STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2020

An Inverted U: Testing for a Swedish Municipal Environmental Kuznets Curve with an Inequality Perspective

Johan Faxner (23861) and Carl Edvin Steinvall (23830)

Abstract: The Environmental Kuznets Curve (EKC) hypothesis suggests a relationship between environmental degradation and economic development where degradation initially increases with economic development until a certain point where additional development is associated with decreasing degradation. In the literature, the effect of economic inequality on this relationship has been studied from both theoretical and empirical viewpoints. In this thesis the EKC and the role of inequality are investigated in the 290 Swedish municipalities in the period 2008-2017. Panel data, all logarithmic, for five different types of emissions, mean earned income, squared mean income, and Gini coefficients are applied to a set of fixed effects models. We find a significant EKC-type relationship for GHGs while not considering the impact of inequality. We find mixed evidence regarding the impact of income inequality; thus, we are cautious drawing strict conclusions from its predicted effects. For robustness purposes, similar regressions are done while instead looking at counties. These regressions display similarly shaped relationships between environmental degradation and economic development, yet with no statistical significance.

Keywords: Environmental Degradation, Environmental Kuznets Curve, Income Inequality, Swedish Municipalities

JEL: D31, Q53, Q56

Supervisor: Elena Paltseva
Date submitted: May 14, 2020
Date examined: May 26, 2020
Discussants: Aron Björk and Denise Shen
Examiner: Johanna Wallenius

Acknowledgements

We would like to direct our greatest thanks to our supervisor, Elena Paltseva, for her invaluable guidance and support. We would also like to thank our friend Edvin Ahlander for his much-appreciated feedback, as well as Örjan Sjöberg for his excellent tutorials.

Any errors or mistakes are entirely our own.

Table of Contents

1. Introduction

According to the UN, quantities of greenhouse gases (GHGs) in the atmosphere have risen to a three-million-year high. Total GHG emissions increase when population sizes, economies, and standards of living advance. The rising concentrations of GHGs have been linked to a rising average temperature, with potentially disastrous effects to the climate (United Nations, n.d.). This means that as economic activity increases on a global scale, there is a possibility that environmental degradation will continue to accelerate. However, the environmental Kuznets curve (EKC) hypothesis—a concept that has been around since the early 1990s—disputes this assertion. In short, the EKC hypothesis suggests that an economy will go through two stages of development. In the first, environmental degradation will increase with economic development. At a certain point (the *EKC turning point*), the relationship will then become inverted and environmental degradation is expected to instead decrease with economic development. Graphically, this development over time can be represented by an inverted U-shape. A wide variety of forces, some of which will be detailed in this thesis, have been hypothesized to drive this relationship.

Sweden is an example of a country where emissions of GHGs have generally been decreasing over the past 30 years. Between 1990 and 2018, GHG emissions decreased by circa 27 percent (Statistics Sweden, 2019e). Meanwhile, real gross domestic product (GDP) per capita in Sweden rose by about 55 percent during the same period (World Bank, 2020). Within this time frame, the share of value added that is attributable to industry, a notoriously emissions-intense activity, has decreased from 24.41 percent to 18.69 percent, while the share attributable to services has increased from 65.2 percent to 72.99 percent (OECD, 2020). These data are consistent with what some proponents of the EKC hypothesis would suggest: highly developed economies will undergo structural shifts that cause them to decrease their emissions (Panayotou, 1993). There is reason to believe that Sweden is an example of such a country.

However, the fact that Sweden has decreased GHG emissions does not necessarily mean that individual Swedish municipalities (Swedish: *kommuner*) have. The amount of emissions can vary greatly between municipalities. This means that the trends in municipalities that are large emitters can overwhelm trends in smaller ones, leading to results at a national level that differ significantly from what can be observed in several municipalities. It is possible that large municipalities have generally decreased their emissions, while smaller municipalities have not, which would fit in with the national data just as well. See figure A1 in appendix A for some examples of municipalities that do not show straight downward slopes for environmental degradation, but rather inverted U-shaped curves.

The Government of Sweden and the Riksdag have a responsibility to enact policies and legislation regarding environmental issues. Municipalities are then responsible for adjusting locally to these national policies and legislation (Swedish Association of Local Authorities and Regions, 2014). National policies can thus lead the country toward sustainability, but local government does have influence on the specific level of environmental degradation that will occur in a specific municipality. This is, perhaps, one of the main reasons why studying the EKC at a municipal level can potentially yield interesting results, especially when allowing for income inequality to differ across municipalities.

In the literature, it has been theorized and empirically investigated how inequality might influence the relationship suggested by the EKC hypothesis. Theoretically, equality might put the median agent in society in a better position, which can increase general awareness of environmental degradation, leading environment-enhancing regulations to be put into place (see e.g. Kaika and Zervas, 2013). Increased economic equality could also increase the willingness to pay for environmental protection (see e.g. Magnani, 2000). Economic inequality, on the other hand might cause increased status consumption that would increase emissions (see e.g. Jorgenson et al., 2017). Empirically, it is somewhat ambiguous how inequality influences the EKC-pattern. E.g. Magnani (2000) found empirical results, limited to OECD countries, supporting the hypothesis that income inequality was negatively associated with expenditure related to research and development for environmental protection. Oppositely, Brännlund and Ghalwash (2008) found, based on Swedish household consumption of non-durable goods, that inequality might have the opposite effect as the relationship between income and emissions from consumption was shown to be concave, implying that, *ceteris paribus*, a more equal income distribution could influence emissions positively.

In this thesis, we will empirically test the EKC hypothesis on a Swedish municipal level. This will be done using a fixed effects approach based on a panel data set, covering 290 Swedish municipalities throughout the period of 2008-2017. We will test EKC models that include and do not include income inequality in order to see whether it has the effects on the EKC that some of the previous literature has found it to have. In our case, we use mean earned income for economic development, the Gini coefficient for economic inequality, and five different pollution metrics (per capita) for environmental degradation. These five environmental degradation indicators are meant to capture two different types of environmental degradation—both in terms of effects on the climate and on local environments. Carbon dioxide, nitrous oxides, and aggregate GHGs capture greenhouse gases that principally have a more global effect on the environment. Sulfur dioxide and particulate matter, on the other hand, have consequences that are relatively more local.

In short, the contribution that we seek to make to the literature is that we test a hypothesis, that has been widely researched with mixed results, on a set of data which has not been analyzed through this specific lens before. We also aim to contribute to the literature by allowing for income inequality in our model, which enables us to tell a more nuanced story of the relationship between economic development and environmental degradation in Sweden.

Our estimation results ultimately leave us unable to draw unequivocal conclusions regarding the Swedish municipal EKC and the effects that economic inequality has on it. GHG emissions is the sole environmental degradation indicator for which we can find a statistically significant EKC shape with the methods used in this thesis. We calculate the Swedish municipal EKC turning point for GHGs to be situated at around 312,000 SEK in 2008 prices. The effect that the Gini coefficient has on the model is likewise rather uncertain. In the case of GHGs, our estimates indicate that higher inequality generally leads to higher emissions, but that it also leads to an EKC turning point that occurs earlier. However, we cannot draw definite conclusions on whether inequality impacts the EKC for GHGs since we obtain significant results only when we apply Driscoll-Kraay standard errors, which are not optimal for the dimensions of our panel data set.

We will begin in section 2 by briefly reviewing the literature on the EKC hypothesis, its theoretical foundations, the impact of economic inequality on the EKC, and some common critiques of the EKC framework. In section 3, we detail our research contribution, including our research questions and limitation of scope. Then, in section 4, we present our methodological approach and construct a set of models that we will test, based on data that is presented in section 5. Section 6 will be dedicated to presenting the results of the models and in section 7 we will perform some robustness checks. In section 8, we discuss our findings, including limitations regarding their interpretability, and compare them to the results in the robustness checks. In section 9, we summarize our findings and conclusions.

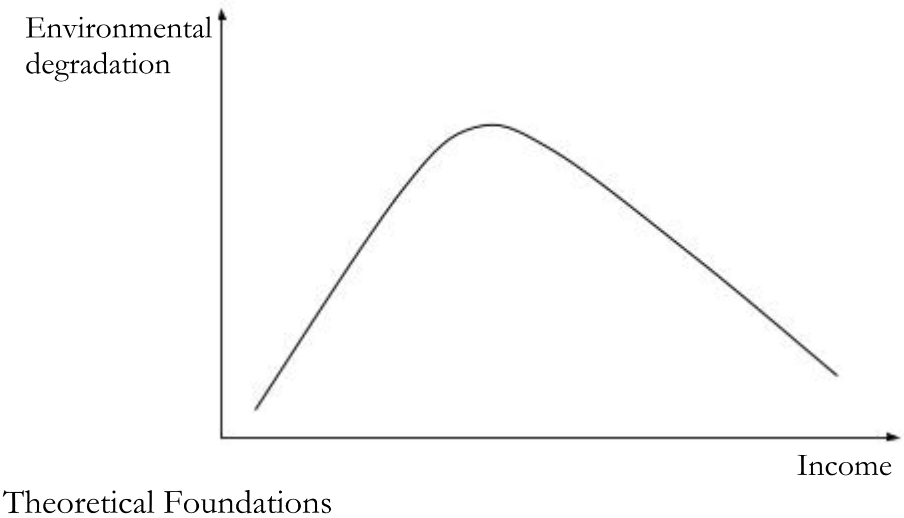
2. Literature Review

In this section, we start by providing a brief overview of the EKC hypothesis. We then review the EKC literature that focuses on municipal data. That is then followed by literature connecting the EKC with income inequality. We also highlight a couple of studies that, regardless of whether the inequality dimension is included or not, have studied the phenomenon in Sweden and found results that may be instrumental for our analysis.

2.1 The Environmental Kuznets Curve

The EKC is a suggested relationship between environmental degradation and economic development. The name of this phenomenon relates to the original Kuznets curve, which suggests there to be an inverted U-shaped relationship between economic inequality and economic development, based on research findings by Kuznets (1955). In a similar fashion, the EKC is suggested to show an inverted U-shaped pattern for the relationship between environmental degradation and economic development. As economies industrialize, emissions and other environmental degradation indicators can be expected to worsen. However, when economies reach a certain point of development, these indicators can be expected to start improving, according to the EKC hypothesis, see Figure 1 for potential graphical interpretation.

Figure 1: Graphical interpretation of a possible EKC (made by the authors)



2.2 Theoretical Foundations

There are several economic factors suggested to affect the relationship between economic development and the environment. Essential factors suggested by Panayotou (1993, p. 2) are "(a) the level of economic activity or size of the economy; (b) the sectoral structure of the economy; (c) the vintage of technology; (d) the demand for environmental amenities; and (e) the conservation and environmental expenditures and their effectiveness." How these factors play out is expected to change as an economy develops. These changes are conducive to an EKC relationship. The intuition behind this is raised by Panayotou (1993). For instance, the eventual downward sloping part of the EKC could be understood as being caused by enhanced technology, more money spent on environmental protection, greater awareness for the environment, and environmental regulations being enforced. These changes are co-occurrent with a transition toward a more "information-intensive" industrial composition (Panayotou, 1993, p. 1).

Related to (d), Shafik and Bandyopadhyay (1992) state that the environment functions both as an input in production as well as a consumption good; thus, income elasticity of demand and supply of environmental amenities will influence how resources are used as the economy develops. Panayotou (1993) reasons that demand for environmental amenities is elastic to income and that it is first at somewhat high incomes where a meaningful budget share is allocated to it.

In a review of the EKC literature, Kaika and Zervas (2013) also raise some potential driving forces of an EKC-pattern. Like Panayotou (1993), they discuss the aspects of size and structure of economic activity, as well as technological aspects. Related to this is the *scale effect* of production on the environment, which refers to a situation where environmental degradation is a consequence of increased production and extraction of natural resources. The scale effect potentially causes the initial increase in environmental degradation. However, the scale effect is suggested to be followed by a *composition effect* and a *technique effect* that eventually cause a decrease in degradation. The composition effect refers to the effect that comes from a transition of the output of the economy where the economy transitions from consisting of industries that intensively use energy and material to service industries having less environmental impact. The technique effect comes, *inter alia*, from technological advancements enabling a more efficient ratio of input to output in industries.

Another potential driving force covered by Kaika and Zervas (2013) is equity in distribution of income. Essentially, the idea is that an equitable income distribution puts the median person in society in a better position; consequently, the general awareness for the degradation of the environment increases and in turn environment-enhancing regulations are put into place. Just like mentioned by Kaika and Zervas (2013, p. 1394-1395), a central issue here is what implications economic development has on distribution of income, i.e. if economic development is followed by income equality or not.

2.3 Evolution of the EKC Literature

The idea of an EKC—while not explicitly using that terminology—was introduced by Grossman and Krueger (1991) in a paper that sought to assess the future environmental impacts of the North American Free Trade Agreement. The authors explored concentration patterns of sulfur dioxide (SO_2), dark matter (fine smoke), and suspended particulate matter (SPM) from a dataset encompassing the period 1977-1988 and several cities in different countries (the number of cities and countries varied depending on year and pollutant) regarding national income per capita. They used random effects models for the three pollutants and found evidence indicating that pollution tends to increase with income, until a certain turning point where pollution starts to decrease with income.

Another early influential study related to the EKC hypothesis was conducted by Shafik and Bandyopadhyay (1992). To assess the nature of the relationship between environmental degradation and income, they tested ten measurements of environmental degradation using three different types of models. One model was linear, one was quadratic, and one was cubic. All three models were expressed so that both environmental degradation and income variables had logarithmic forms. The data used was from the period 1960-1990 and consisted of observations from up to 149 countries. However, the available years and countries varied between different environmental degradation measures. In the basic models, the independent income variable consisted of the logarithm of PPP-adjusted income per capita, as well as a time trend to capture technological advancements and a varying constant term for different countries or cities. The time trend was not used while testing deforestation rates and municipal waste. Moreover, dummies controlling for specific effects were used when running regressions for SPM and SO_2, as well as for dissolved oxygen and fecal coliform in rivers.

The measurements of environmental degradation in Shafik and Bandyopadhyay's (1992) study that significantly followed an EKC-pattern were SPM and SO_2; however, income had a significant impact on almost all measurements but not necessarily in a way that follows an EKC-pattern. Indicators that were concluded to only worsen with income were dissolved oxygen in rivers, generated municipal waste per capita and carbon dioxide (CO_2) emissions. In the case of CO_2, Shafik and Bandyopadhyay (1992) argued that costs that come with the emissions of CO_2 are not internalized where these occur, but rather that they are somewhat evenly distributed globally; consequently, the issue of CO_2 emissions portrays a *free rider problem*. The opposite, however, was the case for SPM and SO_2 which generate external costs geographically close to where the emissions occur but are, in relative terms, costly to abate. Presumably, this will be reflected in the social choices about dealing with environmental degradation. The net benefit of abatement (weighting costs with private and social benefits) will affect whether, and at which point, degradation will be dealt with. SPM and SO_2 are generally dealt with in medium-income countries since at this stage of the economic development, as a result of industrialization and greater energy intensity, these pollutants are assumed to become increasingly troublesome.

Panayotou (1993) conducted another one of the earlier empirical studies on the EKC. As indicators for environmental degradation, Panayotou used deforestation rates as well as per capita emissions of SO_2, nitrogen oxides (NO_X), and SPM in a sample of developed and developing countries. Income data from 1988, SPM data from 1987, and SO_2 and NO_X data from the late 1980s were collected. The dependent variables were in logarithmic form. This also applied to the independent variables—the natural logarithm of per capita income and the squared natural logarithm of per capita income. In the case of deforestation, the dependent variable was also expressed as a function of population density. An ordinary least squares-based approach was used to conduct the analysis and Panayotou found support for the EKC hypothesis for all the environmental degradation indicators that were used.

The EKC hypothesis has received a significant amount of criticism over the years. Especially the earliest studies were subject to criticism for a lack of econometric sophistication. Stern (2004, p. 1420) points out that few studies have considered some of "the statistical properties of the data used—such as serial dependence or stochastic trends in time series—and little consideration has been paid to issues of model adequacy such as the possibility of omitted variable bias." Issues with the way that certain studies have utilized cointegration techniques have also been raised, for example by Wagner (2015). These points of criticism have been partially responsible for how research on the EKC hypothesis has developed over time. While the typical basic version of the EKC model is still predominantly used, Stern (2017) finds that alternative approaches such as non-parametric approaches, decomposition analysis, and convergence analysis have increasingly been applied to the problem as well. The criticism brought forth by Stern (2004) has influenced other work, such as that by Narayan and Narayan (2010) which applied a panel cointegration approach to the EKC hypothesis in order to mitigate purported multicollinearity problems and found conflicting results with regard to EKC evidence. *Carbon leakage*, meaning that environmentally destructive production is moved to other jurisdictions (sometimes labeled *pollution havens*) by means of international trade, has also been suggested as a possibly significant cause of a reported EKC shape by e.g. Steinkraus (2016). However, the empirical evidence in favor of such phenomena is still sparse and somewhat controversial (Franzen and Mader, 2018).

2.4 EKC and Subnational Geographical Units

The EKC hypothesis has been tested a number of times on diverse types of subnational regional units around the globe. Keene and Deller (2013) found evidence in favor of an EKC pattern for $PM_{2.5}$ in U.S. counties. They also found that high degrees of social capital, ruralness, and education rates pulled down the EKC, while inequality and racial fragmentation pushed it upward. Furthermore, a wide array of studies has been made on municipal solid waste growth and $PM_{2.5}$ concentrations in prefecture-level cities and municipalities within an EKC framework, especially in China. See e.g. the work by Cheng et al. (2020) for the former case and Wang and Komonpipat (2020) for the latter. Both studies found evidence in favor of a modified EKC, with an N-shaped pattern rather than an inverted U-shaped one, as a cubic income variable was found to have a significant positive impact on environmental degradation.

Among the previous studies done on a potential Swedish EKC is an article by Marbuah and Amuakwa-Mensah (2017), which examined Swedish municipalities, while also considering spatial dependence, for the years 2005 to 2013. The authors tested for the relationship that emissions of CO_2, SO_2, NO_X, and carbon monoxide, along with particulate matter PM10 and PM2.5 and total suspended particulates, might have with real income per capita. They also tested for spatial effects, seeking to capture any spillover that might exist between different municipalities. Such spillover effects were hypothesized to arise from assumptions such as that neighboring municipalities might, for instance, have similar economic activities, similar meteorological conditions, and co-operation surrounding environmental issues. The results of the article indicate that there is evidence for an EKC in Swedish municipalities for all the tested environmental indicators, apart from carbon monoxide which was not significantly displaying an EKC pattern. They also found significant spatial dependence for all emission types, suggesting that there are spillover effects.

2.5 Income Inequality Within the EKC Context

One early step towards establishing a link between allocation of power, income distribution and environmental degradation was made by Torras and Boyce (1998), who hypothesized that as the distribution of power equalizes among a population, the quality of air and water should improve in the geographical area that is inhabited by that population, holding all else equal. The authors arrive at this hypothesis through a set of assumptions. Perhaps the chief assumption is that income is associated with the benefit gained from activities that generate pollution, meaning that high-income people enjoy a greater benefit from it. This is because higher-income people tend to own more assets and consume more, so they achieve higher surplus on both the producer and consumer side of transactions, including those causing pollution. Since they also assume that power inequality is a function of economic inequality, this will lead to less pressure on governments in highly unequal jurisdictions to adopt stringent environmental regulations than in ones that are more economically equal.

Using OLS, Torras and Boyce (1998) tested five different pollution variables, as well as water potability and sanitation rates, in two different regression models. The independent variables in both models included GDP per capita and geographical characteristics. In one of the models, literacy rates, political rights, Gini coefficients, dummy variables for low- and high-income countries (with 5,000 USD GDP per capita, PPP-adjusted), and interaction terms were added as independent variables. The results from the regressions appear to support the EKC hypothesis for several indicators. Upon comparing the results of both models, they found that the effects that income has on pollution become, in general, less statistically significant when inequality variables are introduced into the model. The significance of the effects that the different inequality measures had on the different pollution variables varied. The authors of the paper interpreted their results as largely backing their hypothesis.

Magnani (2000) approached the EKC by investigating the role of inequality for policy decisions regarding the environment within rich countries. More precisely Magnani looked at public expenditures for research and development related to the protection of the environment. First, Magnani discussed the expression of abatement of pollution as a function of some measure of economic well-being. The change of abatement happens with respect to the demand for environmental quality which in turn change with respect to income per capita. With respect to income per capita, the change of abatement is greater than zero, based on the environmental literature according to Magnani (2000). This can be understood by increased demand for environmental quality when per capita income rises. Second, Magnani discussed a *relative-income effect* upon which the degree of environmental protection depends. The relative-income effect is understood by inequality of income. If the inequality is large, the median voter will have lower income, relatively speaking, and will be more inclined to spend money on consumptive private goods rather than contributing to public expenditures for the environmental quality; hence, the relative income effect explains the willingness to pay for environmental policies.

Magnani's (2000) empirical examination was completed using data from OECD countries in the period 1980-1991. The dependent variable used was the logarithm of public research and development expenditures for environmental protection. Independent variables were GDP per capita (PPP-adjusted) (a linear and a quadratic term), one of two indicators of income equality/inequality, i.e., either a ratio between the percentage of income of the first and the fourth quintile of the population or a Gini coefficient, and for some regressions also a time trend. An interaction term was also added. The regressions supported the hypothesis that inequality affects public choices made about the environment. One conclusion drawn is that the latter part of the EKC-pattern is displayed in rich countries given that economic growth does not cause substantial incremental inequality.

One study looking at the relationship between emissions and income in Sweden, while considering income distribution, was conducted by Brännlund and Ghalwash (2008). In their micro approach, they looked at consumption on a household level. They modeled an equation where a household's emissions are a function of a basket of goods that the household consumes. In turn, each good is a function of price and household income. If household income changes, the content of the basket of goods changes and consequently the household emissions change since some goods cause more emissions than others. From the equation on individual household emissions, an equation for average household emissions was derived. Data on household consumption, income, and characteristics came from surveys in 1984, 1988 and 1996. The authors concatenated emissions data from Statistics Sweden to derive the share and intensity of emissions from different non-durable goods in terms of CO_2, SO_2 and NO_X. Conclusions drawn from the regressions were that, in the area around the observed data, household emissions increased with income. However, the pace of the incremental emissions seemed to slow down with income. Given this concluded shape of the emissions-income relationship, holding everything else equal, a hypothetical redistribution of income from a high-income to a low-income household would increase emissions.

Jorgenson et al. (2017) studied whether CO_2 emissions at the state level in the United States depended on economic inequality within the states. The authors mention three ways that the literature has suggested how inequality possibly could influence emissions. First, they discuss the view that the wealthy benefit more from pollution-generating activities due to ownership of companies (a similar argument to the one presented by Torras and Boyce, 1998) and due to it being relatively easier to protect themselves from the negative consequences that comes with pollution. Second, they discuss an aspect reminiscent of e.g. the study by Brännlund and Ghalwash (2008), namely the aspect of how the "propensity to emit", through consumption, could change with income. Third, Jorgenson et al. discuss the possibility that competition in consumption may increase in unequal societies. This is partly caused by Veblen effects, i.e., that inequality leads to status consumption, which in turn leads to higher levels of emissions. Two different inequality metrics, namely the Gini coefficient and the income share of the top ten percent of earners, were used. The results of the study while controlling for, *inter alia*, GDP per capita, implicated that a higher income share of the highest earners was related to more emissions. The Gini coefficient on the other hand did not seem to have a significant impact.

A recent study, conducted by Ridzuan (2019), empirically tested the effect that the Gini coefficient has on emissions of SO_2 within the EKC context. Ridzuan examined 174 countries, using data from 1991 up to and including 2010 on GDP and SO_2, both per capita, and found evidence in favor of an EKC. An interaction term, defined as the product of GDP per capita and the Gini coefficient, was also introduced. This variable had a positive coefficient, indicating that increased economic inequality contributes to a higher (in terms of income) EKC turning point. Therefore, higher levels of economic inequality were inferred to yield higher levels of SO_2 emissions at any given level of income.

3. Research Contribution

Many studies have aimed at testing the EKC using different methods, data sets and indicators of environmental degradation. Some has also investigated the relevance of inequality to EKC-relationship (see e.g. Torras and Boyce, 1998; Magnani, 2000; Jorgenson et al., 2017). The environmental degradation-income relationship and the EKC have also been looked at in a Swedish context. Marbuah and Amuakwa-Mensah (2017), while considering spatial dependence, found evidence in favor of a Swedish EKC for several emissions using municipal data. We look at similar types of emissions using equivalent income data, yet, we look at a more recent time period and instead of investigating spatial dependence we weigh in potential effects of inequality. Furthermore, Brännlund and Ghalwash (2008) looked at the degradation-income relationship while considering income allocation. This study focused on households, whereas our area of focus are municipality-wide emissions and income.

Using updated data, we aim to see whether a Swedish EKC exists on a municipal level. We also aim at building upon the literature that has examined the role that income inequality plays in the EKC context, such as the works by Torras and Boyce (1998), Magnani (2000), and Jorgenson et al. (2017). Thereby, this thesis will aim at answering the following questions:

3.1 Research Questions

1) Is there a discernible EKC-type relationship between income and environmental degradation in Swedish municipalities?

2) If so, does inequality in the distribution of income affect the nature of that relationship?

3.2 Limitation of Scope

It is important to note that we do not heavily consider much of the criticism that has been put forth by e.g. Stern (2004; 2017) and Wagner (2015). Rather than focusing on novelty of the econometric approach to the EKC hypothesis, we use an approach that is consistent with the general modus operandi of the field, but apply it to a more nuanced, and richer, EKC relationship. In doing so, we hope to yield results that are easily interpretable through an already existing lens. Another limitation is that the only inequality that we consider concerns that of income. Potentially other types of inequality, such as political inequality, could be impactful in the EKC context. Torras and Boyce (1998) suggested that power inequality is a function of not only income inequality and per capita income, but also non-income determinants of power such as political rights and civil liberties. However, this paper only investigates the aspect of economic inequality. With that said, it is possible that different types of inequalities have different effects on the EKC, which may be warranted to also conduct research on. However, in a Swedish context, literacy and political enfranchisement are effectively universal, making it difficult to meaningfully apply such metrics to a Swedish context in the 21st century. Furthermore, we will not seek to analyze any other potential forms of the EKC, such as the cubic (N-shaped) form, that certain others have found evidence for.

4. Method

4.1 Methodological Approach

In order to test whether the municipal EKC exists and whether income inequality plays a key role in its dynamics, we will apply a panel data-based regression analysis using fixed effects (FE) models. The FE approach enables us to regress our dependent variables against our independent variables, while controlling for potential individual factors of the municipalities that are constant over time, i.e. the municipal effects. Potentially, we believe such factors could be where the municipality is located, whether there are natural resources in the area or whether a municipality is urbanized. While using this approach it is essential that there is some variation of the independent variables over time, i.e. for the use of the FE approach to make sense in our case we must assume that income and the Gini coefficients will vary within our ten-year period. As part of a robustness check, we will evaluate whether a random effects (RE) approach would have been more suitable than a FE approach by using correlated random effects (CRE).

4.2 Specification of Models

Table 1: Descriptions of non-logarithmic variables

	Variable description	***Unit***
GHG	Greenhouse gases (CO_2 equivalents)	Tons per capita
CO_2	Carbon dioxide	Tons per capita
SO_2	Sulfur dioxide	Tons per capita
NO_X	Nitrous oxides	Tons per capita
PM10	Particulate matter (<10 μm)	Tons per capita
Inc	Real mean income of persons aged 20+, 2008 prices	SEK, thousands
Gini	Gini coefficient	Continuous, 0-1

Table 1: Descriptions of non-logarithmic dependent and independent variables.

We begin by testing whether there is an EKC for our five environmental degradation metrics. See appendix B for short explanations of the five types of emissions that we test for, including some of their main sources and impacts on health and on the environment. We include income and income squared as independent variables in our fixed-effects regressions.

Based on an overview of the EKC made by Stern (2017), the typical way of specifying the regression model of the EKC in the literature is to model the natural logarithm of the dependent variable, i.e. the natural logarithm of some environmental quality indicator or per capita emissions, against the natural logarithm of GDP per capita and the squared natural logarithm of GDP per capita—including country effects, time effects, and an error term. The consequence of including the country effects is that the income elasticity of emissions is identical across countries at a certain income. In our case, "country effects" should rather be described as "municipal effects." Furthermore, when logarithms are used, estimations of the dependent variable will be greater than 0 and is in most cases a relevant trait. The direct reading of elasticities motivated the logarithmic transformation made by Panayotou (1993). Following these examples, we transform the variables to natural logarithms.

Model 1: Basic EKC model

$$ln(DEG_{it}) = \beta_1 ln(Inc_{it}) + \beta_2 (ln(Inc_{it}))^2 + \delta_1 D_{2009} + \cdots + \delta_9 D_{2017} + a_i + u_{it}$$

where DEG_{it} signifies the tons per capita emissions of GHGs, CO_2, SO_2, NO_X, and PM10, respectively, for municipality i in year t. Inc_{it} is the real mean income; $D_{2009} + \cdots + D_{2017}$ are time intercept dummies for all years in the period, except for 2008; a_i are the individual municipal effects that are constant over time for each municipality; and u_{it} is the error term. The fixed effects approach deals with how the values of the variables deviate across the years within each municipality i from the mean of each variable therein and how those variations may or may not explain variations in the dependent variable.

Five models will be tested: one for each of the five environmental degradation measures. In accordance with the EKC hypothesis, we expect that $\beta_1 > 0$ and $\beta_2 < 0$. This would satisfy the criteria for an EKC. We will then calculate the turning point of the EKC, i.e. at which mean income level environmentally detrimental emissions are expected to begin to decrease. This turning point will only be calculated if we for any emissions find statistically significant parameter estimates that are consistent with the EKC hypothesis.

EKC turning point

$$= \exp[-\beta_1/(2\beta_2)]$$

We then continue by extending our basic EKC model to include the natural logarithm of the Gini coefficient. This is done to test whether the level of income inequality affects the environmental degradation-income relationship, regardless of whether it e.g. stems from equality putting the median agent in society in a better position (see e.g. Kaika and Zervas, 2013), increasing the willingness to pay for environmental protection (see e.g. Magnani, 2000) or inequality causing increased status consumption (see e.g. Jorgenson et al., 2017). We also create an interaction term of Gini and the natural logarithm of income. This variable allows the EKC turning point to move to the left or to the right, depending on the degree of income inequality in a municipality, meaning that we no longer need to worry about the implicit assumption that the EKC is identical across all municipalities (Ridzuan, 2019). Since we are only concerned with how income inequality affects a potential EKC, we will test model 2 for dependent variables producing parameter estimates of an EKC in model 1.

Model 2: Extended EKC model with Gini coefficient included

$$\ln(DEG_{it}) = \beta_1 ln(Inc_{it}) + \beta_2 (ln(Inc_{it}))^2 + \beta_3 ln(Gini_{it}) + \beta_4 Gini_{it} \times ln(Inc_{it}) + \delta_1 D_{2009} + \cdots + \delta_9 D_{2017} + a_i + u_{it}$$

Our hypotheses for the parameter estimates remain the same for this model as for the one before it, insofar that the variables are included in both. We also hypothesize that $\beta_3 > 0$, which we base on, *inter alia*, what was theorized by Torras and Boyce (1998), meaning that higher income inequality is associated with higher per capita emissions. Moreover, our hypothesis is also that $\beta_4 > 0$, meaning that income inequality corresponds with EKC turning points that occur at relatively higher levels of income. The hypothesis for the latter is largely based on the results retrieved by Ridzuan (2019).

5. Data

To test our models and hypotheses, we have collected data for all 290 Swedish municipalities, covering a ten-year period (2008-2017). All data has been accessed from Statistics Sweden. The data has then been compiled into a panel format of 2,900 individual observations. Since we have observations for all 290 municipalities and all ten years, the panel is initially perfectly balanced. The time span applied in this study is mandated by the years for which this specific type of emissions data is available and comparable across the years.

A strength with using municipal data, as we see it, is that we can be confident that the data is consistent regarding definitions and methods of collection. When using cross sections of different countries, differences in such matters may lead to difficulty when comparing the data. In our case, it can safely be assumed that data on population, income, and emissions have been collected in the same way in all 290 municipalities. A downside, however, is that some municipalities may have economies that are dominated by a single, large firm that has located a factory there. In such cases, the effects of the actions of one single factory may produce results, and possibly patterns, that on the surface appear to be endemic to the municipality but in fact could as well be due to the, perhaps capricious, actions of a single firm.

5.1 Dependent Variables

As dependent variables, we use annual air emissions of greenhouse gases (GHGs) in carbon dioxide (CO_2) equivalents, CO_2, sulfur dioxide (SO_2), nitrous oxides (NO_X), and particle matter with a diameter of less than 10 μm (PM10). These variables are largely the same as the ones tested by Marbuah and Amuakwa-Mensah (2017), apart from that we have elected to omit particulate matter with a diameter less than 2.5 µm (PM2.5), total suspended particles, and carbon monoxide from our analysis and added GHGs to it. We also use a different time frame, 2018-2017 rather than 2005-2013. The GHGs used are CO_2, nitrous oxide (N_2O), methane (CH_4), hydrofluorocarbons (HFC), perfluorocarbons (PFC), and sulfur hexafluoride (SF_6). Each gas has been converted into CO_2 equivalents using a Greenhouse Warming Potential factor (Statistics Sweden, 2019c). The emissions data is based on a residential principle, i.e. "emissions arising from activities of Swedish companies and households (resident units), are accounted for regardless of where these emissions actually occur." (Statistics Sweden, 2019b) Furthermore, national emissions are distributed between different subnational geographical units using different distribution keys (for further details, see Statistics Sweden, 2019c).

All emissions data in this thesis have been converted to tons per capita measurements. These per capita measurements are based on population data for each municipality as of November 1 each year. Descriptive statistics of these five per capita variables are found in table 2. As previously mentioned, we are using the natural logarithm of each of these emissions metrics. There were 17 instances where zero tons of SO_2 was observed to be emitted in a municipality in a year. Since it is not possible to logarithmically transform zeroes, we thus have 17 missing values in our dataset. Descriptive statistics of the logarithmic emissions variables are found in table 3.

5.2 Independent Variables

The first independent variable that we use is mean earned income. Income data has been retrieved from Statistics Sweden (2020d) and concerns the mean income of all citizens aged 20 years or above in each municipality, each calendar year. This is the same income variable as the one used by Marbuah and Amuakwa-Mensah (2017). Earned income in this context should be interpreted as "taxable income from employment, business income, pensions, sickness benefit and other taxable transfers." (Statistics Sweden, 2020d) We converted the income data from nominal income to real income in 2008 SEK (expressed in thousands SEK). The conversion was based on consumer price index (CPI) data retrieved from Statistics Sweden (2020a) for which the CPI was indexed so that 1980=100. The natural logarithm of mean income is used. This variable is also used in its squared form. Note that the logarithmic transformation of the income variable occurs before the squaring of it does.

We also include the natural logarithm of the Gini coefficient in each municipality for each year. The Gini coefficient is a measure of the inequality of an income distribution within a given population. It ranges from 0 to 1 where 0 means perfect equality and 1 means perfect inequality. The Gini coefficient is derived from the Lorenz curve, which is a curve that plots the cumulative income share against the cumulative population share, and the line of equality, meaning the Lorenz curve that would have existed if every person in the population had the same income. The Gini is the quotient given by the area between the Lorenz curve and the line of equality divided by the area under the line of equality (see *Figure 2* for illustration).

Figure 2: Line of equality and the Lorenz curve

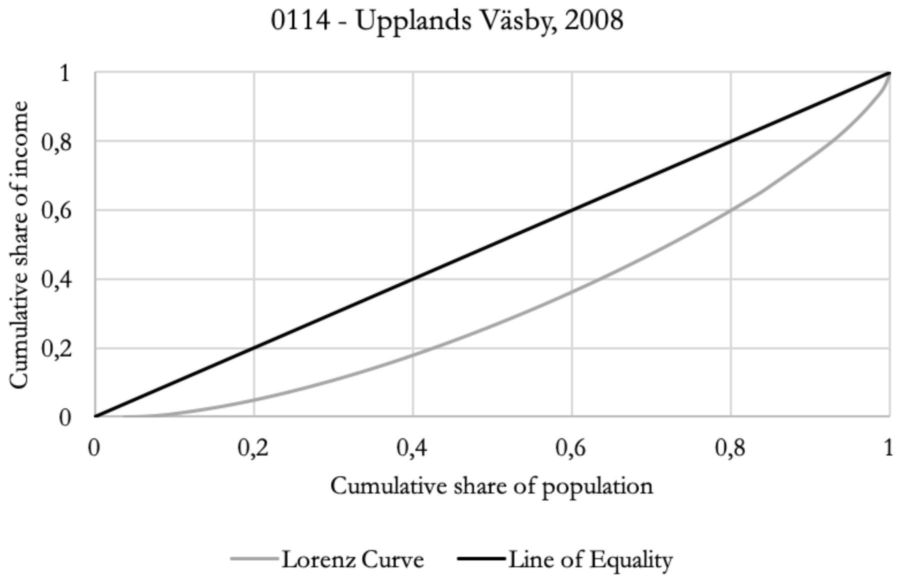


Figure 2: The line equality and the Lorenz curve for the municipality Upplands Väsby in 2008. Graph made by the authors in Excel using income data accessed from Statistics Sweden (2020d).

To generate a Gini coefficient for a given municipality in a given year we have used data from Statistics Sweden on the total sum of earned nominal income and the number of people in different income classes for residents aged 20 years or above. The first income class has an earned income of zero, the following income classes each have a width of 20,000 SEK up to an earned income of 400,000 SEK. Beyond 400,000 SEK the income classes have a span of 100,000 SEK up until an earned income of 600,000 SEK where the spans are 200,000 SEK. Finally, there is an income class that includes all those with an earned income of more than 1,000,000 SEK. One implication of this is that the accuracy is not as high for incomes of 400,000 SEK or more, since people with widely different levels of income are assigned to the same income class. For income classes with observations of fewer than four people, the data has been anonymized by Statistics Sweden. This means that the exact number of people and total sum of income in that specific income class is not known to us, which we elected to solve by replacing that unknown number with a zero. We believe there is a risk that this might cause an underestimation of the Gini coefficients since the missing values generally occur in the two very upper income classes. However, as this only occurs in the cases of fewer than 4 observations, we expect this potential underestimation to be rather limited.

With this data, the first step to generate the Gini for any municipality in any year was to calculate the share of the population and the share of total income that corresponded to respective income class. The second step was to calculate the cumulative income – increasing as one moves to higher income classes. Thirdly, the integral of the Lorenz curve was derived step-by-step by taking the average of two adjacent points of cumulative income along the curve and multiplying their average with the population share of the associated income class. One can imagine a continuous curve, with cumulative income share on the Y-axis and cumulative population share on the X-axis, which is divided up into bar charts where the area of each bar in the bar chart is calculated separately. As described in the previous paragraph, with the integral of the Lorenz curve we derived the Gini coefficient. These calculations were repeated for all 2,900 observations.

5.3 Descriptive Statistics

Table 2: Descriptive statistics of non-logarithmic variables

Variable	*Mean*	*SD*	*Min*	*Max*	*N*
GHG	7.372796	10.93427	1.07932	207.7515	2900
CO_2	5.618703	10.86873	.9774972	207.3935	2900
SO_2	.0038316	.0102928	0	.1672698	2900
NO_X	.0225141	.0260283	.0033603	.4939016	2900
PM10	.0067776	.0055729	.0012626	.0818186	2900
Inc	247.72	34.92618	183.5919	510.6758	2900
Gini	.3306008	.0326549	.2621668	.5179158	2900

Table 2: Descriptive statistics of non-logarithmic dependent and independent variables. Means in table 2 are means for which each observation for respective municipality and year have been equally weighted. Observations are from the 290 Swedish municipalities in the years 2008-2017. Original data retrieved from Statistics Sweden (2019b; 2020d). Emissions metrics have been transformed to per capita measures by the authors based on population data, also retrieved from Statistics Sweden (2019d). Real mean income (2008 prices) has been calculated by the authors based on consumer price index data from Statistics Sweden (2020a). The Gini coefficients have been calculated by the authors, using mean income data from Statistics Sweden (2020d).

Table 3: Descriptive statistics of logarithmic variables

Variable	***Mean***	***SD***	***Min***	***Max***	***N***
ln GHG	1.748803	.577437	.0763311	5.336343	2900
ln CO_2	1.403915	.6015381	-.0227598	5.334618	2900
ln SO_2	-6.901542	1.443133	-10.65827	-1.788147	2883
ln NO_X	-4.045947	.6294384	-5.695714	-.7054189	2900
ln PM10	-5.181368	.5856488	-6.674582	-2.503251	2900
ln Inc	5.503719	.1268816	5.212715	6.235735	2900
$(ln\ Inc)^2$	30.30702	1.420196	27.1724	38.88439	2900
ln Gini	-1.11131	.0926547	-1.338775	-.6579426	2900

Table 3: Descriptive statistics of logarithmic dependent and independent variables. Means in table 3 are means for which each observation for respective municipality and year have been equally weighted. Observations are from the 290 Swedish municipalities in the years 2008-2017. 17 observations of the natural logarithm of sulfur dioxide have been categorized as missing values, due to the impossibility of taking the natural logarithm of zero.

5.4 Model and Data Validity Concerns

A concern central to the FE approach is to ascertain that there is enough within variation in the independent variables. If there is no variation across the years within each municipality, an FE model will not be able to produce parameter estimates for the concerned independent variable. This is due to FE models looking at deviations from the mean of, in our case, the municipalities. Therefore, we have assessed the within variations of each variable using two methods. First, we generated a series of two-way scatter plots. (see figures A2 and A3 in appendix A) Second, we generated tables showing the standard deviations for each variable within the municipalities (see table A1 in appendix A). We found that the Gini coefficient, although exhibiting within variation, did so at a small rate. Nonetheless, we do include the Gini coefficients in model 2, but the estimations should be viewed with some caution.

To see whether there was any heteroskedasticity present in our data we plotted the prediction of the error component against the linear prediction given from models 1 and 2 (see figures A4 and A5 in appendix A). Based on the graphical analysis we could not rule out eventual heteroskedasticity.

Following Drukker (2003), serial correlation leads to standard errors that are biased in models based on panel data and consequently the results are not as efficient. Any potential serial correlation was examined by conducting a series of Wooldridge tests. As the Wooldridge test uses residuals from first-differenced regression (Drukker, 2003), we did not include year dummies while running the tests. The null hypothesis of no first-order autocorrelation was rejected ($p<0.01$) for both models and all degradation indicators. Following the reasoning of Drukker (2003), in the presence of serial correlation consistent estimates of the standard error is generated through clustering at a panel level.

Marbuah and Amuakwa-Mensah (2017) found in their study that there existed spatial dependence in the case of Swedish municipalities. Since we also perform our analysis on municipal data, we therefore found reason to suspect potential cross-sectional dependence in our case. For instance, since the municipalities all exist inside Sweden, they are subject to a considerable extent to shared factors such as similar environmental regulations or economic shocks. Thus, a Pesaran CD test was conducted for both models and the test results indicated that there for GHGs, CO_2, and PM10 might be issues with cross-sectional dependence but not for NO_X. However, as mentioned by De Hoyos and Sarafidis (2006), a failure to reject that there is no cross-sectional dependence of the errors can come from that the test sums positive and negative correlations. Thus, in the case of NO_X we decided also to look at the average absolute value of the correlations which was 0.428 for model 1 and 0.429 for model 2. For SO_2 we were not able to run the test. Therefore, we could not rule out the possibility of cross-sectional dependence.

According to Hoechle (2007, p.281), while using panel data biased statistical inference can be caused from wrongly overlooking potential correlation between errors of different individuals and over time. De Hoyos and Sarafidis (2006, p. 482) highlight, based on literature on panel data, that models using panel data are "likely to exhibit substantial cross-sectional dependence in the errors". This could, *inter alia*, come from common shocks and unobserved components, or spatial dependence. Hoechle (2007, p.282) also argues that for many instances it might not be appropriate to assume that there is no cross-sectional dependence of the errors. However, any unobserved shared factors, given that they do not correlate with the independent variables, will not make the estimates inconsistent for models such as an FE model. Yet, various techniques for estimating robust standard errors will cause biased standard errors. Thus, Hoechle (2007, p.282) highlights the Driscoll-Kraay standard errors that are consistent in case of heteroskedasticity and autocorrelation, but also robust in the case of general forms of dependence—spatial or temporal. It should be kept in mind, though, that this is the case when the time dimension becomes large (Hoechle, 2007).

Given the probable presence of heteroskedasticity, serial correlation and cross-sectional dependence, the regressions will be done using various standard errors taking these issues into account. All FE regressions will be completed with clustered (Rogers) standard errors, clustered at the municipality level. This is done in order to mitigate the effects of heteroskedasticity and serial correlation. In order to make a comparison and take potential cross-sectional dependence into account, Driscoll-Kraay standard errors will be used in addition to the Rogers standard errors. As touched upon previously, we have a panel with a relatively large N compared to T; thus, we need to be cautious when interpreting results from any regression that employs Driscoll-Kraay standard errors. Therefore, we have elected to also use clustered standard errors in order to see whether the results differ drastically.

6. Results

We start by testing the EKC hypothesis in its most basic form, where our five dependent environmental degradation variables are regressed against the income variables. Results are reported in table 4 and table 5.

Table 4: Model 1 results, Rogers SEs (Municipality and time FE, dummies omitted)

Variables	**ln GHG**	**ln CO_2**	**ln SO_2**	**ln NO_X**	**ln PM10**
ln Inc	8.160**	3.588	-1.286	-4.306	-1.827
	(3.819)	(4.021)	(18.90)	(7.152)	(4.706)
(ln Inc)^2	-0.710**	-0.305	0.324	0.425	0.144
	(0.351)	(0.369)	(1.692)	(0.647)	(0.430)
Constant	-21.50**	-8.947	-9.326	6.963	0.573
	(10.46)	(11.03)	(52.99)	(19.81)	(12.93)
Observations	2,900	2,900	2,883	2,900	2,900
R-squared	0.477	0.492	0.146	0.425	0.290
No. of municipalities	290	290	290	290	290
S.E.	Rogers	Rogers	Rogers	Rogers	Rogers

Table 4: Regression output for model 1 and all degradation indicators. Parameter estimates reported for the independent variables. Rogers standard errors in parentheses. ** $p<0.05$. R-squared refers to within R^2.

Table 5: Model 1 results, Driscoll-Kraay SEs (Municipality and time FE, dummies omitted)

Variables	ln GHG	ln CO_2	ln SO_2	ln NO_X	ln PM10
ln Inc	8.160**	3.588	-1.286	-4.306	-1.827
	(2.521)	(2.557)	(7.302)	(2.921)	(1.666)
(ln Inc)²	-0.710***	-0.305	0.324	0.425	0.144
	(0.215)	(0.214)	(0.651)	(0.250)	(0.152)
Constant	-21.50**	-8.947	-9.326	6.963	0.573
	(7.353)	(7.603)	(20.52)	(8.521)	(4.649)
Observations	2,900	2,900	2,883	2,900	2,900
No. of groups	290	290	290	290	290
S.E.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay

Table 5: Regression output for model 1 and all degradation indicators. Parameter estimates reported for the independent variables. Driscoll-Kraay standard errors in parentheses. *** p<0.01, ** p<0.05.

In tables 4 and 5 we can see the, in the most basic form of the EKC, the signs on the parameter estimates are only statistically significant at the 5% level in the case of GHGs. This is the case regardless of whether Rogers or Driscoll-Kraay standard errors are employed. Furthermore, the parameter estimates for GHGs are consistent with the EKC hypothesis. For the remainder of the emissions metrics, we do not see any statistical significance for neither the parameter estimates of the income variables nor the intercept estimates. We do see that the year dummies in general have a statistically significant impact on the models. See tables A2 and A3 in appendix A for more detailed information about the output of the GHG models, including more exact p-values. F tests were performed to determine whether the inclusion of year dummies was the correct course of action, and they yielded affirmative results. However, these are omitted in tables 4 and 5.

These results enable us to calculate an EKC turning point at circa 312,000 SEK[1] (2008 prices) for GHGs. In other words, a generic Swedish municipality would in the presence of an EKC be expected to decrease its emissions of GHGs once the mean real income in said municipality reaches 312,000 SEK in 2008 prices. Since the parameter estimates for the remaining four emissions variables are statistically insignificant, we elect not to calculate any more EKC turning points.

We continue by testing our second model, which includes the natural logarithm of the Gini coefficient as well as an interaction term that is defined as the product of the Gini coefficient and the natural logarithm of real mean income. Just like for the basic EKC model, year dummies are included in this extended model. F-tests have been conducted for this model as well, which have shown that we can reject the null hypothesis that the parameter estimates for the year dummies are

[1] This number is based on the parameter estimates in table A2. Using the parameter estimates in tables 4 and 5 produces another number due to rounding error.

mutually equal to 0. Only the GHG model is tested, since this was the only environmental degradation indicator that showed an EKC pattern in model 1 and we are only interested in how economic inequality affects the EKC, not how economic inequality affects emissions in general.

Table 6: Model 2 results (Municipalities and time fixed effects (dummies omitted in table))

Variables	ln GHG	ln GHG
ln Inc	6.738*	6.738***
	(4.015)	(1.682)
(ln Inc)^2	-0.531	-0.531***
	(0.376)	(0.132)
ln Gini	3.013*	3.013***
	(1.586)	(0.913)
Gini * (ln Inc)	-1.669*	-1.669**
	(0.917)	(0.557)
Constant	-12.71	-12.71***
	(11.76)	(3.637)
Observations	2,900	2,900
R-squared	0.480	
No. of municipalities	290	
S.E.	Rogers	D-K
Number of groups		290

Table 6: Regression output for model 2 and *GHGs*. Parameter estimates reported for the independent variables. Robust standard errors in parentheses—Rogers standard errors in the left column and Driscoll-Kraay standard errors in the right column. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. R-squared refers to within R^2.

In the case of GHGs we still see significant parameter estimates in favor of an EKC shape, i.e. the signs of the parameter estimates for the income terms are consistent with the EKC hypothesis. Yet, it is only when Driscoll-Kraay standard errors are used that the regression exhibits significance ($p<0.05$) for all the independent variables of interest. The logarithmic Gini has a positive parameter estimate, implying that economic inequality is associated with increased emissions. Meanwhile, the parameter estimate for the interaction term is negative, implying that the EKC turning point occurs earlier when the Gini coefficient is considered.

7. Robustness Checks

7.1 Multicollinearity

To evaluate any potential multicollinearity, we generated a correlation matrix (see table A4 in appendix A). The correlation between the Gini coefficient and income, in some but not all cases, was higher than the correlation of these independent variables with the dependent variables. This appeared to be the case for, *inter alia*, GHGs. We are not concerned about the apparently high correlation between the two income terms, as that is a consequence of the model specification. Nor are we concerned about the correlation that the interaction term has with the income and Gini variables, for the same reason. Additionally, the variance inflation factors (VIF) were checked. This was based on OLS regressions on GHGs against the linear term of logarithmic income and the logarithmic Gini coefficient, with year and municipal dummies included. The VIFs were high for the two independent variables (~228 for income and ~53 for Gini).

7.2 Evaluation of Applying a FE Approach

To assess whether a FE approach is suitable in comparison to a RE approach, we decided to follow a correlated random effects (CRE) approach explained by Wooldridge (2018). First, we generated time averages of our independent variables $(ln(Inc_{it})$, $(ln(Inc_{it}))^2, ln(Gini_{it})$ and $Gini_{it} \times ln(Inc_{it}))$ for each municipality, i. Second, we regressed our two models once more, but using random effects while also including the generated means of the variables for each municipality. The regressions were completed using clustered standard errors to take any potential heteroskedasticity and serial correlation into account. CRE for model 2:

$$\begin{aligned}\ln(DEG_{it}) = \beta_1 ln(Inc_{it}) + \beta_2 (\ln(Inc_{it}))^2 + \beta_3 ln(Gini_{it}) + \beta_4 Gini_{it} \times ln(Inc_{it}) + \gamma_1 \overline{ln(Inc)}_i \\ + \gamma_2 \overline{(\ln(Inc))^2}_i + \gamma_3 \overline{ln(Gini)}_i + \gamma_4 \overline{Gini \times ln(Inc)}_i + \delta_1 D_{2009} + \cdots + \delta_9 D_{2017} + a_i \\ + u_{it}\end{aligned}$$

Third, we conducted a test for whether the parameter estimates of these means were mutually equal to zero. A rejection of the null hypothesis that they are mutually equal to zero should be interpreted as RE not being sufficient. Test for model 2:

$$H_0 : \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$$

For all indicators in both models we rejected the null hypothesis of the parameter estimates for the means being equal to zero. Thus, we concluded that the FE approach was suitable.

7.3 County EKC

For further robustness, we have tested an alternative hypothesis to a municipal EKC, namely a county (Swedish: *län*) EKC. This is in order to see whether our findings hold at a larger, but still subnational, scale and whether the county level would have been a better fit with the EKC hypothesis. Several recent studies on the EKC have been conducted either on cross-sections of countries or at a subnational level in countries that are larger than Sweden, e.g. the United States and China. Since a substantial portion of Swedish municipalities have very low levels of emissions for certain gases and particles, it is possible that commercial and political leaders, as well as the citizenry of those municipalities, do not pay attention to the small-scale environmental degradation that the emissions cause. On a county level, however, each unit in the cross-section has a considerable amount of emissions. There are 21 counties in Sweden, meaning that the average county houses circa 13.8 municipalities. Additionally, in larger subnational units the levels of certain types of emissions per capita might not be as heavily influenced by the occurrence of single entities generating large amounts of emissions. For instance, the per capita emissions for a sparsely populated municipality could potentially be high if one large factory is located within that municipality. It is possible that in the aggregate, we can see patterns that are not visible at a small scale.

The county data covers the same time period (2008-2017) as the municipal data. The county data we used have been retrieved from Statistics Sweden. The emissions data on the counties is based on the sum of emissions from all available industrial classifications, i.e. A01-F43 producer of goods, G45-T98 producer of services, unallocated, government and non-profit institutions serving households (NPISH), and private consumption. To calculate the per capita emissions, we used the equivalent county population data to the data used for the municipalities. Just like for the municipalities, the income data was based on mean earned income for which we converted so that it was expressed in real terms in 2008 SEK (thousands). Similarly, the Gini coefficients for the counties were calculated using the same method and the equivalent data as for the municipalities. The same tests for heteroskedasticity, serial correlation, and cross-sectional dependence as earlier have been conducted. These have indicated that heteroskedasticity and serial correlation are present, but not cross-sectional dependence. Therefore, we have elected to report the results using clustered/Rogers standard errors. The number of observations is 210 and there are no missing observations. Furthermore, following the same procedure as for the municipality regressions, it appeared like multicollinearity was present in the county models.

Table 7: Descriptive statistics of variables used in county regressions

Variable	*Mean*	*SD*	*Min*	*Max*	*N*
GHG	8.86844	9.588189	3.884482	50.8764	210
CO_2	7.295639	8.812378	2.692153	45.45771	210
SO_2	.0064164	.0084415	.0004174	.051101	210
NO_X	.0277587	.0279345	.0107531	.1723119	210
PM10	.0065441	.0033368	.0027084	.01954	210
Inc	249.2923	19.72845	212.8	330.9314	210
Gini	.3304448	.0237627	.2904957	.3953163	210
ln GHG	1.939006	.5678471	1.35699	3.929399	210
ln CO_2	1.683008	.6350379	.9903413	3.816782	210
ln SO_2	-5.635355	1.067506	-7.78157	-2.973951	210
ln NO_X	-3.81706	.5785462	-4.532557	-1.758449	210
ln PM10	-5.122617	.4063875	-5.911397	-3.935292	210
ln Inc	5.515671	.0761133	5.360353	5.801911	210
ln Gini	-1.109776	.0695523	-1.236167	-.9280691	210

Table 7: Descriptive statistics of dependent and independent variables on county level. Observations are from the 21 Swedish counties in the years 2008-2017. Original data retrieved from Statistics Sweden (2019a; 2020d). Emissions metrics have been transformed to per capita measures by the authors based on population data, also retrieved from Statistics Sweden (2019d). Real mean income (2008 prices) has been calculated by the authors based on consumer price index data from Statistics Sweden (2020a). The Gini coefficients have been calculated by the authors, using mean income data from Statistics Sweden (2020d).

Table 8: Model 1 using county data (County and time fixed effects (year dummies omitted))

Variables	ln GHG	ln CO_2	ln SO_2	ln NO_X	ln PM10
ln Inc	12.10	6.154	-21.21	-13.42	-15.67**
	(18.76)	(21.44)	(70.21)	(25.41)	(7.350)
(ln Inc)^2	-0.823	-0.340	2.596	1.375	1.338*
	(1.645)	(1.898)	(6.188)	(2.156)	(0.655)
Constant	-39.45	-21.63	33.12	28.68	40.68*
	(53.73)	(60.89)	(200.2)	(74.96)	(22.07)
Obs.	210	210	210	210	210
R-sq.	0.695	0.674	0.458	0.668	0.734
No. of counties	21	21	21	21	21

Table 8: Regression output for model 1 and all degradation indicators on county level. Parameter estimates reported for the independent variables. Rogers standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. R-squared refers to within R^2.

When testing model 1, the signs of the parameter estimates for all five degradation indicators are consistent with the results from the regression on the municipalities, with the exception of the intercept for SO_2 (for comparison: see table 4). Thus, GHGs and CO_2 exhibit an EKC shape, yet without any statistical significance. Thus, we cannot conclude an EKC pattern at the county level. However, the overall F-statistics for all indicators (including year dummies) are significant. Also, we find that PM10 appears to have slightly statistically significant parameter estimates (at the 10%

level) that would be consistent with an inverted EKC, meaning a U-shaped relationship between environmental degradation and income.

Table 9: Model 2 using county data (County and time fixed effects (year dummies omitted))

Variables	ln GHG	ln CO_2
ln Inc	13.12	6.844
	(20.60)	(23.15)
(ln Inc)^2	-1.038	-0.587
	(1.927)	(2.193)
ln Gini	-1.651	-4.251
	(9.103)	(10.67)
Gini * (ln Inc)	0.182	1.459
	(5.408)	(6.316)
Constant	-40.77	-25.41
	(54.81)	(62.76)
Obs.	210	210
R-sq.	0.698	0.678
Number of counties	21	21

Table 9: Regression output for model 2 on county level for *GHGs* and *CO_2*. Parameter estimates reported for the independent variables. Rogers standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. R-squared refers to within R^2.

As the inequality variables are added the EKC pattern, in terms of parameter estimate signs, prevails for GHGs and CO_2. For these indicators, the sign of the parameter estimates for the logarithmic Gini coefficient is negative while the interaction term is positive (in direct contrast to the municipality regressions). However, no significance is depicted for the inequality variables individually. Meanwhile, the overall F-test for the model (including year dummies) is significant for both indicators.

8. Discussion

8.1 Regression Results

The only indicators that, in our original regressions, to some extent depicted a municipal Swedish EKC was GHGs and CO_2. In the basic form the EKC pattern showed significance, i.e. the income variables individually were significant ($p<0.05$), for GHGs using both clustered standard errors and Driscoll-Kraay standard errors. Regarding CO_2, it showed an EKC pattern, yet with no significance for the income parameter estimates individually.

As the model was extended with income inequality variables, the EKC prevailed for GHGs but stayed significant only when Driscoll-Kraay standard errors were employed. When it comes to any potential impact of income inequality on the EKC-type relationship at a municipal level, it appears from our original regressions on GHGs that higher income inequality would increase the degradation level at which a potential turning point occurs. This interpretation is based on the positive parameter estimate for the Gini coefficient. Meanwhile, as the interaction term has a negative parameter estimate, higher income inequality, *ceteris paribus,* would generate an "earlier" turning point in terms of real mean earned income. However, as the Driscoll-Kraay standard errors are not optimal for the dimensions of our panel data set with a small time dimension, we are careful with drawing any determined conclusions here of whether these findings hold.

Upon comparing the results from our municipal models with those from our county models, we can see that, in the cases of GHGs and CO_2, the parameter estimates for the income variables are quite similar, at least in terms of their respective signs. However, we obtain no significant results when we look at the counties. An interesting finding is that PM10 appears to have a slightly statistically significant ($p<0.10$) inverse EKC at the county level.

A potential problem with using the Gini coefficient for county EKCs is that county governments do not have an explicit responsibility for environmental issues, which municipal governments do, apart from within the domains of sustainable health care and local transportation (Swedish Association of Local Authorities and Regions, 2019). The Gini coefficient is supposed to capture not only income inequality, but also the political inequality that could arise from income inequality. If that assumption holds true, then the effects of that inequality on a certain issue could reasonably be suspected to manifest themselves to a greater extent at those administrative levels where the influence on that issue is the greatest. A limitation with the county data is that it only encompasses 210 observations, which is a substantially smaller number than the 2,900 (2,883 for SO_2) observations in the municipal data. Another potential issue could be that the aggregation of municipalities into counties diminishes some of the variation that exists between municipalities. These aspects should all be considered when interpreting these results. Also, there may exist multicollinearity issues that affect the significance of the parameter estimates. We have found evidence suggesting the Gini coefficient to be somewhat correlated to mean earned income.

A somewhat surprising aspect of the results was that we did not find any evidence in favor of an EKC for SO_2 or PM10. Among the dependent variables tested, these two are the ones that primarily have a negative impact on the immediate surroundings of where they are emitted. GHGs, CO_2, and NO_X are connected to climate change, whereas SO_2 and PM10 have more local effects. This difference in their impacts has implications for how the social costs are internalized. It could therefore be expected that people act more vigilantly on a local scale to reduce emissions of SO_2 and PM10, and that government intervention is needed more for the three other pollutants. The fact that this is not clearly visible in the regression results we therefore find rather surprising.

8.2 Limitations

A concern regarding the validity and reliability of model 2 is the presence of suspected multicollinearity. Our results indicate that we may need ignore potential multicollinearity, to some extent, in order to be able to interpret and draw conclusions from model 2. The presence of imperfect multicollinearity causes variance inflation of estimates making the interpretation of the significance for individual variables difficult. We suspect this to be the case since, as model 1 was extended to model 2 and clustered standard errors were employed, the significance for the individual parameter estimates dropped. Potential remedies could e.g. be to add more observations or to drop independent variables. When it comes to adding more observations, we were limited to the period for which we were able to obtain consistent emissions data, i.e. 2008-2017. Furthermore, we did not consider dropping variables as a viable option as we would not be able to determine the effect of income inequality on a potential EKC pattern. Instead, we could perhaps have used some other measure of income inequality (for other types of income inequality measures used see e.g. Magnani, 2000; Jorgenson et al. 2017); however, we do not know for certain that this would have been a remedy for the multicollinearity.

Furthermore, as discussed in section 5.4, we have potential issues with heteroskedasticity, autocorrelation and in some cases also cross-sectional dependence. This damages the confidence we have in drawing any determined conclusions on whether an EKC is present or not, and if income inequality has an impact on that relationship. Another weakness is the short period of time that data was available for. We know that emissions of GHGs have decreased since, at least, the early 1990s in Sweden. However, methods of collecting emissions data have changed over time, which would have made it too difficult to compare the statistics across years. In the future, once the emissions database that we have gathered our data from has been filled with more observations, researchers may be able to draw better conclusions from any potential patterns that can be discerned.

We look at aggregate air emissions, meaning that we only consider the sum of emissions from all sources. A downside with this is that we are unable to see exactly where in society the trends occur. If we had specifically looked at industrial manufacturing or household consumption, we could possibly have seen different patterns. The upside however is that this approach enables us to see larger-scale patterns. The same logic applies for the positive and negative aspects of using GHGs as an independent variable. Including it in our analysis allows us to see an aggregate EKC pattern, which is relevant in order to make society-wide conclusions. However, the variable includes several different types of emissions with different Global Warming Potential factors. Since we have determined that we are unable to produce a statistically significant EKC for CO_2, the EKC pattern that we did find for GHGs might in part come from other pollutants, the identities of which remain unclear. Future studies may want to look solely at different types of GHGs, in order to determine which ones exhibit EKC characteristics, and which ones do not.

Further limitations include the lack of investigation in terms of other potential shapes of the relationship between environmental degradation and income. Even though this eventuality was not within the scope of this thesis, it is important to highlight that the relationship might portray other shapes. Or perhaps, that there is no simple predictable relationship between environmental degradation and income.

Additionally, we originally considered including multiple indicators of inequality, like e.g. those suggested by e.g. Torras and Boyce (1998) while investigating the effects of inequality on the environmental degradation-income relationship and the EKC. However, we decided to solely include one commonly used type of inequality indicator, namely some measure of distribution of income. Based on Torras and Boyce's use of literacy rate (as one of the parameters for capturing power inequalities) we considered including a parameter capturing the share of the population with any post-secondary school education. Yet, it appeared like we had high levels of multicollinearity (partly due to our choice of income variable) and the implications of e.g. a low or a high share of people with post-secondary education on equality were not as clear-cut in comparison to e.g. literacy rate. Hence, we decided to go with a narrower approach for regressing any effects of equality on the environmental degradation-income relationship and the EKC.

Another control variable that we had planned to include was population density. This variable would have been used as a proxy for urbanization, which is commonly used as a control variable in the EKC literature. However, the population density did not change much across the years, which would not have made it an appropriate addition to the models due to our fixed effects approach to answering the research questions; however, it would in random effects approach.

9. Conclusion

This thesis has, using a fixed effects approach, tested the environmental Kuznets curve hypothesis through a lens of economic inequality. This has been done on a dataset consisting of observations from the 290 municipalities of Sweden over a ten-year period (2008-2017). Per capita emissions of greenhouse gases, carbon dioxide, sulfur dioxide, nitrous oxides, and particulate matter have been used as dependent variables. Mean income and the Gini coefficient have been used as independent variables. The main contribution to the literature is that we have tested a hypothesis that has been tested before a great number of times, with mixed results, on a new dataset. The nuance that comes with adding economic inequality enriches the story that can be told about how environmental degradation develops as a function of economic development.

With the methods used, we find to some extent evidence in favor of a Swedish municipal EKC existing for GHGs. The turning point is estimated to be at around 312,000 SEK in 2008 prices. We also find evidence suggesting that a higher Gini coefficient in a municipality is associated with higher levels of emissions of GHGs. However, by using an interaction term, defined as the product of the Gini coefficient and the natural logarithm of mean income, we also find evidence indicating that a higher Gini coefficient shifts the EKC turning point to a lower level of income. When including income inequality, however, statistical significance varies with the standard errors employed. Furthermore, the short time span of data makes the use of Driscoll-Kraay standard errors slightly problematic, inhibiting our ability to draw decisive conclusions from the results of our second model. For the remaining four indicators of environmental degradation, we do not find any evidence in favor of the Swedish municipal EKC. Another interesting finding comes from the robustness checks, where we find evidence with 10% significance in favor of an inverted EKC for PM10 at the county level, indicating a U-shaped relationship between economic development and environmental degradation for that specific indicator.

10. References

Brännlund, R., & Ghalwash, T. (2008). The Income–Pollution Relationship and the Role of Income Distribution: An Analysis of Swedish Household Data. *Resource and Energy Economics, 30*(3), 369-387. doi:10.1016/j.reseneeco.2007.11.002

Cheng, J., Shi, F., Yi, J., & Fu, H. (2020). Analysis of the Factors that Affect the Production of Municipal Solid Waste in China. *Journal of Cleaner Production, 259*. doi:10.1016/j.jclepro.2020.120808

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for Cross-Sectional Dependence in Panel-Data models. *The Stata Journal, 6*(4), 482-496. doi:10.1177/1536867X0600600403

Drukker, D. M. (2003). Testing for Serial Correlation in Linear Panel-Data Models. *The Stata Journal, 3*(2), 168-177. doi:10.1177/1536867X0300300206

Franzen, A., & Mader, S. (2018). Consumption-Based Versus Production-Based Accounting of CO2 Emissions: Is There Evidence for Carbon Leakage? *Environmental Science and Policy, 84*, 34-40. doi:10.1016/j.envsci.2018.02.009

Grossman, G. M., & Krueger, A. B. (1991) Environmental Impacts of a North American Free Trade Agreement. *National Bureau of Economic Research*. NBER Working Paper No. 3914. Retrieved from: https://www.nber.org/papers/w3914

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *The Stata Journal, 7*(3), 281-312. doi:10.1177/1536867X0700700301

Jorgenson, A., Schor, J., & Huang, X. (2017). Income Inequality and Carbon Emissions in the United States: A State-Level Analysis, 1997–2012. *Ecological Economics, 134*, 40-48. doi:10.1016/j.ecolecon.2016.12.016

Kaika, D., & Zervas, E. (2013). The Environmental Kuznets Curve (EKC) Theory—Part A: Concept, Causes and the CO2 Emissions Case. *Energy Policy, 62*, 1392-1402. doi:10.1016/j.enpol.2013.07.131

Keene, A., & Deller, S. C. (2015). Evidence of the Environmental Kuznets' Curve Among US Counties and the Impact of Social Capital. *International Regional Science Review, 38*(4), 358-387. doi:10.1177/0160017613496633

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review, 45*(1), 1-28. Retrieved from http://www.jstor.org/stable/1811581

Magnani, E. (2000). The Environmental Kuznets Curve, Environmental Protection Policy and Income Distribution. *Ecological Economics, 32*(3), 431-443. doi:10.1016/S0921-8009(99)00115-9

Marbuah, G., & Amuakwa-Mensah, F. (2017). Spatial Analysis of Emissions in Sweden. *Energy Economics, 68*, 383-394. doi:10.1016/j.eneco.2017.10.003

Narayan, P. K., & Narayan, S. (2010). Carbon Dioxide Emissions and Economic Growth: Panel Data Evidence from Developing Countries. *Energy Policy, 38*(1), 661-666. doi:10.1016/j.enpol.2009.09.005

Organisation for Economic Co-operation and Development. (2020). *Value Added by activity*. OECD Data. Retrieved April 19, 2020 from https://data.oecd.org/natincome/value-added-by-activity.htm

Panayotou, T. (1993). Empirical Tests and Policy Analysis of Environmental Degradation at Different Stages of Economic Development. *International Labour Organization*. Retrieved from https://ideas.repec.org/p/ilo/ilowps/992927783402676.html

Ridzuan, S. (2019). Inequality and the Environmental Kuznets Curve. *Journal of Cleaner Production, 228*, 1472-1481. doi:10.1016/j.jclepro.2019.04.284

Shafik, N., & Bandyopadhyay, S. (1992). Economic Growth and Environmental Quality: Time Series and Cross-country Evidence (English). Policy, Research Working Papers; no. WPS 904. World Development Report. Washington, DC: World Bank. Retrieved from http://documents.worldbank.org/curated/en/833431468739515725/Economic-growth-and-environmental-quality-time-series-and-cross-country-evidence

Statistics Sweden. (2019a). *Air Emissions by Region (County, NUTS3), NACE Rev. 2 and Subject. Year 2008 – 2017* [Data set]. Statistikdatabasen. Retrieved from http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__MI__MI1301__MI1301B/UtslappLan/#

Statistics Sweden. (2019b). *Air Emissions by Region (Municipality, LAU2) and Subject. Year 2008 - 2017* [Data set]. Statistikdatabasen. Retrieved from http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__MI__MI1301__MI1301B/UtslappKommun/#

Statistics Sweden. (2019c). *Kvalitetsdeklaration: Miljöräkenskaperna - Utsläpp till luft, regionala per år.* Retrieved from ttps://www.scb.se/contentassets/f0d9c7eda5be4b8a96c5827e4bebf513/mi1301_kd_2008_2017_ns_191114.pdf

Statistics Sweden. (2019d). *Population 1 November by Region, Age and Sex. Year 2002 - 2019* [Data set]. Statistikdatabasen. Retrieved from http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101A/FolkmangdNov/#

Statistics Sweden. (2019e). *Utsläpp av växthusgaser i Sverige*. Retrieved April 19, 2020 from https://scb.se/hitta-statistik/sverige-i-siffror/miljo/utslapp-av-vaxthusgaser-i-sverigeut/

Statistics Sweden. (2020a). *Consumer Price Index (CPI), Fixed Index Numbers, Total Annual Average, 1980=100. Year 1980 - 2019* [Data set]. Statistikdatabasen. Retrieved from http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__PR__PR0101__PR0101A/KPIFastAmed/#

Statistics Sweden. (2020b). *Population 16-95+ Years of Age by Region, Level of Education, Age and Sex. Year 2008 - 2019* [Data set]. Statistikdatabasen. Retrieved from https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__UF__UF0506/UtbBefRegionR/#

Statistics Sweden. (2020c). *Population Density per sq. km, Population and Land Area by Region and Sex. Year 1991 - 2019* [Data set]. Statistikdatabasen. Retrieved from https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101C/BefArealTathetKon/?rxid=bd5169ae-f630-42db-8c8e-3ffdbf806a73#

Statistics Sweden. (2020d). *Total Earned Income for Persons Registered in the National Population Register 31 December by Region, Sex, Age and Income Bracket. Year 1991 – 2018* [Data set]. Statistikdatabasen. Retrieved from https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__HE__HE0110__HE0110A/SamForvInk2/#

Steinkraus, A. (2016). Investigating the Carbon Leakage Effect on the Environmental Kuznets Curve Using Luminosity Data. *Economics Department Working Paper Series, 15* Retrieved from https://www.econstor.eu/handle/10419/142235

Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development, 32*(8), 1419-1439. doi:10.1016/j.worlddev.2004.03.004

Stern, D. I. (2017). The Environmental Kuznets Curve After 25 Years. *Journal of Bioeconomics, 19*(1), 7-28. doi:10.1007/s10818-017-9243-1

Swedish Association of Local Authorities and Regions. (2014). *Guide till lokalt arbete med miljömål.* Retrieved from https://skr.se/download/18.27a42f99148a7e81e22ede5/1411977795875/lokalt-arbete-med-miljomal.pdf

Swedish Association of Local Authorities and Regions. (2019). *Öppna jämförelser: Miljöarbetet i regionerna 2019.* Retrieved from https://webbutik.skr.se/bilder/artiklar/pdf/7585-797-8.pdf

Swedish EPA. (2019a). Därför blir det varmare. Retrieved May 13, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Klimat/Darfor-blir-det-varmare/

Swedish EPA. (2019b). Fakta om kväveoxider i luft. Retrieved April 2, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Luftfororeningar/Kvaveoxider/

Swedish EPA. (2019c). Fakta om svaveldioxid i luft. Retrieved April 2, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Luftfororeningar/Svaveldioxid/

Swedish EPA. (2019d). Fakta om partiklar i luft. Retrieved April 3, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Luftfororeningar/Partiklar/

Swedish EPA. (2019e). Marknära ozon. Retrieved April 2, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Luftfororeningar/Marknara-ozon/

Swedish EPA. (2019f). Utsläpp av kväveoxider till luft. Retrieved April 2, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-O/Kvaveoxid-till-luft/

Swedish EPA. (2019g). Utsläpp av svaveldioxid till luft. Retrieved April 2, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-O/Svaveldioxid-till-luft/

Swedish EPA. (2020). Växthuseffekten förstärks. Retrieved May 13, 2020 from https://www.naturvardsverket.se/Sa-mar-miljon/Klimat-och-luft/Klimat/Darfor-blir-det-varmare/Vaxthuseffekten-forstarks/

Torras, M., & Boyce, J. K. (1998). Income, Inequality, and Pollution: A Reassessment of the Environmental Kuznets Curve. *Ecological Economics, 25*(2), 147-160. doi:10.1016/S0921-8009(97)00177-8

Torres-Reyna (2007). Panel Data Analysis Fixed and Random Effects using Stata (v. 4.2). Panel 101. Princeton University. Retrieved from: https://dss.princeton.edu/training/

United Nations. (n.d.). *Climate Change.* Retrieved May 9, 2020 from: https://www.un.org/en/sections/issues-depth/climate-change/

Wagner, M. (2015). The Environmental Kuznets Curve, Cointegration and Nonlinearity. *Journal of Applied Econometrics, 30*(6), 948-967. doi:10.1002/jae.2421

Wang, Y., & Komonpipat, S. (2020). Revisiting the Environmental Kuznets Curve of PM2.5 Concentration: Evidence from Prefecture-Level and Above Cities of China. *Environmental Science and Pollution Research International, 27*(9), 9336-9348. doi:10.1007/s11356-020-07621-x

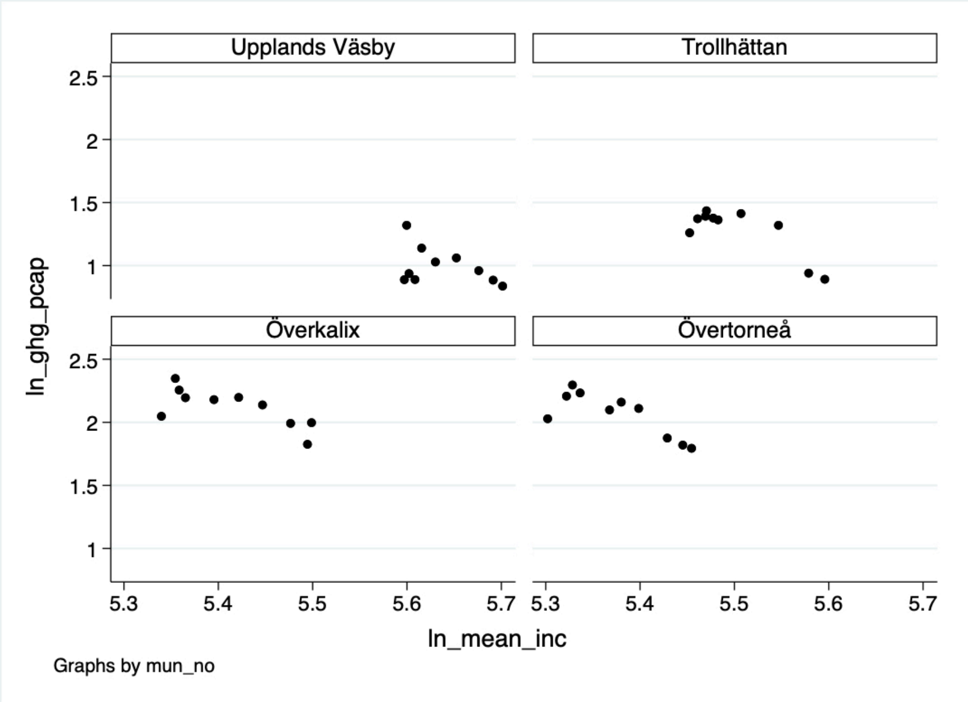
Wooldridge, J. M. (2018). Introductory Econometrics: A Modern Approach. Seventh Edition. *Cengage.*

World Bank. (n.d.). *GDP per Capita (Constant 2010 US$) - Sweden.* World Bank Data. Retrieved April 19, 2020 from https://data.worldbank.org/indicator/NY.GDP.PCAP.KD?end=2018&locations=SE&start=1990

11. Appendices

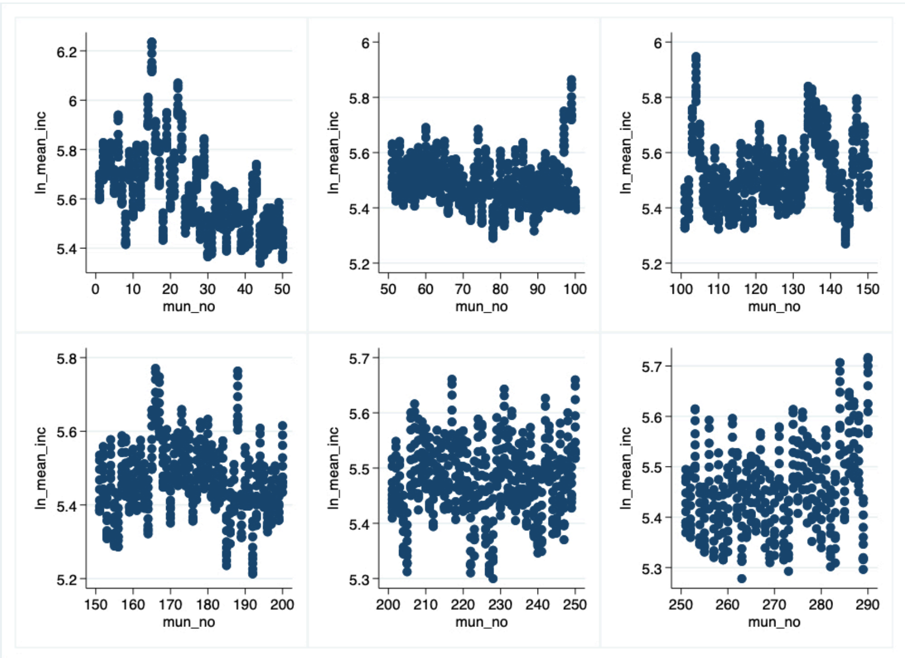
Appendix A – Supplementary Figures and Tables

Figure A1: Logarithmic GHG against Logarithmic mean income for four municipalities



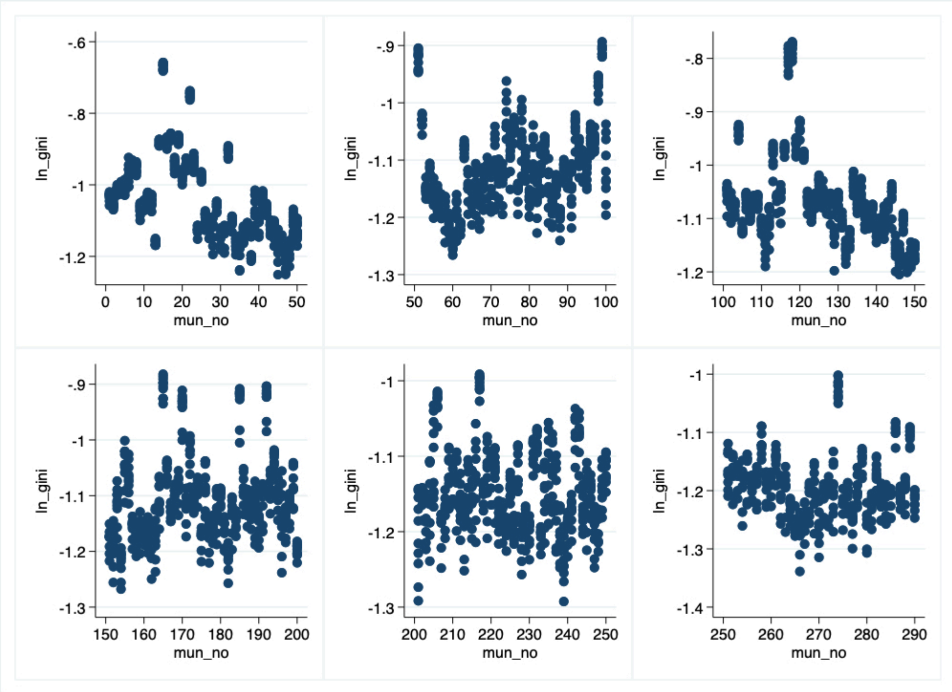
Notes: Logarithmic GHGs per capita on the Y-axis against logarithmic real earned mean income on the X-axis for four different municipalities that potentially depicts a pattern reminding of an inverted U-shaped within the period 2008-2017.

Figure A2: Logarithmic mean income by municipality



Notes: Logarithmic real mean income (2008 prices) for the 290 Swedish municipalities in the years 2008-2017, sorted by municipality, each of which has been assigned a number between 1 and 290.

Figure A3: Logarithmic Gini by municipality



Notes: Logarithmic Gini coefficients for the 290 Swedish municipalities in the years 2008-2017, sorted by municipality, each of which has been assigned a number between 1 and 290.

Table A1: Detailed descriptive statistics

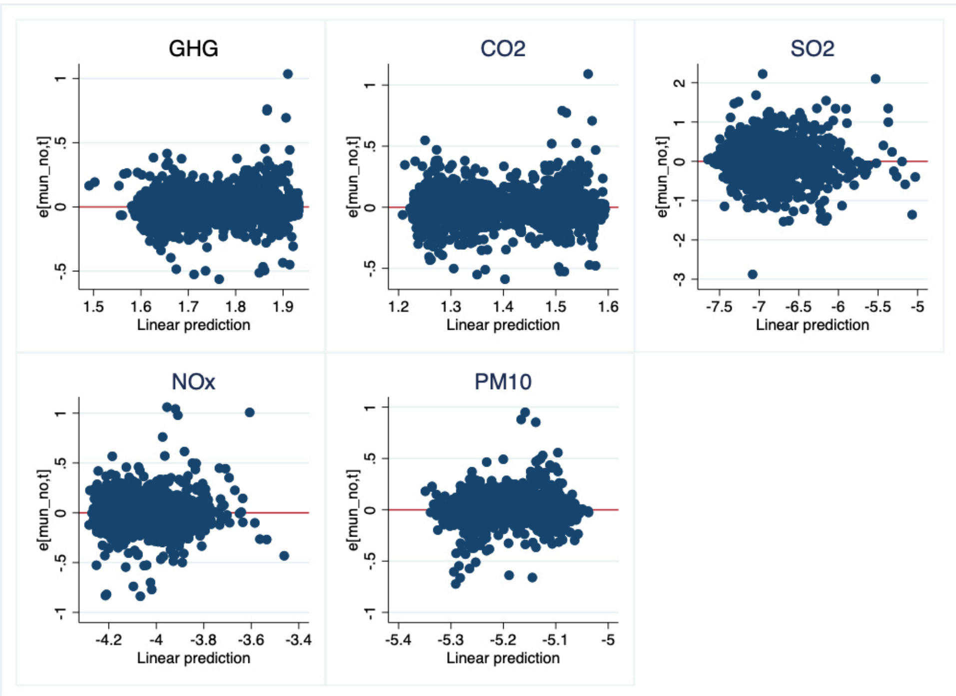
Variable		Mean	Std. Dev.	Min	Max	Observations
mun_no	overall	145.5	83.72973	1	290	N = 2900
	between		83.86	1	290	n = 290
	within		0	145.5	145.5	T = 10
year	overall	2012.5	2.872777	2008	2017	N = 2900
	between		0	2012.5	2012.5	n = 290
	within		2.872777	2008	2017	T = 10
ln_ghg~p	overall	1.748803	.577437	.0763311	5.336343	N = 2900
	between		.5660649	.4236397	4.971154	n = 290
	within		.1183167	1.145451	2.904429	T = 10
ln_co2~p	overall	1.403915	.6015381	-.0227598	5.334618	N = 2900
	between		.5860647	.3237252	4.969384	n = 290
	within		.1394365	.7700859	2.624364	T = 10
ln_so2~p	overall	-6.901542	1.443133	-10.65827	-1.788147	N = 2883
	between		1.411003	-10.41831	-2.881269	n = 290
	within		.3588226	-9.878359	-4.616795	T-bar = 9.94138
ln_nox~p	overall	-4.045947	.6294384	-5.695714	-.7054189	N = 2900
	between		.6124157	-5.487401	-1.822188	n = 290
	within		.1493454	-5.026944	-2.878321	T = 10
ln_pm1~p	overall	-5.181368	.5856488	-6.674582	-2.503251	N = 2900
	between		.5755682	-6.541531	-3.011862	n = 290
	within		.1128462	-5.982809	-4.177341	T = 10
ln_mea~c	overall	5.503719	.1268816	5.212715	6.235735	N = 2900
	between		.1170136	5.270995	6.166811	n = 290
	within		.0494902	5.39955	5.619946	T = 10

Table A1: continued

ln_mea~2	overall	30.30702	1.420196	27.1724	38.88439	N =	2900
	between		1.312876	27.78578	38.03166	n =	290
	within		.5465014	29.15819	31.62654	T =	10
ln_gini	overall	-1.11131	.0926547	-1.338775	-.6579426	N =	2900
	between		.0901543	-1.273257	-.6694163	n =	290
	within		.0219619	-1.210092	-1.027475	T =	10
edu_sh~e	overall	.19836	.0641832	.1037768	.4978873	N =	2900
	between		.0629876	.1205221	.4864317	n =	290
	within		.0128202	.1636299	.234759	T =	10
ln_edu~e	overall	-1.662119	.2888412	-2.265513	-.6973814	N =	2900
	between		.2811417	-2.119386	-.7209289	n =	290
	within		.0680732	-1.835966	-1.508592	T =	10
ln_pop~s	overall	3.349938	1.671152	-1.609438	8.646307	N =	2900
	between		1.673503	-1.609438	8.484869	n =	290
	within		.0288255	2.985019	3.511376	T =	10

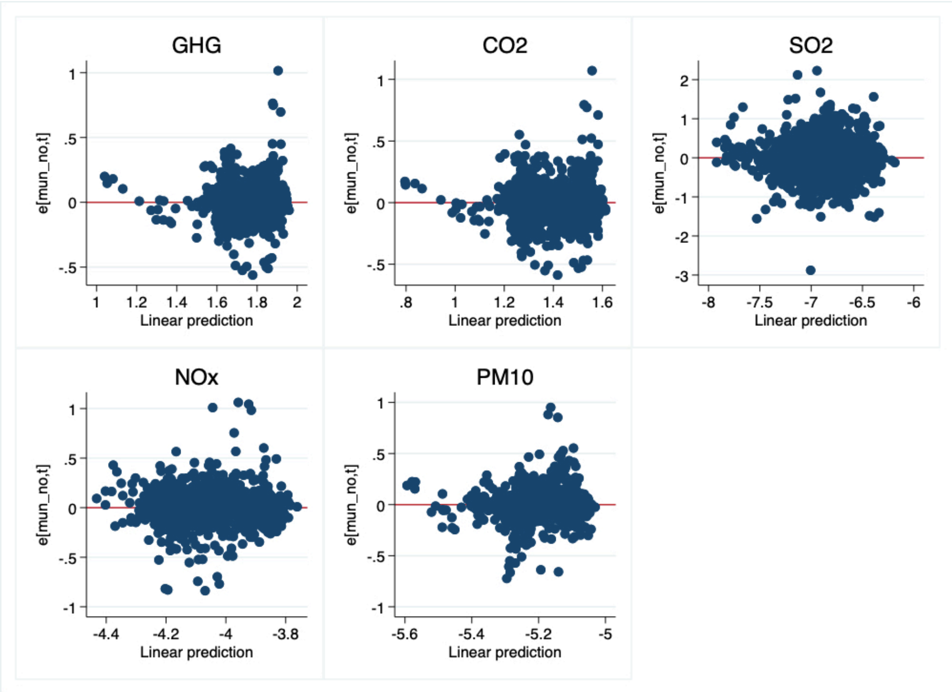
Notes: This table shows the means, standard deviations, minimum values, and maximum values overall, between groups, and within groups. The variables that descriptive statistics are reported for are the municipal numbers (of which the authors have assigned a unique one to each municipality) and the years in the period (2008-2017); the natural logarithm of the five emissions metrics used in the thesis (GHGs, CO_2, SO_2, NO_X, and PM10); the independent variables used in this thesis (the natural logarithm of mean real income (2008 prices) and its square, as well as the municipal Gini coefficients); and lastly, the share of the population that has any post-secondary school education, as well as the natural logarithm of the same, plus the natural logarithm of population density.

Figure A4: Error component against linear prediction (model 1)



Notes: Error component against the linear prediction of model 1 for all degradation indicators separately.

Figure A5: Error component against linear prediction (model 2)



Notes*:* Error component against the linear prediction of model 2 for all degradation indicators separately.

Table A2: Regression output of model 1 for GHGs (clustered standard errors)

```
. xtreg ln_ghg_pcap ln_mean_inc ln_mean_inc2 i.year, fe cluster(mun_no)

Fixed-effects (within) regression               Number of obs     =      2,900
Group variable: mun_no                          Number of groups  =        290

R-sq:                                           Obs per group:
     within  = 0.4772                                         min =         10
     between = 0.0493                                         avg =       10.0
     overall = 0.0036                                         max =         10

                                                F(11,289)         =      65.60
corr(u_i, Xb)  = -0.0912                        Prob > F          =     0.0000

                              (Std. Err. adjusted for 290 clusters in mun_no)
------------------------------------------------------------------------------
             |               Robust
 ln_ghg_pcap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 ln_mean_inc |   8.159569   3.819432     2.14   0.033     .6421373      15.677
ln_mean_inc2 |  -.7104114   .3506181    -2.03   0.044      -1.4005   -.0203225
             |
        year |
       2009  |  -.0490522   .0075929    -6.46   0.000    -.0639965   -.0341078
       2010  |  -.0100217   .0075171    -1.33   0.184    -.0248169    .0047734
       2011  |  -.0490773   .0085257    -5.76   0.000    -.0658576   -.0322969
       2012  |  -.1180025    .011786   -10.01   0.000    -.1411998   -.0948053
       2013  |  -.1591029   .0175844    -9.05   0.000    -.1937127   -.1244932
       2014  |  -.1920484   .0236366    -8.13   0.000    -.2385701   -.1455267
       2015  |  -.2194257   .0323752    -6.78   0.000    -.2831467   -.1557046
       2016  |  -.2582573   .0389164    -6.64   0.000    -.3348528   -.1816618
       2017  |  -.2701012    .043587    -6.20   0.000    -.3558894    -.184313
             |
       _cons |  -21.49621   10.46384    -2.05   0.041     -42.0912   -.9012178
-------------+----------------------------------------------------------------
     sigma_u |  .57329018
     sigma_e |  .09035251
         rho |  .97576316   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Notes: Detailed description of regression output for fixed effects model where the natural logarithm of yearly Swedish municipal per capita GHG emissions is regressed on the natural logarithm of municipal mean real income (2008 prices), the square of the same, and year dummy variables.

Table A3: Regression output of model 2 for GHGs (clustered standard errors)

```
. xtreg ln_ghg_pcap ln_mean_inc ln_mean_inc2 ln_gini interact i.year, fe
cluster(mun_no)

Fixed-effects (within) regression               Number of obs     =      2,900
Group variable: mun_no                          Number of groups  =        290

R-sq:                                           Obs per group:
     within  = 0.4796                                         min =         10
     between = 0.0040                                         avg =       10.0
     overall = 0.0236                                         max =         10

                                                F(13,289)         =      63.94
corr(u_i, Xb)  = -0.0122                        Prob > F          =     0.0000

                               (Std. Err. adjusted for 290 clusters in mun_no)
------------------------------------------------------------------------------
             |               Robust
 ln_ghg_pcap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 ln_mean_inc |   6.737815   4.015071     1.68   0.094    -1.164674     14.6403
ln_mean_inc2 |    -.53113   .3760632    -1.41   0.159      -1.2713    .2090399
     ln_gini |   3.012834   1.585547     1.90   0.058    -.1078495    6.133518
    interact |  -1.668854   .9171469    -1.82   0.070    -3.473989      .13628
             |
        year |
       2009  |  -.0512815   .0080503    -6.37   0.000    -.0671262   -.0354369
       2010  |   -.014122   .0104925    -1.35   0.179    -.0347735    .0065295
       2011  |  -.0536165   .0123764    -4.33   0.000    -.0779758   -.0292571
       2012  |  -.1232807   .0180539    -6.83   0.000    -.1588145    -.087747
       2013  |   -.164351   .0251138    -6.54   0.000    -.2137802   -.1149219
       2014  |  -.1973475   .0339849    -5.81   0.000    -.2642368   -.1304582
       2015  |  -.2253973   .0429067    -5.25   0.000    -.3098465   -.1409482
       2016  |  -.2652879   .0500314    -5.30   0.000    -.3637601   -.1668157
       2017  |  -.2779346   .0536639    -5.18   0.000    -.3835561    -.172313
             |
       _cons |  -12.71203    11.7596    -1.08   0.281    -35.85734    10.43329
-------------+----------------------------------------------------------------
     sigma_u |  .56507733
     sigma_e |   .0901805
         rho |  .97516369   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Notes: Detailed description of regression output for fixed effects model where the natural logarithm of yearly Swedish municipal per capita GHG emissions is regressed on the natural logarithm of municipal mean real income (2008 prices), the square of the same, the natural logarithm of the municipal Gini coefficient, an interaction term between the natural logarithm of municipal mean real income (2008 prices) and the Gini coefficient (non-logarithmic), and year dummy variables.

Table A4: Correlation matrix

	ln GHG	ln CO_2	ln SO_2	ln NO_X	ln PM10	ln Inc	(ln Inc)^2	ln Gini	interact	ln edu share
ln GHG	1.0000									
ln CO_2	0.9334	1.0000								
ln SO_2	0.5456	0.6340	1.0000							
ln NO_X	0.7360	0.7091	0.7515	1.0000						
ln PM10	0.6464	0.5485	0.5159	0.7653	1.0000					
ln Inc	-0.3832	-0.2603	-0.1096	-0.3274	-0.5115	1.0000				
(ln Inc)^2	-0.3852	-0.2621	-0.1118	-0.3290	-0.5113	0.9998	1.0000			
ln Gini	-0.3284	-0.2319	-0.0573	-0.2820	-0.5558	0.4612	0.4670	1.0000		
interact	-0.3742	-0.2600	-0.0744	-0.3182	-0.5867	0.6349	0.6414	0.9724	1.0000	
ln edu share	-0.3345	-0.2254	-0.0395	-0.2843	-0.5449	0.7514	0.7492	0.6525	0.7320	1.0000

Notes: Correlation matrix for all dependent and independent variables used in this thesis, plus the natural logarithm of the rate of higher education attainment. Observations = 2,883.

Appendix B – Environmental degradation indicators

Greenhouse gases, GHG, and Carbon dioxide, CO_2

According to the Swedish EPA (2019a), CO_2 emissions are the single main cause of climate change. Other examples of emitted greenhouse gases, besides CO_2, are methane and chlorofluorocarbons. Human-induced GHG emissions contribute to an augmented greenhouse effect, consequently contributing to a rising temperature (Swedish EPA, 2020).

Sulfur dioxide, SO_2

Sulfur dioxide, SO_2, is a gas emitted through combustion of fossil fuels as well as other substances containing sulfur. Sulfuric acid is created as the gas oxidizes in the atmosphere. Sulfuric acid, in turn, contributes to acidification of soil and water. Inhalation of sulfur dioxide can also cause respiratory problems (Swedish EPA, 2019c). Based on data from the Swedish EPA (2019g), a major source of sulfur dioxide emissions in Sweden, contributing to 13.12 out of the total 17.31 thousand tons emitted in 2018, is industrial combustion and processes.

Nitrous oxides, NO_X

Nitrous oxides, NO_X, are mostly emitted through combustion. Nitrous oxides contribute to acidification of soil and water, and ground-level ozone which in turn contribute to the greenhouse effect. Nitrous oxides are also harmful seen from a perspective of human health (Swedish EPA, 2019b; Swedish EPA, 2019e). The largest sources of nitrous oxides in Sweden are, in falling order based on data for 2018, domestic transportation, industrial combustion and processes, as well as machines used in e.g. forestry and agriculture (Swedish EPA, 2019f).

Particles, PM2.5 and PM10

Sources of human-induced emissions of particles are e.g. combustion and industrial process as well as from the use of studded tires. The mass of particles with a diameter of less than 10 micrometer are measured by PM10. The negative effects of emission of particles relate to health issues—short and long term. Heart diseases and lung cancer are diseases that can, in the longer run, be caused by exposure to particles (Swedish EPA, 2019d).