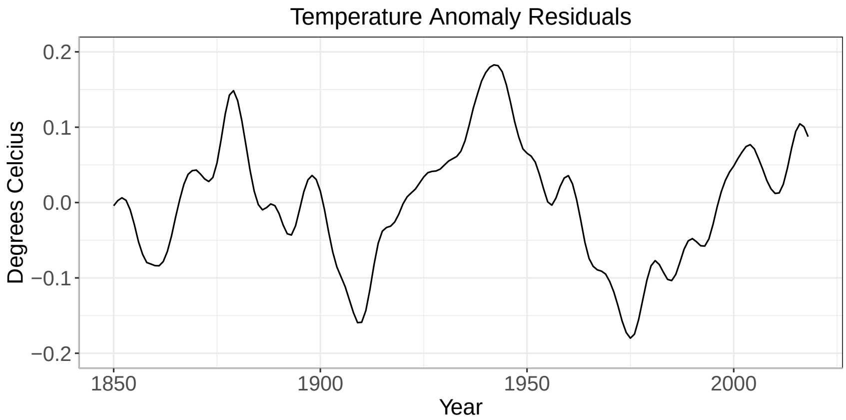


Figure 4.10: The residuals from the model in Equation (4.3), where x_t is temperature anomaly at time t.

time trend has been removed, and the time series looks more stationary: However, the ACF still indicates some persistence, as seen in Figure 4.11. Differencing the

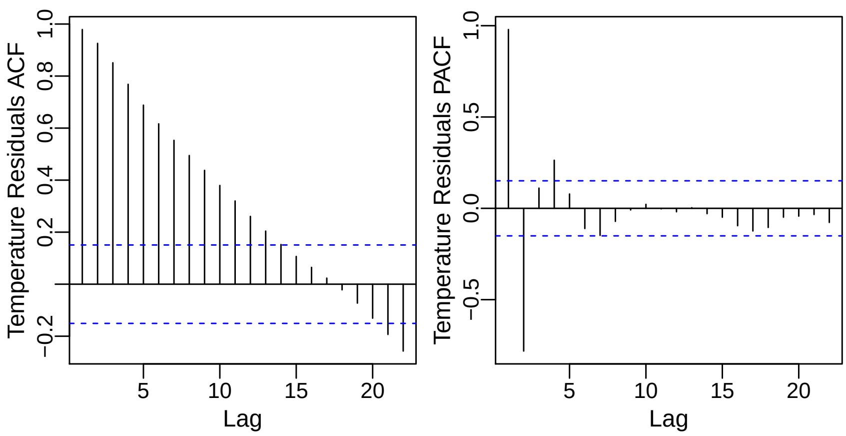


Figure 4.11: ACF and PACF for the residual time series in Figure 4.10.

time series one time makes the time series look over-differenced, which is confirmed by an estimated fractional integration parameter, d, of approximately 0.77. As for the atmospheric CO_2 concentration, it is chosen to fractionally integrate the temperature anomaly residuals with $d = 0.49$. This yields the time series in Figure 4.12. The time series looks stationary, which is confirmed by the ACF and PACF

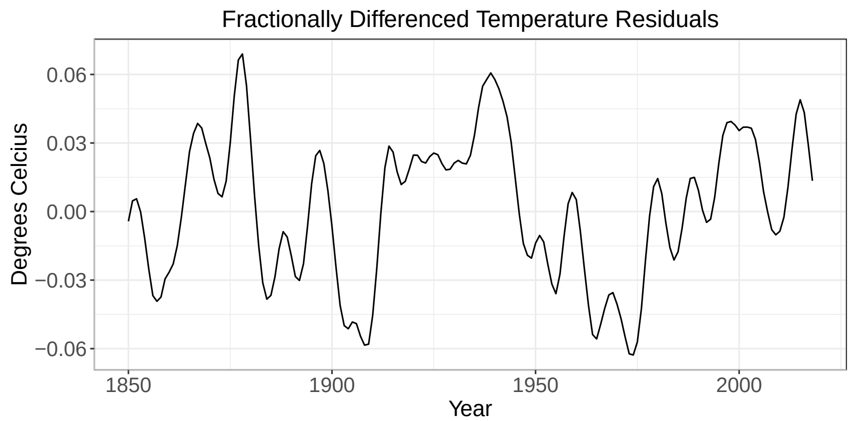


Figure 4.12: The fractionally differenced temperature anomaly residual time series from 1850 to 2016, where $d = 0.49$.

shown in Figure 4.13. The ACF and PACF in the figure could indicate several

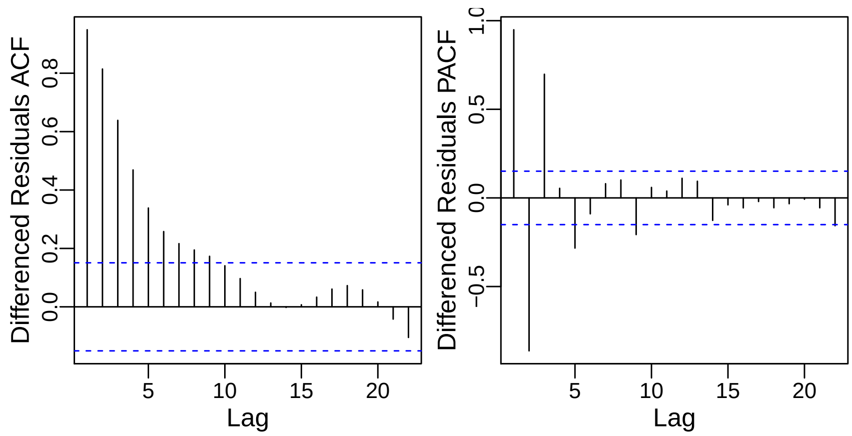


Figure 4.13: ACF and PACF for the differenced residual time series in Figure 4.12.

different models, since it could be argued that the ACF is tailing off and that the PACF cuts off after lags three or five, or it could be argued that both the ACF and PACF are tailing off. After fitting a range of different models and using BIC and residual analysis for choosing the best model, it is concluded that the best model for the fractionally differenced temperature anomaly residuals is an ARMA$(3, 3)$ model with the following significant parameters:

$$x_t = 2.108x_{t-1} - 1.725x_{t-2} + 0.555x_{t-3} + \varepsilon_t + 0.424\varepsilon_{t-1} + 0.678\varepsilon_{t-2} + 0.275\varepsilon_{t-3},$$

meaning that the best model for the temperature anomaly residuals is an ARFIMA$(3, 0.49, 3)$ model.

4.4 GDP per Capita

The first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies are very similar as seen in Figure 2.10. This means, that the process for finding the model that fits each time series best is also similar for the four groups. None of the four time series in Figure 2.10 seem stationary, since they all have an upwards-going trend.

The following models are used to remove the time trend in each time series:

$$\begin{aligned}
&\text{Low-Income:} && x_t = -4.159 + 0.004t^2, \\
&\text{Lower-Middle-Income:} && x_t = -4.244 - 0.093t + 0.006t^2, \\
&\text{Upper-Middle-Income:} && x_t = -4.435 - 0.096t + 0.006t^2, \\
&\text{High-Income:} && x_t = -7.423 + 0.006t^2.
\end{aligned} \tag{4.4}$$

Note that all of the models include a significant intercept, which might seem surprising when looking at Figure 2.10. This is due to the four time series having mean zero and therefore they do not intercept zero at the beginning of the data; however, in Figure 2.10 each time series has been shifted to have initial value equal to zero in order to make it easier to compare their development.

The residuals from the models above are shown in Figure 4.14. The figure illus-

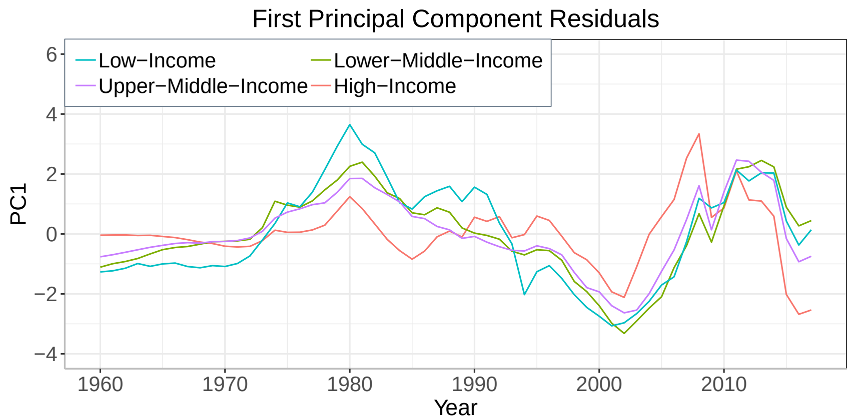


Figure 4.14: The residuals from the models in Equation (4.4) of the first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies from 1961 to 2017.

trates how the models in Equation (4.4) have removed the time trend from each time series, however, they do not look completely stationary. This is confirmed by the ACF and PACF, shown in Figure 4.15 for the residuals of the model for the low-income economies. The ACF and PACF plots for the residuals from the other

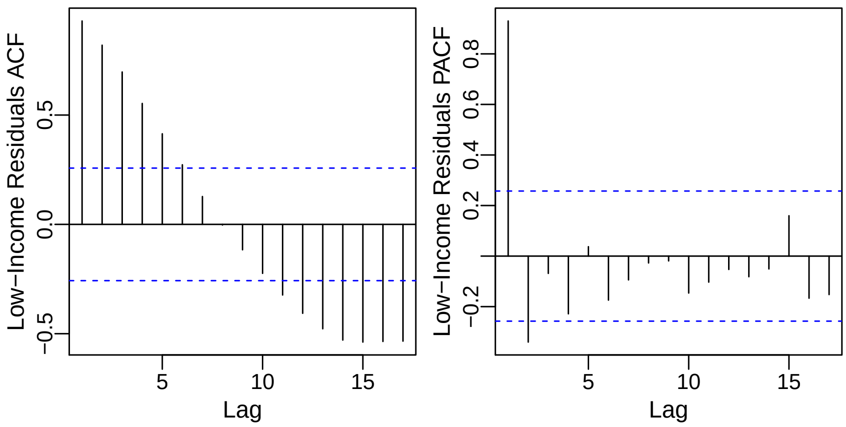


Figure 4.15: ACF and PACF for the residual time series in Figure 4.14 for the low-income economies.

models look similar to those from the low-income economies. Differencing the residual time series yields the results shown in Figure 4.16. The time series seem

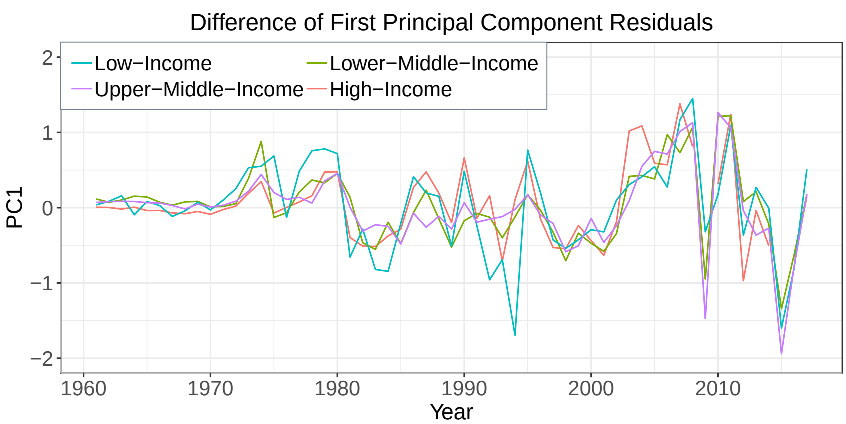


Figure 4.16: The difference of the residual time series in Figure 4.14 from 1961 to 2017.

more stationary now, which is confirmed by by the ACFs and PACFS for the times series, which are shown in Figure 4.17.

In order to determine preliminary values of p and q, these ACFs and PACs are examined.

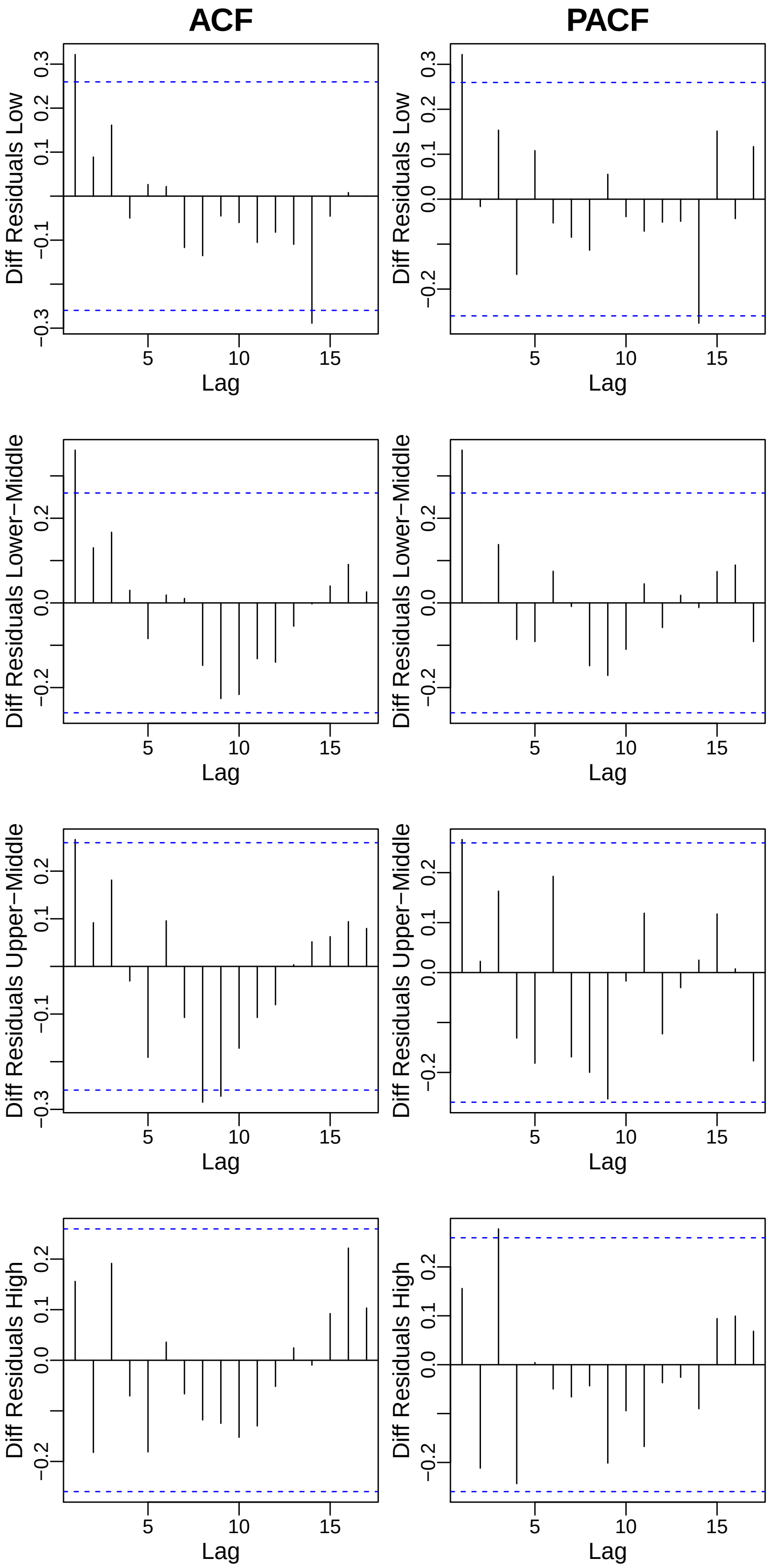


Figure 4.17: ACF and PACF for the differenced first principal component residuals in Figure 4.16 for low-income, lower-middle-income, upper-middle-income, and high-income economies.

4.4.1 Low-Income Economies

For the low-income economies, the ACF and PACF could indicate an AR(1) model or an MA(1) model, since it could be argued that the ACF is tailing off and the PACF cuts off at lag one, or it could be argued that the PACF is tailing off and the ACF cuts off at lag one. Since these two models are only preliminary choices similar models are also fitted and the best model is chosen based on BIC and a residual analysis. This results in the best model for the differenced residuals being an MA(1) model with the following parameter:

$$x_t = \varepsilon_t + 0.376\varepsilon_t,$$

meaning that the best model for the residuals from the model for the first principal component for the low-income economies is an ARIMA$(0,1,1)$ model.

4.4.2 Lower-Middle-Income Economies

For the lower-middle-income economies, the ACF and PACF could also indicate an AR(1) model or an MA(1) model, since it could be argued that the ACF is tailing off and the PACF cuts off at lag one, or that the PACF is tailing off and the ACF cuts off at lag one. After fitting similar models, it is concluded that the best model for the differenced residuals is an AR(1) model with the following parameter:

$$x_t = 0.356x_{t-1} + \varepsilon_t.$$

This means that the best model for the residuals from the model for the first principal component for the lower-middle-income economies is an ARIMA$(1,1,0)$ model.

4.4.3 Upper-Middle-Income Economies

The ACF and PACF for the differenced first principal component residuals for the upper-middle-income economies are very similar to those for the lower-middle-income economies. Therefore, the ACF and PACF for the upper-middle-income economies indicate the same models as for the lower-middle-income economies; that is, an AR(1) model or an MA(1) model. Similarly to the lower-middle-income economies, the best model for the differenced data for the upper-middle-income economies is an AR(1) model with the following parameter:

$$x_t = 0.262x_{t-1} + \varepsilon_t,$$

meaning that the best model for the residuals from the model for the first principal component for the upper-middle-income economies is an ARIMA$(1,1,0)$ model.

4.4.4 High-Income Economies

For the high-income economies, the ACF and PACF could indicate an AR(3) model or an MA(1) model. The best model for the differenced data ends up being a MA(2) model with the following parameters:

$$x_t = \varepsilon_t + 0.318\varepsilon_{t-1} - 0.259\varepsilon_{t-2},$$

which means that the best model for the residuals from the model for the first principal component for the high-income economies is an ARIMA$(0,1,2)$ model.

The following boxes sum up the models chosen for the time series:

Global CO_2 Emissions
$x_t =$ Global CO_2 Emissions at time t
$y_t = \log(x_t) + 5.181 - 4.715 \cdot \mathbb{1}_{(t>1958)} - 0.036t + 0.015 \cdot \mathbb{1}_{(t>1985)} t$
$z_t = y_t - y_{t-1}$
$z_t = \varepsilon_t - 0.149\varepsilon_{t-1}$

Atmospheric CO_2 Concentration
$x_t =$ Atmospheric CO_2 Concentration at time t
$y_t = x_t - 277.3 - 416.4 \cdot \mathbb{1}_{(t>1958)} + 4.399 \cdot \mathbb{1}_{(t>1958)} t - 0.0009t^2 - 0.012 \cdot \mathbb{1}_{(t>1958)} t^2$
$z_t = (1-B)^{0.49} y_t$
$z_t = 1.822z_{t-1} - 0.847z_{t-2} + \varepsilon_t - 0.473\varepsilon_{t-1} + 0.181\varepsilon_{t-2}$

Temperature Anomalies
$x_t =$ Temperature Anomaly at time t
$y_t = x_t + 0.266 + 0.004t - 0.0001t^2$
$z_t = (1-B)^{0.49} y_t$
$z_t = 2.108z_{t-1} - 1.725z_{t-2} + 0.555z_{t-3} + \varepsilon_t + 0.424\varepsilon_{t-1} + 0.678\varepsilon_{t-2} + 0.275\varepsilon_{t-3}$

GDP per Capita for Low-Income Economies
$x_t =$ First Principal Component for Low-Income Economies at time t
$y_t = x_t + 4.159 - 0.004t^2$
$z_t = y_t - y_{t-1}$
$z_t = \varepsilon_t + 0.376\varepsilon_{t-1}$

GDP per Capita for Lower-Middle-Income Economies
$x_t =$ First Principal Component for Lower-Middle-Income Economies at time t
$y_t = x_t + 4.244 + 0.093t - 0.006t^2$
$z_t = y_t - y_{t-1}$
$z_t = 0.356z_{t-1} + \varepsilon_t$

GDP per Capita for Upper-Middle-Income Economies
$x_t =$ First Principal Component for Upper-Middle-Income Economies at time t
$y_t = x_t + 4.435 + 0.096t - 0.006t^2$
$z_t = y_t - y_{t-1}$
$z_t = 0.262z_{t-1} + \varepsilon_t$

GDP per Capita for High-Income Economies
$x_t =$ First Principal Component for High-Income Economies at time t
$y_t = x_t + 7.423 - 0.006t^2$
$z_t = y_t - y_{t-1}$
$z_t = \varepsilon_t + 0.318\varepsilon_{t-1} - 0.259\varepsilon_{t-2}$

Chapter 5

Predictions

Temperature Anomalies

The models from Chapter 4 can now be used to predict the future development in CO_2, temperature, and GDP per capita. First the temperature anomaly model is used to predict when the temperature has increased 1.5°C and 2°C compared to pre-industrial levels. Since the temperature anomaly in 1850 was -0.274°C, the temperature has increased 1.5°C when the temperature anomaly has reached $-0.274 + 1.5 = 1.226$°C, and the temperature has increased 2°C when the temperature anomaly has reached $-0.274 + 2 = 1.726$°C. Using the model fitted to

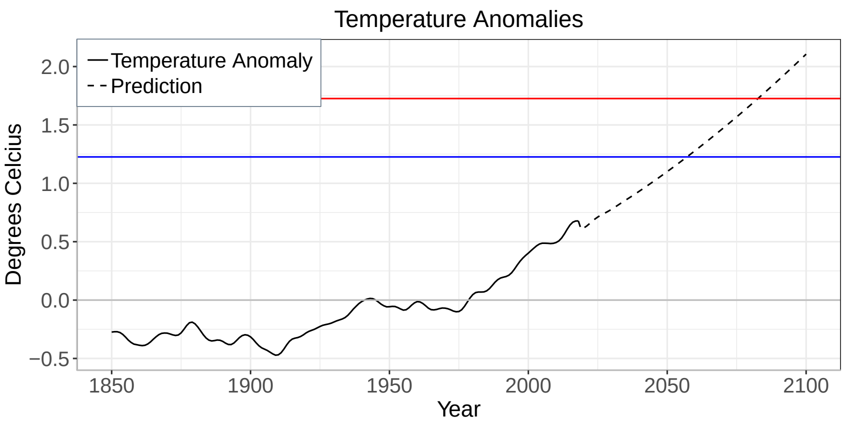


Figure 5.1: Temperature anomalies (black solid line) from 1850 to 2018 and temperature anomaly predictions (black dashed line) from 2019 to 2100. The horizontal grey line illustrates 0°C, the horizontal blue line illustrates 1.226°C, that is, a temperature increase of 1.5°C, and the horizontal red line illustrates 1.726°C, that is, a temperature increase of 2°C.

the temperature anomaly data in Chapter 4 to make predictions yields the temperature anomaly predictions seen in Figure 5.1, where the horizontal blue line illustrates 1.226°C and the horizontal red line illustrates 1.726°C, that is, temperature increases of 1.5°C and 2°C, respectively. The temperature anomaly predictions intersect 1.226°C between 2057 and 2058, meaning that the temperature has increased 1.5°C compared to pre-industrial levels in 2058 according to the model fitted in Chapter 4, and the temperature anomaly predictions intersect 1.726°C between 2082 and 2083, meaning that the temperature has increased 2°C compared to pre-industrial levels in 2083.

After having determined the years in which the temperature has increased 1.5°C and 2°C compared to pre-industrial levels, predictions for the other time series can be made and their behaviour examined in these years.

Global CO_2 Emissions

The predictions of the global CO_2 emissions can be seen in Figure 5.2. The fig-

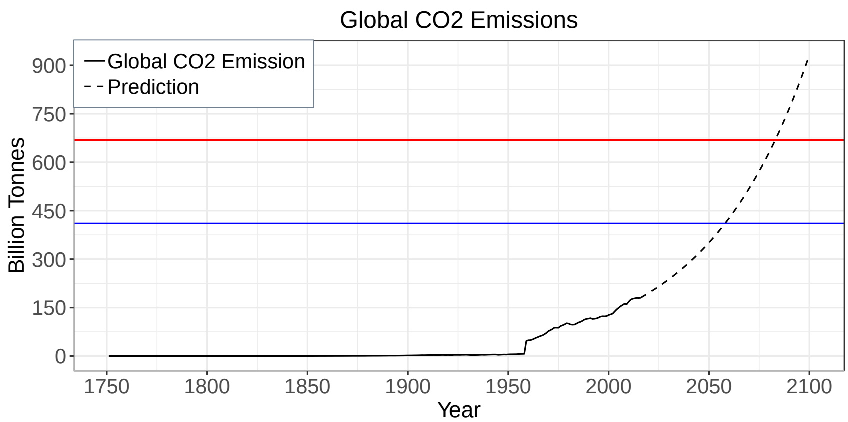


Figure 5.2: Global CO_2 emissions (black line) from 1751 to 2016 and global CO_2 emission predictions (black dashed line) from 2017 to 2100. The horizontal blue line illustrates the global CO_2 emissions in 2058, and the horizontal red line illustrates the global CO_2 emissions in 2083.

ure illustrates how the global CO_2 emissions continue to increase rapidly since the exponential growth trend continues. By 2058 the global CO_2 emissions will have reached approximately 410 billion tonnes and by 2083 they will have reached approximately 669 billion tonnes unless the current development is changed.

Atmospheric CO_2 Concentration

The predictions of the atmospheric CO_2 concentration can be seen in Figure 5.3. The atmospheric CO_2 concentration also continues to increase, and in 2058 the

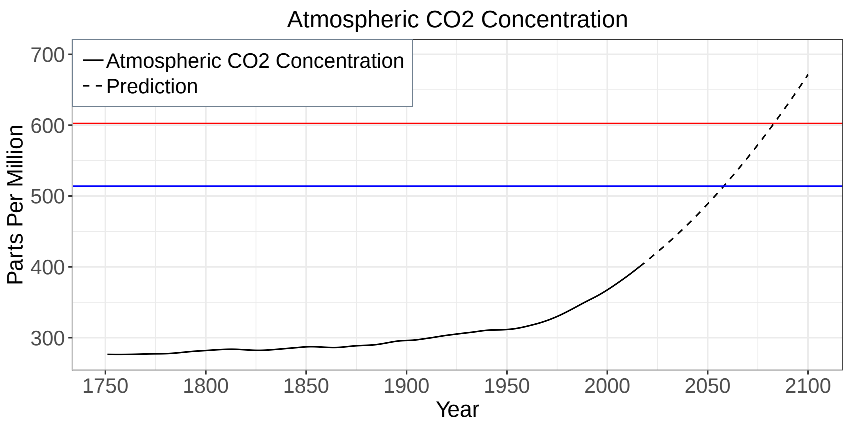


Figure 5.3: Atmospheric CO_2 concentration (black line) from 1750 to 2016 and atmospheric CO_2 concentration predictions (black dashed line) from 2017 to 2100. The horizontal blue line illustrates the atmospheric CO_2 concentration in 2058, and the horizontal red line illustrates the atmospheric CO_2 concentration in 2083.

atmospheric CO_2 concentration will be approximately 511 PPM and in 2083 it will be approximately 599 PPM.

GDP per Capita

The predictions of the development in the first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies can be seen in Figure 5.4. The figure illustrates how the first principal component for all four groups will continue their upwards-going trend, however, the rate of increase seems to be higher for the high-income, upper-middle-income, and lower-middle-income groups and lower for the low-income group. This means that higher-income economies will continue to have larger increases in GDP per capita compared to low-income economies, and that the difference between development in higher-income and low-income economies will continue to grow.

5.1 Regression with ARMA errors

The predictions for each time series made so far are only based on the time series itself and do not incorporate information about the development in the other time

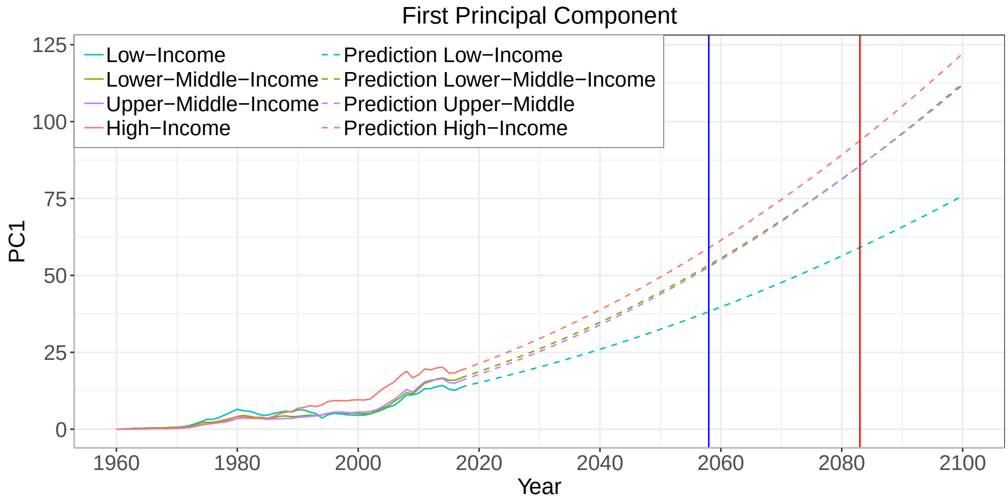


Figure 5.4: First principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies (coloured solid lines) from 1960 to 2017 and predictions of the time series (soloured dashed lines) from 2018 to 2100. The vertical blue line illustrates 2058, and the vertical red line illustrates 2083.

series. Thus, in order to examine the effect of one time series on another, regression models with ARMA errors are used, which where introduced in Section 3.4.

When fitting a regression model with ARMA errors, the time series in the model should be stationary. However, it would be preferred to also include the time trends in each time series, since it is hoped that the time trend in the exogenous time series can help explain the time trend in the dependent time series. Therefore, instead of using the original time series in the regression models, the stationary data that is used to fit the ARMA models in Chapter 4 is used, where each time series' time trend is added again. For example, for the global CO_2 emissions, the data defined as z_t in the box at the end of Chapter 4 is used, where the inverse of the transformations used in order to transform the data from x_t to y_t, which is the step where the time trend is removed, is applied. This means that the first difference still is applied to the data, but the time trend has been added again. This new data is refered to as the stationary data with time trend.

As in Chapter 4, the choice of an ARMA model fitted to the residuals from the regression model is based on BIC and that the residuals from the ARMA model behave as white noise.

Global CO_2 Emissions

In order to determine the effect of global CO_2 emissions on temperature changes, different paths of changes in global CO_2 emissions are predicted. This will aid in determining if reducing global CO_2 emissions will have an effect on climate changes and if temperature changes will have an effect on economic growth. In Figure 5.5 the different prediction paths are showed in the form of the stationary data with time trend. The dashed line (Prediction) illustrates the predictions based

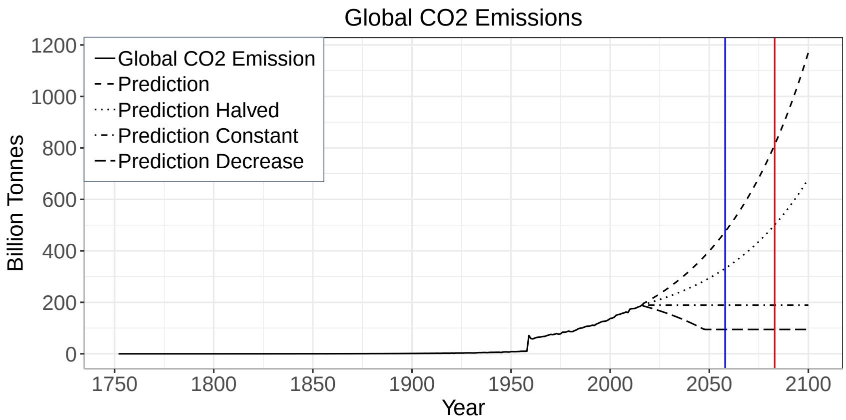


Figure 5.5: Stationary global CO_2 emissions with time trend (black solid line) from 1752 to 2016 and global CO_2 emission predictions from Figure 5.2 based on the ARMA$(0,1)$ model fitted in Chapter 4 with time trend (black dashed line) from 2017 to 2100. Prediction Halved (black dotted line) illustrates the prediction path where the increase in the predictions has been halved. Prediction Constant (black dotted and dashed line) illustrates the prediction path where the global CO_2 emissions continue to be at the same level as in 2016. Prediction Decrease (black long-dashed line) illustrates the prediction path where the global CO_2 emissions decrease to half of the current level. All predictions are from 2017 to 2100.

on the ARMA$(0,1)$ model fitted in Chapter 4. These are the predictions which are also illustrated in Figure 5.2, however, in Figure 5.5 the data and predictions are transformed to be stationary with time trend.

Prediction Halved, which is the second prediction path from the top, illustrated by the dotted line, is constructed using the predictions mentioned above from the ARMA$(0,1)$ model illustrated as the dashed line in the same figure: The increase in the global CO_2 emissions during the prediction period is halved, such that in the space of the stationary data with time trend, the increase in the prediction period is now approximately $680 - 189 = 491$ billion tonnes compared to an increase in the original predictions of approximately $1171 - 189 = 982 = 491 \cdot 2$ billion tonnes.

Prediction Constant, which is the third prediction path from the top, illustrated by the dotted and dashed line, is where the global CO_2 emissions continue at the same level as in 2016, which is the last observation in the data set. In the space of the stationary data with time trend, this level is approximately 189 billion tonnes.

Prediction Decrease, which is the fourth and final prediction path, illustrated by the long-dashed line, is also based on the predictions from the ARMA$(0,1)$ model: The increases in the global CO_2 emissions during the prediction period has been subtracted from the 2016 level, such that the global CO_2 emissions are decreasing instead of increasing, until it reaches half of the level in 2016, which in the space of the stationary data with time trend is approximately $189/2 \approx 95$ billion tonnes.

Atmospheric CO_2 Concentration

After having constructed different prediction paths for the global CO_2 emissions, the effects of this on the atmospheric CO_2 concentration can be examined. Fitting a regression model on the atmospheric CO_2 concentration with global CO_2 emissions as exogenous variable and fitting an ARMA model to the residuals yields the following parameters, where $\{x_t\}$ is the atmospheric CO_2 concentration, $\{y_t\}$ is global CO_2 emissions, and $\{u_t\}$ is white noise:

$$x_t = 288.678 + 0.572y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows an ARMA$(1,1)$ process given by

$$\varepsilon_t = 0.982\varepsilon_{t-1} + u_t - 0.219u_{t-1}.$$

The predictions for the stationary atmospheric CO_2 concentration with time trend including information about the global CO_2 emissions can be seen in Figure 5.6. The grey solid line (Original Prediction) illustrates the predictions based on the ARMA$(2,2)$ model fitted in Chapter 4 in the space of the stationary data with time trend, thus, they do not include information about global CO_2 emissions. The remaining four predictions; Prediction, Prediction Halved, Prediction Constant, and Prediction Decrease are predictions from the regression model with ARMA errors using the four prediction paths for the global CO_2 emissions showed in Figure 5.5.

Figure 5.6 illustrates how changing the global CO_2 emissions will affect the atmospheric CO_2 concentration. Continuing to increase the global CO_2 emissions will cause the atmospheric CO_2 concentration to increase at an even higher rate than in the past. Halving the global CO_2 emissions will only cause the atmospheric CO_2 concentration to increase at a lower rate until approximately 2058, after which it will increase at a higher rate than in the past, causing the prediction path for Prediction Halved to pass the prediction path for Original Prediction at the beginning of the next century. However, limiting the global CO_2 emissions will have a

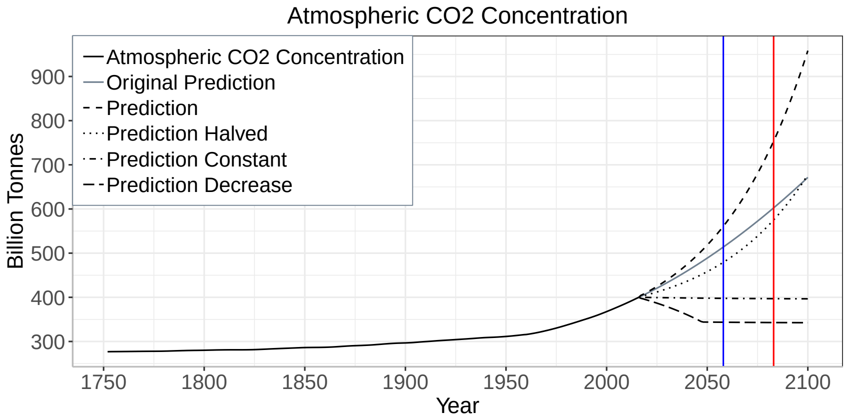


Figure 5.6: Stationary atmospheric CO_2 concentration with time trend (black solid line) from 1752 to 2016 and atmospheric CO_2 concentration predictions from Figure 5.3 based on the ARMA$(2,2)$ model fitted in Chapter 4 with time trend (grey solid line) from 2017 to 2100. Prediction (black dashed line), Prediction Halved (black dotted line), Prediction Constant (black dotted and dashed line), and Prediction Decrease (black long-dashed line) illustrate the predictions made by using the global CO_2 emission prediction paths in the regression model with ARMA errors. All predictions are from 2017 to 2100.

positive effect on the atmospheric CO_2 concentration: Stopping the increase in the global CO_2 emissions or even decreasing global CO_2 emissions will have an instant positive effect on the atmospheric CO_2 concentration. However, as mentioned in Chapter 2 the atmospheric CO_2 concentration will continue to increase even if the global CO_2 emissions are stabilised, and it would take a substantial decrease in the global CO_2 emissions for the atmospheric CO_2 concentration to stabilise. This is not apparent from the regression model with ARMA errors, indicating that the fitted model is too simple for the atmospheric CO_2 concentration.

Temperature Anomalies

Fitting a regression model on the temperature anomalies with atmospheric CO_2 concentration as exogenous variable and fitting an ARMA model to the residuals yields the following parameters, where $\{x_t\}$ is the temperature anomalies, $\{y_t\}$ is atmospheric CO_2 concentration, and $\{u_t\}$ is white noise:

$$x_t = -2.938 + 0.009 y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows an ARMA$(3,3)$ process given by

$$\varepsilon_t = 2.179\varepsilon_{t-1} - 1.821\varepsilon_{t-2} + 0.621\varepsilon_{t-3} + u_t + 0.395u_{t-1} + 0.708u_{t-2} + 0.290u_{t-3}.$$

The predictions for the stationary temperature anomalies with time trend based on the different predictions paths for the atmospheric CO_2 concentration can be seen in Figure 5.7. Again the grey solid line (Original Prediction) illustrates the temperature anomaly predictions based on the ARMA$(3,3)$ model fitted in Chapter 4 in the space of the stationary data with time trend. The four other prediction paths illustrate the predictions based on the regression model with ARMA errors using the four prediction paths for the atmospheric CO_2 concentration showed in Figure 5.6 that are based on the global CO_2 emission predictions showed in Figure 5.5.

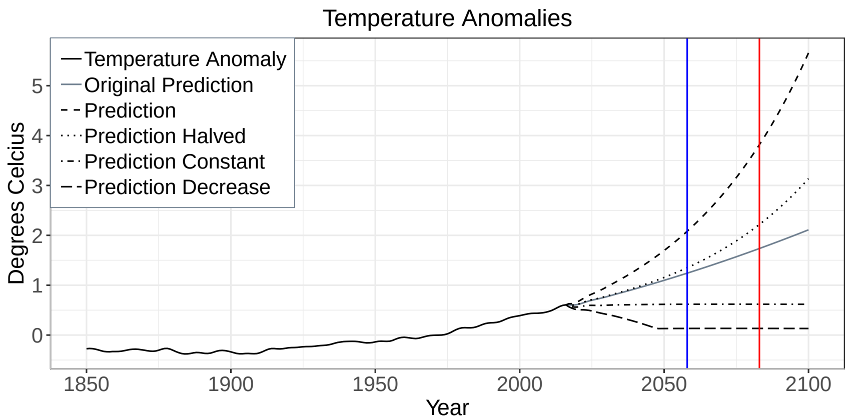


Figure 5.7: Stationary temperature anomalies with time trend (black solid line) from 1850 to 2016 and temperature anomaly predictions from Figure 5.1 based on the ARMA$(3,3)$ model fitted in Chapter 4 with added time trend (grey solid line) from 2017 to 2100. Prediction (black dashed line), Prediction Halved (black dotted line), Prediction Constant (black dotted and dashed line), and Prediction Decrease (black long-dashed line) illustrate the predictions made by using the atmospheric CO_2 concentration prediction paths in the regression model with ARMA errors. All predictions are from 2017 to 2100.

Figure 5.7 illustrates the difference between the temperature predictions that do not include information about atmospheric CO_2 concentration (Original Prediction) and the predictions from the regression model with ARMA errors. It is seen that continuing the current increase in global CO_2 emissions will cause the temperature to increase even faster than in the past. Furthermore, halving the increase in global CO_2 emissions will not limit the increase in the atmospheric CO_2 concentration enough in order to limit the temperature increase; the result will still be that the temperature will rise faster compared to the past.

In the beginning of this chapter, predictions for the temperature anomalies based on the development in the temperature from 1850 to 2018 were made. Based on those predictions, it was concluded that the temperature will have increased 1.5°C compared to pre-industrial levels in 2058 and 2°C in 2083. However, since including information about global CO_2 emissions and atmospheric CO_2 concentration causes the temperature predictions to change, it is likely that the temperature will have increased 1.5°C and 2°C before these years, if global CO_2 emissions are not limited.

The predictions from the regression model with ARMA errors indicate that limiting or decrease the global CO_2 emissions will have a positive effect on temperature changes.

GDP per Capita

Fitting a regression model on the first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies with temperature anomalies as exogenous variable and fitting ARMA models to the residuals yields the following results, where $\{x_t\}$ is the first principal component for the four economy groups, $\{y_t\}$ is temperature anomalies, and $\{u_t\}$ is white noise:

Low-Income Economies:

$$x_t = -3.913 + 16.738 y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows an ARMA$(1,1)$ process given by

$$\varepsilon_t = 0.901\varepsilon_{t-1} + u_t - 0.347 y_{t-1}.$$

Lower-Middle-Income Economies:

$$x_t = -4.956 + 21.094 y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows and AR(1) process given by

$$\varepsilon_t = 0.923\varepsilon_{t-1} + u_t.$$

Upper-Middle-Income Economies:

$$x_t = -5.163 + 21.855 y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows an ARMA$(1,1)$ process given by

$$\varepsilon_t = 0.945\varepsilon_{t-1} + u_t - 0.247 u_{t-1}.$$

High-Income Economies:

$$x_t = -7.075 + 29.941 y_t + \varepsilon_t,$$

where $\{\varepsilon_t\}$ follows an ARMA$(1,1)$ process given by

$$\varepsilon_t = 0.914\varepsilon_{t-1} + u_t - 0.367 u_{t-1}.$$

As expected the intercept decreases from the low-income economies to the high-income economies. Furthermore, the slope parameters associated with the temperature anomalies are all positive and increase from the low-income economies to the high-income economies. In all four groups except the lower-middle-income economies, where the residuals are modelled using an AR(1) model, the residuals from the regression models are modelled using an ARMA$(1,1)$ model.

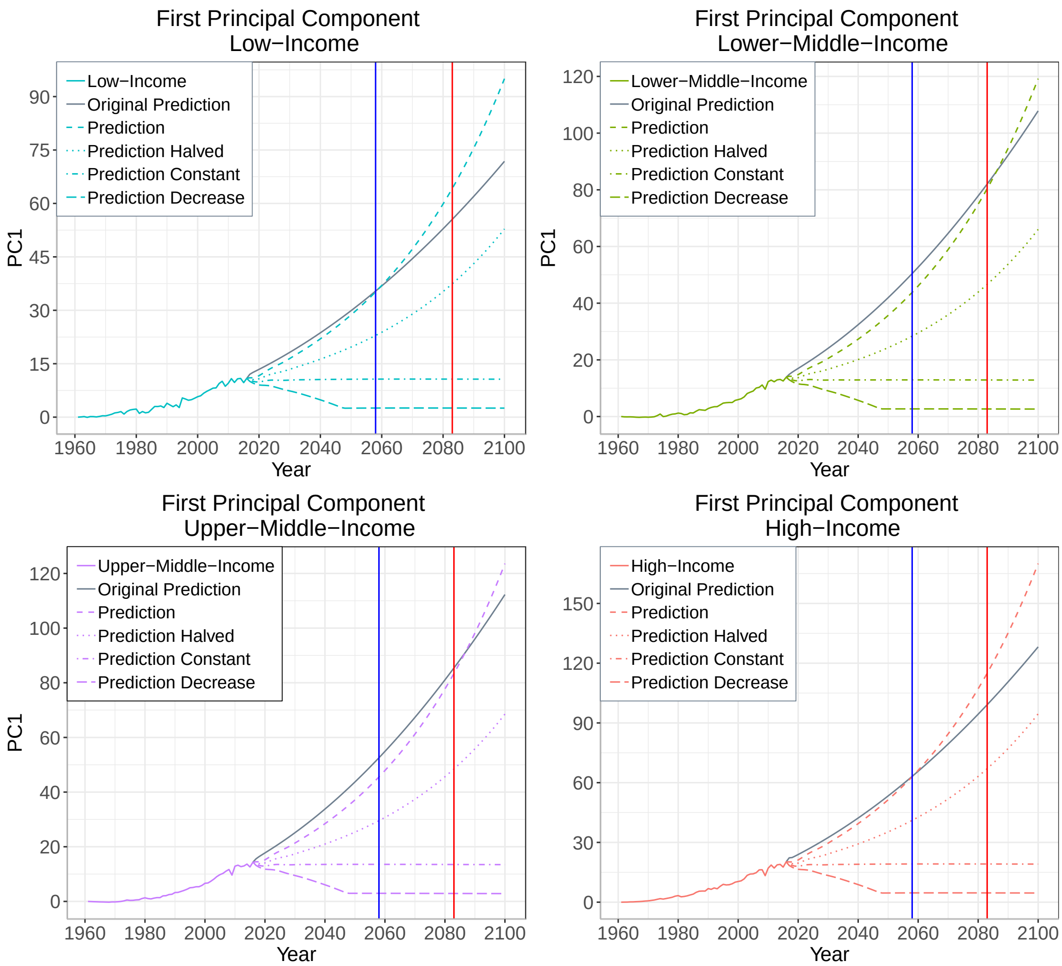


Figure 5.8: Stationary first principal components for the low-income, lower-middle-income, upper-middle-income, and high-income economies with time trend (coloured solid lines) from 1961 to 2016 and first principal component predictions from Figure 5.4 based on the ARMA models fitted in Chapter 4 with added time trend (grey solid lines) from 2017 to 2100. Prediction (coloured dashed lines), Prediction Halved (coloured dotted lines), Prediction Constant (coloured dashed and dotted lines), and Prediction Constant (coloured long-dashed lines) illustrate the predictions made by using the temperature anomaly prediction paths in the regression models with ARMA errors. All predictions are from 2017 to 2100.

The predictions for the stationary first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies with time trends can be seen in Figure 5.8. The figure both illustrates the predictions that are based on the ARMA models fitted in Chapter 4 in the space of the stationary data with time trend (grey solid lines) and the predictions based on the regression models with ARMA errors using the four prediction paths for the temperature anomalies showed in Figure 5.7 that are based on the atmospheric CO_2 concentration and global CO_2 emissions.

In the figure it is seen that the predictions from the regression models with ARMA errors are fairly close to the predictions based only on the first principal component time series for all four groups. The figure suggests that limiting the increase in global CO_2 emissions would limit the increase in the development in the GDP per capita. This makes sense, since growth in a country is associated with the level of CO_2 emissions: Economic growth is defined by the growth of the goods and services produced by an economy over time. Producing more goods will often increase the CO_2 emissions. Moreover, as an economy produces more goods its demand for energy increases, and it therefore needs to increase its energy supply, also leading to an increase in CO_2 emissions, since much of the energy in many countries comes from burning fossil fuels. Furthermore, in some countries, the development is driven by deforestation of forests or jungles in order to increase agriculture. The combustion or burning of the harvested trees is a large contributor to CO_2 emissions, and since trees absorb CO_2 in their photosynthesis, they help remove CO_2 from the atmosphere. As forests are harvested, there are less tress to absorb CO_2 leading to a higher atmospheric CO_2 concentration.

Thus, the regression models with ARMA errors for the first principal components for the GDP per capita captures the positive relationship between CO_2 emissions and economic growth, and not the negative effect that climate change has on some countries. This indicates that the regression model with ARMA errors only including temperature anomalies is too simple, and that other variables should be included in order to understand the effect of temperature increase on growth in GDP per capita. Such variables could be variables measuring the consequences of climate change, such as sea level rise and the cost of preventing and repairing the damage that follows from this, the cost and number of extreme events, such as severe droughts, extreme precipitation, and more frequent and larger storms, expenses following from climate migration, changes in precipitation affecting agriculture, and so on.

Figure 5.9 illustrates the same data and predictions as Figure 5.8, however, instead of illustrating each economy group separately with all predictions, Figure 5.9 illustrates each prediction path separately for all four groups. The four prediction paths illustrated in the four subplots are the predictions from the regression models with

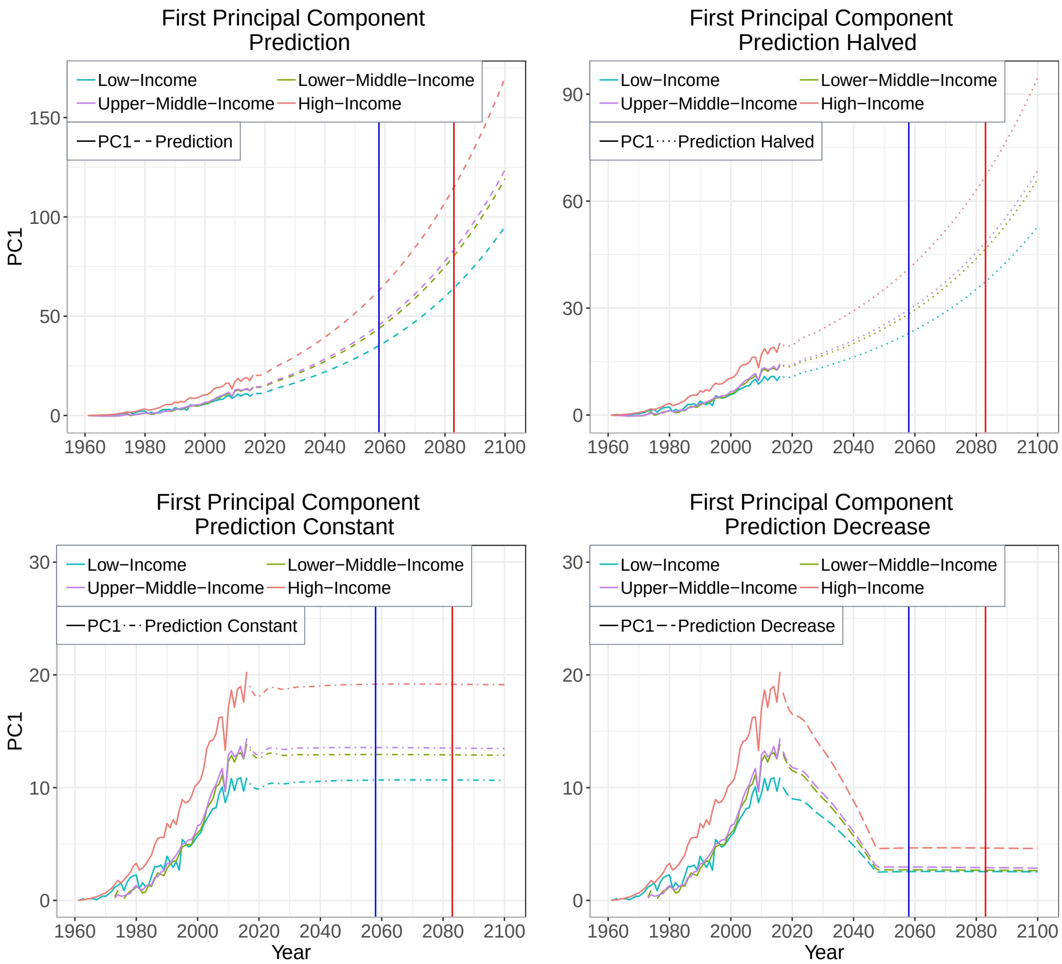


Figure 5.9: Stationary first principal components for the low-income, lower-middle-income, upper-middle-income, and high-income economies with time trend (coloured solid lines) from 1961 to 2016. Prediction (coloured dashed lines), Prediction Halved (coloured dotted lines), Prediction Constant (coloured dashed and dotted lines), and Prediction Constant (coloured long-dashed lines) illustrate the predictions made by using the temperature anomaly prediction paths in the regression models with ARMA errors. All predictions are from 2017 to 2100.

ARMA errors using the four prediction paths for the temperature anomalies in Figure 5.7.

Figure 5.9 indicates that the gap between the development in GDP per capita for different economy groups should become smaller as global CO_2 emissions are limited. This makes sense since many of the countries in the high-income, upper-middle-income, and lower-middle-income economy groups are the ones that are emitting most of the global CO_2, and being forced to reduce the CO_2 emissions would have a larger effect on these countries than the countries in the low-income group, which would decrease the gap between the development in the four economy groups.

The figure also shows how continuing the increase in global CO_2 emissions as in the past will make the gap between development in the four groups much larger in the future, meaning that the global inequality will increase. The Paris Agreement has high emphasis on this inequality, urging developed countries to support developing countries in the implementation of the Paris Agreement. However, the model does not take into account that climate change does not have the same effect on all countries, and that some countries may become uninhabitable in the future due to severe heat events and droughts. This is to a great extend the case for many countries in Africa that are close to the equator.[1] As seen in Figure 2.5, the low-income economy group contains most of the countries in the middle of Africa close to the equator, meaning that many of the people living in these countries may become climate migrants which will affect both the countries they come from and the countries that they will migrate to. This means that the gap between high-income and low-income economies might become even larger than Figure 5.9 indicates.

Furthermore, it has been shown that the consequences of temperature changes due to increased greenhouse gas emissions have increased growth in cool countries, which are mostly high-income economies, and decreased growth in warm countries, which are mostly low-income economies, in the period from 1961 to 2010.[2] Thus, temperature changes have already increased the inequality between low-income and high-income economies. This is also apparent from the parameters in the regression models with ARMA errors fitted in this section: The effect of temperature increases on development in GDP per capita in high-income economies is almost twice as large as the effect on low-income economies. This makes the discussion about high-income economies aiding low-income economies in implementing climate friendly initiatives and fulfilling the Paris Agreement even more relevant, since higher-income economies are responsible for a large part of the global CO_2 emissions that are negatively affecting the growth in low-income economies.

Even though many countries have begun the procedure of phasing out the use of fossil fuels for producing energy, and use renewable energy instead, it is probably infeasible to continue the same growth without increasing the CO_2 emissions in the near future. And since the current level of global CO_2 emissions is affecting the global temperature and causing climate changes which will have serious consequences affecting many economies, such as climate migration, extreme events,

[1] [Torelli, 2017]
[2] [Diffenbaugh and Burke, 2019]

and precipitation changes affecting agriculture, the economic growth of the past is unlikely to be sustainable in the future.

However, some argue that it is possible to obtain the same economic growth in a climate friendly way as the growth that would be obtained by business as usual. This is due to the present economic models not being adequate in relation to the consequences of climate changes. There are many climate friendly business opportunities in areas such as rapid technological innovation, sustainable infrastructure investment, increased resource productivity, energy, cities, food and land use, water, and industry. Furthermore, the present economic models do not take into account the expenses following from continuing to emit CO_2 in order to grow the economy: Disasters triggered by weather- and climate-related hazards were responsible for thousands of deaths and billion dollar losses in 2017. Climate change will lead to more frequent and more extreme events, including floods, droughts, and heat waves. Nor do the economic models include the positive effects of new technological advances, preservation of essential natural capital, and the full health benefits of cleaner air and a safer climate.[3]

To sum up, most of the models applied in this paper are too simple to capture the relationship between CO_2, temperature changes, and GDP per capita, and they should include more known information about the causes and consequences of climate change.

[3] [The New Climate Economy, 2019]

Chapter 6

Summary

After the industrial revolution, the amount of global CO_2 emissions begins to increase, and since approximately 1960 it has increased explosively. This has caused the atmospheric CO_2 concentration to increase, and the atmospheric CO_2 concentration is now higher than it has been in the past 400,000 years. Increased atmospheric greenhouse gas concentration has caused the planet's average surface temperature to rise: The temperature anomalies, which are relative to a 1961-1990 reference period, have increased approximately 1°C from 1850 to 2018.

Instead of examining each country's GDP per capita individually, the countries are divided into four groups: Low-income, lower-middle-income, upper-middle-income, and high-income economies, based on the World Bank Country Groups. The countries for which the GDP per capita data contains more than 20% missing values are excluded from the analysis and the remaining data are imputed.

Many of the countries' GDP per capita in all four groups seem to follow approximately the same growth trends. Furthermore, modelling and making predictions for the GDP per capita for all countries is a lengthy process: Therefore, principal component analysis is used on the countries' GDP per capita in each group. Based on the percentage of variance explained by each principal component, only the first principal component of each group is used in the analysis in this paper.

Time series analysis is used to model and make predictions for the time series in this paper. None of the time series seem stationary. Therefore, their time trend is removed and the residuals for each time series is examined. The residuals for the atmospheric CO_2 concentration and temperature anomalies show long memory, so ARFIMA models are fitted to them with differencing parameter $d = 0.49$. The residuals for the global CO_2 emissions and first principal components are fitted using ARIMA models with $d = 1$.

From predictions from these models the following is concluded: The temperature has increased 1.5°C and 2°C compared to pre-industrial levels in 2058 and 2083, respectively. The global CO_2 emissions continue to increase rapidly and by 2058 and 2083 the global CO_2 emissions will have reached approximately 410 and 669 billion tonnes, respectively. The atmospheric CO_2 concentration also continues to increase, and in 2058 and 2083 it will be approximately 511 and 599 PPM, respectively. The first principal component for all four groups continue their upwards-going trend, however, the rate of increase seems to be higher for the high-income, upper-middle-income, and lower-middle-income groups and lower for the low-income group, meaning that the difference between higher-income and low-income economies will continue to grow.

In order to examine the effect of one time series on another, regression models with ARMA errors are used on the differenced or fractionally differenced residuals where the time trend of each time series has been added again, and in order to determine the effect of global CO_2 emissions on temperature changes, different prediction paths for global CO_2 emissions are constructed.

Predictions of the atmospheric CO_2 concentration including information about global CO_2 emissions indicate that continuing to increase the global CO_2 emissions will cause the atmospheric CO_2 concentration to increase at an even higher rate than in the past. However, as the model does not capture some key characteristics of CO_2 accumulation in the atmosphere, it is likely too simple.

Predictions of temperature anomalies including information about atmospheric CO_2 concentration indicate that the temperature will increase even faster than in the past. Thus, it is possible that the temperature will increase 1.5°C and 2°C compared to pre-industrial levels before 2058 and 2083, respectively.

Predictions of the first principal component for low-income, lower-middle-income, upper-middle-income, and high-income economies including information about temperature anomalies indicate that limiting the increase in global CO_2 emissions would limit the increase in the GDP per capita. This makes sense, since economic growth is defined by the growth of the goods and services produced by an economy over time, and producing more goods will often increase the CO_2 emissions, both due to increased energy demand and the possible need for deforestation. Thus, the model captures the positive relationship between CO_2 emissions and economic growth, and not the negative effect that climate change has on some countries. This indicates that the model is too simple, and that other variables should be included in order to understand the effect of temperature increase on growth in GDP per capita.

The predictions also indicate that continuing the increase in global CO_2 emissions as in the past will make the gap between development in the four groups larger in the future, meaning that the global inequality will increase. However, the model does not take into account that climate change does not have the same effect on all countries, and that some countries may become uninhabitable in the future due to severe heat and droughts. This is the case for many countries in the low-income economy group meaning that the gap between high-income and low-income economies might become even larger.

Since the current level of global CO_2 emissions is affecting the global temperature and causing climate changes which will have serious consequences affecting many economies, the economic growth of the past is unlikely to be sustainable in the future. However, as the present economic models do not take into account the expenses of climate change following from continuing to emit CO_2 in order to grow the economy, or the many climate friendly business opportunities and positive effects of a safer climate, some argue that it is possible to obtain the same economic growth in a climate friendly way as the growth that would be obtained by business as usual.

To sum up, several of the models applied in this paper are too simple to capture the true relationship between CO_2, temperature changes, and GDP per capita, and they should include more known information about the causes and consequences of climate change.

Chapter 7

Discussion

This paper uses long memory models in order to model the atmospheric CO_2 concentration and the temperature anomalies; however, it does not consider the reasons for the data showing long memory. One could examine each time series in greater detail in order to determine, what has caused the data to show long memory, such as structural breaks, for example in global CO_2 emissions and atmospheric CO_2 concentration caused by the industrial revolution, or aggregations, for example in global temperature anomalies, which are compiled by aggregating several temperature anomaly time series from around the globe.

This paper chooses to divide countries into four groups based on the World Bank Country Groups: Low-income, lower-middle-income, upper-middle-income, and high-income economies. This division is based purely on the level of GDP per capita. Instead of using this method for dividing the countries into groups, one could have chosen a number of different divisions, for example division based on the level of development in each country, division based on geographic location, or a mixture of these.

Moreover, since it has been shown that the northern and southern hemisphere are affected differently by climate change, one could examine southern hemisphere temperature anomalies and northern hemisphere temperature anomalies instead of the global temperature anomalies, and combine this with a division of countries based on geographic location.[1]

Regression models with ARMA errors are used to examine the relationship between time series. The model assumes that the explanatory time series are exogenous: However, one would expect that the time series in this paper will have an effect on each other; therefore, one could instead create a system between the time

[1] [Freedman, 2013]

series, and examine how global CO_2 emissions and atmospheric CO_2 concentration affect temperature anomalies, how temperature anomalies affect GDP per capita, and how the development in GDP per capita affects global CO_2 emissions, etc. Such a system, shown in Figure 7.1, could also give an indication of the effect of including climate change in a country's fiscal policies, for example by introducing a tax on CO_2 emissions.

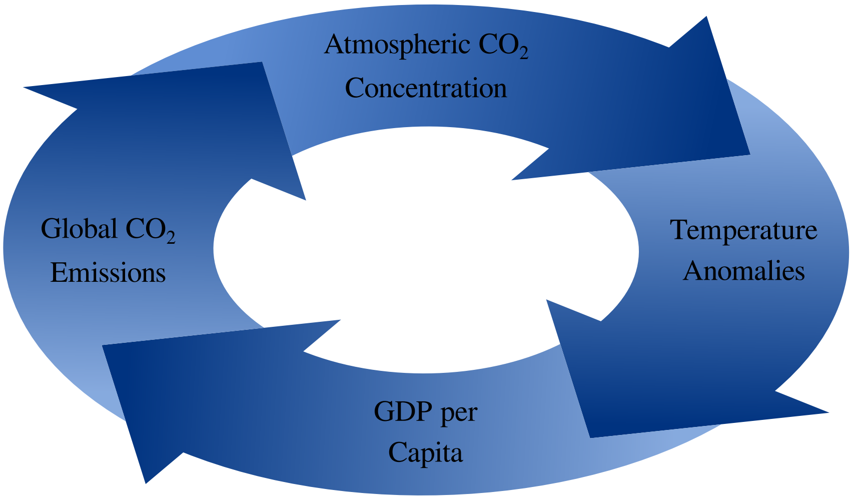


Figure 7.1: System between global CO_2 emissions, atmospheric CO_2 concentration, temperature anomalies, and GDP per capital

The analysis in this paper indicates that the regression model with ARMA errors of the atmospheric CO_2 concentration on global CO_2 emissions is too simple, since a change in global CO_2 emissions has an instant effect on atmospheric CO_2 concentration according to the model. However, since CO_2 persists in the atmosphere for 50 to 200 years, atmospheric CO_2 concentration will continue to increase even though the global CO_2 emissions are stabilised.[2] One solution could be to add more lags of global CO_2 emissions to the regression model with ARMA errors.

Moreover, the analysis in this paper indicates that the regression model with ARMA errors for the first principal component for GDP per capita that only includes temperature anomalies as exogenous variables is too simple to capture the effect of temperature changes on GDP per capita. There are many things affecting the climate and the economy, such as precipitation, natural disasters, extreme events, climate migrants, macroeconomic variables, etc., which should be included in the model in order to understand the true effect of climate change on different economies.

[2] [NASA, 2019c] and [Ritchie and Roser, 2019]

Bibliography

Diffenbaugh, N. S. and Burke, M. (2019). Global warming has increased global economic inequality. URL: https://www.pnas.org/content/116/20/9808.

Fan, J. and Yao, Q. (2017). *The Elements of Financial Econometrics*. Cambridge University Press, second edition.

Freedman, A. (2013). In Warming, Northern Hemisphere is Outpacing the South. URL: https://www.climatecentral.org/news/in-global-warming-northern-hemisphere-is-outpacing-the-south-15850.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*. John Wiley & Sons, INC., second edition.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer Texts in Statistics. Springer, first edition.

Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer, second edition.

Lankham, I., Nachtergaele, B., and Schilling, A. (2016). Linear Algebra as an Introduction to Abstract Mathematics: Lecture Notes for MAT67. URL: https://www.math.ucdavis.edu/~anne/linear_algebra/mat67_course_notes.pdf.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *Journal of Geophysical Research*, 117. D08101, doi:10.1029/2011JD017187.

Moritz, S. (2018). Package 'imputeTS'. URL: https://cran.r-project.org/web/packages/imputeTS/imputeTS.pdf.

NASA (2019a). A blanket around the Earth. URL: https://climate.nasa.gov/causes/.

NASA (2019b). Climate change: How do we know? URL: https://climate.nasa.gov/evidence/.

NASA (2019c). Graphic: The relentless rise of carbon dioxide. URL: https://climate.nasa.gov/climate_resources/24/graphic-the-relentless-rise-of-carbon-dioxide/.

NASA (2019d). How climate is changing. URL: https://climate.nasa.gov/effects/.

Nielsen, M. and Frederiksen, P. H. (2005). Finite Sample Comparison of Parametric, Semiparametric, and Wavelet Estimators of Fractional Integration. *Econometric Reviews*, 24(4). URL: https://dx.doi.org/10.1080/07474930500405790.

Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. Springer, second edition.

Ritchie, H. and Roser, M. (2019). CO_2 and other Greenhouse Gas Emissions. *Our World in Data*. URL: https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions.

Ruppert, D. (2004). *Statistics and Finance: An Introduction*. Springer.

Ruppert, D. and Matteson, D. S. (2015). *Statistics and Data Analysis for Financial Engineering with R examples*. Springer, second edition.

Shumway, R. H. and Stoffer, D. S. (2011). *Time Series Analysis and Its Applications*. Springer, third edition.

Sibbertsen, P. (2004). Long memory versus structural breaks: An overview. *Statistical Papers*, 45(4). URL: https://doi.org/10.1007/BF02760564.

The New Climate Economy (2019). Unlocking the Inclusive Growth Story of the 21st Century: Accelerating Climate Action in Urgent Times. URL: https://newclimateeconomy.report/2018/key-findings/.

The World Bank (2019a). Gdp per capita (current us&). URL: https://databank.worldbank.org/data/indicator/NY.GDP.PCAP.CD/1ff4a498/Popular-Indicators.

The World Bank (2019b). World Bank Country and Lending Groups. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

Torelli, S. M. (2017). Climate-driven migration in Africa. URL: https://www.ecfr.eu/article/commentary_climate_driven_migration_in_africa.

United Nations (2015). Paris Agreement. URL: https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

United Nations (2019). What is the Paris Agreement? URL: https://unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement.

Wikipedia (2019a). 1973 oil crisis. URL: https://en.wikipedia.org/wiki/1973_oil_crisis.

Wikipedia (2019b). Economy of Burundi. URL: https://en.wikipedia.org/wiki/Economy_of_Burundi.

Wikipedia (2019c). Economy of Syria. URL: https://en.wikipedia.org/wiki/Economy_of_Syria.

Wikipedia (2019d). Equatorial Guinea. URL: https://en.wikipedia.org/wiki/Equatorial_Guinea.

Wikipedia (2019e). Financial crisis of 2007–2008. URL: https://en.wikipedia.org/wiki/Financial_crisis_of_2007%E2%80%932008.

Wikipedia (2019f). Industrial Revolution. URL: https://en.wikipedia.org/wiki/Industrial_Revolution.

Appendix A

GDP per Capita Data

Country	Economy Class	NA's (%)
Afghanistan	Low-Income	34.48
Albania	Upper-Middle-Income	41.38
Algeria	Upper-Middle-Income	0.00
American Samoa	Upper-Middle-Income	72.41
Andorra	High-Income	17.24
Angola	Lower-Middle-Income	34.48
Antigua and Barbuda	High-Income	29.31
Argentina	High-Income	3.45
Armenia	Upper-Middle-Income	51.72
Aruba	High-Income	44.83
Australia	High-Income	0.00
Austria	High-Income	0.00
Azerbaijan	Upper-Middle-Income	51.72
The Bahamas	High-Income	0.00
Bahrain	High-Income	34.48
Bangladesh	Lower-Middle-Income	0.00
Barbados	High-Income	24.14
Belarus	Upper-Middle-Income	51.72
Belgium	High-Income	0.00

Table A.1: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Country	Economy Class	NA's (%)
Belize	Upper-Middle-Income	0.00
Benin	Low-Income	0.00
Bermuda	High-Income	6.90
Bhutan	Lower-Middle-Income	34.48
Bolivia	Lower-Middle-Income	0.00
Bosnia and Herzegovina	Upper-Middle-Income	58.62
Botswana	Upper-Middle-Income	0.00
Brazil	Upper-Middle-Income	0.00
British Virgin Islands	High-Income	100.00
Brunei	High-Income	8.62
Bulgaria	Upper-Middle-Income	34.48
Burkina Faso	Low-Income	0.00
Burundi	Low-Income	0.00
Cabo Verde	Lower-Middle-Income	34.48
Cambodia	Lower-Middle-Income	31.03
Cameroon	Lower-Middle-Income	0.00
Canada	High-Income	0.00
Cayman Islands	High-Income	96.55
Central African Republic	Low-Income	0.00
Chad	Low-Income	0.00
Channel Islands	High-Income	82.76
Chile	High-Income	0.00
China	Upper-Middle-Income	0.00
Colombia	Upper-Middle-Income	0.00
Comoros	Low-Income	34.48
Democratic Republic of the Congo	Low-Income	0.00
Republic of the Congo	Lower-Middle-Income	0.00
Costa Rica	Upper-Middle-Income	0.00
Côte d'Ivoire	Lower-Middle-Income	0.00
Croatia	High-Income	60.34
Cuba	Upper-Middle-Income	17.24
Curacao	High-Income	100.00
Cyprus	High-Income	25.86
Czech Republic	High-Income	51.72
Denmark	High-Income	0.00
Djibouti	Lower-Middle-Income	44.83
Dominica	Upper-Middle-Income	29.31
The Dominican Republic	Upper-Middle-Income	0.00
Ecuador	Upper-Middle-Income	0.00
Egypt	Lower-Middle-Income	8.62

Table A.2: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Country	Economy Class	NA's (%)
El Salvador	Lower-Middle-Income	8.62
Equatorial Guinea	Upper-Middle-Income	6.90
Eritrea	Low-Income	65.52
Estonia	High-Income	60.34
Ethiopia	Low-Income	36.21
Faroe Islands	High-Income	68.97
Fiji	Upper-Middle-Income	0.00
Finland	High-Income	0.00
France	High-Income	0.00
French Polynesia	High-Income	37.93
Gabon	Upper-Middle-Income	0.00
Gambia	Low-Income	10.34
Georgia	Upper-Middle-Income	51.72
Germany	High-Income	17.24
Ghana	Lower-Middle-Income	0.00
Gibraltar	High-Income	100.00
Greece	High-Income	0.00
Greenland	High-Income	18.97
Grenada	Upper-Middle-Income	29.31
Guam	High-Income	72.41
Guatemala	Upper-Middle-Income	0.00
Guinea	Low-Income	44.83
Guinea-Bissau	Low-Income	17.24
Guyana	Upper-Middle-Income	0.00
Haiti	Low-Income	0.00
Honduras	Lower-Middle-Income	0.00
Hong Kong	High-Income	0.00
Hungary	High-Income	53.45
Iceland	High-Income	0.00
India	Lower-Middle-Income	0.00
Indonesia	Lower-Middle-Income	12.07
Iran	Upper-Middle-Income	3.45
Iraq	Upper-Middle-Income	27.59
Ireland	High-Income	0.00
Isle of Man	High-Income	62.07
Israel	High-Income	0.00
Italy	High-Income	0.00
Jamaica	Upper-Middle-Income	0.00
Japan	High-Income	0.00
Jordan	Upper-Middle-Income	8.62

Table A.3: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Country	Economy Class	NA's (%)
Kazakhstan	Upper-Middle-Income	51.72
Kenya	Lower-Middle-Income	0.00
Kiribati	Lower-Middle-Income	17.24
North Korea	Low-Income	100.00
South Korea	High-Income	0.00
Kosovo	Lower-Middle-Income	68.97
Kuwait	High-Income	13.79
Kyrgyz Republic	Lower-Middle-Income	51.72
Laos	Lower-Middle-Income	41.38
Latvia	High-Income	60.34
Lebanon	Upper-Middle-Income	48.28
Lesotho	Lower-Middle-Income	0.00
Liberia	Low-Income	68.97
Libya	Upper-Middle-Income	51.72
Liechtenstein	High-Income	18.97
Lithuania	High-Income	60.34
Luxembourg	High-Income	0.00
Macao	High-Income	37.93
North Macedonia	Upper-Middle-Income	51.72
Madagascar	Low-Income	0.00
Malawi	Low-Income	0.00
Malaysia	Upper-Middle-Income	0.00
Maldives	Upper-Middle-Income	34.48
Mali	Low-Income	12.07
Malta	High-Income	17.24
Marshall Islands	Upper-Middle-Income	36.21
Mauritania	Lower-Middle-Income	0.00
Mauritius	Upper-Middle-Income	27.59
Mexico	Upper-Middle-Income	0.00
Micronesia	Lower-Middle-Income	43.10
Moldova	Lower-Middle-Income	60.34
Monaco	High-Income	18.97
Mongolia	Lower-Middle-Income	36.21
Montenegro	Upper-Middle-Income	68.97
Morocco	Lower-Middle-Income	0.00
Mozambique	Low-Income	34.48
Myanmar	Lower-Middle-Income	68.97
Namibia	Upper-Middle-Income	34.48
Nauru	Upper-Middle-Income	81.03
Nepal	Low-Income	0.00

Table A.4: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Country	Economy Class	NA's (%)
Netherlands	High-Income	0.00
New Caledonia	High-Income	37.93
New Zealand	High-Income	0.00
Nicaragua	Lower-Middle-Income	0.00
Niger	Low-Income	0.00
Nigeria	Lower-Middle-Income	0.00
The Northern Mariana Islands	High-Income	72.41
Norway	High-Income	0.00
Oman	High-Income	8.62
Pakistan	Lower-Middle-Income	0.00
Palau	High-Income	68.97
Panama	High-Income	0.00
Papua New Guinea	Lower-Middle-Income	0.00
Paraguay	Upper-Middle-Income	8.62
Peru	Upper-Middle-Income	0.00
Philippines	Lower-Middle-Income	0.00
Poland	High-Income	51.72
Portugal	High-Income	0.00
Puerto Rico	High-Income	1.72
Qatar	High-Income	17.24
Romania	Upper-Middle-Income	46.55
Russia	Upper-Middle-Income	50.00
Rwanda	Low-Income	0.00
Samoa	Upper-Middle-Income	37.93
San Marino	High-Income	67.24
São Tomé and Príncipe	Lower-Middle-Income	70.69
Saudi Arabia	High-Income	13.79
Senegal	Low-Income	0.00
Serbia	Upper-Middle-Income	60.34
Seychelles	High-Income	0.00
Sierra Leone	Low-Income	0.00
Singapore	High-Income	0.00
Sint Maarten	High-Income	100.00
Slovakia	High-Income	51.72
Slovenia	High-Income	60.34
Solomon Islands	Lower-Middle-Income	13.79
Somalia	Low-Income	91.38
South Africa	Upper-Middle-Income	0.00
South Sudan	Low-Income	89.66
Spain	High-Income	0.00

Table A.5: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Country	Economy Class	NA's (%)
Sri Lanka	Lower-Middle-Income	0.00
Saint Kitts and Nevis	High-Income	0.00
Saint Lucia	Upper-Middle-Income	29.31
Saint Martin French part	High-Income	100.00
Saint Vincent and the Grenadines	Upper-Middle-Income	0.00
Sudan	Lower-Middle-Income	0.00
Suriname	Upper-Middle-Income	0.00
Swaziland	Lower-Middle-Income	0.00
Sweden	High-Income	0.00
Switzerland	High-Income	17.24
Syria	Low-Income	17.24
Tajikistan	Low-Income	51.72
Tanzania	Low-Income	48.28
Thailand	Upper-Middle-Income	0.00
Timor-Leste	Lower-Middle-Income	68.97
Togo	Low-Income	0.00
Tonga	Lower-Middle-Income	25.86
Trinidad and Tobago	High-Income	0.00
Tunisia	Lower-Middle-Income	8.62
Turkey	Upper-Middle-Income	0.00
Turkmenistan	Upper-Middle-Income	46.55
Turks and Caicos Islands	High-Income	100.00
Tuvalu	Upper-Middle-Income	51.72
Uganda	Low-Income	0.00
Ukraine	Lower-Middle-Income	48.28
United Arab Emirates	High-Income	25.86
United Kingdom	High-Income	0.00
United States	High-Income	0.00
Uruguay	High-Income	0.00
Uzbekistan	Lower-Middle-Income	51.72
Vanuatu	Lower-Middle-Income	32.76
Venezuela	Upper-Middle-Income	5.17
Vietnam	Lower-Middle-Income	43.10
The United States Virgin Islands	High-Income	17.24
West Bank and Gaza Strip	Lower-Middle-Income	58.62
Yemen	Low-Income	51.72
Zambia	Lower-Middle-Income	0.00
Zimbabwe	Low-Income	0.00

Table A.6: Countries used to construct the principal components of GDP per capita. The table includes the name of the country, the economy class, and the percentage of NA's in each country's GDP per capita. Countries with more than 20% missing values are excluded from the principal component analysis and their NA value is coloured red.

Appendix B

Principal Component Analysis

This appendix is based on [James et al., 2013, pp. 374-377, 380-384], [Jolliffe, 2002, pp. 1-6], and [Fan and Yao, 2017, pp. 275-277].

Principal Component Analysis (PCA) is a dimensionality reduction method that tries to present the original data in a lower-dimensional space, while keeping as much information about the original data as possible.

Let $\mathbf{X}$ be a centred $T \times d$ matrix consisting of T observations of d random variables and let $\mathbf{z}_i = \mathbf{X}\boldsymbol{\phi}_i$ be a linear combination of the columns of $\mathbf{X}$, where $\boldsymbol{\phi}_i$ is called the i'th loading vector. Then the sample variance of $\mathbf{z}_i$ is given by

$$\begin{aligned} \frac{1}{T}\mathbf{z}_i^T\mathbf{z}_i &= \frac{1}{T}(\mathbf{X}\boldsymbol{\phi}_i)^\top(\mathbf{X}\boldsymbol{\phi}_i) \\ &= \boldsymbol{\phi}_i^\top\left(\frac{1}{T}\mathbf{X}^\top\mathbf{X}\right)\boldsymbol{\phi}_i. \end{aligned} \tag{B.1}$$

The main idea of the i'th principal component is to maximise the sample variance of $\mathbf{z}_i$ with respect to $\boldsymbol{\phi}_i$ subject to the the constraint that $\|\boldsymbol{\phi}_i\| = 1$ and $\boldsymbol{\phi}_i^T\boldsymbol{\phi}_j = 0$ for $j = 1, \ldots, i-1$.

The maximisation problem for the first principal component is given by

$$\max_{\boldsymbol{\phi}}\left(\boldsymbol{\phi}^\top\left(\frac{1}{T}\mathbf{X}^\top\mathbf{X}\right)\boldsymbol{\phi}\right), \qquad \text{s.t. } \|\boldsymbol{\phi}\| = 1,$$

which is solved using the Lagrange function defined as

$$\mathcal{L}(\boldsymbol{\phi}, \lambda) = \boldsymbol{\phi}^\top\left(\frac{1}{T}\mathbf{X}^\top\mathbf{X}\right)\boldsymbol{\phi} + \lambda\left(\boldsymbol{\phi}^\top\boldsymbol{\phi} - 1\right).$$

The partial derivatives are then given by

$$\frac{\partial \mathcal{L}(\boldsymbol{\phi}, \lambda)}{\partial \lambda} = \boldsymbol{\phi}^\top \boldsymbol{\phi} - 1$$
$$\frac{\partial \mathcal{L}(\boldsymbol{\phi}, \lambda)}{\partial \boldsymbol{\phi}} = 2\boldsymbol{\phi}^\top \left(\frac{1}{T}\mathbf{X}^\top \mathbf{X}\right) + 2\lambda \boldsymbol{\phi}^\top,$$

and setting them equal to zero yields

$$\boldsymbol{\phi}^\top \boldsymbol{\phi} = 1 \tag{B.2}$$
$$\boldsymbol{\phi}^\top \left(\frac{1}{T}\mathbf{X}^\top \mathbf{X}\right) = \lambda \boldsymbol{\phi}^\top. \tag{B.3}$$

Equation (B.2) is one of the constraints of the optimisation problem, and Equation (B.3) states that $\boldsymbol{\phi}$ is an eigenvector of $\frac{1}{T}\mathbf{X}^\top \mathbf{X}$ with corresponding eigenvalue λ. The first loading vector is the eigenvector with the largest eigenvalue since substituting Equation (B.3) into the sample variance in Equation (B.1) yields

$$\begin{aligned} \boldsymbol{\phi}^\top \left(\frac{1}{T}\mathbf{X}^\top \mathbf{X}\right) \boldsymbol{\phi} &= \lambda \boldsymbol{\phi}^\top \boldsymbol{\phi} \\ &= \lambda. \end{aligned} \tag{B.4}$$

Since $\frac{1}{T}\mathbf{X}^\top \mathbf{X}$ is a symmetric and positive semidefinite matrix then the eigenvector basis is orthogonal[1] and there is at most $\min(T-1, d)$ principal components.[2] It follows from this and Equation (B.4) that the eigenvalues are non-negative, which ensures that the sample variance does not become negative.

Since any scalar multiplication of an eigenvector with corresponding eigenvalue λ is still an eigenvector with corresponding eigenvalue λ, the scalar multiplication of an eigenvector still satisfies Equation (B.3). This means that Equation (B.2) can be satisfied by scaling the eigenvector; that is, if $\|\boldsymbol{\phi}\| = c \neq 1$ then choose the eigenvector $\frac{1}{\sqrt{c}}\boldsymbol{\phi}$ as the loading vector.

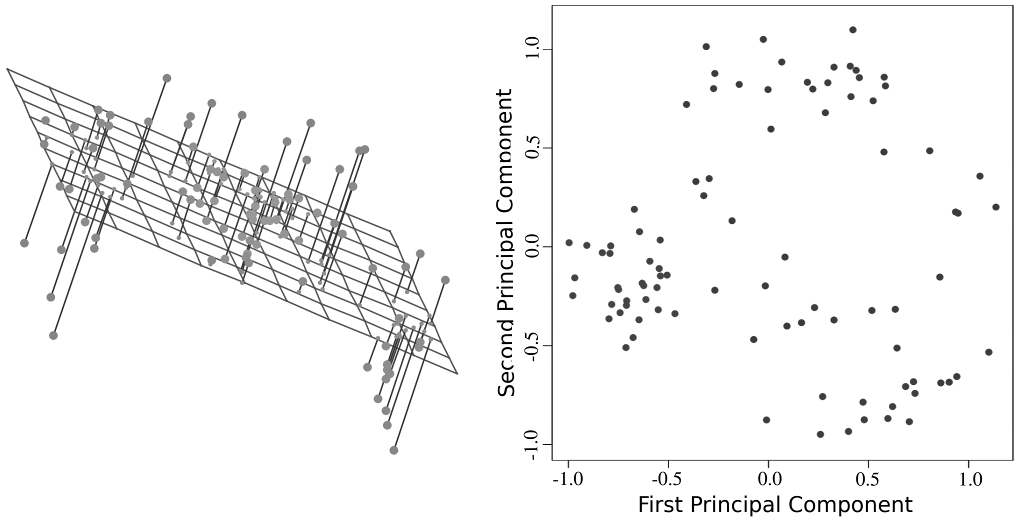
Thus, finding the i'th loading vector amounts to finding the scaled eigenvector with the i'th largest eigenvalue since this ensures that this i'th loading vector maximises the sample variance given that it has to be orthogonal to the previous $i-1$ loading vectors.

Geometric Interpretation and Proportion of Variance Explained

One way of visualising the projection of the original data onto the subspace spanned by the loading vectors is as follows: The k-dimensional hyperplane spanned by the

[1]See [Lankham et al., 2016, p. 151].
[2]See [James et al., 2013, p. 377].

k first loading vectors provides the best k-dimensional approximation of the original data in terms of the squared Euclidean distance.[3] Thus, the loading vector corresponding to the first principal component defines the line in the d-dimensional space that is closest to the T observations, and the first principal component is then the vector containing the projected points rotated onto the real line.



(a) **(b)**

Figure B.1: Figure B.1(a) shows 90 simulated points together with the plane spanned by the loading vectors of the first two principal components. Figure B.1(b) shows the two first principal components plotted against each other. These figures are taken from [James et al., 2013, p. 380].

Figure B.1 illustrates with $d = 3$ how the first two principal components loading vectors span the two-dimensional hyperplane that minimises the Euclidian distance to the simulated points.

A common method for choosing the number of principal components is to measure the percentage of variance explained by each component. Let K be the maximum number of principal components and let $\lambda_1 > \cdots > \lambda_K$ be the ordered eigenvalues of $\frac{1}{T}\mathbf{X}^\top\mathbf{X}$. The variance of the k'th principal component is λ_k and thus the percentage of variance explained by the k'th principal component is

$$\text{PVE}_k = \frac{\lambda_k}{\sum\limits_{i=1}^{K} \lambda_i},$$

and the variance explained by the first k principal components is simply the cumulative sum of the percentage of variance explained by these principal component,

[3] See [James et al., 2013, p. 379].

that is,

$$\mathrm{CPVE}_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{K} \lambda_i}.$$

Thus, in order to determine the number of principal components to use one can plot CPVE_k for $k = 1, \ldots, K$ and choose to use the smallest number of principal components that are required in order to explain a sizeable amount of variation in the data.

Finally, it is worth mentioning that since it is undesirable for the principal components to depend on the scaling of the variables, one typically scales each variable to have mean zero and standard deviation one before performing PCA.

Appendix C

First Principal Components

Low-Income Economies Country	Loading Vector	Mean	Standard Deviation
Benin	0.238	384	250
Burkina Faso	0.242	283	186
Burundi	0.208	162	72
Central African Republic	0.199	282	131
Chad	0.222	341	284
Democratic Republic of the Congo	0.100	307	129
Gambia	0.171	399	231
Guinea-Bissau	0.228	268	184
Haiti	0.234	345	234
Madagascar	0.216	284	101
Malawi	0.229	192	115
Mali	0.238	309	247
Nepal	0.231	244	213
Niger	0.213	247	87
Rwanda	0.240	270	204
Senegal	0.237	740	340
Sierra Leone	0.217	264	141
Syria	0.220	1098	583
Togo	0.239	322	156
Uganda	0.226	274	175
Zimbabwe	0.186	662	297

Table C.1: The loading vector of the first principal component along with the means and standard deviations measured in thousand US dollars used for standardising the GDP per capita for the countries in the low-income economy group.

Lower-Middle-Income Economies Country	Loading Vector	Mean	Standard Deviation
Bangladesh	0.193	376	326
Bolivia	0.198	995	850
Cameroon	0.180	761	444
Republic of the Congo	0.186	1036	836
Côte d'Ivoire	0.180	781	397
Egypt	0.196	1033	960
El Salvador	0.194	1422	1145
Ghana	0.191	587	557
Honduras	0.193	1023	655
India	0.199	498	492
Indonesia	0.199	999	1142
Kenya	0.198	464	377
Kiribati	0.185	822	444
Lesotho	0.196	451	388
Mauritania	0.197	563	366
Morocco	0.197	1258	958
Nicaragua	0.193	838	559
Nigeria	0.185	913	867
Pakistan	0.200	497	404
Papua New Guinea	0.196	916	765
Philippines	0.200	966	814
Solomon Islands	0.181	890	566
Sri Lanka	0.196	960	1153
Sudan	0.190	660	664
Swaziland	0.192	1390	1160
Tunisia	0.191	1795	1357
Zambia	0.186	631	443

Table C.2: The loading vector of the first principal component along with the means and standard deviations measured in thousand US dollars used for standardising the GDP per capita for the countries in the lower-middle-income economy group.

Upper-Middle-Income Economies Country	Loading Vector	Mean	Standard Deviation
Algeria	0.185	2071	1543
Belize	0.184	2245	1610
Botswana	0.192	2574	2444
Brazil	0.191	3685	3615
China	0.182	1495	2423
Colombia	0.194	2272	2261
Costa Rica	0.193	3311	3345
Cuba	0.193	2812	2067
The Dominican Republic	0.194	2173	2062
Ecuador	0.193	2213	1742
Equatorial Guinea	0.176	4063	6926
Fiji	0.191	2110	1479
Gabon	0.179	4400	2879
Guatemala	0.193	1413	1110
Guyana	0.187	1195	1236
Iran	0.177	2530	1998
Jamaica	0.187	2323	1608
Jordan	0.191	1715	1162
Malaysia	0.195	3525	3343
Mexico	0.185	4215	3405
Paraguay	0.192	1790	1705
Peru	0.193	2038	1891
South Africa	0.188	3068	2069
Saint Vincent and the Grenadines	0.189	2623	2397
Suriname	0.186	2788	2638
Thailand	0.193	1939	1949
Turkey	0.194	3784	3885
Venezuela	0.191	4589	4171

Table C.3: The loading vector of the first principal component along with the means and standard deviations measured in thousand US dollars used for standardising the GDP per capita for the countries in the upper-middle-income economy group.

High-Income Economies Country	Loading Vector	Mean	Standard Deviation
Andorra	0.145	18 540	14 555
Argentina	0.135	5081	3926
Australia	0.145	20 676	19 052
Austria	0.148	20 735	17 307
The Bahamas	0.140	14 269	11 191
Belgium	0.148	19 898	15 742
Bermuda	0.146	35 952	32 828
Brunei	0.137	16 192	12 644
Canada	0.148	20 395	15 797
Chile	0.143	4723	4759
Denmark	0.148	25 776	20 778
Finland	0.147	20 886	16 839
France	0.147	19 053	14 341
Germany	0.147	20 349	15 256
Greece	0.142	10 493	8949
Greenland	0.148	18 144	16 043
Hong Kong	0.144	15 693	14 118
Iceland	0.143	24 038	19 554
Ireland	0.145	20 871	21 934
Israel	0.146	13 933	11 630
Italy	0.145	16 279	12 948
Japan	0.135	21 589	16 462
South Korea	0.147	8819	9394
Kuwait	0.136	18 062	14 081
Lichtenstein	0.148	58 892	56 046
Luxembourg	0.148	41 361	39 116
Malta	0.148	8678	8200
Monaco	0.148	75 407	55 053
Netherlands	0.148	21 865	17 819
New Zealand	0.146	15 003	12 946
Norway	0.147	33 540	31 838
Oman	0.145	7185	6588
Panama	0.135	3954	3730
Portugal	0.147	8948	8156
Puerto Rico	0.147	11 309	9921
Qatar	0.138	28 284	25 935
Saudi Arabia	0.130	9529	7132
Seychelles	0.146	5443	4876
Singapore	0.146	17 893	18 572
Spain	0.147	12 590	11 078
Saint Kitts and Nevis	0.147	6269	6009
Sweden	0.148	24 885	18 574
Switzerland	0.147	33 853	27 313
Trinidad and Tobago	0.141	6746	6197
United Kingdom	0.146	19 208	16 042
United States	0.147	24 785	17 998
Uruguay	0.137	4892	4825
The United States Virgin Islands	0.144	17 546	14 320

Table C.4: The loading vector of the first principal component along with the means and standard deviations measured in thousand US dollars used for standardising the GDP per capita for the countries in the high-income economy group.