CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Bachelor thesis

2019 **Martin Moldan**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Martin Moldan

Relationship of Economic Growth and Pollution in the Czech Republic

Bachelor thesis

Prague 2019

Author: Martin Moldan

Supervisor: RNDr. Michal Červinka, Ph.D.

Academic Year: 2018/2019

Bibliographic note

MOLDAN, Martin. *Relationship of Economic Growth and Pollution in the Czech Republic*. Prague, 2019. 46 pp. Bachelor Thesis (Bc.) Charles University, Faculty of Social Sciences, Institute of Economic Studies. Thesis supervisor RNDr. Michal Červinka, Ph.D.

Abstract

The Environmental Kuznets Curve (EKC) is a hypothesized relationship between GDP per capita and pollution. It suggests that the relationship has a shape of a concave quadratic function—i.e. that firstly, with increasing GDP per capita, levels of pollution increase. And then, from some level of GDP per capita, as GDP per capita rises, levels of pollution decrease. This bachelor thesis examines whether the EKC holds for the Czech Republic or not. It uses panel data on air pollution for the period 1995–2017, in particular concentrations SO_2 and NO_x. This analysis is conducted using the fixed effects method.

The results of this bachelor thesis suggest that for the case of SO_2, there is a relationship between GDP per capita and the pollutant's concentrations. However, this relationship does not change over time significantly. Moreover, for the case of NO_x, the relationship between the pollutant's concentrations and GDP per capita is not significant, hence, the EKC hypothesis can be rejected for both examined pollutants.

Keywords

Environmental Kuznets Curve, air pollution, Czech Republic, economic growth, panel data

Abstrakt

Environmentální Kuznetsova křivka (EKC) je hypotetický vztah mezi HDP na obyvatele a znečištěním. Podle něj může být tento vztah graficky znázorněn jako konkávní kvadratická funkce—tedy, že nejprve s rostoucím HDP na obyvatele roste i míra znečištění. Následně, od určité hodnoty HDP na obyvatele, s rostoucím HDP na obyvatele míra znečištění klesá. Tato bakalářská práce zkoumá, zda hypotéza EKC platí pro Českou republiku. Práce využívá panelových dat znečištění ovzduší pro období 1995–2017. Konkrétně byla použita data koncentrací SO_2 a NO_x. Analýza panelových dat byla uskutečněna pomocí metody pevného efektu.

Výsledky této bakalářské práce ukazují, že pro SO_2 vztah mezi HDP na obyvatele a koncentracemi tohoto polutantu existuje. Tento vztah se ovšem v čase nijak významně nemění. Z tohoto důvodu nemůžeme hypotézu EKC pro tento polutant potvrdit. Výsledky analýzy dále ukazují, že pro NO_x neexistuje statisticky významný vztah mezi koncentrací tohoto polutantu a HDP na obyvatele. Výsledky této bakalářské práce tedy odmítají hypotézu EKC pro obě zkoumané znečišťující látky.

Klíčová slova

Environmentální Kuznetsova křivka, znečištění ovzduší, Česká republika, ekonomický růst, panelová data

Declaration of Authorship

I hereby proclaim that I wrote my bachelor thesis on my own under the leadership of my supervisor and that the references include all resources and literature I have used.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, 3 July 2019

Signature

Acknowledgment

I would like to express my appreciation and gratitude to RNDr. Michal Červinka, Ph.D. His knowledge and willingness to advise me at any occasion has helped me to a great extent in writing my thesis.

This thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 681228.

The Bachelor's Thesis Proposal

Author: Martin Moldan

Supervisor: RNDr. Michal Červinka Ph.D.

Topic: Relationship of Economic Growth and Pollution in the Czech Republic

---

Description

In this thesis we shall examine to what extent is pollution influenced by the economic growth for the case of the Czech Republic. A part of this phenomenon can be described by the so-called Kuznets curve. In this thesis, we shall determine whether Kuznets curve for the Czech Republic has similar properties as those known for other countries. The Kuznets curve hypothesis postulates an inverted U-shape relationship between GDP per capita and pollutant emissions or concentrations.

The term "pollution" is rather inexact, therefore we shall first choose a sample of pollutants for which data can be obtained and, at the same time, such that the sample represents the overall level of pollution sufficiently well. In the practical part of the thesis we shall use regression analysis to examine the relationship of economic growth and pollution in the Czech Republic. While constructing the regression analysis model, we shall be aware of other factors not included in the model that can cause our model to be inaccurate. One of such possible factors could be, e.g., new legislative regarding environmental protection.

Several authors suggest that economic growth results in increased environmental pollution while others argue otherwise. Many papers offer empirical test of the Kuznets curve hypothesis by analyzing the relationship between economic growth and level of certain pollutants, e.g. carbon dioxide emissions or concentrations. For the purposes of this thesis, we have chosen to analyze panel data of concentrations of NO_x and SO_2. To our knowledge there has not been any such study on this topic for the Czech

Republic. This thesis aims to provide such an analysis.

Outline

1. Introduction and literature review
2. Dataset description
3. Model specification
4. Empirical results
5. Conclusion

Main literature

1. Jeffery Wooldridge. *Introductory Econometrics: A Modern Approach*. Nelson Education, 2015.
2. Gene M. Grossman and Alan B. Krueger. “Environmental Impacts of a North American Free Trade Agreement”. *Technical report of National Bureau of Economic Research* (1991).
3. Simon Kuznets. “Economic growth and income inequality”. *The American economic review* 45.1 (1955), pp. 1–28.
4. Nemat Shafik. “Economic development and environmental quality: an econometric analysis”. *Oxford economic papers* 46.4 (1994), pp. 757–774.

Contents

1 Introduction **1**

2 Literature review **5**

3 Pollutants **9**

3.1 Description of pollutants 9

3.2 Choice of pollutants 11

4 Dataset description **13**

4.1 Pollution 13

4.2 GDP per capita 18

4.3 Population density 20

4.4 Unemployment 22

4.5 Policy variable 24

4.6 Share of mining in the economy 25

5 Model Specification **27**

5.1 Panel data analysis—econometric methods 27

5.2 Regression Model of Pollution Determinants 30

6 Empirical Results **34**

6.1 Testing of model assumptions 34

6.2 Estimation of models without year dummy variables 35

6.3 Estimation of models with year dummy variables 38

6.4 Models with lags 42

6.5 Further analysis of the SO_2 model 42

7 Conclusion **46**

References **47**

List of tables and figures **50**

1 Introduction

Air pollution is defined by the OECD as the existence of pollutant substances in the air that harm human health or welfare, or create other damaging environmental effects. The following history part is based on Brimblecombe and Makra 2005 and Brunekreef and Holgate 2002.

The first exposure of humans to air pollution is considered to be the smoke generated by burning various fuels to heat caves in which humans lived a couple of thousands years ago. The fact that their caves were not spacious enough for the pollution to disperse led to serious breathing problems of humans living there. On the other hand, the smoke protected them from mosquitoes.

During the ancient times, air pollution was almost exclusively problem of cities. Namely, people were concerned with smoke and unpleasant smells resulting from decaying organic waste or human feces. There were also first laws regarding air pollution introduced in some cities.

The most important milestone for the occurrence of air pollution is the Industrial Revolution. While bringing economic prosperity, the Industrial Revolution also led to a serious increase of air pollution. Most of the air pollution was generated by burning of coal, which was the essential source of energy on which the growing economy extensively relied. The degree of impact the pollution had on people's health and the public concerns associated with it led to the introduction of emission control regulations.

Another important event in the history of air pollution is the widespread of internal combustion engines (ICE) during the end of the 19$^{\text{th}}$ and the beginning of the 20$^{\text{th}}$ century. Today, regulations tackling air pollution are one of the most strict in the area of automobile emissions.

Both the Industrial Revolution and widespread car ownership is associated with an increase in wealth. Therefore, it is natural to think of a possible relationship between economic growth and air pollution.

There has been many attempts to study this hypothesized relationship. The most well-known theory regarding this topic developed from a study,

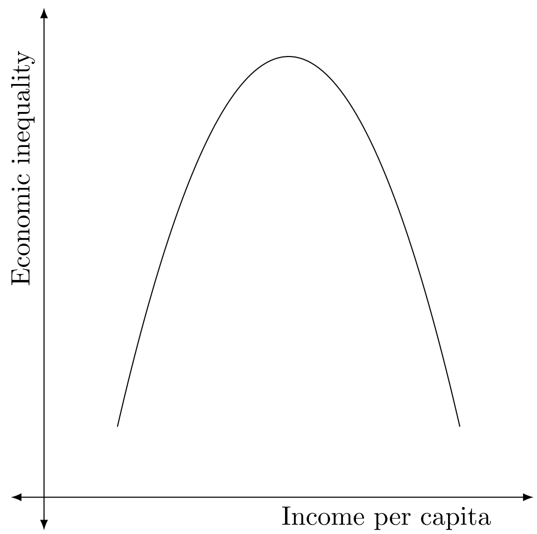

Figure 1: Kuznets curve

which was originally meant to study an entirely different topic.

This relationship is the so-called Kuznets curve, which was firstly mentioned in Kuznets 1955. It is a hypothesized relationship between income per capita and economic inequality in a given country. Graphically, the Kuznets curve is a quadratic concave function. This can be seen in Figure 1.

From the Kuznets curve the so-called Environmental Kuznets curve (EKC) developed. It was firstly mentioned in Grossman and Krueger 1991. The EKC shows a relationship between income per capita and the level of environmental degradation in a given country. Graphically, as in the case of the original Kuznets curve, the EKC has a concave shape. This can be seen in Figure 2.

Both income per capita and environmental degradation can be represented by a number of variables. Income per capita is often represented by GDP per capita—real, nominal or adjusted for purchasing power parity—PPP (e.g. Shafik 1994). Environmental degradation can be measured using various indicators. For instance, by deforestation (e.g. Koop and Tole 1999), water pollution (e.g. Torras and Boyce 1998), CO_2, NO, NO_2, SO_2 emissions (e.g. Cole, Rayner and Bates 1997) or CO_2, NO, NO_2, SO_2 concentrations—also called immissions (e.g. Torras and Boyce 1998). Usually, logarithm of environmental degradation appears in the model, as well as a quadratic form

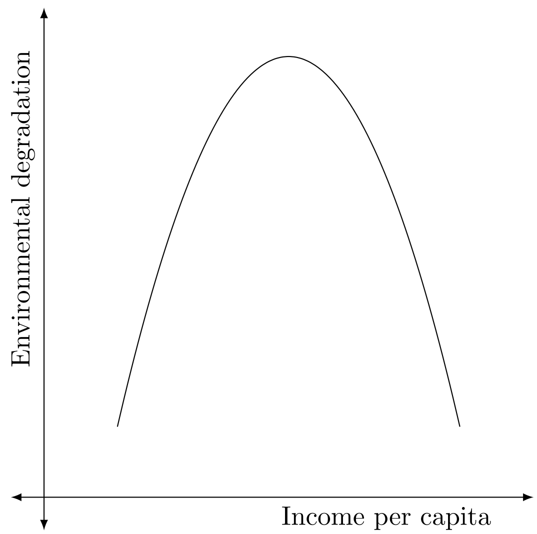

Figure 2: Environmental Kuznets curve

of a logarithm of income per capita (e.g. Wang et al. 2015). In most studies, other explanatory variables are added to the model, for instance location dummies (e.g. Shafik 1994), income inequality (e.g. Torras and Boyce 1998), population density (e.g. Grossman and Krueger 1991) or policy variables (e.g. Panayotou 1997).

The EKC assumes that in developing countries, the scale effect resulting from economic growth outweighs environmental degradation reduction efforts, while from some point of income per capita, the situation is the other way round. This can be supported by the fact that most policies, the goal of which is to address pollution problems, are taking place in developed countries, as is, for example, mentioned in Dasgupta et al. 2002.

In this bachelor thesis, we shall analyze air pollution data, economic growth data, as well as data for other variables with the potential to influence air pollution, in order to determine whether the EKC holds for the case of the Czech Republic or not.

In the first section, we provide an overview of literature studying the EKC topic and describe how our analysis further develops this area. In the second section, we introduce four of the most well-known substances used to measure air pollution. Thirdly, we introduce variables which will be included in our model and the methodology used to obtain data on them. In the fourth section, we describe econometric methods used in our analysis and introduce our model. In the last section, we present results of our analysis as well as the interpretation of them.

2 Literature review

As we have already mentioned, the EKC hypothesis was firstly formulated in Grossman and Krueger 1991. The purpose of their research paper was to assess the environmental impact of the North American Free Trade Agreement (NAFTA). Using a cross-sectional analysis, they concluded that a reduction in trade barriers will affect the environment by increasing the level of economic activity, by changing the composition of economic activity, and by changing the techniques of production. They concluded that for low levels of national income, concentrations of air pollutants increase with increasing GDP per capita, and that for higher levels of national income, concentration of air pollutants decrease with increasing GDP per capita. The turning point, where the EKC has its peak, was, according to their findings, \$4,772–5,965 per year. They added into their model variables for population density, trend and locational dummy variables.

Another influential research paper in this field is Panayotou 1993. Unlike Grossman and Krueger 1991, it studied air pollution by analyzing emissions of pollutants, rather than concentrations. Moreover, as a measure of environmental degradation, Panayotou also analyzed deforestation data. As well as Grossman and Krueger, Panayotou used a cross-sectional analysis. According to his findings, the EKC has its peak for income per capita of \$3,137.

Theodore Panayotou published another research paper on the EKC topic—Panayotou 1997. As well as in his 1993 study, he used a cross-sectional approach. But he made other significant changes in his methodology. Instead of analyzing emissions, he decided to analyze air pollutants' concentration data. Moreover, unlike in his previous study, Panayotou added into his model policy variables and a variable for population density. With this new approach, he concluded that the EKC turning point is significantly higher than in his previous research paper—\$5,965, which is the same as the upper bound of the interval found in Grossman and Krueger 1991.

Another influential study in this field is Shafik 1994. This study, unlike the two previously mentioned studies, used panel data, rather than cross-sectional. As well as in Grossman and Krueger 1991, there are pollutants' concentrations data used. Shafik concluded that the EKC turning point is $4,379. Shafik added to his model locational dummies and time trend. Unlike the two previously mentioned studies, Shafik used GDP PPP data.

Moreover, the EKC topic is also covered by Selden and Song 1994. Using pool cross-sectional data Selden and Daqing studied mainly emissions of air pollutants. According to this study, the turning point of the EKC is $10,391–10,620. Furthermore, they concluded that the turning point of the EKC tends to be higher in the case of emissions than in the case of concentrations. This is, according to their findings, because of the following reasons:

- Urban air quality is of a more immediate importance to city inhabitants and therefore to city policymakers than emissions.
- Improvements in urban air quality can be made at significantly lower costs than a reduction in emissions. For instance, requiring taller smokestacks would decrease air pollutants' concentrations, but would not reduce their emissions.
- Rising urban rents would cause factories, which are responsible for a significant proportion of pollution, to move out of the urban area. This, again, would only decrease air pollutants' concentrations and not their emissions.
- Urban residents tend to have higher income than is the national average, hence, they tend to have more political influence.

Their model included a variable for population density. As well as Shafik 1994, they used GDP PPP data.

The EKC concept was also studied for the case of Canada by Ward et al. 2016. They conducted their analysis based on time-series data. As a measurement of environmental degradation, they used emissions of CO_2. Their main goal was to analyze whether decoupling energy demand and economic growth is plausible. In their findings they suggest that the EKC relationship does not hold for the case of Canada and that GDP growth and pollution cannot be decoupled.

Furthermore, the EKC hypothesis is investigated by Andreoni and Levinson 2001. In their study they used an approach which does not take into consideration any other variables apart from abatement technologies. They studied this variable—abatement technologies and their costs—rather than analyzing immissions or concentrations. In their findings, they suggest that with increasing GDP abatement costs decline due to the economies of scale. This, according to them, consequently causes pollution levels to decrease.

There is also a few research papers the topic of which is the EKC which are focused on the Czech Republic, for example Brůha and Ščasný 2005. This research paper uses data of various air pollutants, including CO_2, SO_2 or NO_x collected during the 1992–2003 period. However, due to the lack of suitable data, they suggested the EKC examination for further research and focused mainly on the impact of regulation on pollution, for which they had necessary data. Their findings were that especially for emissions of SO_2, legislation has a significant effect.

Moreover, there is one research paper focused on the Czech Republic which tries to go beyond the year 1989. It is Kreuz, Lisa and Šauer 2017. This analysis was on the basis of 1975–2012 time series data of emissions of CO_2, NO_x, SO_2 and particulate matter (PM). However, they had to deal with various problems resulting from using data from the period before the year 1989. They only used data of emissions from stationary sources, which do not represent all emissions. Also, for all pollutants they used different periods, since for each of them data for different periods were available. Furthermore, GDP data had to be adjusted since before 1989 a different methodology was used. From their study, they concluded that the turning point for NO_x, SO_2 and PM was before 1989, but the exact year was not found, due to the lack of quality of the before 1989 data. The turning point for CO_2 was not found. Moreover, based on the after 2000 data, they stated that the validity of the whole EKC concept may be in question for the case of the Czech Republic.

This bachelor thesis aims to provide an analysis based on panel data, which has not been used in the case of the Czech Republic yet. Moreover, in the case of the Czech Republic, the previous analyses were based on emissions data. This thesis shall be, on the other hand, based on data of pollutants' concentrations.

3 Pollutants

3.1 Description of pollutants

In this section, we shall present the most well-known pollutants and what is the reasoning behind choosing those which will be included in our model. There are 4 main air pollutants: CO_2, SO_2, NO_x and particulate matter (PM). In this section, there are 4 subsections devoted to each of the previously mentioned air pollutants. There are certainly many other pollutants which can be used to measure air pollution (e.g. heavy metals or polycyclic aromatic hydrocarbons). They are, however, not mentioned often in the literature regarding the EKC topic and data on them with consistent methodology for a long enough period are, in the case of the Czech Republic, difficult to obtain. For these reasons, we shall not consider them for the purposes of our analysis. Moreover, there is one subsection at the end of this section which explains which of the air pollutants we choose for our regression analysis and the reasoning for that.

CO_2

Carbon dioxide (CO_2) is likely to be the most well-known pollutant to the general public. It is a greenhouse gas, which means that its higher concentrations lead to the greenhouse effect, which subsequently leads to global warming. As CO_2 is naturally present in our atmosphere, measuring CO_2 concentrations for the case of air pollution is rather problematic. Therefore, emission calculations are used in this case. Emissions of CO_2 influence the climate merely on a global scale—the local air quality is not affected. Therefore, policies which try to tackle CO_2 emissions mostly result from supranational policies. The main sources of CO_2 emissions are the energy sector, manufacturing and transportation. An example of a study regarding the EKC, which uses CO_2 data, is Grossman and Krueger 1991.

$\mathbf{SO_2}$

Sulfur dioxide (SO_2) is a toxic gas, which is mostly emitted during the process of burning fossil fuels containing a certain amount of sulfur—mostly brown coal (lignite). The main source of SO_2 emissions is burning of coal, natural gas and oil. Most of these emissions result from electricity generation. Smaller proportion of them is generated by other industrial processes. They can be measured either by calculating emissions, similarly as in the case of CO_2, or by directly measuring SO_2 concentrations in the air. According to the European Commission (EC), SO_2 concentration shall not exceed $350\mu g/m^3$ when calculating an hourly average and $125\mu g/m^3$ when calculating a daily average. If these SO_2 concentration limits are exceeded, it can cause health issues to the citizens exposed to it. Examples of studies regarding the EKC, which use SO_2 data, are Shafik and Bandyopadhyay 1992 or Selden and Song 1994.

$\mathbf{NO_x}$

NO_x is a generic term for two nitrogen oxides: nitric oxide (NO) and nitrogen dioxide (NO_2). NO is further oxidized to NO_2 rapidly in the air. These gases are mostly produced by fossil fuels car engines, especially diesel engine cars. A minor share of NO_x emissions is produced by burning coal in high temperatures. Similarly to SO_2, NO_x can be measured either by calculating emissions or by directly measuring NO_x concentrations in the air. According the the EC, NO_2 concentration should not exceed $200\mu g/m^3$ when calculating an hourly average and $40\mu g/m^3$ when calculating an annual average. The time periods for which the limits are set vary for the pollutants due to the fact that health problems associated with exceeded limits of concentrations occur after a different length of exposure to them. Higher NO_x concentrations, are, similarly to SO_2, likely to lead to respiration problems. Examples of studies regarding the EKC, which use NO_x data, are Lim 1997 or Selden and Song 1994.

PM

Particulate matter (PM) are liquid or solid microscopic matter dispersed in the atmosphere. As well as for the case of SO_2 or NO_x, PM can be measured either by calculation of their emissions, or by measuring their concentrations. There are 3 main subgroups of PM: $PM_{1.0}$, which is defined as particles the diameter of which is less than 1μ, $PM_{2.5}$, which is defined as particles the diameter of which is less than 2.5μ, and PM_{10}, which is defined as particles the diameter of which is less than 10μ. Generally speaking, the smaller PM are, the easier and deeper they can get into one's respiratory system. According to the EC, $PM_{2.5}$ concentration shall not exceed $25\mu g/m^3$ when calculating an annual average. Moreover, the concentration of PM_{10} should not exceed $50\mu g/m^3$ when calculating a daily average and $40\mu g/m^3$ when calculating an annual average. Examples of studies regarding the EKC, which use PM data, is Grossman and Krueger 1991.

3.2 Choice of pollutants

In the previous subsections, we have described the 4 most well-known pollutants. Our objective is to choose two of them, which will be included in our analysis. Firstly, we shall note that methodology of obtaining data on PM concentrations is not fully consistent during our observed period. This is due to the fact that it was technologically unattainable to reliably measure smaller types of PM ($PM_{2.5}$). In order to choose which of the remaining three pollutants to analyze, we shall also clarify whether we will analyze emissions or concentrations (also called immissions). Our dataset does not contain any data prior to the year 1989, as the methodology changed significantly after that year. The study Kreuz, Lisa and Šauer 2017 tried to go beyond that year, however, they concluded that their results are not statistically significant, because of the need to estimate and adjust data due to their methodology disparity.

For these reasons, we shall, to some extent, try to overcome the relative shortness of the time period for which we can obtain pollution data by using panel data, i.e. data for each year in our observed period and for each region in the Czech Republic. Breakdown to regions is, however, only available for the concentrations data. Hence, we shall choose concentrations data. As discussed in the previous subsections, CO_2 is measured, for the purposes of pollution inspection, merely by calculating its emissions. Hence, we shall use, for the purposes of our analysis, data of SO_2 and NO_x.

4 Dataset description

In this section, we shall describe the variables we will use in this analysis, their source and the methodology using which they were obtained.

4.1 Pollution

As we have already mentioned, for the purposes of our analysis we shall use data for concentrations of SO_2 and NO_x. These data were obtained from the Czech Hydrometeorological Institute. Concentrations of these pollutants are measured on a daily basis in 46 climate stations in the Czech Republic. Firstly, we shall be aware of the fact that each climate station tends to use its own methodology. Moreover, these methodologies tend to vary even for a given climate station over time. In total, there are 7 methodologies that have been used in the Czech climate stations during our observed period to collect data on the two variables of our interest. The following summary is based on Linscheid 1999 and Valcárcel 2012.

- West-Gaeke Method—this method is used to determine the levels of SO_2 concentrations in the air by analyzing a sample of air from which SO_2 is absorbed in a solution of sodium tetrachloromercurate (Na_2HgCl_4).

- Chemiluminescence—this method is used to determine the levels of NO_x concentrations in the air by measuring the light emitted as a result of a reaction between oxides of nitrogen and ozone.

- Ultraviolet fluorescence—this method is used to determine the levels of SO_2 concentrations in the air by analyzing the natural ultraviolet (UV) radiation of SO_2 through measurement of UV radiation of an air sample.

- Thorin spectrophotometric method—this method is used to determine the levels of SO_2 in the air by using thorin to determine the amount of barium ions in a solution of sulfuric acid which was created by a reaction of a sample of air and hydrogen peroxide.

- Ion chromotography—this method is used to determine the levels of SO_2 in the air by determining the amount of SO_2 in alkaline formaldehyde solution, in which SO_2 was absorbed from the air sample, by ion chromotography.
- Guaiacol method—this method is used to determine the levels of NO_x in the air. It is based on the natural reactivity of guaiacol (naturally-occurring organic compound) and NO_x.
- Coulometry—this method is used to determine the levels of SO_2 in the air by measuring the amount of electricity produced by an electrolysis reaction of SO_2 and a filter (a given organic or an inorganic gas).

Even though data on pollution in our dataset are collected by using various methodologies, they are comparable to each other. This is because of the following reasons. Firstly, they all use the same units, secondly, levels of the measurements are the same, and, thirdly, the individual methodologies did not change during our observed period.

The data in our dataset are panel data—there is both the time dimension (1995–2017) and the geographical dimension (46 climate stations). It would be possible to have more detailed data than yearly data of SO_2 and NO_x concentrations, as they are measured every day, but since our explanatory variables will not have this detailed structure, we shall use yearly data. The geographical dimension has to be adjusted, as we need to have pollution data on the 14 regions of the Czech Republic. In order to do that, we shall calculate averages of SO_2 and NO_x concentrations of all climate stations measuring SO_2 and NO_x in each of the 14 Czech Republic regions. This methodology is consistent with the methodology of calculating averages across time (in the case of calculating yearly data from daily data) used by the Czech Hydrometeorological Institute.

The obstacle we had to overcome is that in 2000 the regional division of the Czech Republic was changed. Before that date, there were only 7 regions. In order not to allow this change to affect our analysis, we looked at the data of respective climate stations, rather than data by regions, and

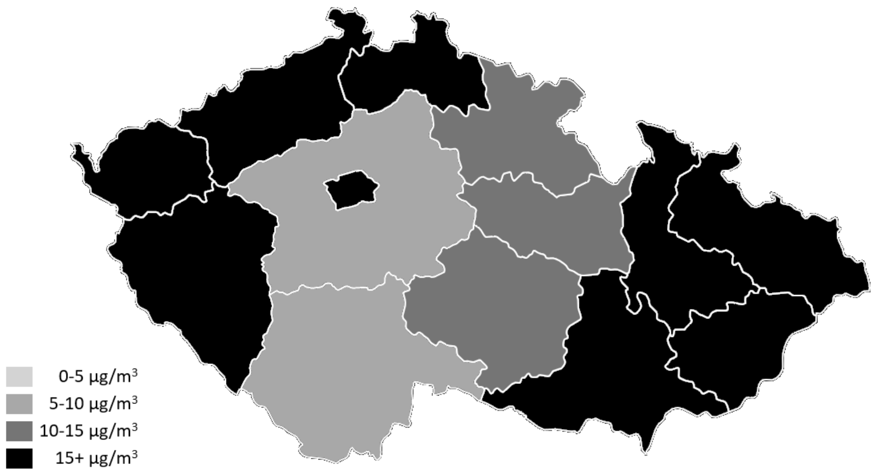

Figure 3: Concentrations of SO_2 in regions of the Czech Republic, 1995

then assigned to them the region to which they belong according to the current classification.

In order to examine the main trend of the data we shall look at Figure 3 and Figure 4 which illustrate change of SO_2 concentrations between 1995 and 2017.

We can clearly observe that, generally speaking, SO_2 pollution significantly declined between 1995 and 2017. While in 1995, there were no regions with SO_2 yearly concentrations below $5\mu g/m^3$, in 2017 there were 8 such regions. On the other hand, in 1995 there were 9 regions with SO_2 yearly concentrations above $15\mu g/m^3$, while in 2017 there were no such regions.

Now, let us look at Figure 5 and Figure 6 which illustrate NO_x concentrations between 1995 and 2017.

We can see that, similarly to the case of SO_2, concentrations of NO_x were significantly lower in 2017 than in 1995. For example, in 1995 there were 4 regions with SO_2 concentrations above $35\mu g/m^3$, while in 2017 there was only 1 such region.

In order to further illustrate the development of NO_x and SO_2 emissions, let us look at Figure 7, where we can see yearly averages of all 14 regions in the Czech Republic for both SO_2 and NO_x.

We can observe that even though overall, both the SO_2 and the NO_x con-

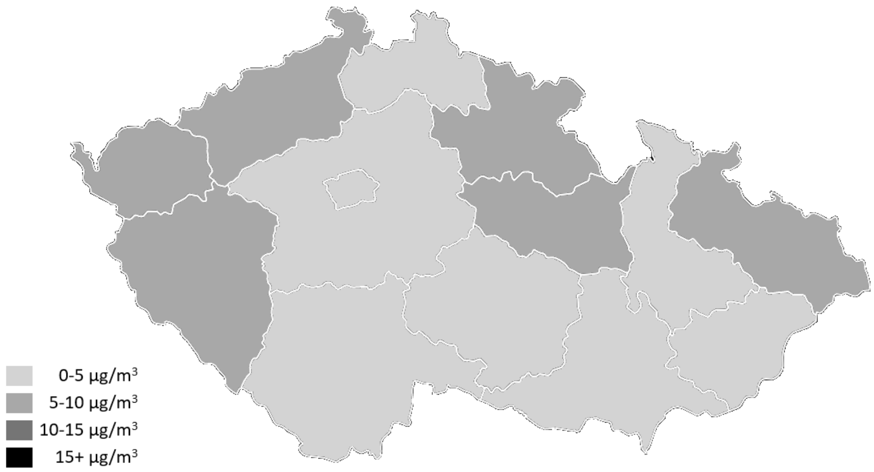

Figure 4: Concentrations of SO_2 in regions of the Czech Republic, 2017

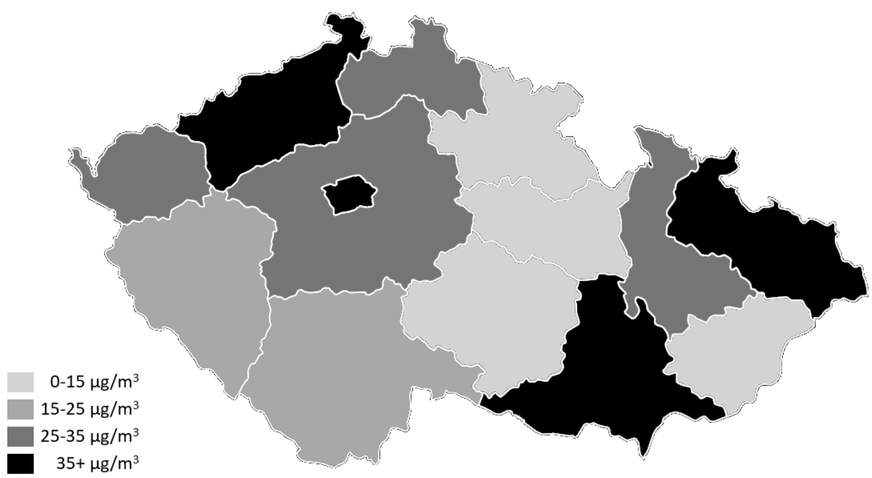

Figure 5: Concentrations of NO_x in regions of the Czech Republic, 1995

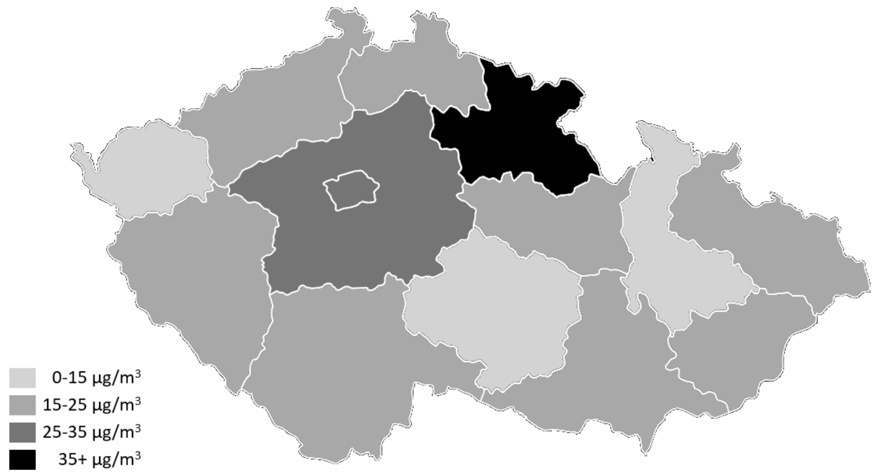

Figure 6: Concentrations of NO_x in regions of the Czech Republic, 2017

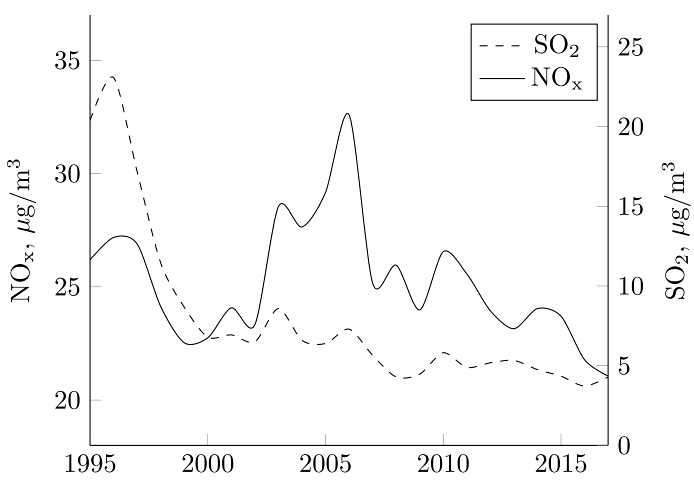

Figure 7: SO_2 and NO_x concentrations in the Czech Republic between 1995 and 2017

centrations decreased during the observed period, the development of their concentrations was very different. While the SO_2 concentrations experienced a peak of $23.07\mu g/m^3$ in 1996 and then continued to decrease relatively steadily, the NO_x concentrations experienced their peak of $32.59\mu g/m^3$ in 2006, after which they decreased steadily as well.

To conclude, the concentrations of both our gases of interest fell during our observed period with the decrease of the SO_2 immissions being more significant than the NO_x immissions decrease.

4.2 GDP per capita

In this bachelor thesis, we use the most well-known measurement of economic growth—Gross Domestic Product (GDP). It provides information on the final value of all goods and services produced within one year in a given geographical unit. In our case, we shall use GDP per capita data with regional breakdowns. For the purposes of our analysis, GDP per capita is more useful as it is the commonly used metric when analyzing the EKC. The reason for that is that the EKC is based on the hypothesis that with increasing wealth, firstly, pollution levels increase, and then, from some level of wealth, they decrease. And wealth is generally measured with respect to population, as otherwise it would not take into consideration the number of people living in a given geographical unit contributing to the overall GDP. In our analysis, we cannot use GDP PPP data, as GDP data adjusted for the purchasing power parity with respect to regions are not available. However, the difference between purchasing power parity across regions of the Czech Republic is, in most cases, less significant than differences between purchasing power parities across countries, which was the case of most analyses regarding the EKC. The source of both the GDP per capita data and the population data is the Czech Statistical Office (CSO).

Now, let us briefly analyze the development of GDP per capita in the Czech Republic which is depicted in Figure 8.

We can see that during our observed period, GDP per capita showed a

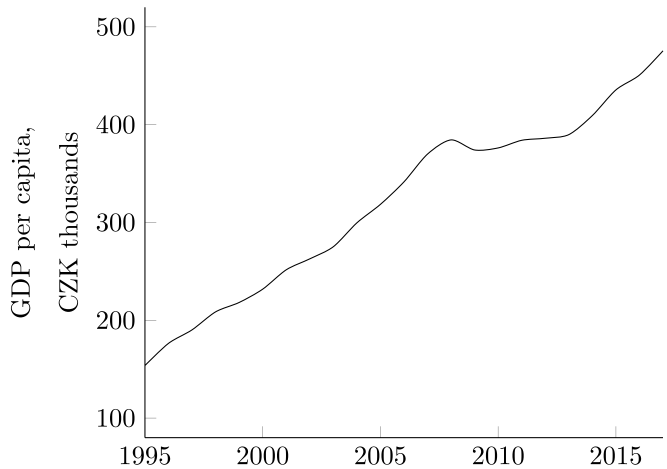

Figure 8: GDP per capita in the Czech Republic between 1995 and 2017

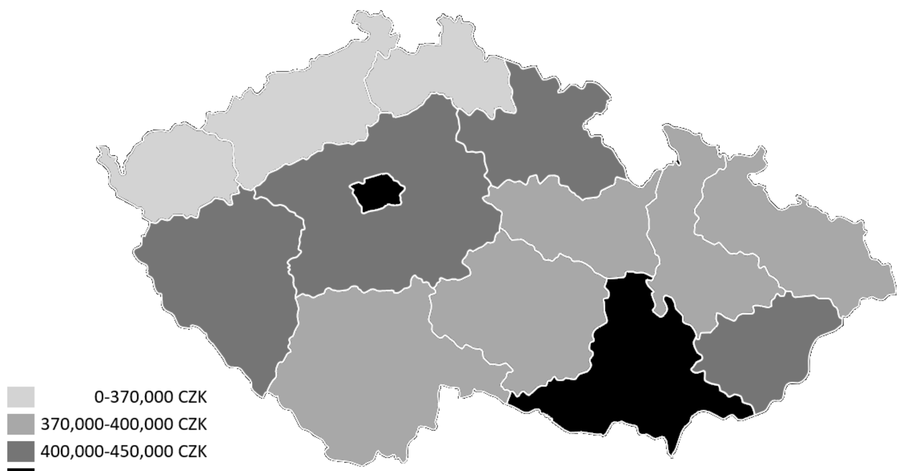

Figure 9: GDP per capita in the Czech Republic, 2017

steady growth, with the exception of the 2007–2008 economic crisis.

Now, let us compare the regions of the Czech Republic by their respective 2017 GDP per capita. This can be seen in Figure 9.

We can see that the regions with the highest GDP per capita are Hlavní město Praha and Jihomoravský kraj, which are also regions in which there are the two largest cities of the Czech Republic—Praha and Brno, respectively.

When we combine Figure 7 and Figure 8, we can see that the immissions of both pollutants at some point in time reached their peak from which they to a certain extent steadily declined, while the GDP per capita steadily increased during the entire observed period, with the exception of the period during and immediately after the 2007–2008 economic crisis. This would, so far, give us hints that the idea of the EKC shall not be rejected at this point, as the GDP per capita could have reached the turning point of SO_2 and NO_x immissions from which they declined. Of course, further analysis is needed in order to provide empirical evidence either for or against the EKC hypothesis. Now, let us add additional variables to our model.

4.3 Population density

We have already introduced two variables in our model—pollution (dependent variable) and economic growth (independent variable). Now, let us add another independent variable into our model—population density. The reasoning behind why it might be a relevant variable for our model is that when we are analyzing immissions, the more dense a given region is the higher pollutants' concentrations there will be in the region (ceteris paribus). This is because concentrations are measured in cubic units, hence, the larger the area, and consequently volume, on which a given weight of these gases is spread, the lower concentrations of these gases. Now, let us look at Figure 10 and Figure 11 in order to see how regions of the Czech Republic differ in terms of population density and how their population density changed during our observed period.

From Figure 10 and Figure 11 we can see that most regions belonged to the same population density category in 2017 as they did in 1995. As of 2017, the most densely populated regions were Hlavní město Praha, Jihomoravský kraj and Moravskoslezský kraj. Therefore, pollution in these regions is likely to be worse than in other less densely populated regions. The problem here is that people tend to concentrate in places with higher GDP per capita, which, according to the EKC hypothesis, also influences the level of environmental degradation. If the Czech Republic already reached the EKC turning point, then, in regions with higher GDP per capita would be lower

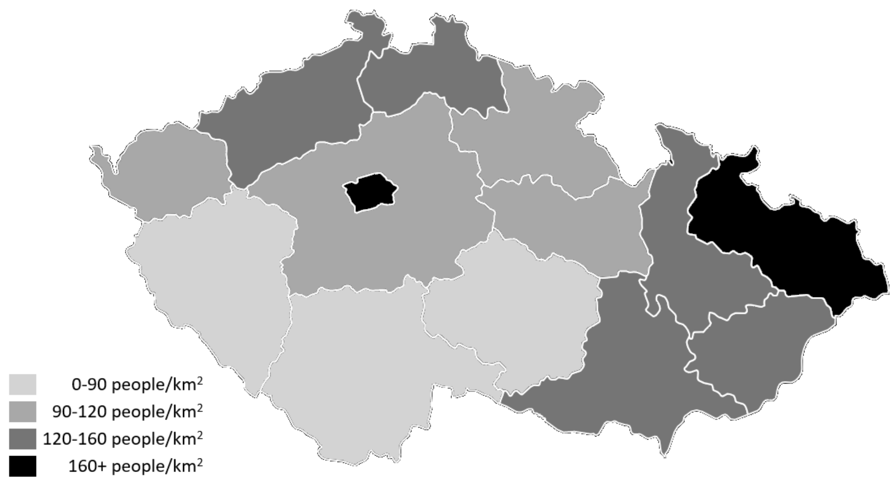

Figure 10: Population Density in the Czech Republic, 1995

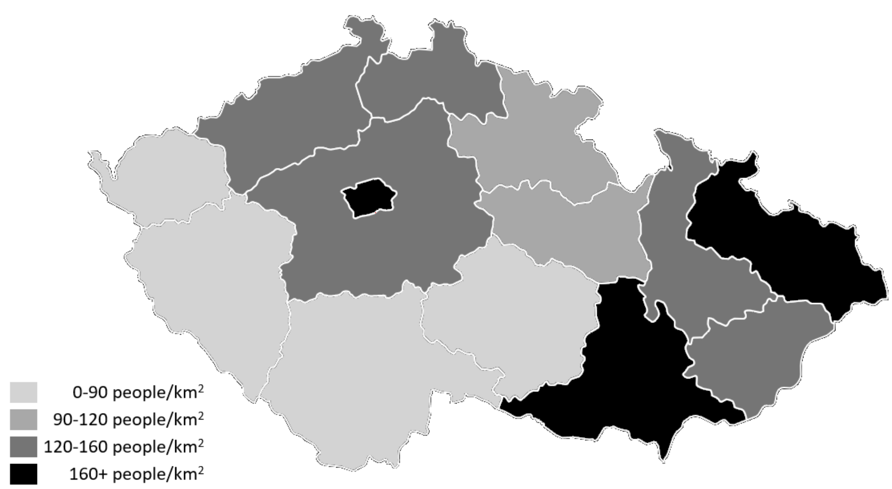

Figure 11: Population Density in the Czech Republic, 2017

concentrations of pollutants. When we compare results from Figure 9 and Figure 11, we can see that regions with the highest GDP per capita are also regions with the highest population density, which supports the hypothesis that people concentrate in regions with higher GDP per capita. These two variables (GDP per capita and population density) are, therefore, likely to have opposite effects on pollutants' immissions.

4.4 Unemployment

Now, we shall introduce another dependent variable—unemployment. Even though unemployment is not usually part of models examining the EKC, some research papers consider it to be a factor which has the potential to influence pollution in a given country (e.g. Braniš and Linhartová 2012). The main reasoning of why unemployment could influence pollution is that people's behaviour differs significantly when they are employed and when they are not. The effect of unemployment can turn out to be positive or negative, as there are effects associated with higher unemployment, which have both positive and negative effect on air pollution.

Examples of such effects are lower demand for transportation (which is one of the main contributors to local pollution) or lower demand for goods and services (as a result of lower income). On the other hand, unemployment can also have a negative effect on the environmental quality in a given city or a country. When a person is unemployed, it can lead to a less environmentally responsible behaviour, as the person is likely to have, in their eyes, more pressing issues to deal with (such as finding a job or feeding their family). Moreover, the lower income resulting from being unemployed may cause the given person to prefer cheaper alternatives of transportation or heating regardless of their emission production (such as driving in an older car, which produces more emissions than a newer one, or burning lower-quality coal).

The source of the panel data of unemployment is also the CSO. The methodology of the CSO is that the unemployment rate represents a ratio of the number of the unemployed and the number of people between the age

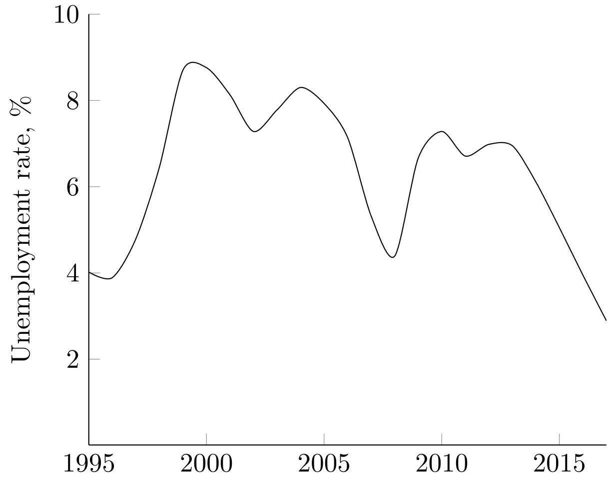

Figure 12: Unemployment rate in the Czech Republic between 1995 and 2017

of 15 and 64. Before 2011, the methodology was different—the denominator of the ratio was the number of people in labor force. This change of methodology is not a problem for our analysis, since the CSO adjusts the older unemployment data according to the new methodology. Firstly, before we look at data with the regional breakdown, let us look at the overall development of the unemployment rate in the Czech Republic presented in Figure 12.

We can see from Figure 12 above that the unemployment rate is rather volatile. During our observed period (1995–2017), the unemployment rate reached its peak in 2000 (8.76%). Now, let us look at Figure 13, where we can find comparison of regions of the Czech Republic with respect to their unemployment rate (as of 2017).

From Figure 13, we can see that the lowest unemployment rate is in Hlavní město Praha and in Plzeňský kraj, whereas the highest unemployment rate is in Liberecký kraj, Zlínský kraj and Moravskoslezský kraj. On the first sight, there does not seem to be any pattern in terms of relationship between the unemployment rate and pollution. For example, even though Plzeňský kraj has one of the lowest unemployment rates in the Czech Republic and Moravskoslezský kraj has one of the highest unemployment rates in the Czech Republic, they are both in the same category of both the SO_2 and the

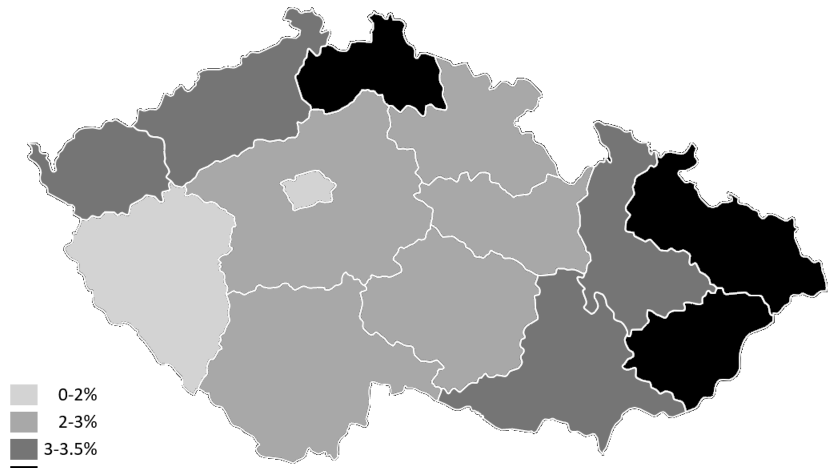

Figure 13: Unemployment rate in regions of the Czech Republic, 2017

NO_x immissions. Of course, further analysis is needed in order to statistically test whether there is any significant effect of unemployment on the SO_2 and the NO_x concentrations.

4.5 Policy variable

Now, we shall introduce the first variable which will, unlike those previously mentioned, not differ across regions. A dummy policy variable was, for example, added into the model in Panayotou 1997. This variable is a dummy variable which represents the introduction of an essential legislative which has the potential to influence the SO_2 and the NO_x concentrations.

This dummy variable will represent the effectiveness of emission limits set in the law no. 309 from 1991 (Act No. 89/2012 Coll., §14 (3)).

> Emission limits which are set for current sources of pollution are based on the lowest achievable level of emissions using the given technical devices when they comply to the conditions set for their operations. Authorities responsible for protection of air quality will set the limits until the 31^{st} of December 1994 and will, at the same time, set the date until when the set limits must be achieved. This time must be set with respect to the used technology and must

be, at latest, the 31st of December 1998.

The importance of the limits set with effectiveness from the beginning of the year 1999 is, for example, mentioned in Hofman 2015 or Brůha and Ščasný 2005. The main focus of this relatively strict regulation was on coal power plants.

4.6 Share of mining in the economy

Last variable we shall add in our model is the share of mining on the total sum of value added in the Czech economy. For this variable, the breakdown with respect to regions is not available. In most studies regarding the EKC, this variable was not directly mentioned, however, a similar ones were. For example, in Kaufmann et al. 1998, there was steel production added in the model, which is closely connected to coal mining. This is due to the fact that steel production requires large amounts of coal in order to have sufficient temperature in furnaces, and coal producers rely on steel producers as one of their main clients. Moreover, coal can be used in other ways which harm the environment—heating households or producing electricity.

It can be argued that this variable does not take into consideration exports and imports of coal and other natural resources. This is true, however, exports and imports, in the case of coal, do not represent a significantly large share relative to total coal production in the Czech Republic. The CSO collects data on imports, exports and total production of brown coal, which is the main natural resource in the Czech Republic. As of 2016, exports and imports shares relative to total production in the Czech Republic were 2.2% and 0.2% respectively. Hence, we can conclude that it can be considered a relevant variable.

Moreover, mining natural resources (not only coal) may serve as a proxy for the level of environmental degradation, as most of mining represents a burden for the environment.

The natural hypothesis here is that with decreasing importance of coal the level of pollution decreases. Share of mining on the total sum of value

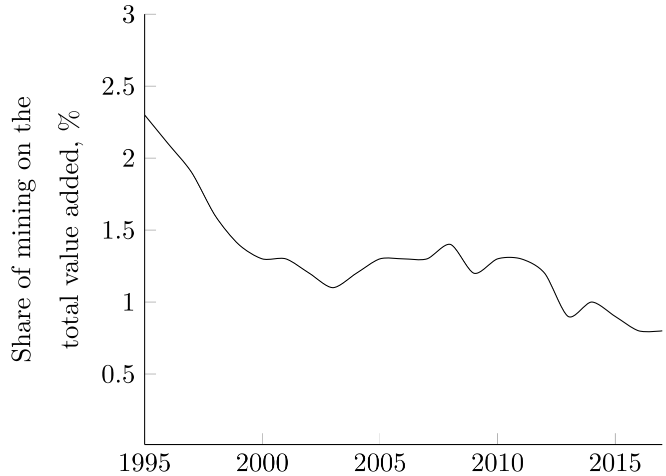

Figure 14: Share of mining on the total sum of value added in Czech economy between 1995 and 2017

added in the economy between 1995 and 2017 can be seen in Figure 14.

From Figure 14 we can see that share of mining on the total share of value added in the Czech Republic overall declined throughout our observed period. On the first sight, there seems to be a positive correlation between concentrations of SO_2 and NO_x and the share of mining on the total share of value added in the Czech Republic, as all those variables declined significantly during the observed period. However, further analysis is needed in order to either prove or disprove the hypothesized correlation.

5 Model Specification

5.1 Panel data analysis—econometric methods

As we have already mentioned, for the purposes of our analysis we collected data, which have two dimensions—the cross-sectional dimension (in our case 14 regions of the Czech Republic) and the time dimension (from 1995 to 2017, i.e. 23 years). This means that we are using panel data. In this subsection, we shall summarize the most essential methods for econometric analysis of panel data. Most of this subsection will be based on Wooldridge 2015. If the reader is already familiar with econometric methods of panel data analysis, they can skip this subsection.

Panel data

Firstly, let us present the basic concepts of panel data analysis. Panel data (also called longitudinal data) are data which have both the cross-sectional and the time dimension. The panel data model is represented by the following system of equations

$$\begin{aligned} y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_{k-1} x_{it(k-1)} + \beta_k x_{itk} + a_i + u_{it}, \\ i = 1, \ldots, N, \ t = 1, \ldots, T, \end{aligned} \tag{1}$$

where N represents the size of the cross-sectional dimension (in our case 14) and the number of time periods in our dataset (in our case 23) is represented by T. The coefficients determining the effects of respective dependent variables are represented by $\beta_1, \ldots, \beta_k$, the term a_i and the term u_{it} represents factors influencing y_{it}, which are not included in the model individually. The difference between a_i and u_{it} is that a_i is time-invariant.

Pooled Cross-sections

An independently pooled cross-section data are data which are collected by randomly sampling from a population at different time periods. The important distinguishing feature of polled cross-sections and panel data is that the cross-sectional units based on which the data were collected, differ across time. There is a possibility of having a bias in the pooled cross-sectional model in case some of the explanatory variables are correlated with the time-invariant error term.

First differencing

The main purpose of using first differencing for analysis tends to be the fact that it is possible to allow the time-invariant unobserved effects (a_i) to be correlated with explanatory variables. Generally, the first differencing equation, for the case of two periods, can be written as

$$(y_{i2} - y_{i1}) = \delta_0 + \beta_1(x_{i2} - x_{i1}) + (u_{i2} - u_{i1}), \qquad i\text{=1,...,}N,$$

or

$$\Delta y_i = \delta_0 + \beta_1 \Delta x_i + \Delta u_i, \qquad i\text{=1,...,}N.$$

Here, Δ stands for the difference between the second and the first period and δ_0 stands for the difference between the intercepts of the models of the two periods. We can see that a_i is not present in the equation. The reason for that is that the first differencing equation is a result of subtraction of the first period equation from the second period equation.

The first differencing approach can be analogously applied to the case of more than two periods by differencing the adjacent periods. This approach does not allow any of the explanatory variables to be constant over time, as such variables would be subtracted away (as in the case of a_i).

Fixed effects

This method was named in this way because its model parameters are fixed. Fixed effects estimation model can be obtained by, firstly, having the equation (1). Then, by averaging by time (t) for each i, we get

$$\bar{y}_i = \beta_0 + \beta_1\bar{x}_{i1} + \beta_2\bar{x}_{i2} + \ldots + \beta_{k-1}\bar{x}_{i(k-1)} + \beta_k\bar{x}_{ik} + \bar{u}_i, \qquad i{=}1,\ldots,N$$

where $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ and so on. Now, let us subtract the averaged equation from the original model:

$$\begin{aligned} y_{it} - \bar{y}_i = &\beta_1(x_{it1} - \bar{x}_{i1}) + \beta_2(x_{it2} - \bar{x}_{i2}) + \ldots + \\ &+ \beta_{k-1}(x_{it(k-1)} - \bar{x}_{i(k-1)}) + \beta_k(x_{itk} - \bar{x}_{ik}) + (u_{it} - \bar{u}_i), \\ &\qquad\qquad i = 1,\ldots,N,\ t = 1,\ldots,T. \end{aligned}$$

This expression can also be written using the following notation as

$$\ddot{y} = \beta_1\ddot{x}_{it1} + \beta_2\ddot{x}_{it2} + \ldots + \beta_{k-1}\ddot{x}_{it(k-1)} + \beta_k\ddot{x}_{itk} + \ddot{u}_{it}, \qquad i{=}1,\ldots,N,\ t{=}1,\ldots,T,$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$, $\ddot{x}_{it1} = x_{it1} - \bar{x}_{i1}$, $\ddot{x}_{it2} = x_{it2} - \bar{x}_{i2}$, $\ldots$, $\ddot{x}_{it(k-1)} = x_{it(k-1)} - \bar{x}_{i(k-1)}$, $\ddot{x}_{itk} = x_{itk} - \bar{x}_{ik}$.

In case the strict exogeneity assumption is violated, it is possible to have a bias close to zero by including a significant number of observed periods in the model.

Random effects

This method was named in this way because certain model parameters are considered to be random variables. This approach is based on estimation of the following equation:

$$\begin{aligned} y_{it} - \theta\bar{y}_i = &\beta_0(1-\theta) + \beta_1(x_{it1} - \theta\bar{x}_{i1}) + \beta_2(x_{it2} - \theta\bar{x}_{i2}) + \\ &+ \ldots + \beta_{k-1}(x_{it(k-1)} - \theta\bar{x}_{i(k-1)}) + \\ &+ \beta_k(x_{itk} - \theta\bar{x}_{ik}) + (v_{it} - \theta\bar{v}_i), \\ &\qquad\qquad i = 1,\ldots,N,\ t = 1,\ldots,T. \end{aligned}$$

where $\theta = 1 - [\sigma_u^2/(\sigma_u^2 + T\sigma_a^2)]^{1/2}$, $\sigma_u^2 = Var(u_{it})$, where $\sigma_a^2 = Var(a_i)$ and $v_{it} = u_{it} + a_i$. Moreover, if $\theta = 1$, the random effects model is the same as the fixed effects model. Therefore, the fixed effects model is a special case of the random effects model.

One of the main advantages of the random effects model is that it can contain a time constant. The model is based on the assumption that a_i is not correlated with any of the dependent variables (for all $i,\ t$).

5.2 Regression Model of Pollution Determinants

In this subsection, we shall examine models determining factors influencing concentrations of SO_2 and NO_x (each time, we shall have two models for each of the pollutants). Moreover, this subsection aims to investigate and discuss whether assumptions of the respective models are satisfied or not.

Model without dummies for years

Firstly, we shall introduce a model which does not contain any dummy variables for respective years in our observed period.

As we have already mentioned, there will be models for both SO_2 and NO_x. Both of these pollutants will be our dependent variables. As in the case of other research papers focusing on the EKC topic, our dependent variable will be in a logarithmic form. This is due to the fact that we are mainly interested in a relative change of pollutants' concentrations, rather that an absolute change. If the dependent variable was not in a logarithmic form, then a relative change of pollution for lower pollution values would be underestimated compared to the same relative change for higher pollution values, even though their importance and difficulty to achieve them is comparable.

The first, and the most important, explanatory variable in the model is GDP per capita. This variable will be in our model, similarly to the case of the dependent variable, in a logarithmic form. Reasoning for that is analogical to the case of the explained variable, i.e. we are more interested in the relative change than in the absolute change. Moreover, quadratic form of this variable will be used. This is typical for research covering this topic, for the underlying EKC hypothesis is that the relationship between pollution and GDP per capita can be graphically depicted as a concave quadratic function.

Other independent variables, which were already introduced in Section 4, will be included in a standard, non-logarithmic, form. These variables are population density, unemployment, policy variable (dummy variable) and share of mining in the economy.

Therefore, our first model for the case of SO_2 immissions is written as

$$log(y_{i,t}) = \beta_1 \cdot [log(b_{i,t})]^2 + \beta_2 \cdot c_{i,t} + \beta_3 \cdot d_{i,t} + \beta_4 \cdot e_t + \beta_5 \cdot f_t + a_i + u_{i,t},$$
$$i = 1, \ldots, N, \; t = 1, \ldots, T,$$

where y stands for concentrations of SO_2 in μg/m^3, variable b stands for GDP per capita (in CZK millions), variable c stands for population density (in people per km^2), variable d stands for unemployment rate (in %), variable e stands for the emission policy (dummy variable which takes values 0 if before 1999 and 1 if in 1999 or after 1999), variable f stands for the share of mining on the total value added in the Czech economy (in %), variable a_i represents all factors not included in the model individually which do not change over time and variable u stands for all factors not included in the model individually which change over time. We can note that GDP per capita is included in the model in a quadratic form and not in a linear form. This is due to the fact that the linear form of GDP per capita was found to be statistically insignificant so we did not include it in the model.

Analogously, we can define our initial model for the case of the NO_x concentrations as

$$log(z_{i,t}) = \beta_1 \cdot [log(b_{i,t})]^2 + \beta_2 \cdot c_{i,t} + \beta_3 \cdot d_{i,t} + \beta_4 \cdot e_t + \beta_5 \cdot f_t + a_i + u_{i,t},$$
$$i{=}1,\ldots,N, \; t{=}1,\ldots,T,$$

where variable z stands for concentrations of NO_x in μg/m^3 and all other notation is analogous to the case of SO_2.

Model with dummies for years

Now, let us define models for both the SO_2 and the NO_x concentrations, which will include dummy variables for years. There will be 23 dummy variables in our model, as we have 23 years in our observed period. Each dummy variable will be equal to 1 for the year which it represents and 0 otherwise. Year dummies will help us to assess the ceteris paribus impact of respective years on pollution, rather than the effect of one specific event, as is the case of the policy dummy variable of the previous models.

For the case of the SO_2 concentrations, the model with year dummy variables is written as

$$log(y_{i,t}) = \beta_1 \cdot [log(b_{i,t})]^2 + \beta_2 \cdot c_{i,t} + \beta_3 \cdot d_{i,t} + \beta_4 \cdot f_t + \delta_t + a_i + u_{i,t},$$
$$i{=}1,\ldots,N,\ t{=}1,\ldots,T,$$

where δ_t represents time-fixed effects. Other notation is the same as for the previously described models.

For the NO_x concentrations, the model with year dummies will be analogical to the SO_2 model with year dummies:

$$log(z_{i,t}) = \beta_1 \cdot [log(b_{i,t})]^2 + \beta_2 \cdot c_{i,t} + \beta_3 \cdot d_{i,t} + \beta_4 \cdot f_t + \delta_t + a_i + u_{i,t},$$
$$i{=}1,\ldots,N,\ t{=}1,\ldots,T,$$

where the notation is the same as in the case of the SO_2 model with year dummy variables. We shall notice that the policy dummy variable is not included in this model. The reason for that is that having dummies for years in the model and, at the same time, including policy dummy variable would violate the assumption of no linear dependence between the dependent variables.

Choice of method of analysis

Now, it is essential to choose which method of panel data analysis to use. It is important to note that there is a high probability of a_i including factors influencing pollution which can be correlated with independent variables included in the model. An example of such factors, which may be contained in a_i, may be factors, which are characteristic for specific regions and which do not change over time, such as geographical conditions. Geographical conditions can significantly influence immissions of pollutants, as their dispersion to a large extent depends on the geographical type of the given region, since in areas surrounded by mountains it takes longer for these pollutants to disperse. Geographical conditions can also possibly be correlated with explanatory variables included in our model. For example, the geographical nature of a given region may affect the economic prosperity of the given region, as some types of geography can be beneficial for a certain type of businesses.

Having said that, we can rule out the random effects method, as it requires a_i to be treated as random. However, for the reasons stated in the previous paragraph, this is not possible. Hence, we shall decide between the first differences method and the fixed effects method. It is possible that there is feedback between u_{it} and future outcomes of our dependent variable (the SO_2 or the NO_x concentrations). It is likely that factors which vary over time which are not included in our model individually affect future values of our explained variables. For example, subsidies for either transportation, manufacturing or power plants with lower emissions of our pollutants of interest are not included in our model, as comprehensive statistic of them is not available, but they are very likely to decrease concentrations of both SO_2 and NO_x. For this reason, we shall use the fixed effects method rather than the first difference method.

6 Empirical Results

In this section, we shall provide results of regressions of models which we introduced in the previous section. Moreover, various tests will be computed in order to determine whether or not our data possess certain properties, which are necessary to correctly assess validity of our results and whether the underlying assumptions of our models do or do not hold.

Firstly, we shall discuss the assumption that the expected value of the idiosyncratic error of our models is, in all time periods, equal to zero, i.e. the strict exogeneity assumption. This assumption is needed to be fulfilled in order for our estimators to be unbiased. In our case, we shall assume that this assumption is likely not to be violated, for there does not seem to be any pattern of error terms tending to have, on average, non-zero mean for any values of the explanatory variables, since there is no variable not included in our model which would be likely to cause such a pattern. The factors, which can possibly affect concentrations of our pollutants of interest, do not seem to be likely to have any patterns which would cause the error terms to be non-zero for any of our observed regions.

6.1 Testing of model assumptions

Testing for heteroskedasticity

Secondly, there is the assumption of homoskedasticity for any given region for all periods in our dataset. For this purpose, we shall use the Breusch-Pagan test. The estimation results can be seen in Table 1.

We can see that all our observed models turned out to possess heteroske-

Model	Breusch-Pagan test p-value
SO_2 model without year dummy variables	< 0.01
NO_x model without year dummy	0.01196
SO_2 model with year dummy variables	0.02245
NO_x model with year dummy variables	0.03913

Table 1: Breusch-Pagan test p-values

dasticity. Therefore, we shall bear that in mind and when estimating our models we shall obtain heteroskedasticity-robust results.

Testing for serial correlation

Another assumption of our models is the assumption of no serial correlation in our data. We shall use the Breusch-Godfrey test in order to determine whether this assumption holds or not. The Breusch-Godfrey test is designed for that. It is based on the idea of regressing residuals obtained from the original regression (for all time periods) on explanatory variables of the original regression and lags of those residuals. The null hypothesis is that there is no serial correlation, the alternative hypothesis is that serial correlation is present.

We can see from Table 2 that for all our 4 models the p-value calculated by the Breusch-Godfrey test is lower than 0.001. Therefore, we can reject the null hypothesis of no serial correlation at a significance level lower than 0.1%.

Since in all our four models the assumption of no serial correlation is very likely to be violated, we shall use a robust test statistic.

6.2 Estimation of models without year dummy variables

SO_2 model without year dummy variables estimation

In the previous subsections, we have observed that all of our 4 models violate the assumption of homoskedasticity and the assumption of no serial correlation. Therefore, we shall use a robust test statistic, which will provide us with consistent coefficients.

Model	Breusch-Godfrey test p-value
SO_2 model without year dummy variables	< 0.001
NO_x model without year dummy	< 0.001
SO_2 model with year dummy variables	< 0.001
NO_x model with year dummy variables	< 0.001

Table 2: Breusch-Godfrey test p-values

Using the robust test statistic, the SO_2 model without year dummy variables is estimated. The regression results can be seen in Table 3. We can see that all explanatory variables, with the exception of unemployment, are significant on a significance level lower than 5%. The effect of GDP per capita turned out to have the highest t-statistic (in absolute terms) of all independent variables in our model, which means that it is the most significant variable. The coefficient of GDP per capita is negative, which is consistent with the EKC theory, as the EKC is graphically a quadratic concave function, which is attributed to a parabola with a negative slope coefficient. The value of this coefficient is approximately -0.42, which means that when GDP per capita changes by a certain percentage, the SO_2 concentrations, ceteris paribus, change by the same percentage by which GDP per capita changed to the power of 2 and multiplied by -0.42 (as both the SO_2 concentrations and the GDP per capita are included in the model in a logarithmic form and, additionally, the logarithmic form of GDP per capita is included in a quadratic form). The coefficient of share of mining is approximately 0.26, which means that when the share of mining on the total value added in the Czech economy decreases by 1% (which is the unit in which the variable is included in the model) the SO_2 concentrations decrease by 26%.

The coefficient of the policy variable is approximately -0.44, which is interpreted in the following way. Ceteris paribus, concentrations of SO_2 with

Variable	Estimate	Std. Error	t-value	p-value
GDP per capita	-0.42448594	0.05311624	-7.9916	< 0.001
Share of mining	0.25577564	0.07765731	3.2936	< 0.01
Policy variable	-0.43987549	0.07792170	-5.6451	< 0.001
Population density	-0.00197570	0.00052706	-3.7485	< 0.001
Unemployment	0.00111719	0.01008129	0.1108	0.911834
N	14			
T	23			
R^2	0.77138			
Adjusted R^2	0.7578			

Table 3: Estimation results of the SO_2 model without year dummy variables

the policy being effective are lower by 44% when compared to the situation without this law being effective.

The coefficient of population density is approximately -0.002, which can be interpreted in the way that when population density decreases by 1 person per km^2, holding other factors fixed, the SO_2 concentrations decrease by 0.2%. The negative sign of this coefficient is relatively unexpected, as we discussed in Section 4 that higher population density is likely to lead to higher emissions of pollutants, which shall lead to higher concentrations of these pollutants. One of the possible reasons for the sign of this coefficient to be negative may be that government efforts to reduce pollution are likely to be targeted primarily in areas with higher population density, as it can, for similar costs, benefit more people than these efforts in areas with lower population density. One could argue that in order to assess whether this reason is valid or not, we shall add variable for government efforts with breakdown to regions to our model. This is, however, not feasible, as the CSO does not collect any data on government efforts to reduce pollution with the necessary regional description. The last variable in the model is the unemployment rate. This variable is not significant at a sufficiently low significance level, therefore, we shall not present the interpretation of its coefficient

NO_x model without year dummy variables estimation

Now, let us estimate the NO_x model without year dummy variables. The results of this estimation can be seen in Table 4.

From Table 4 we can see that there is only 1 significant variable in our model, which is population density. Therefore, we cannot make any conclusions based on the model. The lack of suitability of the model can be also indicated by the low R^2 (0.055), which indicates that this model does not explain any significant share of variation of the NO_x concentrations.

6.3 Estimation of models with year dummy variables

SO_2 model with year dummy variables estimation

Now, let us estimate the SO_2 model with year dummy variables. The estimation results can be seen in Table 5.

From Table 5 we can see that the adjusted R^2 increased in comparison with the SO_2 without year dummies model. From that, we shall conclude that the model with year dummies is more suitable for our analysis, as adjusted R^2 is constructed in the way that it penalizes for adding insignificant variables into the model. Furthermore, we can observe that significance of variables which were also included in the SO_2 model without year dummy variables in general decreased. This may be due to the fact that variation of the SO_2 concentrations is to some extent explained by respective year dummy variables, rather than by the original explanatory variables. At a 5% significance level, unemployment and share of mining on the total value added of the Czech economy are not significant, but GDP per capita and population density are significant. Moreover, coefficients of GDP per capita and population density both did not change their sign. In Figure 15, we can see coefficients of dummy variables for respective years.

From Figure 15, we can see that, in general, coefficients of year dummy variables gradually increase over time from negative values to values closer

Variable	Estimate	Std. Error	t-value	p-value
GDP per capita	0.09745281	0.09790198	0.9954	0.3203302
Share of mining	0.23034452	0.14727366	1.5641	0.1188478
Policy variable	−0.00064281	0.08168083	−3.7705	< 0.001
Population density	−0.00201323	0.00053394	−2.6455	< 0.01
Unemployment	0.02160681	0.01244998	1.7355	0.0836709
N	14			
T	23			
R^2	0.054738			
Adjusted R^2	−0.001416			

Table 4: Estimation results of the NO_x model without year dummy variables

Variable	Estimate	Std. Error	t-value	p-value
GDP per capita	−0.4162023	0.2086198	−1.9950	0.0469986
Share of mining	3.7478564	2.1098084	1.7764	0.0767419
Population density	−0.0016688	0.0004987	−3.3463	< 0.001
Unemployment	−0.0002991	0.0073983	−0.0404	0.9677806
Year 1995	−5.0820991	3.3714055	−1.5074	0.1328204
Year 1996	−4.0379878	2.9266043	−1.3798	0.1687522
Year 1997	−3.5363281	2.5130873	−1.4072	0.1604758
Year 1998	−2.9241522	1.9866175	−1.4719	0.1421523
Year 1999	−2.2133095	1.4596550	−1.5163	0.1305537
Year 2000	−2.0242638	1.1675324	−1.7338	0.0840432
Year 2001	−1.9925765	1.1752212	−1.6955	0.0910822
Year 2002	−1.8062870	1.0763568	−1.6781	0.0944219
Year 2003	−0.9772479	0.8515512	−1.1476	0.2520992
Year 2004	−1.6235414	1.1355396	−1.4298	0.1538909
Year 2005	−1.7704516	1.1974005	−1.4786	0.1403649
Year 2006	−1.8057203	1.3148955	−1.3733	0.1707522
Year 2007	−1.9603911	1.2503375	−1.5679	0.1180243
Year 2008	−2.5909232	1.4580772	−1.7769	0.0766513
Year 2009	−1.9911063	1.1405708	−1.7457	0.0819465
Year 2010	−1.7674500	1.1512971	−1.5352	0.1258565
Year 2011	−2.2240979	1.3182385	−1.6872	0.0926717
Year 2012	−1.5977442	0.9992365	−1.5990	0.1109442
Year 2013	−0.4097216	0.3568782	−1.1481	0.2519087
Year 2014	−1.0202431	0.6287899	−1.6226	0.1057991
Year 2015	−0.7052943	0.4279632	−1.6480	0.1004572
Year 2016	−0.92919458	0.5985250	−1.5525	0.1217846
Year 2017	−1.18224713	0.7485130	−1.5795	0.1154530
N	14			
T	23			
R^2	0.83913			
Adjusted R^2	0.81753			

Table 5: Estimation results of the SO_2 model with year dummy variables

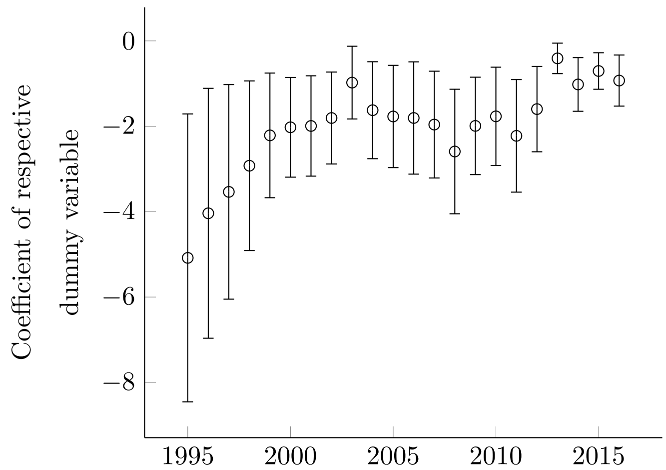

Figure 15: Estimated coefficients of year dummy variables with their respective standard errors

to zero, but still negative. From that, we can conclude that year effects tend to be lower and lower over time (in absolute terms).

NO_x model with year dummy variables estimation

Now, let us estimate the NO_x model with year dummy variables. The estimation results can be seen in Table 6.

From Table 6 we can see that both R^2 and adjusted R^2 increased in comparison to the NO_x model without year dummy variables. However, they

Variable	Estimate	Std. Error	t-value	p-value
GDP per capita	0.27645210	0.31926927	0.8659	0.38728
Share of mining	−1.20381120	1.10620889	−1.0882	0.27742
Population density	−0.00138892	0.00062935	−2.2069	0.02812
Unemployment	0.00235326	0.01620321	0.1452	0.88463
Year 1995	2.54808165	2.25522863	1.1299	0.25949
Year 1996	2.38446155	1.95645580	1.2188	0.22395
Year 1997	2.08946237	1.70723739	1.2239	0.22201
Year 1998	1.68521536	1.38004661	1.2211	0.22305
Year 1999	1.25627449	1.08148563	1.1616	0.24637
Year 2000	1.09451977	0.92692795	1.1808	0.23867
Year 2001	1.10816339	0.88210844	1.2563	0.21006
Year 2002	0.91905781	0.80386731	1.1433	0.25388
Year 2003	1.00669662	0.66845441	1.5060	0.13318
Year 2004	1.16985320	0.77091217	1.5175	0.13026
Year 2005	1.13054972	0.77899085	1.4513	0.14780
Year 2006	1.24698295	0.81295075	1.5339	0.12617
Year 2007	0.94957323	0.75870499	1.2516	0.21176
Year 2008	1.12171459	0.84840301	1.3221	0.18719
Year 2009	0.84602621	0.68836092	1.2290	0.22008
Year 2010	0.98430806	0.70669940	1.3928	0.16477
Year 2011	1.01086616	0.76067913	1.3289	0.18495
Year 2012	0.77078352	0.60377592	1.2766	0.20279
Year 2013	0.35763579	0.27806233	1.2862	0.19943
Year 2014	0.51401308	0.38630985	1.3306	0.18440
Year 2015	0.35853802	0.25203495	1.4226	0.15596
Year 2016	0.47649603	0.34352980	1.3871	0.16653
Year 2017	0.36409097	0.28628930	1.2718	0.20449
N	14			
T	23			
R^2	0.15356			
Adjusted R^2	0.039899			

Table 6: Estimation results of the NO_x model with year dummy variables

are still too low (especially adjusted R^2), therefore, we cannot regard this model as useful for us. In both cases of the NO_x models, GDP per capita is of too little significance, hence, we cannot conclude that there would be, according to our model, any significant relationship between the NO_x concentrations and GDP per capita. For this reason, we shall conclude that for the NO_x concentrations, in our observed period in the Czech Republic, the EKC is not a suitable concept.

6.4 Models with lags

Another option of constructing our model is to include lags of either our dependent variables or any of our independent variables. The results of this estimation (not presented in full) have shown us that neither of them is significant for the case of both the SO_2 and the NO_x model. For this reason, we shall not include lags of any variables in our models.

6.5 Further analysis of the SO_2 model

We have so far concluded that NO_x do not seem to have any significant relationship with GDP per capita. On the other hand, our SO_2 models have showed us that the relationship with GDP per capita is significant for the SO_2 concentrations. The fact that it is significant, however, does not suffice itself to either prove or disprove the EKC theory for the Czech Republic. In order to do that, we shall add interaction terms into our model, which shall show us the development of the effect of GDP per capita on the SO_2 concentrations over time.

The SO_2 model with interaction terms will be constructed as

$$\begin{aligned} log(y_{i,t}) = &\beta_1 \cdot [log(b_{i,t})]^2 + \beta_2 \cdot c_{i,t} + \beta_3 \cdot d_{i,t} + \beta_4 \cdot f_t + \delta_t + \delta_t \cdot [log(b_{i,t})]^2 + \\ &+ a_i + u_{i,t}, \qquad\qquad i = 1, \ldots, N, \; t = 1, \ldots, T. \end{aligned}$$

This model shall allow us to test how the effect of GDP per capita on the SO_2 immissions developed in time (during our observed period). Recall Figure 2. Here, it is important to note that the horizontal axis of the EKC represents GDP per capita and not time. It is, however, common to talk about the EKC concept with respect to time, as, in most cases, GDP per capita grows over time. The underlying hypothesis of the EKC is that, as countries develop over time, the level of environmental degradation firstly increases with time (and with GDP per capita) and then, from some point of GDP per capita (and also from some point in time), it decreases with time.

As we discussed in section 4, GDP per capita grew steadily in the Czech Republic during our observed period (with the exception of the 2007–2008 crisis). Therefore, in this case, examining change of GDP per capita effect on SO_2 concentrations with change of GDP per capita may be substituted for examining GDP per capita effect on the SO_2 concentrations over time. This shall provide additional evidence which will help us to determine whether the EKC theory is suitable for the case of the Czech Republic or not.

Firstly, we shall test whether the homoskedasticity assumption holds for this model. For this, we shall use the Breusch-Pagan test. The p-value calculated by this test is 0.01839. Therefore, we can reject the null hypothesis of homoskedasticity at a 1.839% significance level.

Secondly, we will test next assumption—the assumption of no serial correlation. For this, we shall use the Breusch-Godfrey test. The p-value calculated by this test is lower than 0.001, which means that we can reject the null hypothesis of no serial correlation at a significance level lower than 0.1%.

As in the case of the four previously mentioned models, we shall use a robust test statistic in order to deal with heteroskedasticity and serial correlation.

Now let us estimate this model. The results of this estimation can be seen in Table 7.

From Table 7, we can see that adjusted R^2 is lower than in the case of the SO_2 model with year dummies without interaction terms. This is an indicator of the fact that the interaction terms are not significant. Now, we can see that none of the interaction terms has p-value so that it could

Variable	Estimate	Std. Error	t-value	p-value
GDP per capita	−0.21232488	0.24199653	−0.8774	0.3811
Share of mining	4.10006714	4.99338293	0.8211	0.4123
Population density	−0.00303090	0.00331658	−0.9139	0.3616
Unemployment	−0.00086981	0.01033422	−0.0842	0.9330
Year 1995	−5.09240278	8.27511679	−0.6154	0.5388
Year 1996	−3.98573732	7.23385175	−0.5510	0.5821
Year 1997	−3.62833535	6.18167696	−0.5870	0.5577
Year 1998	−3.43817236	4.97198531	−0.6915	0.4899
Year 1999	−2.50656347	3.80402581	−0.6589	0.5105
Year 2000	−1.92374071	3.22067679	−0.5973	0.5508
Year 2001	−2.06571520	3.44885278	−0.5990	0.5497
Year 2002	−2.22102044	3.15088501	−0.7049	0.4815
Year 2003	−1.87208345	2.57006783	−0.7284	0.4670
Year 2004	−2.38044815	3.22287714	−0.7386	0.4608
Year 2005	−2.49886524	3.23216117	−0.7731	0.4401
Year 2006	−2.25822079	3.52816688	−0.6401	0.5227
Year 2007	−2.45167682	3.29464170	−0.7441	0.4575
Year 2008	−3.01399898	3.65974958	−0.8236	0.4109
Year 2009	−2.18061065	2.70462607	−0.8063	0.4208
Year 2010	−1.95734420	2.85609848	−0.6853	0.4937
Year 2011	−2.70238417	3.19008306	−0.8471	0.3977
Year 2012	−1.73471715	2.44585948	−0.7092	0.4788
Year 2013	−0.42180052	0.99028563	−0.4259	0.6705
Year 2014	−1.37294353	1.58633959	−0.8655	0.3876
Year 2015	−0.87605977	1.07177872	−0.8174	0.4145
Year 2016	−1.12641988	1.21608571	−0.9263	0.3592
Year 2017	−1.26829112	0.82812037	−1.5315	0.1325
GDP per capita ·Year 1995	0.03068820	0.33806555	0.0908	0.9277
GDP per capita ·Year 1996	0.04526115	0.35552650	0.1273	0.8988
GDP per capita ·Year 1997	−0.00664009	0.36159221	−0.0184	0.9854
GDP per capita ·Year 1998	−0.15315387	0.36637188	−0.4180	0.6763
GDP per capita ·Year 1999	−0.11103664	0.39121301	−0.2838	0.7768
GDP per capita ·Year 2000	0.00832023	0.40396911	0.0206	0.9836
GDP per capita ·Year 2001	−0.03677972	0.47409407	−0.0776	0.9382
GDP per capita ·Year 2002	−0.15910417	0.47783450	−0.3330	0.7394
GDP per capita ·Year 2003	−0.34721646	0.45637452	−0.7608	0.4475
GDP per capita ·Year 2004	−0.28440536	0.44423885	−0.6402	0.5226
GDP per capita ·Year 2005	−0.27264867	0.40288560	−0.6767	0.4992
GDP per capita ·Year 2006	−0.15138453	0.38764422	−0.3905	0.6965
GDP per capita ·Year 2007	−0.16930596	0.32057213	−0.5281	0.5979
GDP per capita ·Year 2008	−0.12253052	0.22259932	−0.5505	0.5825
GDP per capita ·Year 2009	−0.04741225	0.16521330	−0.2870	0.7744
GDP per capita ·Year 2010	−0.04351738	0.16551203	−0.2629	0.7928
GDP per capita ·Year 2011	−0.16172735	0.18705821	−0.8646	0.3881
GDP per capita ·Year 2012	−0.02650402	0.18354370	−0.1444	0.8853
GDP per capita ·Year 2013	−0.01681981	0.19329144	−0.0870	0.9307
GDP per capita ·Year 2014	−0.15287547	0.16379832	−0.9333	0.3515
GDP per capita ·Year 2015	−0.07336495	0.11690845	−0.6275	0.5309
GDP per capita ·Year 2016	−0.01743232	0.08445107	−0.2064	0.8366
GDP per capita ·Year 2017	−0.04829103	0.18291829	−0.2640	0.7929
N	14			
T	23			
R^2	0.84421			
Adjusted R^2	0.8084			

Table 7: Estimation results of the SO_2 model with year dummy variables and with interaction terms

be perceived as significant. Therefore, since the interaction terms are not significant, we can claim that it is very likely that the effect of GDP per capita on the SO_2 concentrations does not vary over time (for our observed period) in the Czech Republic, hence, the EKC hypothesis is not suitable for the case of the SO_2 concentrations.

7 Conclusion

In this bachelor thesis, we have, firstly, introduced the concept of the EKC. After that, we explained reasons for choosing, as our pollutants of interest, NO_x and SO_2 and why we decided to analyze concentrations, rather than emissions. Secondly, we described our dataset and introduced the main reasons of why the variables in our dataset have the potential to influence concentrations of the chosen pollutants. The following section was dedicated to the construction of the model itself. After that, we provided tests to determine properties of our model and then, we finally estimated our results. The outcome of our result is that, for our observed period, we do not have enough evidence that the EKC hypothesis would be valid either for concentrations of SO_2 or for concentrations of NO_x.

This conclusion is in line with Kreuz, Lisa and Šauer 2017, as their findings were that, for the case of the Czech Republic, there is a lack of empirical evidence of existence of the relationship between pollution and economic growth suggested by the EKC hypothesis. The EKC hypothesis for the case of a single country was also conducted for Canada—Ward et al. 2016. Their findings were similar to ours, i.e. that the hypothesized relationship does not seem to exist. However, their approach was different than ours—they analyzed emissions of CO_2.

On the other hand, the study Shen, Hashimoto et al. 2004, in which the case of China was studied using a similar approach to ours, i.e. analyzing panel data for provinces of China, suggests, that a similar relationship between SO_2 and income per capita exists. However, they question the entire concept of the EKC as they claim that the relationship between the two variables is too complicated as to be analyzed by a single equation.

References

[1] James Andreoni and Arik Levinson. "The simple analytics of the environmental Kuznets curve". *Journal of Public Economics* 80.2 (2001), pp. 269–286.

[2] Martin Braniš and Martina Linhartová. "Association between unemployment, income, education level, population size and air pollution in Czech cities: Evidence for environmental inequality? A pilot national scale analysis". *Health place* 18.5 (2012), pp. 1110–1114.

[3] Peter Brimblecombe and Laszlo Makra. "Selections from the history of environmental pollution, with special attention to air pollution. Part 2*: From medieval times to the 19th century". *Int. J. Environment and Pollution J. Environment and Pollution* 23.4 (2005), pp. 351–367.

[4] Jan Brůha and Milan Ščasný. "Economic Analysis of Driving Forces of Environmental Burden during the Transition Process: EKC hypothesis testing in the Czech Republic". *Paper presented at the 6^{th} International Conference of the European Society for Ecological Economics, Lisbon* (2005).

[5] Bert Brunekreef and Stephen T. Holgate. "Air Pollution and Health". *The Lancet* 360.9341 (2002), pp. 1233–1242.

[6] Matthew A. Cole, Anthony J. Rayner and John M. Bates. "The environmental Kuznets curve: an empirical analysis". *Environment and development economics* 2.4 (1997), pp. 401–416.

[7] Susmita Dasgupta et al. "Confronting the environmental Kuznets curve". *Journal of economic perspectives* 16.1 (2002), pp. 147–168.

[8] Gene M. Grossman and Alan B. Krueger. "Environmental Impacts of a North American Free Trade Agreement". *Technical report of National Bureau of Economic Research* (1991).

[9] Tomáš Hofman. "Právní úprava ochrany ovzduší v České republice a její význam pro česko-polské pomezí v oblasti Krkonoš". *Master's thesis at Západočeská univerzita v Plzni* (2015).

[10] Robert K. Kaufmann et al. "The determinants of atmospheric SO2 concentrations: reconsidering the environmental Kuznets curve". *Ecological economics* 25.2 (1998), pp. 209–220.

[11] Gary Koop and Lise Tole. "Is there an environmental Kuznets curve for deforestation?" *Journal of Development economics* 58.1 (1999), pp. 231–244.

[12] Jaroslav Kreuz, Aleš Lisa and Petr Šauer. “Environmentální Kuznetsova křivka v podmínkách České republiky v období 1975-2012”. *Politická ekonomie* 65.1 (2017), pp. 119–130.

[13] Simon Kuznets. “Economic growth and income inequality”. *The American economic review* 45.1 (1955), pp. 1–28.

[14] Jaekyu Lim. “The Effects of Economic Growth on Environmental Quality”. *Seoul Journal of Economics* 10.3 (1997), pp. 273–296.

[15] M Linscheid. “Analytical Chemistry: R. Kellner, J.-M. Mermet, M. Otto, HM Widmer”. *Angewandte Chemie International Edition* 38.13 (1999), pp. 1856–2070.

[16] Theodore Panayotou. “Demystifying the environmental Kuznets curve: turning a black box into a policy tool”. *Environment and development economics* 2.4 (1997), pp. 465–484.

[17] Theodore Panayotou. “Empirical tests and policy analysis of environmental degradation at different stages of economic development”. *Technical report of International Labour Organization* (1993).

[18] Thomas M. Selden and Daqing Song. “Environmental quality and development: is there a Kuznets curve for air pollution emissions?” *Journal of Environmental Economics and management* 27.2 (1994), pp. 147–162.

[19] Nemat Shafik. “Economic development and environmental quality: an econometric analysis”. *Oxford economic papers* 46.4 (1994), pp. 757–774.

[20] Nemat Shafik and Sushenjit Bandyopadhyay. *Economic growth and environmental quality: time-series and cross-country evidence*. World Bank Publications, 1992.

[21] Junyi Shen, Yoshizo Hashimoto et al. “Environmental Kuznets curve on country level: evidence from China”. *Discussion Papers in Economics and Business* 4.9 (2004), pp. 1–11.

[22] Mariano Torras and James K. Boyce. “Income, inequality, and pollution: a reassessment of the environmental Kuznets curve”. *Ecological economics* 25.2 (1998), pp. 147–160.

[23] Miguel Valcárcel. *Principles of analytical chemistry: A textbook*. Springer Science & Business Media, 2012.

[24] Lijun Wang et al. “An empirical study of the environmental Kuznets curve for environmental quality in Gansu province”. *Ecological Indicators* 56 (2015), pp. 96–105.

[25] James D. Ward et al. "Is decoupling GDP growth from environmental impact possible?" *PLOS One* 11.10 (2016), e0164733. DOI: https://doi.org/10.1371/journal.pone.0164733.

[26] Jeffery Wooldridge. *Introductory Econometrics: A Modern Approach*. Nelson Education, 2015.

List of tables and figures

List of Tables

1 Breusch-Pagan test p-values 34

2 Breusch-Godfrey test p-values 35

3 Estimation results of the SO_2 model without year dummy variables 36

4 Estimation results of the NO_x model without year dummy variables 38

5 Estimation results of the SO_2 model with year dummy variables 39

6 Estimation results of the NO_x model with year dummy variables 41

7 Estimation results of the SO_2 model with year dummy variables and with interaction terms 44

List of Figures

1 Kuznets curve 2

2 Environmental Kuznets curve 3

3 Concentrations of SO_2 in regions of the Czech Republic, 1995 15

4 Concentrations of SO_2 in regions of the Czech Republic, 2017 16

5 Concentrations of NO_x in regions of the Czech Republic, 1995 16

6 Concentrations of NO_x in regions of the Czech Republic, 2017 17

7 SO_2 and NO_x concentrations in the Czech Republic between 1995 and 2017 17

8 GDP per capita in the Czech Republic between 1995 and 2017 19

9 GDP per capita in the Czech Republic, 2017 19

10 Population Density in the Czech Republic, 1995 21

11 Population Density in the Czech Republic, 2017 21

12 Unemployment rate in the Czech Republic between 1995 and 2017 23

13 Unemployment rate in regions of the Czech Republic, 2017 . 24

14 Share of mining on the total sum of value added in Czech economy between 1995 and 2017 26
15 Estimated coefficients of year dummy variables with their respective standard errors 40