LUND UNIVERSITY

Eco-Efficiency Analysis of Swedish Regions: A second-stage DEA approach

Lisa Grabo

Josefine Olsson

Supervisor

Pontus Hansson

Master's Thesis

MSc in Economics

The Institution of Economics

Spring 2018

Abstract

This paper uses a two-stage data envelopment analysis (DEA) approach to measure and evaluate the relative eco-efficiency and the influencing factors of 17 out of 21 regions in Sweden during the period of 2008 and 2015. In the first stage of the DEA, a standard CCR model is used to calculate the relative eco-efficiency of the regions given multiple inputs and outputs. The model contains the variables labour, capital and energy consumption as desirable inputs, regional GDP as a desirable output and Co2 emissions as an undesirable output. On average, the results show that four regions are eco-efficiency leaders, eight are eco-efficiency followers and the remaining five are eco-efficiency moderates. The main finding from the first stage of the DEA is that the regions in general should concentrate on energy consumption and capital stock reduction to improve eco-efficiency. However, the results differ slightly across the regions. The second stage of the DEA constitutes a fractional logit regression model, which is used to analyse the factors influencing differences in eco-efficiency across regions. The results indicate that enterprises' production value of goods has a significant and negative effect on eco-efficiency, whereas environmental awareness and environmental sector have positive and significant effects on eco-efficiency. Furthermore, net income and population density are insignificant. Thus, there is potential to improve eco-efficiency in regions via a reduction in energy consumption and in capital stock while keeping the output constant. A concluding remark is that guidance of national legislation should be better implemented at the regional level. Alternatively, regional legislation based on the regions' environmental prerequisites should be implemented to a greater extent.

Key words: eco-efficiency, two-stage DEA model, fractional logit regression model, regional environmental performance, undesirable output

List of abbreviations

CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision-Making Unit
ETFEE	Ecological Total Factor Energy Efficiency
EU	the European Union
OECD	the Organisation for Economic Co-operation and Development
SBM	Slacks-Based Model
SCB	Statistiska Centralbyrån/Statistics Sweden
SKL	Sveriges Kommuner och Landsting/the Swedish Association of Local Authorities and Regions
TFP	Total Factor Productivity
UNCED	United Nations Conference on Environment and Development
WBCSD	the World Business Council for Sustainable Development

Table of Content

1 Introduction

While industrialization is important to economic growth and development of the society, there is a worldwide growing concern about the environmental burden associated with industries and businesses. A relatively new concept in the public discussion on environmental policy is the concept of eco-efficiency. Eco-efficiency has become a popular tool towards the transition to sustainable development, as it does not only account for efficiency with respect to productivity but also efficiency from an environmental point of view (Caiado et al., 2017; Luptáčik, 2001).

Eco-efficiency can be applied on many levels, such as the macroeconomic (national economy), the mesoeconomic (regional) and the microeconomic (organisational) levels (Mickwitz et al., 2006). According to Rosenström and Palosaari (2000), regional efficiency is "the efficiency with which ecological resources are used to meet human needs," thereby expressing how efficient the production process is with regards to natural resources (Zhang et al., 2008). The approach thus enables regional decision-makers to monitor regional changes, using economic-environmental ratio indicators and simultaneously obtain information on the social progress taking place.

As an effective tool to calculate relative efficiency, data envelopment analysis (DEA), developed by Charnes et al. (1978), has gained extensive attention in the literature. DEA evaluates the relative efficiency of entities called decision-making units (DMUs) and thus provides a ranking of the DMUs chosen to observe. This study examines regions as DMUs and can therefore provide a ranking of which regions are efficient and which regions are not.

Existing literature on regional eco-efficiency is relatively scarce, and is mainly concentrated to China. To our knowledge, literature on regional eco-efficiency in European countries has so far only included Poland (Masternak-Janus and Rybaczewska-Błażejowska, 2017). Hence, there is no existing literature covering Sweden's economic and environmental performance at regional level.

Sweden has a reputation of being a global leader in environment degradation prevention, with extensive work in reducing greenhouse gas emissions and using renewable energy sources (OECD, 2014). The Organisation for Economic Co-operation and Development (OECD) refers to Sweden as a "front runner in many fields of environmental policy" in their latest Environmental performance review of Sweden. In 2017, Sweden was ranked as the leader in RobecoSAM's country sustainability ranking, which covers 65 countries around the world (RobecoSAM, 2017). However, as also mentioned by the OECD, there is a lack of guidance from the central government to regional and local authorities on how to implement environmental policies. This could result in differences in eco-efficiency across regions. There is also a need of evaluation in terms of environmental and economic efficiency of various policy instruments (OECD, 2014).

This paper makes multiple new contributions, providing Swedish regional decision-makers with more extensive information on regional performance as well as providing a better understanding of factors that influence the regional differences in eco-efficiency. Thus, this study aims to answer the following research question: *How does eco-efficiency vary among Swedish regions in the period 2008 to 2015, and what are the factors influencing the regional differences in eco-efficiency?*

In order to answer the research question, regional eco-efficiency will be measured using a two-stage DEA approach. At the first stage of the analysis, the eco-efficiency scores for each region will be obtained using a basic DEA model. At the second stage, a regression analysis will be carried out in order to identify which factors can influence the differences in eco-efficiency across regions.

The paper is organized as follows. Section 2 discusses the previous literature on eco-efficiency and different measurement and evaluation approaches. Section 3 provides the first stage of the data envelopment analysis, with a theoretical review of the methodology and the data selection, as well as the empirical study of eco-efficiency and its results. The fourth section examines the second stage of the DEA and analyses the variation in relative efficiency using a regression model. Finally, section 5 provides a discussion and a conclusion of the results previously obtained.

2 Literature review

The literature on environmental effects of production traces back to Leontief (1970) who developed a new input-output model that accounted for externalities, such as pollution. Hence, Leontief's approach raises the undesirable environmental effects of technology and uncontrolled economic growth. In later years, pursuing sustainable development has become more important and debated amongst economies.

A common way to measure eco-efficiency at the national level is using the DEA approach. DEA was first introduced by Farell (1957) and later developed into the so-called CCR model, by Charnes et al. (1978), and has since been expanded greatly. Charnes et al. (1978) introduced the measure of efficiency as a concept that could be used in the evaluation of public programs with multiple inputs and outputs. These outputs were viewed as desirable, whereas Färe et al. (1989)

were interested in comparing multilateral productivity across firms when some of the outputs were undesirable. Thus, Färe et al. (1989) were first to conduct an analysis addressing undesirable outputs.

Different DEA models have been used in the literature depending on the area of interest. Two models are especially common and they are the basic DEA models: the CCR and BCC model (Charnes et al., 1978; Banker et al., 1984). The BCC model is especially useful when dealing with imperfect markets and is therefore sometimes preferred over the CCR model. The CCR model, however, is popular because of its simplicity. Several papers that have examined regional eco-efficiency have also used another DEA model, namely the slacks-based model (SBM) (Li and Hu, 2012; Wang et al., 2010). The SBM does not assume equi-proportional reductions of inputs or outputs and therefore it is a non-oriented model. However, the model measures efficiency of the units using slack variables only, whereas the models developed by Charnes et al. (1978) and Banker et al. (1984) provide both a ranking of efficient and inefficient units and a slacks result.

The two-stage DEA has attracted many authors due to its simplicity and due to the way efficiency is described and interpreted (Raheli et al., 2017, Song et al., 2013) The second stage can involve different types of regressions, however most researchers use the two-limit Tobit approach with limits at zero and unity, first suggested by McDonald (2009). Given the bounded nature of the DEA scores, Hoff (2007) and McDonald (2009) considered using a fractional regression model, established by Papke and Wooldrigde (1996). This approach has later been adopted by a few authors, such as Ramalho et al. (2010) and Raheli et al. (2017), but remains relatively trivial in the second-stage DEA literate.

Moreover, research on eco-efficiency at regional level, using a two-stage DEA, is very scarce and concentrated to China. Song et al. (2013) reviews the environmental efficiency and its influencing factors in China, using the SBM to measure the efficiency and a Tobit regression to test the influencing factors. The paper shows that the variables GDP per capita dependent on foreign capital and trade, environmental awareness and population density have positive and significant impacts on environmental efficiency, while the proportion of the secondary industry in the GDP shows a significant and negative impact on the environmental efficiency. Wang et al. (2010) also studies regions in China, using the SBM and Luenberger productivity indicator to measure environmental efficiency as well as environmental total factor productivity (TFP) growth. The main findings of the paper are that the overuse of energy and the excess of emission are the main sources of environmental inefficiency. Gross regional product per capita, foreign direct investment, structure factor, the capacity for environmental management of government and firms, environmental protection awareness of people proved to have a varying effect on the environmental efficiency and environmental TFP growth. Furthermore, Li and Hu (2012) analyses the ecological total-factor energy efficiency (ETFEE) in China using the SBM following by a truncated regression model. A finding from the DEA is that China should concentrate on both energy saving and emission reduction to improve its energy efficiency. The truncated regression results argue that energy efficiency tends to improve with economic development. A higher ratio of research and development expenditure to GDP and a higher degree of foreign dependence contribute to a higher ETFEE. Moreover, a higher ratio of secondary industry to GDP and a higher ratio of government subsidies for industrial pollution treatment to GDP contribute to a lower ETFEE.

Other papers that examine eco-efficiency at regional level have used a one-stage DEA, and are also concentrated to China. There is only one paper, to our knowledge, that examines regional eco-efficiency in the European region, which is the Masternak-Janus and Rybaczewska-Błażejowska (2017). The paper uses a CCR model and concludes that the model is a valuable alternative approach for the calculation of eco-efficiency. A finding from the analysis is that promotion of development of intellectual and technological innovations, leading to the restructuring of the industry, employing new business models and introducing new organizational forms aimed at increasing the efficiency of the use of resources should be promoted to increase eco-efficiency.

A concluding remark is that even though there exist studies with similar evaluation methods to analyse comparable problems, their theoretical cores may vary significantly and are concentrated to China especially. So far, there are no studies on regional eco-efficiency that use a fractional regression model in the second stage of the DEA. Moreover, there is no existing literature that evaluates the eco-efficiency of Sweden and its regions. Thus, this study is an addition to the literature on regional eco-efficiency.

3 First stage: Data envelopment analysis

Relative efficiency is a general measurement that has in recent years gained popularity in various fields of investigation. Relative efficiency might be used for example to measure banks' efficiency performance, environmental efficiency among enterprises and hospitals' efficiency performance. A common method to measure relative efficiency is the data envelopment analysis approach, which evaluates the relative efficiency of a set of comparable entities called decision-making units (DMUs) with multiple inputs and outputs.

One approach to measure environmental performance at the national level is using either the one-stage or two-stage DEA approach. In the one-stage model, efficiency is measured while controlling for the influence of exogenous variables, whereas in the two-stage model variation in efficiency is attributed to variation in the exogenous variables (Fried et al., 1993). Thus, using a two-stage DEA approach allows for a comprehensive evaluation of DMUs' efficiency and which factors that influence the differences between the DMUs over time. For the purpose of this study, Swedish regions will be considered as DMUs where the environmental performance of each region will be estimated using the two-stage DEA approach.

This chapter will first review the concept of DEA and the model chosen. Thereafter, the concept of inefficiency, data used and sensitivity of the results will be discussed as well as a brief description of the regions will be given. Finally, the DEA results will be presented.

3.1 Methodology

Data envelopment analysis is an approach, which entails the relative production efficiency for a set of comparable decision-making units. The DMUs exhibit one or multiple inputs and outputs, where inputs are minimized and outputs are maximized in order to obtain optimal efficiency. The DMUs that exhibit optimal efficiency form a technological frontier (reference set) where the shape of it depends on the assumptions made about the returns to scale of the DMUs. The distance to the technological frontier is used to compute the relative performance of each DMU. Thus, DEA is a benchmarking tool that can only be used to observe relative efficiency among a set of DMUs and not absolute efficiency.

A considerable advantage of DEA is that it does not require any distinct functional relationship, for example a Cobb-Douglas production function, between inputs and outputs. Hence, no assumption regarding a production function has to be made. Another advantage of the method is that it can treat inputs and outputs with different units and therefore no transformation of the units is required. Therefore, it is a simple method to estimate efficiency among a group of entities such as enterprises, cities, regions or even countries.

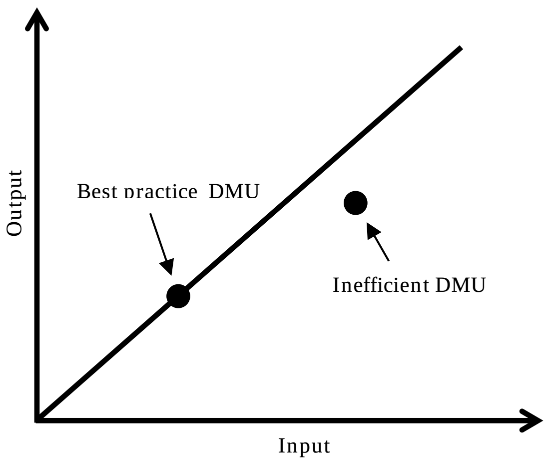
In terms of choosing between the CCR and the BBC model, there are several aspects to consider. The two models are similar but with one important difference, that is, the former assumes constant returns to scale while the latter assumes varying returns to scale (Charnes et al., 1978; Banker et al., 1984). Both have been used in analysing eco-efficiency, but with different results. The BCC model is especially useful when dealing with imperfect markets as it assumes that each DMU can exhibit different returns to scale from another DMU. However, the CCR model is popular because of its simplicity. It has been argued that as long as the DMUs have rather homogenous prerequisites, the CCR model can be used successfully. Therefore, the input-oriented CCR model is used for this study and will be further introduced in section 3.2.

A DEA model can either be input-oriented or output-oriented, where the former treats how inefficient DMUs can achieve efficiency through a reduction of its inputs while its outputs are held constant. The output-oriented approach requires instead an increase in output, while holding its inputs constant, to achieve efficiency. The CCR model will yield the same eco-efficiency scores regardless of the orientation. However, since the model in this study will have fewer outputs than inputs, the input-oriented approach is being considered solely. It is also a useful approach when dealing with undesirable outputs, which is being discussed in section 3.2.

3.2 The CCR model

The CCR model was first developed by Charnes et al. (1978) and assumes constant returns to scale, imposing that an increase in inputs will give a proportional increase in outputs. The model measures the overall technical efficiency and corresponds to a DMU that produces maximum output given a specific set of inputs.

Figure 1. The technological frontier of the input-oriented CCR model



Source: Created by authors

If n DMUs (DMU_j, j=1,2,...,n) are assumed to be present, where each consumes m different inputs (x_{ij}, i=1,2,...,m) to produce s different outputs (y_{rj}, r=1,2,...,s). The matrix of inputs is then X={x_{ij}, i=1,2,...,m} and the matrix of outputs is Y={y_{rj}, r=1,2,...,s}. As the inputs and outputs have different levels of significance for each DMU, weights must be assigned to each input and output. However, no subjective assumptions regarding the size of the weights have to be made. The weights are instead obtained by solving the following problem:

$$\max \sum_{r-1}^{s} u_r y_{rq} / \sum_{i-1}^{m} v_i x_{iq}$$

$$st. \sum_{r=1}^{s} u_r y_{rj} / \sum_{i=1}^{m} v_i x_{ij} \leq 1,\ j = 1,2,\dots,n \tag{1}$$

$$u_r \geq 0,\ r = 1,2,\dots,s$$
$$v_i \geq 0,\ i = 1,2,\dots,m$$

where u_r is the weight of rth output (r=1,2,..,s), v_i is the weight of the ith input (i=1,2,...,m), y_{rq} is production of the rth output (r=1,2,...,s) for the qth DMU, x_{iq} is consumption of the ith input (i=1,2,...,m) for the qth DMU, y_{rj} is production of the rth output (r=1,2,...,s) for the jth DMU (j=1,2,...,n), x_{ij} is consumption of the ith input (i=1,2,...,m) for the jth DMU (j=1,2,...,n).

The CCR model can then, by transformation, be obtained. The ordinary linear program can take the form:

$$\theta^* = \min[\theta - \varepsilon(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+)]$$

$$st. \sum_{j=1}^{n} x_{ij}\lambda_j + S_i^- = \theta X_{ij},\ i = 1,2,\dots,m \tag{2}$$

$$\sum_{j=1}^{n} Y_{rj}\lambda_j - S_r^+ = Y_{rj},\ r = 1,2,\dots,s$$

$$\theta, \lambda, S_i^-, S_r^+ \geq 0;\ j = 1,2,\dots,n$$

where θ is the efficiency of the DMU, S_{r^+} and S_{r^-} are slack variables, ε is non-Archimedean constant (10^{-6} or 10^{-8}) and λ_j is a weight assigned to the DMU_j (j=1,2,...,n).

In terms of eco-efficiency, one common approach is to distinguish between desirable and undesirable inputs and outputs in the DEA models. For example, undesirable outputs may be pollutants such as waste, CO2 and sulphur dioxide (Li et al. 2013). In order to be eco-efficient a DMU should minimize undesirable outputs along with its inputs. However, as the CCR model does not distinguish between desirable and undesirable inputs and outputs, some adjustments to the model must be made. One way is to transform the data of the undesirable output according to the multiplicative inverse such as f(U) = 1/U (Golany and Roll, 1989). However, this approach is complicated as the transformation becomes non-linear. Another approach is to use the ADD approach introduced by Koopmans (1951), which transforms the undesirable output to f(U)=-U. This may, however, result in data becoming negative, and it is not straightforward to define efficiency scores for negative data. A third approach is to treat the undesirable output as a desirable input (Dyckhoff and Allen, 2001). If only operational efficiency is investigated from this point of view, there is no need to distinguish between inputs and outputs, but only minimum and maximum. For the purpose of this study, the undesirable output is simply treated as a desirable input.

3.3 Interpretation of inefficiency

A DMU with an efficiency score lower than one is identified as inefficient. The total inefficiency can be reduced by first reducing all inputs proportionately up to $1-\theta^{*}$, then making a further reduction in slacks of certain inputs. Slacks are the leftover portion of inefficiencies and are only present for inefficient DMUs. Reducing the slack is necessary if a DMU cannot reach the technological frontier despite the proportional reductions in all inputs.

3.4 Data in the DEA

The sample in this study comprises of 17 out of the 21 counties in Sweden, for the time period 2008 to 2015. For simplicity, counties are in this study called regions. Skåne, Östergötland and Jönköping have been excluded since data for capital stock is missing for all time periods. Stockholm is identified as an outlier and has therefore been excluded as well. Even though four regions have been excluded, the rule of thumb established by Golany and Roll (1989), which states that the number of DMUs should be at least twice the number of inputs and outputs, is fulfilled to make sure that some discriminatory power of the model exists. The input and output variables that are included in the data set have been collected from Statistics Sweden (Swedish: Statistiska Centralbyrån, SCB) and the Swedish Association of Local Authorities and Regions (Swedish: Sveriges Kommuner och Landsting, SKL). All data has been mean normalized in order to ensure it is of the same or similar magnitude so that no imbalances will arise.

3.4.1 Choice of variables

To measure eco-efficiency accurately, all input and output variables should be considered to the extent that the data is available. The desirable inputs used in this study are labour, capital stock and energy consumption. The variables have been chosen, as they all appear in previous literature and are thus well established in DEA reserach (Huang et al., 2014; Luptáčik 2001; Song, 2013). Due to lack of data at regional level, only energy consumption could be used in this study as an environmental input. Variables that could have been used additionally are waste management, water consumption and land use.

Labour is defined as the total number of employees in a certain region in a specific year. Similarly, the capital stock is measured as the total capital stock a given year in a specific region. As the data is measured in million SEK it has been transformed into 2010 constant prices to avoid any measurement errors due to inflation. Furthermore, regional data on energy consumption was missing for some regions. To solve for this problem, data was instead collected on a local municipality level and added up to account for total regional energy consumption. Some data was however missing on the local level as well due to confidentiality. To correct for this, existing data for previous or future year's data was used. An illustration of this is if data for 2015 existed but not for 2014, then the same data for the two years was simply used. In those cases where data for both previous and future years but not current year existed, the average of the two values was estimated to receive a value for the current year. It should also be noted that some municipalities have not registered their use of energy consumption for some years, meaning that energy consumption takes the number of zero for those years. This may affect the results of this study. However, since the municipalities that experienced this sort of problem are small in size and in population, they should not account for a significant part of the total energy consumption in the region.

A desirable output that has been frequently used in similar studies is the regional GDP and has therefore been included in this study as well (Song et al., 2013; Huang et al. 2014; Wang, 2014). The regional GDP is measured in million SEK and has been transformed into 2010 constant prices. As an undesirable output, environmental pollutants such as CO2 emissions have been chosen. It is a common approach to use emissions in general as undesirable outputs and therefore CO2 emissions is used in this study (Yang & Pollitt, 2009; Liu et al., 2017). Since data for other pollutants at the regional level was only accessible for every second year, only CO2 emissions are used and the data is measured in kiloton. Other variables that could be treated as undesirable outputs are waste and other pollutants. However, due to lack of data at regional level, these variables could not be used in this analysis.

3.4.2 Description of regions

The paper uses the Classification of Territorial Units for Statistics (NUTS), established and regulated by the European Union (EU), to divide Sweden into different subdivisions. For each EU member, there are three levels of NUTS. In Sweden, NUTS 1 consists of three country regions, NUTS 2 of eight national areas and NUTS 3 of 21 counties. This paper will examine the 17 NUTS 3 regions viewed in table 1.

Table 1. Classification of regions

NUTS 2	**NUTS 3**
Östra Mellansverige	Uppsala Södermanland Örebro Västmanland
Småland med öarna	Kronoberg Kalmar Gotland
Sydsverige	Blekinge
Västsverige	Halland Västra Götaland
Norra Mellansverige	Värmland Dalarna Gävleborg
Mellersta Norrland	Västernorrland Jämtland
Övre Norrland	Västerbotten Norrbotten

Table 2 gives a brief overview of the situation in Sweden excluding Stockholm, Jönköping, Östergötland and Skåne. As can be seen, the total labour input has grown by 11 percent from 2008 to 2015, and the capital stock has increased by approximately 147 percent at the national level. The energy consumption and Co2 emission, however, have decreased by almost 3 percent and 15 percent respectively, while regional GDP has increased by 55 percent. These figures indicate a high economic growth rate that has been accompanied by a high growth rate in capital stock and lower growth rate of labour.

Table 2. Summary of input and output indicators at national level

Variable	**Unit**	**Annual quantity**		**Growth rate %**
		2008	2015	
Input				
Labour	Thousand persons	3 689	4 094	11.00
Capital stock	Million SEK	26 274	64 772	146.52
Energy consumption	Mega Watt Hours	187 673 176	182 339 740	-2.84
Desirable Output				
Regional GDP	Million SEK	2 646	4 112	55.36
Undesirable Output				
CO2	Thousand tons	51 483	43 897	-14.73

Table 3 shows the average values of input and output indicators at the NUTS 2 level in Sweden. As can be viewed, input values for Östra Mellansverige and Västsverige are far higher than the other NUTS 2 areas. This might indicate that the production process in Östra Mellansverige and Västsverige is more energy intensive than in other regions. The level of CO_2 emissions is significantly higher in Västsverige, which might be another indicator of a heavier production process in this area than elsewhere. Övre Norrland does as well emit quite large amounts of CO_2, which might be a result of an energy intensive industry.

Table 3. The average values of input and output indicators at the major regional level

Variable	Östra Mellansverige	Småland	Sydsverige	Västsverige	Norra Mellansverige	Mellersta Norrland	Övre Norrland
Input							
Labour	527	228	69	948	365	175	247
Capital Stock	7 819	4 568	1 352	7 864	2 845	1 843	3 853
Energy consumption	58 184 631	37 952 923	655 109	74 025 402	1 984 131	2 136 644	2 592 315
Desirable output							
RGDP	398	160	48	744	265	123	182
Undesirable output							
CO2	6 391	3 958	540	14 474	4 020	1 933	6 427

Notable is that there are some differences in variables both across time and across NUTS 2. When studying the regions, it is for example notable that Västra Götaland has about five times as much average regional GDP compared to the region with the second highest value of average regional GDP, Uppsala. Västra Götaland has 34 percent of the total average regional GDP whereas Gotland has less than 1 percent. However, if a region has high values in both inputs and outputs, this does not necessarily affect the relative efficiency scores, as they will balance out.

3.5 Sensitivity analysis

Measurement errors and random variations are negative aspects of the basic DEA model since it does not account for any uncertainty in data. To make sure that the model is stable, a sensitivity analysis is conducted to check for extreme values in the data. One method for doing this is jackknifing, which in practice means excluding DMUs from the set one by one and checking how the efficiency rates are affected by doing this. One version of the jackknifing method is to first exclude a DMU in the reference set and test for efficiency of the remaining population. Next, another DMU in the reference set is excluded without adding the previously excluded DMU, making the population set now smaller by two DMUs. The efficiency of the remaining population set is then measured once again. This procedure is repeated until all DMUs in the reference set have been excluded from the population set. All efficiency scores are finally compared to check for any major variances. Another version of the jackknifing method is to first exclude a DMU located on the technological frontier and to then measure the efficiency of the remaining population of DMUs. These efficiency scores are then compared to the original efficiency scores to check for any differences. Thereafter, the DMU is added to the data set again and another DMU in the reference set is excluded, repeating the same procedure. For the purpose of this study, both versions of the jackknifing method are used.

There exist several methods to test for differences between model specifications. Yet, DEA is a complicated measurement from a statistical point of view and hence it might be hard to ensure that the method chosen is stable. One method that can be used is the Spearman rank correlation test, which can define whether one model performs better than another. The test estimates the correlation between two variables by defining a monotone function of the two variables. A perfect score of +1 or -1 implies that there exist a perfect Spearman correlation between the two and hence one variable is a perfect monotone function of the other. In this study, three alternative models are tested against the original model where one input is removed from each model. If the models are highly correlated according to the Spearman rank correlation test, it can be concluded that the choice of model does not affect the result of this study. The three models tested can be viewed in table 4.

Table 4. Alternative models for model specifications

Variable	Original model	Model 1	Model 2	Model 3
Inputs				
Labour	X		X	X
Capital stock	X	X		X
Energy consumption	X	X	X	
Output				
Regional GPD	X	X	X	X
Undesirable Output				
CO2 emissions	X	X	X	X

3.6 Results of the DEA

The results from the data envelopment analysis can be observed in table A4 in appendix and indicate that the overall efficiency score across regions is on average 0.91. It indicates that there is a capacity of improving the efficiency among the regions up to 9 percentage points on average. Based upon the estimated average eco-efficiency scores for all regions, they have for the purpose of this study been divided into the following groups to easily interpret the results:

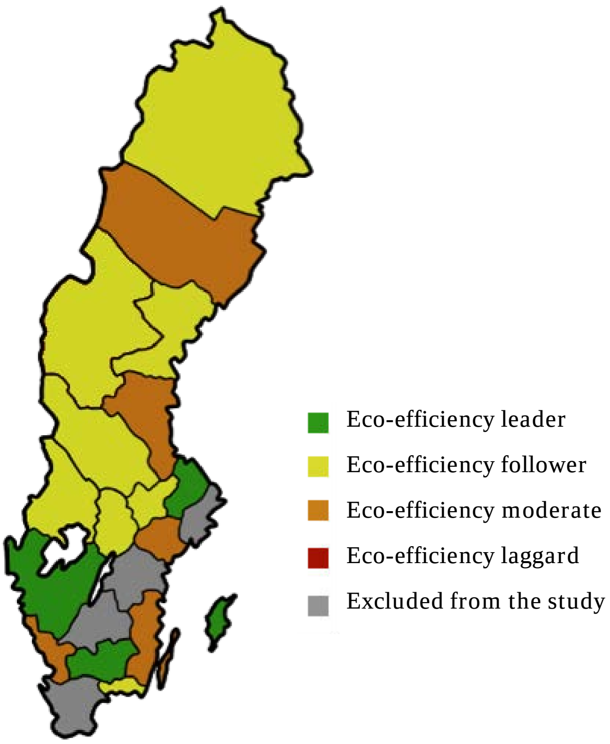
- Eco-efficiency leaders, having eco-efficiency scores of 1
- Eco-efficiency followers, having eco-efficiency scores of 0.90 to 0.99
- Eco-efficiency moderates, having eco-efficiency scores of 0.80 to 0.89
- Eco-efficiency laggards, having eco-efficiency scores lower than 0.80

A summary of the results can be observed in table 5, which describes whether a region has experienced slacks in an input in one or more periods from 2008 to 2015. As the results suggest, regions often experienced slacks in capital stock and energy consumption while it is less likely that slacks have been present in labour and emissions.

Table 5. Summary results

Region	Average score	Slacks			
		Labour	Capital stock	Energy consumption	Emissions
Leaders					
Kronoberg	1		X	X	
Gotland	1				
Västra Götaland	1				
Uppsala	1				
Followers					
Dalarna	0.98			X	
Värmland	0.98	X		X	
Västernorrland	0.95		X	X	
Örebro	0.94		X	X	
Västmanland	0.93		X	X	
Norrbotten	0.92		X	X	X
Blekinge	0.91		X	X	
Jämtland	0.90	X	X	X	X
Moderates					
Gävleborg	0.89		X	X	
Södermanland	0.88		X	X	X
Västerbotten	0.88		X	X	
Halland	0.86		X	X	
Kalmar	0.86		X	X	X

Figure 2. Division of eco-efficiency groups



3.6.1 Uppsala

When observing the average eco-efficiency scores, Uppsala is in the top as an eco-efficiency leader. This means that on average, Uppsala did not need to reduce any inputs in order to remain an eco-efficiency leader during the time period 2008 to 2015. However, in an in-depth analysis for each year it is notable that Uppsala did not achieve maximum score across all years. In 2010 and 2015, Uppsala received an efficiency score of 0.98 and 0.99 respectively, meaning that the region can improve its efficiency proportionately by reducing its inputs up to 2.5 percent in 2010 and up to 1 percent in 2015. In addition to the proportional reduction, the region needed to make further reductions in its capital stock and energy consumption to reach the technological frontier.

3.6.2 Södermanland

In terms of average eco-efficiency, Södermanland performs relatively poorly as an eco-efficiency moderate. Its score of 0.88 means that it needed to improve its efficiency proportionately by reducing its inputs on average up to 12 percent across all years. If looking at the results each year, Södermanland performed relatively well in the beginning of the time period and then worsened in 2013 and onwards. In 2008 up to 2010, the region could have made further reductions in its energy consumption and pollution. However, in 2o11 there is a shift in which area further reductions could have been made. According to the slack results, the focus after 2011 was on reducing its pollution and capital stock rather than its pollution and energy consumption.

3.6.3 Kronoberg

Kronoberg is an eco-efficiency leader across all years in this study and thus has an eco-efficiency score of one throughout the entire period. This indicates that it did not need to improve its eco-efficiency to remain relatively efficient.

3.6.4 Kalmar

The region performs, on average, worse out of all regions studied in this paper. Its score remains relatively stable across the time period, with scores shifting between 0.81 and 0.91. On an average, Kalmar can improve its efficiency proportionately by reducing its inputs up to 14 percent across all years. The slack results indicate that further focus should especially be on reducing its capital stock and energy consumption. In 2009 and 2013, the slack results instead implied a shift from a reduction in the capital stock to a reduction in pollution.

3.6.5 Gotland

Gotland is an eco-efficiency leader across all years in the period and thus has an eco-efficiency score of one throughout the entire period. This indicates that it did not need to improve its eco-efficiency to remain relatively efficient.

3.6.6 Blekinge

The region obtained a score of 0.91 on average, meaning that it performs as a follower on average. Blekinge was efficient in the first year of the time period, but has since then received a score between 0.84 and 0.92. To reach the technical frontier, it needed to reduce its inputs with nine percent on average, and make additional reductions in the capital stock and energy consumption.

3.6.7 Halland

Halland ranks as an eco-efficiency moderate on average, with a score of 0.86. Its scores have shown some variation over the period, ranging between 0.79 and 0.92. The region has dropped continuously in the ranking since 2013. Besides the proportional 14 percent reduction in all inputs, on average, Halland needed to make further reductions in its capital stock and especially in its energy consumption.

3.6.8 Västra Götaland

The region is an eco-efficiency leader across all years in this study and thus has an eco-efficiency score of one throughout the entire period. This indicates that it did not need to improve its eco-efficiency to remain relatively efficient.

3.6.9 Värmland

The region performs, on average, as a follower with a score of 0.98. It reaches efficiency in the first year of the period and in the last two years, but not the years in between. On an average, Värmland can improve its efficiency proportionately by reducing its inputs up to only two percent across all years. The slack results indicate that a further reduction could have been made in the energy consumption in the period 2009 and 2013. In 2011 and 2013, further reduction could also have been made in labour.

3.6.10 Örebro

The region is on average an eco-efficiency follower with a score of 0.94. It reaches efficiency in 2012, but the scores other years lay between 0.85 and 0.98. On an average, Örebro can improve its efficiency proportionately by reducing its inputs up to six percent across all years. Further reduction is suggested by the slack results, especially in its energy consumption and its capital stock. An exception to this is in 2008 where the slack results indicate that no further improvement is needed despite not being efficient that year.

3.6.11 Västmanland

In terms of average eco-efficiency, Västmanland ranks as a follower with a score of 0.93. It does not reach efficiency in any given year, but it is very stable across time with scores lying between 0.91 and 0.95. Thus, on average, it needed to improve its efficiency proportionately by reducing its inputs up to seven percent across all years. The slack results indicate that further reduction can be made in its energy consumption and capital stock.

3.6.12 Dalarna

Dalarna is a leader in the beginning of the time period, but from 2014 and onwards it drops its ranking and becomes an eco-efficiency follower, which results in a position as a follower on average. On average, the region scores 0.98, which means that it only needs a proportional reduction on 2 percent in all inputs on average each year. Despite not being efficient in 2014, no further reduction is required according to the slack results. However, in 2015, the results suggest that further improvements are needed in the energy consumption.

3.6.13 Gävleborg

The region is on average an eco-efficiency moderate with a score of 0.89. It never reaches total efficiency in any given year, but it is very stable across time with scores lying between 0.84 and 0.94. On an average, Gävleborg can improve its efficiency proportionately by reducing its inputs up to 11 percent across all years. In 2008 no further improvements are necessary. In 2009 and 2010, further reduction can be made in energy consumption, whereas in 2011 to 2014 further reductions are needed in the capital stock. In 2015, a further reduction is suggested in both the capital stock and in the energy consumption.

3.6.14 Västernorrland

Västernorrland scores on an average 0.95 and can therefore be defined as an eco-efficiency follower. It never reaches efficiency in any given year, however it is very stable in a range between 0.87 and 0.98, with only one score below 0.95. The lowest score was registered in 2008, but despite the low score no further improvements other than the proportional reduction was needed in order to achieve efficiency. Following years, the slack results suggest that further improvements are needed in capital stock and in energy consumption.

3.6.15 Jämtland

Jämtland is on average an eco-efficiency follower, with a score of 0.90. The scores vary a bit over time, ranging between 0.80 and 1. The region achieved efficiency in 2010, 2014 and 2015. Further improvements, besides the proportional reduction in all inputs, needed to be done in different inputs across time. In 2008, a reduction in the capital stock only was required to reach the frontier. In 2009, this changed to a result that indicated a reduction in the energy consumption and pollution. In 2012 and 2013, a reduction in energy consumption was suggested by the slack results. However, the results also suggest that a reduction in labour is necessary to become efficient.

3.6.16 Västerbotten

In terms of average eco-efficiency, Västerbotten ranks as an eco-efficiency moderate with a score of 0.88. The region's scores are relatively stable across the time period, with a range between 0.83 and 0.90. Thus, on average, Västerbotten needs to reduce its inputs proportionately with 12 percent each year. Moreover, despite not being a leader in 2008, it did not need to do any further reductions to become efficient. Further reducing is required to achieve efficiency, however, which input that should have been further reduced differs across time. On average, the slack results indicate further reductions in energy consumption and capital stock.

3.6.17 Norrbotten

Norrbotten is on average an eco-efficiency follower, with a score of 0.92. It was a leader in 2008, to then drop in the ranking quite a lot in 2009. It managed to rise in ranking in 2010 and was a stable leader in 2010 to 2012. After 2012, it begun to drop in the ranking again, which it has done continuously each year after this. In 2009, the slacks results show that, in addition to reduce all inputs proportionately, it could further reduce its capital stock, energy consumption and pollution. Looking at the slack results in general, Norrbotten has a bigger need to reduce its pollution compare to other regions. Expect for reduction in pollution, the slack results for Norrbotten also imply that further reduction is needed in the capital stock. In 2015, a reduction in the energy consumption is also required to reach the frontier.

3.7 Results of the sensitivity analysis

The results of the sensitivity analysis can be observed in the appendix, table A1 and A2. Four DMUs are considered eco-efficiency leaders and are hence included in the reference set. These are Kronoberg, Gotland, Västra Götaland and Uppsala. They have all been excluded from the population set according to the two versions of the jackknifing method described in section 3.6. The results show that there are small differences between the efficiency scores when removing DMUs in the reference set from the population set. When removing Gotland from the population set, the differences between the efficiency scores are a bit higher compared to the other cases when other DMUs are being excluded. However, as these differences are small it can be concluded that the chosen DEA model is stable.

The Spearman rank correlation test showed that there is a relatively high correlation between the original DEA model and the other models stated in this analysis. Model 2 has the highest correlation with the original model, with a correlation of 0.89, while Model 3 has a correlation coefficient of 0.73 with the original model. Model 1 has a Spearman rank correlation of 0.67 with the original model. These scores imply that depending on which model is chosen, the efficiency scores might vary within the regions and across the years, which is undesirable. However, as the original DEA model only accounts for three inputs, the effect of removing one input might be higher than in the case where the number of inputs is much larger. As data on regional level for other inputs is missing, these inputs cannot be included at this stage. Therefore, it can be concluded that the model specification should perform well enough and that it is reasonably robust, even though the inclusion or exclusion of some inputs might improve the robustness even further.

4 Second stage: Regression analysis

In order to improve the regional environmental performance, it is important to understand which other components than those treated in chapter 3 can influence the variation received in the eco-efficiency. Chapter 3 thus provided the first stage of the DEA, where the eco-efficiency scores were presented. Chapter 4 is the second stage of the two-stage DEA approach, which aims to explain which exogenous factors can influence differences in relative eco-efficiency among the regions. First, the chapter introduces the regression model used for this study. Secondly, it will give a description of the data and why it was chosen. The last part of chapter provides the regression results and how to interpret them.

4.1 Regression model

To gain a deeper understanding of what causes differences in relative eco-efficiency among regions, a regression analysis is performed for the time period 2008 to 2015. A common econometric model used in this step of the analysis is the Tobit regression model (Merkert and Hensher, 2011; Wang and Huang, 2007; Kirjavainen and Loikkanent, 1998). However, several authors have argued that the efficiency data is not censored by nature, as efficiency scores can only take values on the unit interval [0,1]. Therefore no censoring process is needed as no scores take a value outside the given interval (Simar and Wilson, 2007; McDonald, 2009). Thus, a Tobit model might not be appropriate for the second-stage DEA analysis.

The fractional model developed by Papke and Wooldridge (1996) has been argued to be better suited when the dependent variable is continuous and defined on the unit interval. To ensure that the independent variables are on the unit interval, a probit, logit or heteroskedastic probit must be used. For the purpose of this study, the fractional logit regression model is therefore used.

To better understand the fractional logit regression model, assume a continuous dependent variable y in the unit interval and a vector of independent variables x_i. To fit a regression with these conditions, the mean of the dependent variable y conditional on the independent variable x, $E(y|x)$, is considered. The fractional logit model accomplishes this by using equation 3.

$$E(y|x) = e^{x\beta}/(1 + e^{x\beta}) \tag{3}$$

The model is a quasi-likelihood estimator and as a consequence of this, the true distribution of the entire model does not have to be known to receive consistent parameter estimates. The only information needed is the correct specification of the conditional mean. Thus, consistent point estimates can be received while asymptotically efficient standard errors cannot be received, if the conditional mean of the model is the same as the conditional mean of another model. For example, if the conditional mean of the model is the same as the conditional mean of a probit when the model used is not a probit. Furthermore, as no assumptions about the distribution of the unobserved elements are made in the model, the standard errors are not efficient. Thus, robust standard errors are used in order to receive unbiased estimators

4.2 Data in the regression analysis

As observed in chapter 3, several regions were inefficient and where slacks in inputs were one reason for this. However, in order to explain which factors influence exogenous differences in eco-efficiency, other variables than those used in the first stage of the DEA will be observed in a regression model. The regression analysis accounts for both economic and environmental aspects of the eco-efficiency measure, using two indicators for economic development and environmental protection respectively. The regression analysis also provides one indicator for regional factors. Table 5 gives a brief overview of the variables chosen for the regression analysis.

Table 5. Description of indexes

Indexes	**Explanatory variable**	**Abbreviation**	**Units**	**Meaning**
Economic development level	Net income	NI	SEK in thousand	Average net income in households
	Enterprises	EP	%	The proportion of enterprises' production value of goods to the regional GDP
Environmental protection	Environmental sector	ES	%	The proportion of workers in the environmental sector to the total amount of workers
	Environmental awareness	EA	%	The proportion of the population with an education level above high school
Regional factors	Population density	PD	People/km^2	The proportion of the total population to the regional area by the end of the year

The eco-efficiency scores retrieved in chapter 3 are relative values while a regression analysis in first hand analyses absolute values. Thus, the explanatory variables have been recalculated into indexes to make sure that the results retrieved are consistent. Kronoberg was therefore set as a reference region to the other regions.

Table 6. Descriptive statistics of transformed data

Variable	Mean	Standard deviation	Min	Max
Net income (NI)	0.98	0.06	0.83	1.14
Enterprises (EP)	0.99	0.14	0.59	1.29
Environmental sector (ES)	0.70	0.15	0.49	1.22
Environmental awareness (EA)	0.99	0.12	0.85	1.35
Population density (PD)	1.25	0.91	0.11	3.06

4.2.1 Economic development level

The environmental Kuznets curve discusses that higher environmental quality is more desired than income and employment when households' living standard has improved above a certain income level (Shafik, 1994). This leads to a tendency of promoting environmental quality and efficiency in highly developed countries and the opposite in underdeveloped countries. As a result, many studies using the two-stage DEA approach for regions have included regional GDP in the regression analysis. However, including regional GDP in the regression model at the second stage of the DEA might give biased estimation results. As regional GDP was used as an output in the first stage, autocorrelation might appear in the second stage. Therefore, regional GDP is not used in this study. Instead, net income (NI) is chosen as one economic development level indicator. Net income (NI) is hence expected to have a positive effect on eco-efficiency according to the theory of the environmental Kuznets curve. Data for net income (NI) has been collected from SCB and is the sum of all individuals' taxable and non-taxable incomes excluding taxes, at the age 16 and above. One possible disadvantage of the net income (NI) variable is that it is not possible to trace whether income has been earned within the region or in neighbouring regions. However, due to lack of data, net income (NI) is still considered as the best option for the purpose of this study, and should not affect the results significantly.

The second indicator for economic development level is the proportion of enterprises' production value of goods of each region to the regional GDP (EP). Thus, two variables are representing the economic development level: net income (NI) and the enterprises' production value of goods to the regional GDP (EP). The enterprises production value of goods (EP) is a good indicator of the economic development level, as it covers the industrial structure. This in turn is correlated with CO_2 emissions (Song et al., 2013), part of the eco-efficiency measure. The enterprises' production value of goods (EP) is thus expected to have a negative effect on eco-efficiency, as pollution and water-use will increase as a result of an increasing consumption demand and hence a rising exploitation of resources. Data for the enterprises (EP) variable has been collected from The Swedish Agency for Economic and Regional Growth (Swedish: Tillväxtverket).

4.2.2 Environmental protection

The government plays an important role in reducing the CO2 emissions by increasing the environmental awareness and creating incentives to fulfil environmental protection regulations among households and enterprises. Mass media is one tool to increase the awareness while education is another. Therefore, environmental awareness (EA) is investigated as one aspect of the environmental protection indicator. It has as well been argued that environmental awareness increases with the level of education (Song et al., 2013; Aminrad et al., 2011; Shobeiri et al., 2007). Hence, awareness is measured as the proportion of the population with an education level above high school. Thus, the same measurement but for each region is used in this study. The effect of environmental awareness (EA) on eco-efficiency is expected to be positive, which indicates that when awareness (EA) increases efficiency should as well increase.

The environmental sector (ES) has been included in the analysis as the second environmental protection indicator. The environmental sector (ES) measures the strength of the environmental protection in each region and reflects the proportion of workers in the environmental sector to the total amount of workers. The environmental fields within the sector ranges from waste management and renewable energy sources to environmental consultants and recycled materials, and so forth (SCB, 2015). Hence, the environmental sector (ES) is an important tool to decrease environmental degradation by different means. To increase the efficiency it is therefore important that investments are made into the environmental sector (ES) and that the environmental enterprises are favoured over other means of production. As a result, the environmental sector (ES) is expected to have a positive effect on eco-efficiency. Data for both environmental awareness (EA) and environmental sector (ES) has been collected from SCB.

4.2.3 Regional factors

Regional factors such as geographical location, population, climate and resources might cause differences in abilities in acquiring technology and information in each region, which further impact the eco-efficiency scores. Regional factors can be expressed by population density (PD), which measures the proportion of the total population to the regional area (Song et al., 2013). The data over population density has been collected from SCB. The effects of population density (PD) can be many. For example, a higher population may have a positive effect on eco-efficiency because more people are going into research on how to reduce the environmental impact. Also, a high population density (PD) might imply a more developed infrastructure, which in turn implies that more people use other means of transportation than cars. However, a higher population density could also have a negative effect on eco-efficiency since it could cause more environmental pressure. Thus, the results may be difficult to interpret. Although, from the point of view of this study, population density (PD) is expected to have a weakly negative effect on eco-efficiency, since total environmental pressure caused by a high density may be more valued than the benefits of a high population density.

4.3 Results of the regression analysis

As previously mentioned, an advantage with the fractional logit regression model is that it is robust, and therefore corrects for both heteroskedasticity and autocorrelation in the model. Despite this, tests indicate that heteroskedasticity is present in the model but autocorrelation is not. However, this does not exhibit a problem as previously suggested. A test was also conducted to check if multicollinearity is present in the data, in which statistical inferences made about the data might not be reliable. The results indicated no detection of multicollinearity, thus there are no high inter-correlation among the independent variables.

The result of the fractional logit regression model can be viewed in table 7. The result suggests that there are only two variables that are significant at the 5 percent significance level and which are enterprises' production value of goods (EP) and environmental awareness (EA). Environmental sector (ES) is significant at the 10 percent significant level while net income (NI) and population density (PD) are not significant at the 10 percent significance level. As expected, enterprises' production value of goods (EP) that represent the economic development level has a negative effect on eco-efficiency while both environmental sector (ES) and environmental awareness (EA) have a positive effect on eco-efficiency.

Table 7. Result of fractional logit regression

Variable	Coefficient	Standard errors	P-value
Net income (NI)	-3.01	2.65	0.26
Enterprises (EP)	-2.09	0.90	0.02
Environmental sector (ES)	1.19	0.65	0.07
Environmental awareness (EA)	2.49	0.89	0.01
Population density (PD)	0.10	0.13	0.45
Constant	4.32	2.62	0.10
BIC: -628.14 AIC: 0.46			

Note: Robust standard errors

The average marginal effects, dy/dx, of each variable, which are reported in table 8, make it easier to interpret the results. The result shows that a 1 percent increase in enterprises' production value of goods (EP) will reduce eco-efficiency by 13 percent. A 1 percent increase in environmental sector (ES) and environmental awareness (EA) will increase eco-efficiency by approximately 50 percent and 15 percent respectively.

Table 8. Marginal effect of the covariates on the outcome

Variable	dy/ex	Standard errors	P-value
Net income (NI)	-0.18	0.16	0.26
Enterprises (EP)	-0.13	0.06	0.02
Environmental sector (ES)	0.50	0.26	0.06
Environmental awareness (EA)	0.15	0.53	0.01
Population density (PD)	0.01	0.01	0.44

The link test, attached in the appendix table A3, shows whether the model estimated is robust or not. As the result of the test suggests, the model is robust since the prediction squared has no explanatory power at the 10 percent significance level.

5 Conclusions and discussion

Previous chapters have gone through the results from the two-stage DEA approach. Thus, this chapter will give a brief overview of the research question of the paper and a summary of previous results. Thereafter, a discussion will be held about the results obtained, as well as a conclusive discussion. Lastly, the chapter will give go through recommendations to decision-makers.

5.1 The aim of the study

This study aims to give a better understanding of regional eco-efficiency in Sweden and the factors influencing the regional differences in eco-efficiency. The ambition is to provide Swedish decision-makers with more information about regional performance in terms of economic and environmental factors, as there is a need for better central guidance to the regions from the government. Thus, this study is an addition to the existing literature on Swedish environmental performance that can facilitate policy makers with a new extensive tool.

This study has measured the environmental and economic performance of Swedish regions through a two-stage DEA analysis. In the first stage, eco-efficiency scores for each region were obtained and which measured the relative performance among all regions given existing resources. In the second stage, a regression analysis was performed to investigate deeper into the regional differences in eco-efficiency scores. All but four regions were included in this study. Three regions did not fulfil the data requirements and one region acted as an outlier, and hence they were excluded from the analysis.

5.2 Discussion of results

Relative efficiency of environmental and economic performance was obtained by using a standard input-oriented CCR model. The chosen inputs were labour, capital stock and energy consumption, while outputs were regional GDP and CO2 emissions. The latter was considered as an undesirable output, which implies that it should be minimized instead of maximized in order to obtain efficiency.

As section 3.8 and figure 2 indicate, there are no obvious geographical patterns for which regions are eco-efficient or not. However, both Västra Götaland and Uppsala, which are conurbations, are on average eco-efficiency leaders. This might be a result of several factors. For example, households tend to be bigger in major cities while living area per person tends to be smaller. This should result in lower energy consumption as the consumption is divided between more people. Major cities also tend to have several means of transportation, which would imply proportionally lower carbon emissions. Another factor could be that unemployment tends to be lower in cities compared to rural areas. Even though it is not clear whether geographical location matters for the improvement of eco-efficiency, it could have had an impact. As can be seen from figure 2 no northern regions are considered eco-efficiency leaders. A possible explanation for this could be that few other means of transportations than cars are used in the rural areas of northern Sweden, which would result in higher rate of CO2 emissions.

For some Swedish regions, eco-efficiency and input slacks change over time. The changes, which are difficult to derive, could be for several reasons. For example new investments or new policy rules could have been implemented into reducing the inputs where slacks were present. Another reason could be of a temporary nature, such as shocks in energy prices causing the energy consumption to go down. Even though Swedish regions are considered to be relatively homogenous, these changes could be a result of heterogeneity among the regions. This is an important implication since regions tend to have different environmental prerequisites, which would make them respond differently to national policy legislation. Therefore, one notion of this is that policies with the objective of improving eco-efficiency should be implemented to a greater extent on a regional level than previously.

Many regions experienced slacks in capital stock, which could be attributed to the fact that a large capital stock requires more material to replace the worn out materials. This implies that there is a greater need in general amongst regions to invest in more sustainable capital stock in terms of long-term use and more eco-friendly materials. This might be a tough challenge since for example old buildings usually require big investments in terms of materials usage to make them in today's perspective eco-friendly. Old machines might as well be harmful to the environment in terms of pollution. However, replacing these could mean lower eco-efficiency because of the fact that repairing the old machine would require fewer materials than the amount of materials required to build a new machine. Capital stock is thus probably the input that is the most difficult to minimize out of the inputs analysed in this study.

The exogenous factors that seem to have an impact on eco-efficiency, according to the regression analysis in chapter 4, are enterprises' production value of goods (EP), environmental awareness (EA) and environmental sector (ES). All the three variables mentioned have the expected effect on eco-efficiency. The environmental sector contains several means of environmental protection, which once again could be attributed to conurbations. A well-developed waste management system is required for major cities as well as water facilities. Major cities also usually attract research companies, which also would improve eco-efficiency through the environmental sector (ES) variable. The positive effect of environmental awareness (EA) on eco-efficiency suggests that education is one way of reaching efficiency for a region. However, environmental awareness might have several effects, which may not be solely attributed to environmental awareness through education. The variable might show other effects such as corporate structure of companies within the regions. What effect this may have on the analysis is difficult to estimate though.

According to theory, there should be a relationship between eco-efficiency and net income (NI) and therefore the results might seem surprising. The reasons for this might be many, however, one explanation for net income (NI) being insignificant, might be that people with high income consume more goods and services compared to those with low income. The Kuznets curve, on the other hand, suggests that at a certain living standard environmental quality becomes more important than employment and income. Therefore, these two effects might cancel each other out. The results also indicate that population density and eco-efficiency have no relationship which might, similarly to the previous case, seem surprising. However, as discussed in section 4.2.3, population density may have several effects, which could balance each other out and causing the variable to become insignificant. This seems logical as it would mean that population density has several effects with equal weights.

Even though the DEA approach has been widely used in terms of environmental performance, one should interpret the results with caution. There exist few studies that evaluate the DEA methods and the actual performance of the model, which gives room for possible misinterpretations. Furthermore, lack of data at the regional level in Sweden is another constraint, which might affect the results of this study, as inclusion or exclusion of other inputs may give different relative scores. Inclusion of inputs such as land use, waste management and water consumption can give more comprehensive results, which would be of greater use for policy-decision makers. Another issue that might have affected the results is the problem of misspecification. The DEA model used in this study assumes constant returns to scale, which may not reflect the reality well enough as few markets are perfect. A model that assumes varying returns to scale might therefore be better suited for this purpose. However, as previously discussed, a constant returns to scale should be more appropriate in this study, as it accounts for the overall technical efficiency which makes it easier to interpret and is less sensitive to data transformations.

5.3 Conclusions

The eco-efficiency scores obtained in the analysis indicate that there is capacity of improving efficiency among the regions of on average 9 percentage points. As previously noted, this does not indicate that efficient regions cannot improve. Four regions were considered eco-efficiency leaders, eight were eco-efficiency followers and five were eco-efficiency moderates. Investments have to be made in reducing slacks in mainly energy consumption and capital stock for inefficient regions. There is as well a need for some regions to reduce labour and CO_2 emissions in some years and thus, it is important to note that differences across regions are present.

The variation in eco-efficiency across regions can be further explained by the influencing factors in the second stage of the DEA. Improvements can thus be made in terms of environmental awareness (EA), environmental sector (ES) and enterprises' production value of goods (EP). Both net income (NI) and population density (PD) were unexpectedly insignificant. One interpretation of this could be that both variables could have several effects on eco-efficiency and which were balancing each other out, causing them to become insignificant.

The eco-efficiency of the Swedish regions changes in some cases over time, where the causes of these are hard to derive. Regional changes could be a product of new investments, new policy rules with different effects in different regions or temporary shifts in climate causing natural reduction in some inputs. As there is no distinct pattern in changes across regions, a conclusion could be made that Swedish regions are not fully homogenous. Therefore, national environmental policy implementations could have different effects on the Swedish regions.

5.4 Recommendations

The results from the first stage of the DEA indicate that policy makers in the regions should focus especially on reducing the energy consumption and the capital stock given its output. As can be seen in the results and figure 2, there is a big variation in eco-efficiency across the country. Thus there should be a further focus on decreasing the variation and make regions more equal in terms of eco-efficiency. We believe that this requires further efforts in involving actors across different levels, including small and enterprises, non-governmental organisations and municipalities, in the process of minimizing the environmental pressure. Therefore, instead of having a national action plan, we suggest better guidance on levels other than the national level. We argue that there is an information and knowledge gap on these levels, which have a capacity to be minimized. This could partly explain the differences in eco-efficiency across the regions, especially by the awareness variable. This study favours two ways of involving more actors in taking environmental initiatives and minimizing the gaps. One is to introduce incentives to implement well-developed environmental management practices across all actors. The other way in favour is to establish a support tool on compliance and environmental friendly activities developed especially for sectors that induce a great risk to the environment. A general guidance on compliance with legislation and regulations related to environmental performance is also believed to be useful to decrease the variation in eco-efficiency across regions. Further support on eco-innovations and promotion of the environmental sector is also suggested as the regression analysis proves a positive relationship between the environmental sector and eco-efficiency.

References

Aminrad, Z., Sayed Zakaria, S.Z.B. & Hadi, A.S. (2011). Influence of age and level of education on environmental awareness and attitude: Case study on Iranian students in Malaysian Universities, *Social Sciences*, 6, 15-19.

Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science*, 30, 1078-1092.

Calabrese, R. (2012). Regression Model for Proportions with Probability Masses at Zero and One, working paper 201209, Geary Institute, University College Dublin.

Charnes, A.W., Cooper, W.W. & L. Rhodes, E. (1978). Measuring The Efficiency of Decision Making Units, *European Journal of Operational Research*, 2, 429-444.

Dyckhoff, H. & Allen, K. (2001). Measuring ecological efficiency with data envelopment analysis (DEA), *European Journal of Operational Research*, 132, 312-325.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series*, 120, 253–281.

Fried, H., Lovell, A.K & Schmidt, S. (1993). The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York.

Färe R., Grosskopf S. & Lovell CAK. (1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach, *The Review of Economics and Statistics*, 71, 90–98.

Golany, B. & Roll, Y. (1989). An application procedure for DEA, *Omega*, 17, 237-250.

Goyannes Gusmão Caiado R., de Freitas Dias, R., Veiga Mattos L., Gonçalves Quelhas, OL. & Leal Filho W. (2017). Towards sustainable development through the perspective of eco-efficiency - A systematic literature review, *Journal of Cleaner Production*, 165, 890-904.

Hoff, A. (2007). Second stage DEA: comparison of approaches for modelling the DEA score, *European Journal of Operational Research*, 181, 425—435.

Huang, J., Yang, X., Cheng, G. & Wang, S. (2014). A comprehensive eco-efficiency model and dynamics of regional eco-efficiency in China, *Journal of Cleaner Production*, 67. 228–238.

Kirjavainen, T. & Loikkanent, HA. (1998). Efficiency differences of Finnish senior secondary schools: An application of DEA and Tobit analysis, *Economics of Education Review*, 17, 377-394.

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities, in Koopmans, T. C. (Ed.): *Activity Analysis of Production and Allocation, Proceeding of a Conference*, 33-97, John Wiley and Sons Inc., London.

Leontief, W. (1970). Environmental Repercussions and the Economic Structure - An Input-Output 359 Approach, *The Review of Economics and Statistics*, 52, 262-271.

Li, L.B. & Hu, J.L., (2012). Ecological total-factor energy efficiency of regions in China, *Energy Policy*, 46, 216-224.

Liu, X., Chu, J., Yin, P. & Sun, J. (2017). DEA cross-efficiency evaluation considering undesirable output and ranking priority: a case study of eco-efficiency analysis of coal-fired power plants, *Journal of Cleaner Production*, 142, 877-885.

Luptáčik M. (2001). An Eco-Efficiency Analysis of an Economy, *IFAC Proceedings Volumes*, 34, 355-359.

Martic, MM., Novakovic, MS. & Baggia, A. (2009). Data envelopment analysis - Basic Models and their Utilization, *Organizacija*, 42. 37-43.

Masternak-Janus, A. & Rybaczewska Błażejowska, M. (2017). Comprehensive Regional Eco-Efficiency Analysis Based on Data Envelopment Analysis: The Case of Polish Regions, *Journal of Industrial Ecology*, 21, 180-190.

McDonald, J. (2009). Using least squares and Tobit in second stage DEA efficiency analyses, *European Journal of Operational Research*, 197, 792-798.

Merkert, R. & Hensher, DA. (2011). The impact of strategic management and fleet planning on airline efficiency – A random effects Tobit model based on DEA efficiency scores, *Transportation Research Part A: Policy and Practice*, 45, 686-695.

Mickwitz, P., M. Melanen, U. Rosenström, & J. Seppälä. (2006). Regional eco-efficiency indicators — A participatory approach, *Journal of Cleaner Production*, 14, 1603–1611.

OECD (2014). Environmental performance review of Sweden. Available online: http://www.oecd.org/environment/country-reviews/33843590.pdf [Accessed 11 May 2018]

Oggioni, G., Riccardi, R., & Toninelli, R. (2011). Eco-efficiency of the world cement industry: A data envelopment analysis, *Energy Policy*, 39, 2842–2854.

Papke, LE., & Wooldridge, JM. (1996). Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates, *Journal of Applied Econometrics*, 11, 619–632.

Raheli, H., Rezaei R.M., Mehri Raei Jadidi M.R., Mobtaker H.G. (2017). A two-stage DEA model to evaluate sustainability and energy efficiency of tomato production, *Information Processing in Agriculture*, 4, 342-350.

Ramalho, E. A., J. J. S. Ramalho, and P. D. Henriques (2010), Fractional regression models for second stage DEA efficiency analyses, *Journal of Productivity Analysis*, 34, 239–255.

RobecoSAM (2017). Country Sustainability Ranking Update – November 2017. Available online: http://www.robecosam.com/images/Country_Ranking_Update_October_2017.pdf [Accessed 11 May 2018]

Rosenström, U. & Palosaari, M., (Eds.) 2000. Signs of Sustainability. Finland's Indicators for Sustainable Development 2000. Finnish Environment, Environmental Policy, Helsinki.

Schaltegger, S. & Sturm, A. (1990). Öologische Rationalität (German/in English: Environmental rationality). *Die Unternehmung*, 4, 117–131.

Shafik, N. (1994). Economic development and environmental quality: an econometric analysis, *Oxford Economic Papers*, 46, 757-773

Shobeiri, S. M., Omidvar, B. & Prahallada, N.N.. (2007). A comperative study of environmental awareness among secondary school students in Iran and India, *International Journal of Environmental Research*, 1, 28-34.

Simar, L. & Wilson, PW. (2007). Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136, 31-64.

Song, ML., Fisher, R. & Wang, JL. (2016). Environmental performance evaluation with big data: theories and methods, *Annals of Operation Research*, doi: 10.1007/s10479-016-2158-8.

Song, M., Song, Y., An, Q. & Yu, H. (2013). Review of environmental efficiency and its influencing factors in China: 1998–2009, *Renewable and Sustainable Energy Reviews*, 20, 8-14,

Statistics Sweden. (2018). Statistical database. Stockholm: Statistiska centralbyrån. Available online: http://www.statistikdatabasen.scb.se/ [Accessed 16 May]

Statistics Sweden. (2015). Miljösektorns omsättning oförändrad men exporten ökar. Available online: https://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Miljo/Miljoekonomi-och-hallbar-utveckling/Miljorakenskaper/38164/38171/Behallare-for-Press/411541/ [Accessed 14 May]

The Swedish Agency for Economic and Regional Growth. (2018). Statistical database. Stockholm: Tillväxtverket. Available online: https://tillvaxtverket.se/statistik.html [Accessed 16 May]

The Swedish Association of Local Authorities and Regions. (2018). Statistical database. Stockholm: Sveriges Kommuner och Landsting. Available online: https://skl.se/ekonomijuridikstatistik.63.html [Accessed 16 May]

Tone, K., (2001). A slacks-based measure of efficiency in data envelopment analysis, *European Journal of Operational Research*, 130, 498-509.

Wang, EC. & Huang, W. (2007) Relative efficiency of R&D activities: A cross-country study accounting for environmental factors in the DEA approach, *Research Policy*, 36, 260-273.

Wang, K., Shiwei Y. & Zhang, W. (2013). China's regional energy and environmental efficiency: A DEA window analysis based dynamic evaluation, *Mathematical and Computer Modelling*, 58, 1117-1127.

Wang, B., Wu, Y. & Yan, P., (2010). Environmental efficiency and environmental total factor productivity growth in China's Regional Economies, *Economic Research Journal*, 45, 95-109.

Zhou, G., Chung, W. & Zhang, Y. (2014). Measuring energy efficiency performance of China's transport sector: A data envelopment analysis approach, *Expert Systems with Applications*, 41, 709–722.

Zhang, B., Bi J., Fan Z., Yuan Z., & Ge, J. (2008). Eco-efficiency analysis of industrial system in China: A data envelopment analysis approach, *Ecological Economics*, 68, 306–316.

Yang, H. & Pollitt, M. (2009). Incorporating both undesirable outputs and uncontrollable variables into DEA: The performance of Chinese coal-fired power plants, *European Journal of Operational Research*, 197, 1095-1105.

Yang, L. & Zhang, X. (2018). Assessing regional eco-efficiency from the perspective of resource, environmental and economic performance in China: A bootstrapping approach in global data envelopment analysis, *Journal of Cleaner Production*, 173, 100-111.

Appendix

Table A1. Results of the sensitivity analysis, method 1

Region	Original	Excl. U	Excl. U & K	Excl. U, K & G	Excl. U, K, G & VG
Uppsala	1.000	-	-	-	-
Sörmland	0.899	0.886	0.886	0.886	0.912
Kronoberg	0.815	1.000	-	-	-
Kalmar	0.783	0.886	0.902	0.902	0.905
Gotland	0.853	1.000	1.000	-	-
Blekinge	0.804	0.918	0.988	0.988	0.988
Halland	0.923	0.893	0.978	0.978	0.978
Västra Götaland	0.848	1.000	1.000	1.000	-
Värmland	0.768	0.982	0.999	0.999	0.999
Örebro	0.715	0.946	0.973	0.973	0.973
Västmanland	0.739	0.936	0.949	0.949	0.952
Dalarna	0.705	0.981	0.991	0.991	1.000
Gävleborg	0.716	0.902	0.935	0.935	0.936
Västernorrland	0.707	0.954	0.956	0.956	0.966
Jämtland	0.792	0.911	0.993	0.993	0.993
Västerbotten	0.694	0.886	0.929	0.929	0.932
Norrbotten	0.717	0.928	0.928	0.928	0.958

U=Uppsala, K=Kronoberg, G=Gotland, VG=Västra Götaland

Table A2. Results of the sensitivity analysis, method 2

Region	Original	Excl. U	Excl. K	Excl. G	Excl. VG
Uppsala	1.000	-	0.999	0.997	0.997
Sörmland	0.899	0.886	0.882	0.882	0.903
Kronoberg	0.815	1.000	-	1.000	1.000
Kalmar	0.783	0.886	0.860	0.858	0.867
Gotland	0.853	1.000	1.000	-	1.000
Blekinge	0.804	0.918	0.976	0.906	0.907
Halland	0.923	0.893	0.912	0.863	0.864
Västra Götaland	0.848	1.000	1.000	1.000	-
Värmland	0.768	0.982	0.987	0.978	0.978
Örebro	0.715	0.946	0.967	0.939	0.948
Västmanland	0.739	0.936	0.942	0.929	0.937
Dalarna	0.705	0.981	0.991	0.981	1.000
Gävleborg	0.716	0.902	0.926	0.894	0.902
Västernorrland	0.707	0.954	0.949	0.947	0.958
Jämtland	0.792	0.911	0.972	0.898	0.899
Västerbotten	0.694	0.886	0.922	0.879	0.887
Norrbotten	0.717	0.928	0.924	0.924	0.944

U=Uppsala, K=Kronoberg, G=Gotland, VG=Västra Götaland

Table A3. Performance of a link test in Stata

Iteration 0: log psuedolikelihood = 183.74064

Generalized linear models		No. of obs. =	136
Optimization: ML		Residual df =	133
		Scale parameter =	0.0040153
Deviance	= 0.5340368534	(1/df) Deviance =	0.0040153
Pearson	= 0.5340368534	(1/df) Pearson =	0.0040153
Variance function: V(u) = 1	[Gaussian]		
Link function: g(u) = u	[Identity]		
		AIC	-2.657951
Log pseudolikelihood	= 183.7406409	BIC	-652.8491

Ecoefficiency	Coef.	Robust std. err.	z	P>|z|	[95% Conf. interval]	
_hat	-0.2034706	0.1621421	-1.25	0.210	-0.5212632	0.1143221
_hatsq	0.0454039	0.0283493	1.60	0.109	-0.0101596	0.1009675
_cons	1.146143	0.2282647	5.02	0.000	0.6987525	1.593534

Table A4. Eco-efficiency of the 17 regions in period 2008 to 2015

	2008	2009	2010	2011	2012	2013	2014	2015	Average
Östra Mellansverige									
Uppsala	1.000	1.000	0.975	1.000	1.000	1.000	1.000	0.999	**1.00**
Södermanland	0.950	0.829	0.930	0.925	0.977	0.879	0.816	0.753	**0.88**
Örebro	0.912	0.905	0.904	0.993	1.000	0.976	0.971	0.854	**0.94**
Västmanland	0.908	0.929	0.927	0.953	0.914	0.934	0.916	0.950	**0.93**
Average	**0.943**	**0.916**	**0.934**	**0.968**	**0.973**	**0.947**	**0.926**	**0.889**	**0.94**
Småland med öarna									
Kronoberg	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.00**
Kalmar	0.908	0.815	0.857	0.908	0.863	0.844	0.837	0.832	**0.86**
Gotland	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.00**
Average	**0.969**	**0.938**	**0.952**	**0.969**	**0.954**	**0.948**	**0.946**	**0.944**	**0.95**
Sydsverige									
Blekinge	1.000	0.903	0.891	0.911	0.837	0.868	0.915	0.921	**0.91**
Average	**1.000**	**0.903**	**0.891**	**0.911**	**0.837**	**0.868**	**0.915**	**0.921**	**0.91**
Västsverige									
Halland	0.918	0.859	0.940	0.875	0.826	0.863	0.840	0.786	**0.86**
Västra Götaland	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.00**
Average	**0.959**	**0.929**	**0.970**	**0.937**	**0.913**	**0.932**	**0.920**	**0.893**	**0.93**
Norra Mellansverige									
Värmland	1.000	0.905	0.962	0.999	0.989	0.965	1.000	1.000	**0.98**
Dalarna	1.000	1.000	1.000	1.000	1.000	1.000	0.974	0.875	**0.98**
Gävleborg	0.907	0.927	0.902	0.842	0.880	0.882	0.938	0.870	**0.89**
Average	**0.969**	**0.944**	**0.955**	**0.947**	**0.956**	**0.949**	**0.971**	**0.915**	**0.95**
Mellersta Norrland									
Västernorrland	0.876	0.956	0.962	0.985	0.954	0.921	0.961	0.964	**0.95**
Jämtland	0.860	0.825	1.000	0.837	0.795	0.869	1.000	1.000	**0.90**
Average	**0.868**	**0.890**	**0.981**	**0.911**	**0.875**	**0.895**	**0.981**	**0.982**	**0.92**
Övre Norrland									
Västerbotten	0.902	0.879	0.902	0.896	0.898	0.862	0.857	0.834	**0.88**
Norrbotten	1.000	0.780	1.000	1.000	1.000	0.945	0.886	0.783	**0.92**
Average	**0.951**	**0.829**	**0.951**	**0.948**	**0.949**	**0.904**	**0.872**	**0.809**	**0.90**