TECHNISCHE
UNIVERSITÄT
DARMSTADT

Nonparametric Efficiency Analysis in the Presence of Undesirable Outputs

Dissertation

Approved as fulfilling the requirements for the degree of

doctor rerum politicarum (Dr. rer. pol.)

by the Department of Law and Economics of the
Darmstadt University of Technology

by

Diplom-Volkswirt Benjamin Hampf
(born in Bad Soden am Taunus)

First supervisor:	**Prof. Dr. Jens Krüger**
Second supervisor:	**Prof. Dr. Michael Neugart**
Date of defense:	**28.01.2013**

Place of publication:	**Darmstadt**
Year of publication:	**2013**

D 17

Acknowledgements

During the process of conducting my doctoral research and writing this dissertation many people have encouraged and supported me in various ways. First and foremost I am indebted to my supervisor, Prof. Dr. Jens Krüger, who has not only guided me patiently through the pitfalls of economic research but also created a working atmosphere that was both inspiring and focusing. I am also grateful to Prof. Dr. Michael Neugart, my second supervisor, for helping me to finish my doctoral studies successfully. Moreover, I thank Prof. Dr. Volker Caspari for chairing the examination committee as well as Prof. Dr. Ruth Stock-Homburg and Prof. Dr. Dirk Schiereck for providing the view of business administration at my doctoral defense.

At Darmstadt University of Technology many people have provided scientific and moral support. I am especially grateful to Vanessa Belew for proof-reading the various rough drafts of this dissertation. Furthermore, I would also like to thank my former and current colleagues at the chair for empirical economics: Dr. Andreas Dietrich, Dr. Nicolas Gatzke, Stephan Hitzschke, Julian Hoss, Clemens Otte and Maike Schmitt. For supporting me during my tasks as a teaching assistant I am also grateful to Dr. Sabine Eschenhof-Kammer.

During my dissertation research I benefited much from discussions with other researchers at workshops and conferences. I appreciate the valuable comments I received from the participants of the III and IV Halle Workshop on Efficiency and Productivity Analysis (HAWEPA), the XII European Workshop on Efficiency and Productivity Analysis (EWEPA), the DEA Symposium 2012 and the workshops of the Center for Statistics and the Research Group on Innovation and Growth of the TU Darmstadt. Many thanks also to Dr. Heike Wetzel for comments on my work and the opportunity to present parts of this dissertation at the EWI in Cologne.

Of course, this dissertation could not have been written without the continuous support of my friends and my family. I am especially grateful to my parents for supporting me in my studies of economics and for accompanying me during this journey.

Contents

List of tables

List of figures

List of abbreviations

AE	Abatement efficiency
CDF	Cumulative distribution function
CNLS	Convex Nonparametric Least Squares
CO_2	Carbon dioxide
COLS	Corrected Ordinary Least Squares
CRS	Constant returns to scale
DDF	Directional distance function
DEA	Data Envelopment Analysis
DGP	Data generating process
DMU	Decision making unit
EAE	Environmental allocative efficiency
EDF	Empirical distribution function
EEM	Environmental efficiency measure
EIA	U.S. Energy Information Administration
EPA	U.S. Environmental Protection Agency
EU-BSA	European Burden Sharing Agreement
FDH	Free Disposal Hull
FERC	Federal Energy Regulatory Commission
FGD	Flue gas desulfurization unit
FSD	First-order stochastic dominance
GHG	Greenhouse gas emissions
GWh	Gigawatt hours
HW	Handy-Whitman index
MBC	Materials balance condition
ML	Malmquist-Luenberger index
NDRS	Non-decreasing returns to scale
NGE	Natural gas engine
NIRS	Non-increasing returns to scale
PE	Production efficiency
SBM	Slacks-based measure
SE	Stage efficiency
SFA	Stochastic Frontier Analysis
SO_2	Sulfur dioxide
SSD	Second-order stochastic dominance
StoNED	Stochastic Non-smooth Envelopment of Data
SUV	Sports Utility Vehicle
VRS	Variable returns to scale

1 Introduction

1.1 Environmental pollution and climate change

"There is still time to avoid the worst impacts of climate change, if we take strong action now." (Stern (2007, p. vi)). With this statement, Nicholas Stern summarizes the main results of his review on the economic consequences of changes in the global climate due to human activities.[1] He essentially captures the consensus of a majority of climate researchers (see e.g. Oreskes (2004)) by emphasizing that climate change is an inevitable phenomenon which at this stage can not be completely prevented anymore. Most researchers also agree (see Brekke and Johansson-Stenman (2008)) that it is caused by mankind through the increase of greenhouse gas emissions (GHG) to the atmosphere. Crowley (2000) uses historic data to decompose the effects of different sources (e.g. solar irradiance, volcanic activity) on changes in global temperature. He shows that the increase in temperature due to greenhouse gas emissions since the beginning of the industrialization can not be explained within the natural variability observed in the preceding 1000 years. Therefore, the anthropogenic contribution to greenhouse gas emissions significantly increases global temperature. To visualize the trend of global pollution in the second half of the 20$^{\text{th}}$ century and the beginning of the 21$^{\text{st}}$ century, figure 1.1 shows the evolution of the global emissions of carbon dioxide (CO_2) between 1950 and 2008. While CO_2 is not the main driver of the natural greenhouse effect it is the most significant contributor to anthropogenic greenhouse gas emissions (see Davison (2007)).[2] The data on carbon dioxide have been obtained from Boden et al. (2010) and account for the carbon dioxide production from fossil-fuel burning, cement manufacture and gas flaring.

Emissions increased from 5.98 billion tons in 1950 to 32.08 billion tons in 2008 with 40% of the emissions in 2008 being produced by the two largest polluting countries, China (7.04 billion tons) and the United States (5.68 billion tons). It is also visible that the growth rate of emissions has seriously increased since the year 2000. This is very notable given the rising awareness of the problems associated with climate change and the increased political actions like the Kyoto Protocol to limit carbon dioxide emissions. Canadell et al. (2007) have identified three sources for the development of emissions since the year 2000. It is driven by an increase in global economic growth combined with an increase of carbon dioxide intensity (measured in kilogramm CO_2 per Dollar of gross world product). While this intensity declined on average by 1.3% per year between 1970 and 2000 it has since then increased by 0.3% per year. Besides this increased production of CO_2, Canadell et al. (2007) also find that the efficiency of the natural sinks that absorbe carbon dioxide has declined. An example for these natural sinks are the oceans which according to Raven et al. (2005) show increasing acidification because of the increase in anthropogenic CO_2 emissions. Due to this acidification the ability to absorbe emissions detoriates. Hence, the increased production in emissions leads to a decreased absorbence of emissions and thus the visible increase in growth of CO_2.

[1] See Pielke (2005) for a critical discussion of whether the term "climate change" should include natural effects or focus solely on the effect of man on climate.

[2] Karl and Trenberth (2003) provide an overview of the relationship between greenhouse gases, natural and anthropogenic greenhouse effects and climate change.

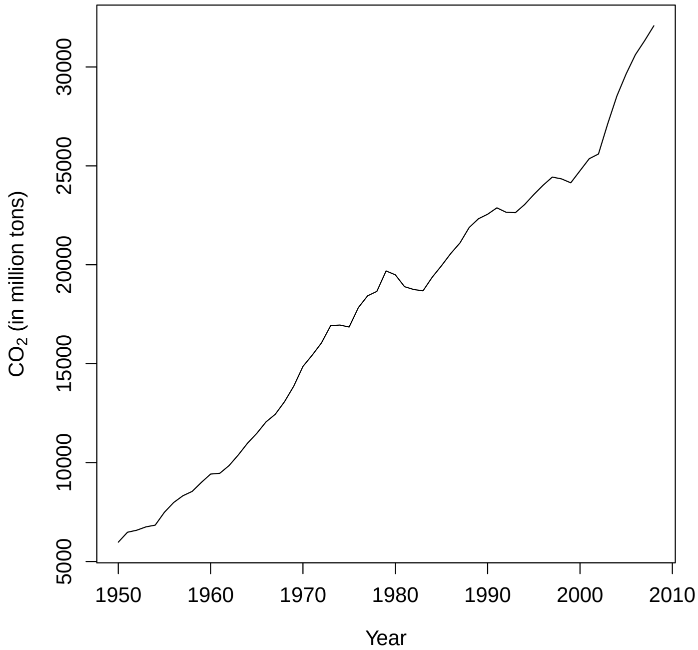

Figure 1.1: Historical trends in CO_2 emissions

Regarding the consequences of this increase of anthropogenic emissions Stern (2007, p. vi) estimates that the global temperature will rise by more than 2°C by 2035 given that the emission of greenhouse gases will increase as before (the "business-as-usual" scenario). This increase in temperature will result in severe global problems including rising sea levels associated with coastal flooding (see Nicholls and Cazenave (2010)), longer periods of drought (see Dai (2011)) which are affecting agriculture (see Cline (2007)) and risks to human health by the increased spreading of infectious diseases (see McMichael et al. (2006)). According to Stern (2007, p. 143) the costs of these effects resulting in the "business-as-usual" scenario will be between 5% and 10% of the global GDP by the end of this century and hence he recommends significant political actions to limit the increase of emissions. In contrast to the discussion of the existence and the causes of climate change there is less agreement between researchers on the impacts of it. Many authors like Mendelsohn (2006), Nordhaus (2007) and Tol (2006) criticize that Stern overestimates the costs associated with climate change and underestimates the costs of abatement activities (see Dietz and Stern (2008) for a reply to the critics). The estimation of the costs is difficult due to the uncertainties about climate change (see Tol (2003)) and the sensitivity of the results to the assumed discount rate (see Quiggin (2008)) as well as the used damage function (see Botzen and van den Bergh (2012)). Therefore, it is still an ongoing discussion how large the costs of climate change are and which political actions are appropriate to encounter it.

In contrast, less controversial results in the literature are the distributional aspects of climate change. Most of the studies on the costs of climate change summarized in Tol (2009) estimate the largest costs (in terms of percentage decrease of GDP) will result for the African countries. Therefore, the distributional problem arises that countries which have contributed the least to climate change will face the largest impacts (see Tol et al. (2004) for a detailed analysis of the distributional aspects of climate change). This consequence of climate change shows the severe problem of external effects through environmental pollution. External effects (or externalities) result if economic activities do not only affect the subjects that are directly engaged in it but also affect third parties without the market being a mediating mechanism (see Endres (2011, p. 326)). The effects of damages to the environment are an example of negative external effects while positive externalities are, for example, given by spillovers from research and development between companies. Given the existence of external effects through pollution a decrease in the production of emissions does not only lead to benefits for the subject that reduces emissions but also for third parties.[3]

The costs associated with climate change as well as the distributional issues underline the importance of accounting for environmental factors in economic analyses. With regard to production economics two strings of environmental research can be distinguished. One part of the economic literature focuses on innovations regarding new technologies and products that are less emission-intensive (see Jaffe et al. (2003) and Popp et al. (2010) for surveys on environmental aspects of technical change). A second part is concerned with the efficient use of existing technologies to reduce the production of emissions (see Tyteca (1996) for an overview). This dissertation links to both parts of the literature. The focus is on the measurement of efficiency of decision making units (DMUs) in the presence of environmental factors.[4] Technical change is addressed when analyzing dynamic changes in the productivity of DMUs.

In the following we present a short overview of different methods to estimate the efficiency of DMUs. The section highlights the advantages and disadvantages of these approaches and discusses their applicability to account for environmental factors.

1.2 The measurement of efficiency

Following Fried et al. (2008, pp. 7-8) the productivity of a DMU can be defined as the ratio of its output to its input. In the case of multiple outputs and/or multiple inputs aggregation weights are needed to render both, the numerator and the denominator, a scalar. Efficiency is defined in a technical sense as a comparison of observed outputs to optimal outputs for given inputs or as a comparison of observed inputs to optimal inputs for given outputs. Economic efficiency extends this concept to quantities like costs or revenues. In the above given definition of efficiency we purposely did not replace “optimal outputs” by “maximal outputs” and “optimal inputs” by “minimal inputs” because in the presence of environmental factors optimality is not that simply defined as will be shown in the following of this introduction and the subsequent chapters.

To estimate the efficiency of DMUs various methods have been proposed in production economics (see Murillo-Zamorano (2004) for an overview). One possibility to divide them into groups of

[3] Of course, the unequal distribution of costs and benefits of the reduction of emissions creates additional problems. For example, with regard to the implementation of a stable international environmental agreement (see Wagner (2002)).

[4] In the literature on efficiency and productivity analysis “decision making unit” is a general term for an economic subject, e.g. a firm or a country.

similar approaches is to classify them as parametric, semi-parametric and nonparametric models. The parametric and the nonparametric models can be further divided into deterministic and stochastic approaches. Figure 1.2 provides an overview of this structure and presents examples for each class.

Parametric models assume a specific functional form of the production function to evaluate the efficiency of DMUs. Aigner and Chu (1968) propose a deterministic approach which estimates a production function using linear and quadratic programming techniques subject to the constraint that all data points have to lie on or below the estimated frontier. Given the estimated function all deviation from it is attributed to inefficiency. Hence, this approach does not account for any random effects which in usual regression analysis are captured in the error term. A similar approach is the "Corrected Ordinary Least Squares" (COLS) method first discussed by Winsten (1957). This model consists of two steps. In the first step the parameters of the production function are estimated using all data points. In a second step the estimated function is shifted so that all data points lie on or below the frontier similar to the approach by Aigner and Chu (1968). Again, this model does not account for any random error.

This drawback of the deterministic parametric approach is encountered by the "Stochastic Frontier Analysis" (SFA) proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) which is also a parametric approach but allows for disturbance in the estimation of the frontier function via a two-part error term. This error term consists of a noise term that accounts for deviation from the frontier due to measurement error and an efficiency term which captures deviations due to inefficiency of the DMUs. However, the performance of this method to analyze efficiency depends like all parametric approaches on whether the appropriate functional form is defined by the researcher. Moreover, it depends on the correct assumption of the distribution of the inefficiency term which is commonly assumed to be half-normal or exponentially distributed (see Kumbhakar and Lovell (2000)). The analysis of environmental aspects of efficiency is very limited in the literature of SFA due to the need for a specific functional form. The existing literature provides only basic incorporations of environmental aspects. They are limited to the inclusion of emissions as inputs (see e.g. Reinhard et al. (2000)) which results in a production theoretically questionable model as will be discussed in the subsequent chapters.

The semi-parametric approach "Stochastic Non-smooth Envelopment of Data" (StoNED) by Kuosmanen and Kortelainen (2012) addresses the problem of dependency on the correct functional form. In this approach, efficiency is estimated using a two-step procedure. In the first step the frontier of the technology is estimated using a nonparametric estimation technique ("Convex Nonparametric Least Squares" (CNLS)) that does not rely on any assumption on the functional form of the production function. In a second step the residuals of this estimation are used to estimate the efficiency of the DMUs. Similar to the SFA approach it is assumed that these residuals are a combination of random error and the inefficiency of the DMUs. Hence, this step exhibits the same problem as the SFA model because specific distributional assumptions regarding the error term have to be imposed rendering this step parametric. Since the model combines a nonparametric and a parametric step it is labeled "semi-parametric". A serious disadvantage of this method is given by its inability to include multiple outputs in the analysis. Hence, it is not suitable for environmental analyses which in general include at least one good

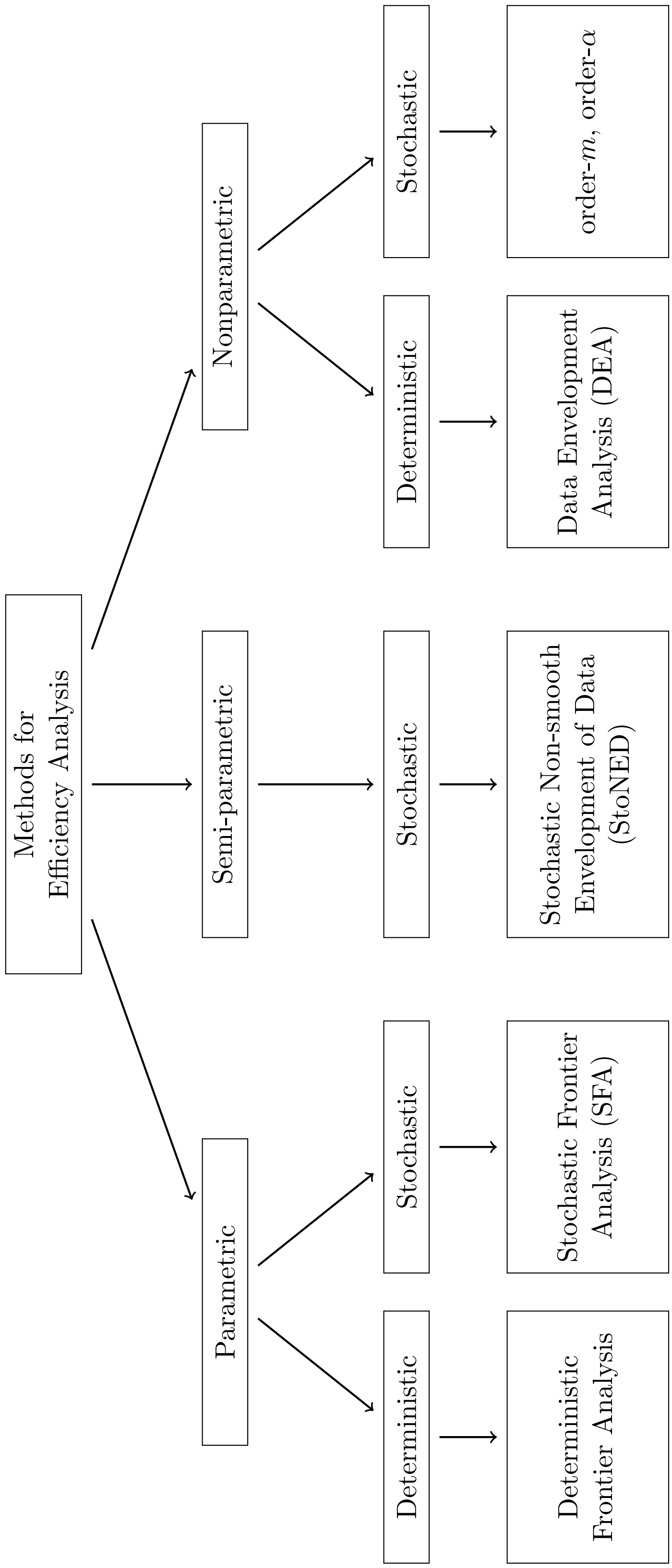

Figure 1.2: Overview of methods for efficiency analysis

and one bad output. Moreover, since this method is rather new its performance compared to existing approaches is still an open question. First results of Monte-Carlo studies (see Andor and Hesse (2012)) do not show a generally better performance.

In contrast to parametric and semi-parametric models, nonparametric approaches do not rely on the choice of a specific functional form of the production function nor do they impose distributional assumptions on the efficiency term. The deterministic nonparametric “Data Envelopment Analysis” (DEA) was proposed by Charnes et al. (1978). In this paper the authors address the problem of estimating the productivity of a DMU measured as the weighted amount of outputs divided by the weighted amount of inputs if no weights are exogenously given. An example for exogenously given weights are prices for the inputs and outputs which, for example, are not obtainable for non-market goods. Using the ideas by Charnes and Cooper (1962) to transform fractional programs into linear models Charnes et al. (1978) show how the aggregation weights can be endogenously determined by solving a linear programming problem. The solution to this problem also provides a measure of the efficiency of a DMU. This so-called “multiplier form” of the DEA has a dual which is referred to as “envelopment form”.[5] This form connects the approach of Charnes et al. (1978) to earlier published literature on production technologies and efficiency analysis.[6] The approach to model a technology by linear combinations of observations has been established as “activity analysis” in Koopmans (1951). Farrell (1957) proposed to measure the efficiency of a DMU as the distance of the DMU to the frontier of a technology. Shephard (1970) provided a set of mathematically and economically reasonable axioms on production technologies.[7] Given these axioms it is possible to estimate a technology and a frontier function by using empirical data without knowing the functional form of the frontier function. One possibility to estimate this frontier function is the “envelopment form” of DEA which constructs a piecewise linear envelopment of the data.[8] Observations in a dataset may be located on this frontier, hence can be identified as efficient, or in the interior of the technology and thus can be identified as inefficient. The Farrell (1957) measure of technical efficiency or the Shephard (1970) distance function which is the reciprocal of the Farrell measure can be applied to estimate the distance of a point in the technology to the frontier.[9] Both measures are radial which means that depending on the chosen direction of the efficiency analysis (input or output orientation) all inputs or all outputs are scaled by the same proportion. They can be obtained as the solution to the linear programming problem of the “envelopment form” of DEA. Since the seminal paper by Charnes et al. (1978) a large amount of papers presenting both empirical applications as well as theoretical extensions to DEA have been published. For example, the bibliography by Emrouznejad et al. (2008) lists more than 4000 publications and Liu et al. (2013) conduct a survey that identifies the most important (in terms of citation) papers in the field.

[5] See e.g. Luenberger and Ye (2008) for an introduction to duality in linear programming.

[6] For a more detailed historical overview of the main ideas of data envelopment analysis and especially the authors involved in it see Førsund and Sarafoglu (2002).

[7] These axioms are discussed in detail in the following chapter. Hence, they are not presented here.

[8] Since this approach is analogous to the approach of “activity analysis” the DEA technology estimation is sometimes also referred to as “activity analysis model” (see e.g. Färe et al. (2001, p. 388)).

[9] Some authors also call this measure “Debreu-Farrell measure of technical efficiency” referencing the work by Debreu (1951). However, ten Raa (2008) notes that the “Farrell measure of technical efficiency” and Debreus’s “coefficient of resource utilization” do not measure the same type of efficiency.

From a statistical point of view one of the most important extensions to the literature on DEA is given by the analysis of both the constructed technology as well as the efficiency measures as estimators of the true but unknown quantities (see Daraio and Simar (2007) for an introduction to the statistical foundation of nonparametric efficiency analysis). Kneip et al. (1998) have proven that under variable returns to scale of the technology the DEA efficiency scores converge with rate $n^{2/(m+s+1)}$ to the true but unknown quantities where n denotes the number of observations and m (s) denotes the number of inputs (outputs) used in the analysis. For the case of constant returns to scale Park et al. (2010) have proven that the rate is $n^{2/(m+s)}$. Comparing this rate with the convergence rate of most parametric estimators like OLS which is $n^{1/2}$ (see e.g. Greene (2008)) shows that DEA suffers from the "curse of dimensionality", hence the precision of the estimates declines if the number of inputs and/or outputs is increased. Moreover, DEA is a biased estimator but Simar and Wilson (1998) have developed bootstrap techniques to correct the results for this bias. Besides the "curse of dimensionality" and the bias, DEA estimates are also sensitive to outliers since the DEA technology envelopes all data points of a sample and hence also the outliers. Several methods to detect outliers in nonparametric technologies have been proposed (see Wilson (1993, 1995) and Simar (2003)).

To overcome these problems, stochastic nonparametric models have been developed. The order-m approach by Cazals et al. (2002) and the order-α approach by Aragon et al. (2005) only partially envelop the data and therefore are more robust against outliers in the sample. Moreover, the efficiency measures converge with rate $n^{1/2}$ and hence do not suffer from the "curse of dimensionality". However, despite their theoretical advantages they also show several disadvantages compared to DEA. Krüger (2012) shows using Monte-Carlo simulations that these estimators are only superior to traditional methods if outliers are a severe problem in the dataset. They perform worse than DEA and SFA if outliers are absent or only present to a small extend. Furthermore, while the literature on DEA provides various extensions that allow for detailed analyses of inefficiency, e.g. network models, the literature on order-m and order-α is limited on basic models of radial efficiency. Finally, the DEA estimations allow for an intuitive interpretation making them very suitable for practical applications. For example, DEA and SFA are applied by many countries for an incentive-based regulation of the energy sector (see Jamasb and Pollitt (2001)). In contrast, the interpretation of order-m and order-α results is more complex. Due to these drawbacks of the stochastic nonparametric approaches this dissertation relies on DEA and its extensions for environmental analyses. Of course, the problems of DEA, in particular the bias of the estimates, are addressed in detail in the subsequent chapters.

1.3 Efficiency and environmental factors

Combining the fields of environmental economics and nonparametric efficiency analysis leads to questions which provide challenges for both theoretical and applied research.

- Should environmental factors be included in efficiency analysis?
- How can environmental factors be included in efficiency analysis?
- How can efficiency be measured in the presence of environmental factors?

The first question aims at the relevance of environmental aspects for efficiency analysis. The valuation of whether a certain pollutant is relevant in terms of damaging the environment is usually left to researchers in natural science. An example of such a debate is the question whether global warming potentials are an adequate measure to aggregate greenhouse gases (see O'Neill (2000) for a summary). For economists it is of interest whether the implications of an efficiency analysis for decision making units change if emissions are included (see e.g. Färe et al. (2012)). For example, a DMU might be found to be efficient in an analysis ignoring emissions but very inefficient if emissions are accounted for.

If one decides to include emissions in the analysis the second question becomes important. Appropriate ways to include emissions are proposed in numerous publications (see Scheel (2001), Zhou et al. (2008a) and Liu et al. (2010) for surveys on this topic). This still ongoing discussion results from the critical view of production economists on the way environmental economists model production processes including environmental factors. In standard environmental economic production models (see e.g. Cropper and Oates (1992)) emissions are included as an additional input. This follows, because similar to an analysis of conventional inputs (e.g. capital) it is regarded as an improvement in the production process if emissions are reduced holding all other factors constant. However, this treatment of emissions although sometimes also used in nonparametric efficiency analysis (see e.g. Hailu and Veeman (2001)) is criticized by production economists (see Färe and Grosskopf (2003)). They focus on modeling emissions in a way that reflects the true production process while respecting physical limitations of the technology. Both points are violated when including emissions as inputs. The first one because emissions are the results of a production process. The second point is violated because treating emissions as inputs and imposing the axioms by Shephard (1970) allows for implausible substitution possibilities and physically impossible input-output combinations.[10] The intention to conduct efficiency analysis without violating basic physical laws connects this literature not only with natural science but also to the economic discussion of physical limits to growth (see e.g. the discussion between Daly (1997) and Solow (1997)).

If emissions are included in the analysis in a suitable way the third question becomes relevant. As presented above efficiency in nonparametric models is often measured by using radial distance functions. However, these distance functions were proposed at a time when environmental factors were not an aspect of economic research. Since they are very inflexible in the way efficiency is measured (all inputs or outputs are scaled equiproportional), they are not easily applied in the case of environmental efficiency analysis. Therefore, more general measures of efficiency are needed which are introduced in the next chapter.

[10] These points will be discussed in more detail in the following chapters.

1.4 Overview of the dissertation

The above stated questions form the basis of the research conducted in this dissertation which comprises four pieces of economic analysis. In these studies empirical and theoretical questions, both micro- and macroeconomically oriented, regarding the analysis of efficiency of decision making units in the presence of undesirable outputs are addressed.

The production theoretical background of these analyses is discussed in the following chapter 2. In this chapter we present formal descriptions of the production technologies as well as static measures of efficiency and show how they can be estimated using nonparametric techniques. Furthermore, dynamic approaches to analyze the productivity changes of DMUs allowing to decompose efficiency and technical changes are introduced. The focus of this chapter lies on the discussion of possibilities to extend methods applied to conventional inputs and outputs to the analysis incorporating environmental factors.

Chapter 3 presents a study on the technical efficiency of automobiles which analyzes whether the inclusion of carbon dioxide emissions has a significant effect on the efficiency of automobiles. This inclusion is of particular importance because the emissions of automobiles have recently gained large interest in the public debate regarding the limitation of carbon dioxide emissions. Our analysis is conducted for an overall sample of 3961 cars. Dividing the sample into several groups of cars (e.g. SUVs) and comparing the results for these groups we gain insights into their relative performance and evaluate whether this changes if emissions are incorporated. Using a separation approach allows us to decompose inefficiencies into group specific and individual effects. By including emissions the analysis extents the existing literature on nonparametric efficiency analysis of cars which does not account for them. Furthermore, we extend previous research by correcting for the bias in nonparametric efficiency results and show that the previously obtained results tend to underestimate inefficiencies by focusing solely on special groups of cars.

The fourth chapter presents a new approach to estimate environmental efficiency of DMUs in a network DEA model. The proposed network technology consists of two stages, a production and an abatement stage, with the emissions that result from the production of good outputs in the production stage being reduced in the abatement stage. The inclusion of a physical constraint on the technology, the so-called "materials balance condition" enables the analysis of this model even if no data on the emissions before abatement are available. Moreover, a new measure for environmental efficiency is proposed which can be used together with the two-stage model to decompose environmental efficiency into efficiency of the production and the abatement stage. Furthermore, inefficiencies due to network and stage effects can be distinguished. We also show how this new measure is related to an existing approach to environmental efficiency which does not account for network technologies or abatement activities. A comparison of our approach to similar network models shows that they exhibit several drawbacks in the way the abatement technology is modeled and the efficiency of the stages is measured. The applicability of this new model is demonstrated by an analysis of coal-fired power plants in the United States with regard to their emissions of sulfur dioxide (SO_2).

Chapter 5 presents and compares different methods to obtain optimal directions for efficiency analyses. In contrast to the above mentioned radial distance functions, the directional distance function (DDF) allows to specify a direction of the efficiency measurement separately for each input and/or output by incorporating a directional vector. However, this vector has to be chosen by the researcher and hence gives room for a large extend of subjectivity. In this study we modify an existing static approach to obtain directional vectors endogenously for an environmental efficiency analysis. Moreover, we propose an extension to an analysis in a dynamic setting. Applying these methods to an efficiency evaluation of the major greenhouse gas emitting countries, we explore reduction potentials for greenhouse gases due to existing inefficiencies and moreover discuss how these potentials vary with the method to obtain the directional vectors.

Chapter 6 analyzes the influence of the Kyoto Protocol on the macroeconomic productivity of European countries. Critics of the protocol argue that it allows for a "business-as-usual" strategy to achieve the reduction targets and hence the countries do not face significant costs associated with signing the protocol. In this chapter we address this critique by comparing the productivity patterns of European countries. The European implementation of the Kyoto Protocol, the EU Burden Sharing Agreement (EU-BSA), allows to differentiate two groups of European countries with regard to whether they are assigned reduction targets or not. A comparison of the productivity patterns for the groups before and after the signing of the agreement is used to identify effects of the Kyoto Protocol. By focusing on the EU15 countries this study addresses a shortcoming of previously conducted analyses which find statistically significant differences in productivity growth due to the Kyoto Protocol. However, since these works are based on very heterogeneous groups of countries we evaluate whether this finding pertains in the more homogeneous group of European countries.

These four chapters all address unique research questions with regard to the efficiency of DMUs in the presence of environmental factors. The analyses share a common methodological basis which uses nonparametric techniques to estimate the production technologies and the efficiency respectively productivity measures. Moreover, while the focus of all studies is concentrated on environmental pollution they address this pollution as one part of a production process that accounts for multiple inputs and outputs. Finally, the included pollutants are all associated with pollution of the atmosphere. Therefore, all studies are related to the phenomenon of climate change discussed above.[11]

[11] In chapters 3, 5 and 6 greenhouse gases and carbon dioxide emissions are analyzed. In chapter 4 the empirical example includes sulfur dioxide emissions which do not only cause acid rain but also contribute to climate change (see Ward (2009)).

2 General concepts

In this section we present the theoretical foundations and the methodology of nonparametric efficiency analysis. We start by defining the conventional and the environmental production technologies and show how they can be constructed using nonparametric techniques. Afterwards we present measures that can be used together with the nonparametric technologies to estimate the efficiency of DMUs. Furthermore, we present methods for dynamic analyses of efficiency and productivity. Our following presentation only gives a short summary of the most important concepts. For more detailed discussions there exists a large amount of literature on the microeconomic foundation of nonparametric efficiency analysis. Coelli et al. (2005) presents an introductory treatment, while Färe et al. (1985) provides a more thorough presentation of the theoretical concepts. Hackman (2008) combines theoretical discussions with examples from engineering. The following presentation gears to Hackman (2008) regarding the production theoretical methods and to Simar and Wilson (2008) for the nonparametric estimations.

2.1 Modeling the production technology

Consider a production process where m inputs $\boldsymbol{x} \in \mathbb{R}_+^m$ are used to produce s outputs $\boldsymbol{y} \in \mathbb{R}_+^s$. The technology set of this process $T \subset \mathbb{R}_+^{m+s}$ is the set of all feasible input-output combinations $(\boldsymbol{x}, \boldsymbol{y})$ and can be defined as

$$T = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}_+^{m+s} : \boldsymbol{x} \text{ can produce } \boldsymbol{y}\}. \tag{2.1}$$

This technology can be equivalently represented by the input correspondence $L : \mathbb{R}_+^s \to 2^{\mathbb{R}_+^m}$ with images

$$L(\boldsymbol{y}) = \left\{\boldsymbol{x} \in \mathbb{R}_+^m : (\boldsymbol{x}, \boldsymbol{y}) \in T\right\} \tag{2.2}$$

and the output correspondence $P : \mathbb{R}_+^m \to 2^{\mathbb{R}_+^s}$ with images

$$P(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}_+^s : (\boldsymbol{x}, \boldsymbol{y}) \in T\}. \tag{2.3}$$

The images of the input correspondence, $L(\boldsymbol{y})$, are referred to as input requirement sets. They contain all input combinations which are capable of producing the output vector $\boldsymbol{y}$. Analogously, the images of the output correspondence, $P(\boldsymbol{x})$, are called output sets. An output set comprises all output vectors $\boldsymbol{y}$ which can be produced from a given input vector $\boldsymbol{x}$.

The technology set is assumed to satisfy the following axioms (see Färe and Primont (1995) for a discussion of these axioms):

1. No free lunch: $(\boldsymbol{x}, \boldsymbol{y}) \notin T$ if $\boldsymbol{x} = \boldsymbol{0} \wedge \boldsymbol{y} \geq \boldsymbol{0}$.
   It is not possible to produce positive amounts of any output without using positive amounts of at least one input.[12]

[12] Note that here and in the following "$\geq$" means that at least one element of the vector satisfies strict inequality while "$\geqq$" means that all elements of the vector can satisfy equality.

2. Strong disposability of inputs: If $(\boldsymbol{x}, \boldsymbol{y}) \in T$ and $\boldsymbol{x}' \geqq \boldsymbol{x}$ then $(\boldsymbol{x}', \boldsymbol{y}) \in T$.
   For any given combination $(\boldsymbol{x}, \boldsymbol{y})$ the same amount of output is attainable by using more inputs.
3. Strong disposability of outputs: If $(\boldsymbol{x}, \boldsymbol{y}) \in T$ and $\boldsymbol{y}' \leqq \boldsymbol{y}$ then $(\boldsymbol{x}, \boldsymbol{y}') \in T$.
   For any given combination of $(\boldsymbol{x}, \boldsymbol{y})$ it is possible to produce less output holding $\boldsymbol{x}$ constant.
4. Convexity: T is convex.
   Convex combinations of observations are attainable, e.g. if $(\boldsymbol{x}_1, \boldsymbol{y}_1)$ and $(\boldsymbol{x}_2, \boldsymbol{y}_2) \in T$ then $(\alpha \boldsymbol{x}_1, \alpha \boldsymbol{y}_1) + ((1-\alpha)\,\boldsymbol{x}_2, (1-\alpha)\,\boldsymbol{y}_2) \in T \; \forall \alpha \in [0,1]$.
5. Closeness: T is a closed set.
   The boundary of T is also part of the technology set.

For evaluating the efficiency of a DMU the upper boundary of the technology is of special interest. It can be defined as

$$\partial T = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}_+^{m+s} : (\delta \boldsymbol{x}, \boldsymbol{y}) \notin T \; \forall \, 0 \leqq \delta < 1 \wedge (\boldsymbol{x}, \gamma \boldsymbol{y}) \notin T \; \forall \, \gamma > 1\}. \tag{2.4}$$

This upper boundary is also called the frontier of the technology.

To estimate this technology nonparametric methods namely Data Envelopment Analysis can be applied. Given a sample of n DMUs with input-output combinations

$$\mathcal{X}_n = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\,, i = 1, \ldots, n\} \tag{2.5}$$

the DEA estimation of the technology set satisfying the axioms stated above reads as

$$\widehat{T} = \left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}_+^{m+s} : \boldsymbol{x} \geqq \boldsymbol{X}\boldsymbol{\lambda}, \boldsymbol{y} \leqq \boldsymbol{Y}\boldsymbol{\lambda}, \boldsymbol{\lambda} \geqq \boldsymbol{0}\right\}. \tag{2.6}$$

In this formulation $\boldsymbol{X}$ represents the $m \times n$ matrix of inputs and $\boldsymbol{Y}$ represents the $s \times n$ matrix of outputs. $\boldsymbol{\lambda}$ denotes a $n \times 1$ vector of weight factors with $\boldsymbol{\lambda}$ positive but otherwise unrestricted implying constant returns to scale (CRS) of the production process. To model variable returns to scale (VRS) of the production technology the additional restriction $\boldsymbol{1}^T\boldsymbol{\lambda} = 1$ needs to be imposed on the estimation. Moreover, non-decreasing returns to scale (NDRS) may be imposed by assuming $\boldsymbol{1}^T\boldsymbol{\lambda} \geqq 1$ and non-increasing returns to scale (NIRS) by imposing $\boldsymbol{1}^T\boldsymbol{\lambda} \leqq 1$ (see Banker et al. (1984)).[13]

The above defined technology describes a production process which transforms conventional inputs into conventional outputs. However, in the case of an environmental efficiency analysis the technology must also comprise the (unintended) production of factors which are damaging the environment. In the following we refer to these factors as bad (or undesirable) outputs.

Consider the case in which the production process is associated with the emission of r pollutants $\boldsymbol{u} \in \mathbb{R}_+^r$, e.g. carbon dioxide. These bad outputs can not be simply incorporated as another conventional output. The assumption of strong disposability of all outputs implies that

[13] If the assumption of convexity of the technology set is removed, the DEA model collapses to the Free Disposal Hull (FDH) model by Deprins et al. (1984).

if $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T$ then $(\boldsymbol{x}, \boldsymbol{y}, \mathbf{0}) \in T$ and hence the resulting technology would allow to produce good outputs without producing any bad outputs. The possibility of complete abatement of emissions without any costs is very unrealistic raising the question why emissions are observed anyway. To exclude this possibility environmental economists often treat emissions as an additional input (see Cropper and Oates (1992)). As Färe and Grosskopf (2003) note, this leads to a physically impossible technology. Given a fixed amount of conventional inputs and good outputs the assumption of strong disposability of inputs would allow to increase emissions without limits. This, for example, would imply that a given amount of coal could be used to produce infinite amounts of carbon dioxide. Moreover, as noted by Førsund (2009) the assumption of strong disposability of inputs allows for substitution possibilities between conventional inputs and emissions which are modeled as inputs. However, the possibility to increase the use of coal while decreasing carbon dioxide emissions seems implausible.[14]

To overcome this problem different ways to incorporate emissions have been developed (see e.g. Scheel (2001) for a summary). These can be roughly divided into methods transforming the data and methods transforming the technology. The approaches that transform the data aim at incorporating emissions in a way that the standard axioms on the technology presented above not need to be modified. For example, Lovell et al. (1995) include the inverse of the emissions as a conventional output. However, a more convenient way to model emissions transforms the technology by introducing new axioms suitable to allow the technology to incorporate bad outputs.

The environmental production technology $T^{\text{Env}} \subset \mathbb{R}_+^{m+s+r}$ that accounts for the production of undesirable outputs contains all feasible combinations of $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u})$ and therefore can be defined as

$$T^{\text{Env}} = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in \mathbb{R}_+^{m+s+r} : \boldsymbol{x} \text{ can produce } (\boldsymbol{y}, \boldsymbol{u})\right\} \tag{2.7}$$

with input requirement sets

$$L(\boldsymbol{y})^{\text{Env}} = \left\{\boldsymbol{x} \in \mathbb{R}_+^{m} : (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T^{\text{Env}}\right\} \tag{2.8}$$

and output sets

$$P(\boldsymbol{x})^{\text{Env}} = \left\{(\boldsymbol{y}, \boldsymbol{u}) \in \mathbb{R}_+^{s+r} : (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T^{\text{Env}}\right\}. \tag{2.9}$$

Shephard (1970) introduces the axiom of weak disposability of outputs. This axiom can be used (see e.g. Färe and Grosskopf (1983)) to include bad outputs into the production technology.

8. Weak disposability of undesirable outputs:
   If $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T^{\text{Env}}$ and $\eta \boldsymbol{u} \leqq \boldsymbol{u}$ with $0 \leqq \eta \leqq 1$ then $(\boldsymbol{x}, \eta\boldsymbol{y}, \eta\boldsymbol{u}) \in T^{\text{Env}}$.

Given the weak disposability assumption it is only possible to produce less of the undesirable outputs if the amount of desirable outputs is decreased by the same proportion. Therefore, the reduction of emissions is costly.

[14] In the subsequent analysis including abatement processes in a network model we will show that such substitution possibilities can exist indirectly.

Together with the assumption of weak disposability of bad outputs, the good and bad outputs are assumed to be null-joint:

9. Null-jointness: If $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T^{\text{Env}}$ and $\boldsymbol{u} = \boldsymbol{0}$ then $\boldsymbol{y} = \boldsymbol{0}$.

The null-jointness assumption simply states that it is impossible to produce positive amounts of desirable outputs without producing any undesirable outputs.

Given a sample of n DMUs with data on inputs, good outputs and bad outputs

$$\mathcal{X}_n^{\text{Env}} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i), i = 1, \ldots, n\} \tag{2.10}$$

the DEA technology considering emissions as undesirable outputs can be estimated as

$$\widehat{T}^{\text{Env}} = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in \mathbb{R}_+^{m+s+r} : \boldsymbol{x} \geqq \boldsymbol{X\lambda}, \boldsymbol{y} \leqq \boldsymbol{Y\lambda}, \boldsymbol{u} = \boldsymbol{U\lambda}, \boldsymbol{\lambda} \geqq \boldsymbol{0}\right\} \tag{2.11}$$

where $\boldsymbol{U}$ is the $r \times n$ matrix of undesirable outputs with the equality constraint indicating weak disposability. Assuming constant returns to scale allows to set the scaling factor η to 1 (see Färe and Grosskopf (2003)). For a discussion of environmental DEA technologies assuming weak disposability and variable returns to scale see Färe and Grosskopf (2009).

The above defined DEA technology exhibits null-jointness of good and bad outputs if the dataset fulfills the conditions by Kemeny et al. (1956) applied to bad outputs (see Färe and Grosskopf (2004)). These conditions are given by

$$\begin{array}{lll} \boldsymbol{U}_{j.}\boldsymbol{1} & > 0 & j = 1, \ldots, r \\ \boldsymbol{1}^T\boldsymbol{U}_{.i} & > 0 & i = 1, \ldots, n \end{array} \tag{2.12}$$

with $\boldsymbol{1}$ denoting a $r \times 1$ vector of ones. The first condition states that every bad output needs to be to be produced by at least one DMU. The second condition states that every DMU has to produce at least one bad output. To see that this assures null-jointness assume $\tilde{\boldsymbol{u}} = \boldsymbol{0}$. Since by the above given conditions each row and each column of $\boldsymbol{U}$ exhibits at least one strict positive element the only feasible solution for the $\boldsymbol{\lambda}$-values given $\tilde{\boldsymbol{u}}$ is $\boldsymbol{0}$.[15] Therefore, $\boldsymbol{y} = \boldsymbol{Y\lambda} = \boldsymbol{0}$ as well. Whether these conditions are met by the dataset can be simply checked by inspecting the rows and columns of $\boldsymbol{U}$ (see Färe (2010)).

To illustrate the difference between the strong disposability and the weak disposability approach, figure 2.1 depicts 3 DMUs (A, B, C). These DMUs are producing one good (y) and one bad (u) output. Furthermore, it is assumed that all DMUs are using the same amounts of inputs. Therefore, they can be included in one output set. Assuming strong disposability of both outputs, the boundary of the output set is given by the horizontal and vertical extension through point B and the output set itself is in the south-west direction of this boundary. Under weak disposability of the undesirable output the boundary is given by $\overline{0AB}$ and the vertical extension to B.

[15] This holds for the case of constant returns to scale. Given variable returns to scale the scaling parameter η has to be set equal to zero.

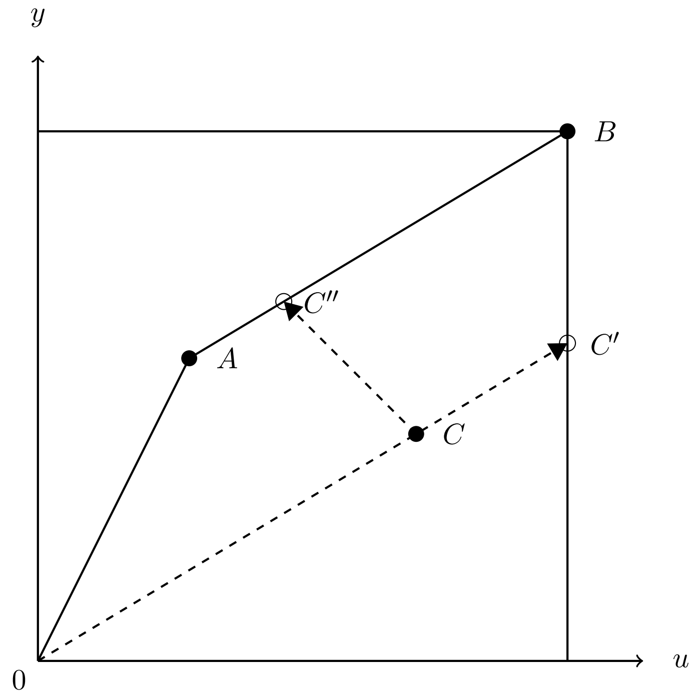

Figure 2.1: Strong and weak disposable output sets

2.2 Distance functions and efficiency evaluation

In contrast to neoclassical theory (see e.g the growth accounting approach to productivity measurement) the above provided definition and estimation of the production technology does not assume that all observed DMUs are located on the frontier of the technology. For example, in the graphical example DMU C is located within the output set and hence can be regarded as inefficient because given inputs more good outputs could be produced without increasing the amount of bad outputs. DMU A is located on the frontier if the assumption of weak disposability of bad outputs is imposed and therefore operates efficiently. Given the assumption of strong disposability of bad outputs DMU A is located inside the technology and therefore operates inefficiently.

To measure the efficiency of DMUs given a nonparametric technology various measures have been proposed (see e.g. Cook and Seiford (2009) for an overview). One of the most frequently applied measures given conventional nonparametric technologies are the radial Shephard distance functions (see Shephard (1970)) which are the inverse of the the Farrell measures of technical efficiency (see Farrell (1957)). The Shephard input distance function and the Farrell input measure of technical efficiency can be defined as

$$\frac{1}{D_I(\boldsymbol{x}, \boldsymbol{y})} = \theta(\boldsymbol{x}, \boldsymbol{y}) = \inf\{\theta : (\theta\boldsymbol{x}, \boldsymbol{y}) \in T\}. \tag{2.13}$$

Similarly, the Shephard output distance function and the Farrell measure of output efficiency can be defined as

$$\frac{1}{D_O(\boldsymbol{x}, \boldsymbol{y})} = \phi(\boldsymbol{x}, \boldsymbol{y}) = \sup\{\phi : (\boldsymbol{x}, \phi\boldsymbol{y}) \in T\} \tag{2.14}$$

where $D_I(\boldsymbol{x}, \boldsymbol{y})$ ($D_O(\boldsymbol{x}, \boldsymbol{y})$) denotes the Shephard input (output) distance function while $\theta(\boldsymbol{x}, \boldsymbol{y})$ ($\phi(\boldsymbol{x}, \boldsymbol{y})$) denotes the Farrell input (output) measure of technical efficiency.

These functions measure efficiency radially. Hence, given input orientation all inputs are reduced equiproportional until the frontier is reached. Similar, in the case of an output-oriented measurement all outputs are increased equiproportional. A point in the technology is considered input efficient (inefficient) if $\theta(\boldsymbol{x}, \boldsymbol{y}) = 1$ (< 1) and output efficient (inefficient) if $\phi(\boldsymbol{x}, \boldsymbol{y}) = 1$ (> 1). Moreover, assuming constant returns to scale of the technology

$$\theta(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\phi(\boldsymbol{x}, \boldsymbol{y})} \tag{2.15}$$

holds (see Färe and Lovell (1978)).

Note that this definition of technical efficiency relies on the definition by Farrell (1957). Hence, a DMU is efficient if no radial increases of outputs or decreases of inputs are possible. However, a Farrell efficient point may not be efficient in the terms of Pareto (1909) and Koopmans (1951). Given their definition a DMU is technical efficient if

$$\left\{(\boldsymbol{x}, \boldsymbol{y}) \in T \ \wedge \ \left(\boldsymbol{x}', \boldsymbol{y}\right) \notin T \ \forall \boldsymbol{x}' \leq \boldsymbol{x} \ \wedge \ \left(\boldsymbol{x}, \boldsymbol{y}'\right) \notin T \ \forall \boldsymbol{y}' \geq \boldsymbol{y}\right\} \tag{2.16}$$

holds. In the literature of efficiency analysis the first concept is often referred to as weak efficiency while the second is referred to as strong efficiency (see e.g. Briec (2000)). The differences in inputs or outputs of a weak efficient point to a strong efficient point in the technology are called slacks. They are addressed in non-radial models like the slacks-based measure (SBM) by Tone (2001).

The distance functions specified above can be estimated using parametric and nonparametric methods. For a comparison of these approaches see e.g. Coelli and Perelman (1999). The DEA estimate of the Farrell output measure of technical efficiency $\widehat{\phi}(\boldsymbol{x}_i, \boldsymbol{y}_i)$ for DMU $i \in \mathcal{X}_n$ can be obtained by solving the linear programming problem

$$\begin{array}{lrcl} \max\limits_{\phi, \boldsymbol{\lambda}} & \phi & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \phi \boldsymbol{y}_i & \leqq & \boldsymbol{Y\lambda} \\ & \boldsymbol{\lambda} & \geqq & \boldsymbol{0} \\ & \phi & \geqq & 0. \end{array} \tag{2.17}$$

Analogously, the DEA estimate of the Farrell input measure of technical efficiency $\widehat{\theta}(\boldsymbol{x}_i, \boldsymbol{y}_i)$ can be calculated as

$$
\begin{aligned}
\min_{\theta,\boldsymbol{\lambda}} \quad & \theta & & \\
\text{s.t.} \quad & \theta \boldsymbol{x}_i & \geqq\ & \boldsymbol{X}\boldsymbol{\lambda} \\
& \boldsymbol{y}_i & \leqq\ & \boldsymbol{Y}\boldsymbol{\lambda} \\
& \boldsymbol{\lambda} & \geqq\ & \boldsymbol{0} \\
& \theta & \geqq\ & 0.
\end{aligned}
\tag{2.18}
$$

Given that the observations in $\mathcal{X}_n$ fulfill the conditions by Kemeny et al. (1956) for all inputs and outputs both linear programming problems are well-behaved and can be solved with the conventional simplex algorithm.

The above defined distance functions can be used to evaluate efficiency in the case of conventional inputs and outputs. In contrast, for measuring the efficiency in the presence of undesirable outputs these methods are less suitable. To see this consider figure 2.1. Given that the efficiency of DMU C is measured output-oriented a radial measure projects the DMU on point C' on the frontier of the output set. This point is associated with a larger amount of good outputs but also with more bad outputs. Therefore, if environmental aspects are taken into account reaching this point does not indicate an efficiency enhancement.

A more flexible approach that allows to measure technical jointly with environmental efficiency is the directional distance function (DDF) which has been proposed by Chambers et al. (1996) based on the work by Luenberger (1992). While in Chambers et al. (1996) the DDF is presented for an input-oriented measurement of efficiency, Chung et al. (1997) have extended it to an output-oriented analysis including undesirable outputs. In this specification a vector

$$
\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g_y} \\ \boldsymbol{g_u} \end{pmatrix} \in \mathbb{R}^{s+r} \tag{2.19}
$$

is introduced that defines the direction of the efficiency measurement. Using this directional vector the directional distance function is defined as

$$
\beta\left(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}; \boldsymbol{g}\right) = \vec{D}_O\left(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}; \boldsymbol{g}\right) = \sup\left\{\beta : \left(\boldsymbol{x}, \boldsymbol{y} + \beta \boldsymbol{g_y}, \boldsymbol{u} - \beta \boldsymbol{g_u}\right) \in T\right\}. \tag{2.20}
$$

where β is the efficiency measure stating by how much the desirable outputs can be increased in the direction $\boldsymbol{g_y}$ and simultaneously the undesirable outputs (e.g. emissions) can be decreased in the direction $\boldsymbol{g_u}$, holding inputs constant, in order to reach the frontier. A DMU can be classified as efficient if $\vec{D}_O\left(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}; \boldsymbol{g}\right) = 0$ and inefficient if $\vec{D}_O\left(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}; \boldsymbol{g}\right) > 0$. Given that bad outputs are ignored and the direction of measurement is chosen as $\boldsymbol{g_y} = \boldsymbol{y}$ it can be shown (see Chung et al. (1997)) that

$$
\vec{D}_O\left(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{y}\right) = \frac{1}{D_O\left(\boldsymbol{x}, \boldsymbol{y}\right)} - 1. \tag{2.21}
$$

Therefore, it can be also concluded that

$$
1 + \vec{D}_O\left(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{y}\right) = \phi\left(\boldsymbol{x}, \boldsymbol{y}\right). \tag{2.22}
$$

For a DMU $i \in \mathcal{X}_n^{\text{Env}}$ the DEA estimate $\widehat{\beta}\left(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i; \boldsymbol{g}\right)$ of the directional distance function can be computed by solving the linear programming problem

$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \boldsymbol{y}_i + \beta \boldsymbol{g}_y & \leqq & \boldsymbol{Y\lambda} \\ & \boldsymbol{u}_i - \beta \boldsymbol{g}_u & = & \boldsymbol{U\lambda} \\ & \boldsymbol{\lambda} & \geqq & \boldsymbol{0} \\ & \beta & \geqq & 0. \end{array} \tag{2.23}$$

The directional vectors $\boldsymbol{g_y}$ and $\boldsymbol{g_u}$ are not predetermined but have to be chosen by the researcher. In most application of the environmental directional distance function the used vectors are $\boldsymbol{g_y} = \boldsymbol{y}_i$ and $\boldsymbol{g_u} = \boldsymbol{u}_i$. Thus, the directional vectors for DMU i are given by the observed amounts of good and bad outputs of this DMU. Therefore, all good and bad outputs are assigned the same weight for the efficiency analysis and the reduction of bad outputs is regarded as an equally important target as the increase of good outputs.

To illustrate the directional distance function consider again figure 2.1. As mentioned above DMU C is located in the interior of the technology and therefore classified as inefficient. Using directions $g_y = y_C$ and $g_u = u_C$ the reference point for DMU C is C''. Hence, to become efficient C has to increase y and decrease u until the frontier point C'' is reached.

2.3 Dynamic analysis of efficiency and productivity

In the last section we have presented the nonparametric analysis given a sample of n DMUs observed at one point in time. Dynamic approaches use panel data to estimate and analyze changes in the productivity of DMUs over time. These changes may occur because DMUs exhibit differences in their efficiency and/or the frontier of the technology has shifted.

To visualize this situation figure 2.2 depicts three DMUs (A, B, C) for two periods t and $t+1$ that produce two good outputs y_1 and y_2. To keep the example as simple as possible it is again assumed that all DMUs use the same amounts of inputs and in addition it is assumed that no changes in the inputs have occurred between the two periods. Note that since we visualize the production of two good outputs both outputs are assumed to be strong disposable. Analyzing DMU C we observe that the productivity increased from period t to $t+1$ because in period $t+1$ the DMU produces a larger amount of both outputs given a constant amount of inputs. Moreover, we find that the frontier has shifted outwards and hence technical progress has occurred. Changes in the relative position of C to the frontiers indicate changes in the efficiency.

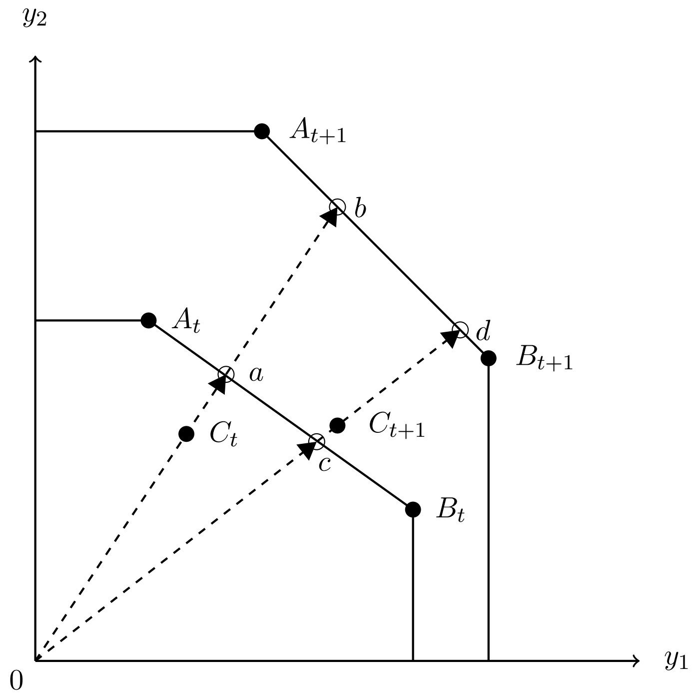

Figure 2.2: Productivity, efficiency and technical change

To analyze productivity changes and their sources using distance functions Caves et al. (1982) have proposed the Malmquist index which name refers to earlier works on index numbers by Malmquist (1953). Caves et al. (1982) developed two different versions of the index to measure changes in the productivity by comparing the distance of a DMU to the frontier functions of two periods (t and $t+1$). The two versions of the index differ with regard to whether the frontier of period t or the frontier of period $t+1$ is used as the benchmark to analyze productivity changes. The first version of the Malmquist index uses the technology of period t to construct the reference frontier and can be defined as

$$\mathrm{M}^t\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right) = \frac{D_O^t\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^t\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)}. \tag{2.24}$$

In this notation the superscript for the index and the distance functions refers to the time period of the reference technology while the subscripts for the inputs and outputs refer to the time period of the analyzed input-output combination. Therefore, a mixed-period distance function is, for example, given by $D_O^t\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)$. The index exhibits values larger (less) than 1 if the productivity has increased (decreased) between two periods. In the above depicted example for DMU C this index can be calculated as

$$\mathrm{M}^t\left(\boldsymbol{x}_{t,C}, \boldsymbol{y}_{t,C}, \boldsymbol{x}_{t+1,C}, \boldsymbol{y}_{t+1,C}\right) = \frac{\overbrace{\overline{0C_{t+1}}/\overline{0c}}^{>1}}{\underbrace{\overline{0C_t}/\overline{0a}}_{<1}} > 1. \tag{2.25}$$

The second version of the Malmquist index uses the technology of period $t+1$ to construct the frontier. This Malmquist index can be defined as

$$\mathrm{M}^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right) = \frac{D_O^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)}. \tag{2.26}$$

For DMU C in the above depicted example this index can be calculated as

$$\mathrm{M}^{t+1}\left(\boldsymbol{x}_{t,C}, \boldsymbol{y}_{t,C}, \boldsymbol{x}_{t+1,C}, \boldsymbol{y}_{t+1,C}\right) = \frac{\overbrace{\overline{0C_{t+1}}/\overline{0d}}^{<1}}{\underbrace{\overline{0C_t}/\overline{0b}}_{<1}} > 1. \tag{2.27}$$

Since the distance of point C_{t+1} to the frontier of period $t+1$ is smaller than the distance of point C_t the ratio of the distance functions is larger than one.[16]

The two versions of the Malmquist index do not necessarily provide the same results and it is not clear which frontier function is relevant for the analysis. Therefore, Färe et al. (1992) have proposed to use the geometric mean of the two indices (similar to the Fisher (1922) ideal index) to measure productivity changes. This version of the Malmquist index reads as:

$$\mathrm{M}^{t,t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right) = \left[\frac{D_O^{t}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)} \cdot \frac{D_O^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)}\right]^{1/2} \tag{2.28}$$

Färe et al. (1992) also provide a decomposition of this index into technical change

$$\mathrm{MTech}^{t,t+1} = \left[\frac{D_O^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)}{D_O^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)} \cdot \frac{D_O^{t}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}\right]^{1/2} \tag{2.29}$$

which captures shifts in the frontier and efficiency change

$$\mathrm{MEff}^{t,t+1} = \frac{D_O^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}\right)}{D_O^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t\right)}. \tag{2.30}$$

which captures changes of the relative position to the frontier. Improvement of efficiency and technical progress are indicated by values larger than one of the associated measures while values lower than one indicate efficiency decrease and technical regress. For DMU C this decomposition would read

$$\mathrm{M}^{t,t+1}\left(\boldsymbol{x}_{t,C}, \boldsymbol{y}_{t,C}, \boldsymbol{x}_{t+1,C}, \boldsymbol{y}_{t+1,C}\right) = \underbrace{\left[\frac{\overline{0C_t}/\overline{0a}}{\overline{0C_t}/\overline{0b}} \cdot \frac{\overline{0C_{t+1}}/\overline{Oc}}{\overline{0C_{t+1}}/\overline{0d}}\right]^{1/2}}_{>1} \cdot \underbrace{\frac{\overline{0C_{t+1}}/\overline{0d}}{\overline{0C_t}/\overline{0a}}}_{<1} > 1. \tag{2.31}$$

The productivity of DMU C has increased although its efficiency declined because technical progress overcompensated the efficiency decrease.

Given the assumption of variable returns to scale several further decompositions of the Malmquist

[16] In terms of the Shephard output distance function a smaller distance leads to a larger value which is closer to one.

index are possible (see Grosskopf (2003) for an overview). However, in this case the mixed-period distance functions may not have a solution and hence the Malmquist index may not be computable.

For a dynamic analysis of productivity in the the presence of undesirable outputs Chung et al. (1997) have proposed an extension of the conventional Malmquist index and named it Malmquist-Luenberger index. This index is based on the directional distance function presented above and can be defined as

$$\mathrm{ML}^{t,t+1} = \left[\frac{1 + \vec{D}_O^t \left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1 + \vec{D}_O^t \left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \cdot \frac{1 + \vec{D}_O^{t+1} \left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1 + \vec{D}_O^{t+1} \left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \right]^{1/2}. \tag{2.32}$$

Similar to the Malmquist index it can be decomposed into technical change

$$\mathrm{MLTech}^{t,t+1} = \left[\frac{1 + \vec{D}_O^{t+1} \left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1 + \vec{D}_O^t \left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)} \cdot \frac{1 + \vec{D}_O^{t+1} \left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)}{1 + \vec{D}_O^t \left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \right]^{1/2} \tag{2.33}$$

and efficiency change

$$\mathrm{MLEff}^{t,t+1} = \frac{1 + \vec{D}_O^t \left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1 + \vec{D}_O^{t+1} \left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)}. \tag{2.34}$$

In order to shorten the notation, the arguments of the ML index and its components $\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t, \boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_t, \boldsymbol{g}_{t+1}\right)$ are suppressed in these expressions.

Depending on the specified directional vector the directional distance function may not have a solution (see e.g. Briec and Kerstens (2009)). Hence, in contrast to the Malmquist index the Malmquist-Luenberger index may not be computable even if the technology exhibits constant returns to scale. Estimations of both indices can be obtained by replacing the distance functions with their nonparametric estimators.

3 Technical efficiency of automobiles†

3.1 Motivation

The problem of global warming due to anthropogenic emissions of greenhouse gases as presented in the introduction of this dissertation has gained large public attention in recent years leading to a thrust in political activities to set environmental targets. Sustaining leadership in environmental standards, the European Commission started a program to limit average emissions of carbon dioxide of new cars to 130 g/km (Commission of the European Communities (2007)). This program recognizes the important contribution of the transport sector to the overall emissions of greenhouse gases. According to the European Environment Agency (2011b) 17.5% of all greenhouse gas emissions in Europe in 2010 were caused by road transportation. In particular, the use of passenger cars contributes to 12% of the overall carbon dioxide emissions (see Commission of the European Countries (2007, p. 2)). Moreover, from a dynamic point of view the development of emissions is very problematic. While the CO_2 emissions of other sectors like energy production or manufacturing have decreased in Europe since 1990 (see Krautzberger and Wetzel (2012)) emissions of road transportation increased by 23% between 1990 and 2009 (see European Environment Agency (2011b)).

These numbers show the importance of accounting for the emission of carbon dioxide when evaluating and comparing the efficiency of automobiles. However, when analyzing the efficiency of different cars the literature often presents comparisons of ratios of the emitted amount of CO_2 to a single specific car characteristic. For example, Sullivan et al. (2004) use the ratio of carbon dioxide to the weight of the car to compare gasoline and diesel automobiles. Zervas (2009) presents an analysis of CO_2 and various car characteristics but again different cars are only compared with regard to the ratios of the emissions to each of the characteristics. This approach therefore does not present an analysis of the overall efficiency of automobiles when accounting for the emission of CO_2.

In contrast, nonparametric efficiency analysis allows to evaluate the performance of different cars accounting for multiple inputs and outputs. Papahristodoulou (1997), Fernandez-Castro and Smith (2002), Staat et al. (2002), Oh et al. (2010) and Cantner et al. (2012) provide analyses of car efficiency but limit their focus on conventional inputs (e.g. fuel consumption, price etc.) and outputs (e.g. top speed, engine power etc.).[17] Environmental factors are included in the nonparametric analyses of sport utility vehicles (SUVs) by Kortelainen and Kuosmanen (2007) and Kuosmanen and Kortelainen (2007). However, both studies focus solely on environmental efficiency and do not present an analysis of technical efficiency. Los and Verspagen (2009) present a dynamic analysis of cars in the United Kingdom including carbon dioxide but due to data limitations they include only a single technical feature (engine capacity) of the cars. González et al. (2012) use DEA to analyze automobiles in Spain including carbon dioxide emissions. They aim at finding the lowest price given the technical characteristics of the car under evaluation. Hence, this analysis is conducted focusing solely on the market performance of cars.

† This chapter is based on Hampf and Krüger (2010).

[17] Studies that conduct nonparametric analyses of other vehicle types also ignore emissions. See Odeck and Hjalmarsson (1996) for an efficiency analysis of trucks and Stokes and Claar (2004) for an analysis of tractors.

In this study we follow the previous literature by conducting a nonparametric analysis of the efficiency of automobiles incorporating multiple inputs and outputs. However, our approach differs in several important aspects. Firstly, we estimate the efficiency including carbon dioxide emissions as a weak disposable output as discussed in the previous chapter. This allows to measure efficiency taking account of both technical as well as environmental targets and thereby obtaining a more differentiated view of the efficiency of cars. Secondly, we use bootstrapping methods to correct for the bias in nonparametric efficiency estimates (see Simar and Wilson (1998)) and we show that ignoring this bias as done in the existing literature leads to an underestimation of the inefficiency of cars. Thirdly, to gain more insights into the relative performance of specific groups of cars which differ with regard to certain characteristics e.g. with respect to the engine type (gasoline, diesel or natural gas) we apply the frontier separation approach first introduced by Charnes et al. (1981). This approach allows to disentangle group specific and individual specific inefficiencies while comparing groups of automobiles.

In the following section we present the theoretical background of our analysis. This section extends the discussion of the general concepts in the previous chapter by introducing the frontier separation approach, a test to compare the group results and the bootstrapping methodology to correct the bias in nonparametric efficiency estimates.

3.2 Theoretical concepts for the analysis of automobiles

3.2.1 The frontier separation approach

One possibility to rank the efficiency of different groups of observations is to use the efficiency results obtained by an analysis of the overall dataset and compare them across groups or alternatively between a group and the remaining observations. A drawback of this approach is that the results of the analysis can be due to either within-group or between-group differences.

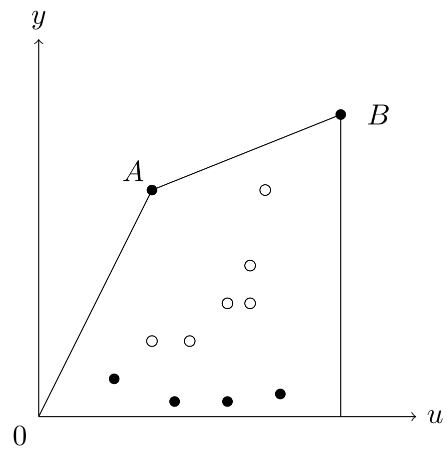

Figure 3.1: Weak disposable output set with two groups of DMUs

To illustrate this point, figure 3.1 shows a weak disposable output set that consists of two groups of observations (group 1 denoted by ∘ and group 2 denoted by •). Since the frontier is constructed only by observations from group 2 (the points A and B) one could characterize group 2 as being more efficient than group 1. But if one compares the overall efficiency measures this result might be overruled since several observations from group 2 are located far below the frontier while the observations of group 1 lie close to the boundary. Hence, although no DMU of group 1 is classified as efficient this group might be indicated as more efficient.

To evaluate whether differences in the performance of a group depend on group or individual aspects we apply the frontier separation approach of Charnes et al. (1981) in a variant proposed by Portela and Thanassoulis (2001). In this approach, the overall efficiency score of a DMU estimated using the whole dataset is split into two components, the managerial efficiency and the program efficiency. While the managerial efficiency indicates the inefficiency of a DMU relative to its group frontier (the individual inefficiency), the program efficiency indicates the difference between the group frontier and the overall frontier (the group specific inefficiency).[18] The efficiency scores of the DMUs can therefore be decomposed as

$$\phi(\boldsymbol{x},\boldsymbol{y})_{\mathrm{Ov}} = \phi(\boldsymbol{x},\boldsymbol{y})_{\mathrm{Ma}} \cdot \phi(\boldsymbol{x},\boldsymbol{y})_{\mathrm{Pr}} \tag{3.1}$$

$$1+\beta(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u};\boldsymbol{g})_{\mathrm{Ov}} = (1+\beta(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u};\boldsymbol{g})_{\mathrm{Ma}}) \cdot (1+\beta(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u};\boldsymbol{g})_{\mathrm{Pr}}) \tag{3.2}$$

where the subscripts “Ov”, “Ma” and “Pr” denote the overall, managerial and program efficiency. Note that we have presented the decomposition for the Farrell output measure of technical efficiency and the directional output distance function because these functions will be applied in the subsequent analysis. The addition of unity to the directional distance function serves to make the results comparable with those of the Farrell measure (see the chapter on general concepts).

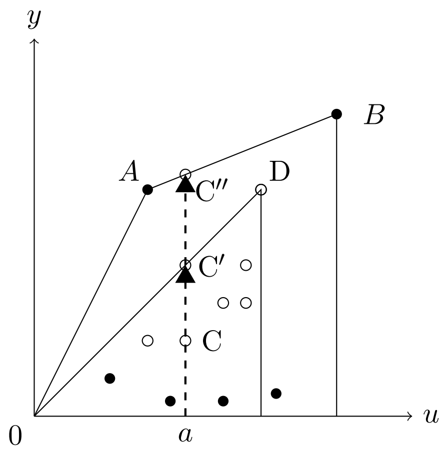

Figure 3.2: Frontier separation approach

To illustrate this approach figure 3.2 shows the weak disposable output set of figure 3.1. The depicted analysis includes an undesirable output u therefore the efficiency is measured using

[18] We use the terminology of managerial and program efficiency because it is common in the literature of the frontier separation approach (see Thanassoulis et al. (2008)) although it does not completely apply to the analysis of cars.

the directional distance function. The arrows indicate that the efficiency is measured using the direction $g_y = y$ and $g_u = 0$. The boundary of the overall output set is given by $\overline{0AB}$ and the vertical extension to B. Thus, the overall efficiency result of C is $\frac{\overline{aC''}}{\overline{aC}}$. If C is compared only to its own program, the DMUs belonging to group 1 become irrelevant and the boundary is given by $\overline{0D}$ and the vertical extension to D. The efficiency score for C is then given by $\frac{\overline{aC'}}{\overline{aC}}$. This is called the managerial efficiency since it expresses all inefficiency that is not based on program differences. The program efficiency is given by the difference between the two boundaries and can be estimated as the residual of the overall and the managerial efficiency, $\frac{\overline{aC''}}{\overline{aC}} \cdot \frac{\overline{aC}}{\overline{aC'}} = \frac{\overline{aC''}}{\overline{aC'}}$.

The above presented decomposition of the efficiency results allows for a more detailed comparison of the different groups of automobiles. To evaluate whether potential differences are statistically significant we apply a test for stochastic dominance which is presented in the next section.

3.2.2 Testing for stochastic dominance

The statistical literature provides numerous tests for group comparison e.g. parametric ones like the standard t-test or nonparametric ones like the Wilcoxon ranksum test (see e.g. Sheskin (2007) for an overview). Usually, the efficiency estimates are non-normally distributed since e.g. the Farrell efficiency measures are bounded at unity and a spurious mass of observations can be found at this boundary. Therefore, the standard t-test can not be applied validly to nonparametric efficiency estimates. Other authors (e.g. Brockett and Golany (1996)) apply nonparametric, rank-based tests for the comparison. The use of these tests in the context of nonparametric efficiency analysis is criticized by Simpson (2005, 2007). Simar and Zelenyuk (2006) propose an application of the test by Li (1996) to a DEA analysis but this test does only compare whether two distributions are equal or not. It does not provide evidence whether the efficiency scores for one group are larger or lower than the results of another group. Therefore, we apply a test of stochastic dominance proposed by Barrett and Donald (2003) which is not rank-based, does not impose distributional assumptions on the efficiency estimates and allows to analyze groups in terms of lower or larger efficiency results.

Stochastic dominance plays an important role in risk and decision theory but is of general applicability for one-sided comparisons of random variables. Several notions of stochastic dominance can be distinguished (see Levy (1992)). Here, we rely on the concepts of first-order stochastic dominance (FSD) and second-order stochastic dominance (SSD). According to first-order stochastic dominance a (real-valued) random variable V stochastically dominates another random variable W if the cumulative distribution function (CDF) of V is completely below that of W over the whole support. Formally, $V \succ_{\text{FSD}} W$ if $F_V(z) \leq F_W(z)$ at all points z in the common support of V and W with strict inequality for some z. Second-order stochastic dominance is less strict in relying on the area below the CDF of V (up to a certain upper bound t) being smaller than the area below the CDF of W (up to the same t). If this requirement is satisfied for all t in the common support of V and W then we say that V second-order stochastically dominates W. Formally, $V \succ_{\text{SSD}} W$ if $\int_{-\infty}^{t} F_V(z)dz \leq \int_{-\infty}^{t} F_W(z)dz$ for all $t \in \mathbb{R}$ and strict inequality for some t.

From these definitions it is clear that FSD implies SSD but not vice versa. Since FSD is a very demanding concept it is important to have a less demanding one such as SSD.[19] Examples of first- and second-order stochastic dominance are given in figure 3.3. In the left graph $V \succ_{\text{FSD}} W$ because $F_V(z)$ lies completely below $F_W(z)$. In the right graph $V \succ_{\text{SSD}} W$ because the area in which $F_V(z)$ lies above $F_W(z)$ is smaller than the area in which $F_V(z)$ lies below $F_W(z)$.

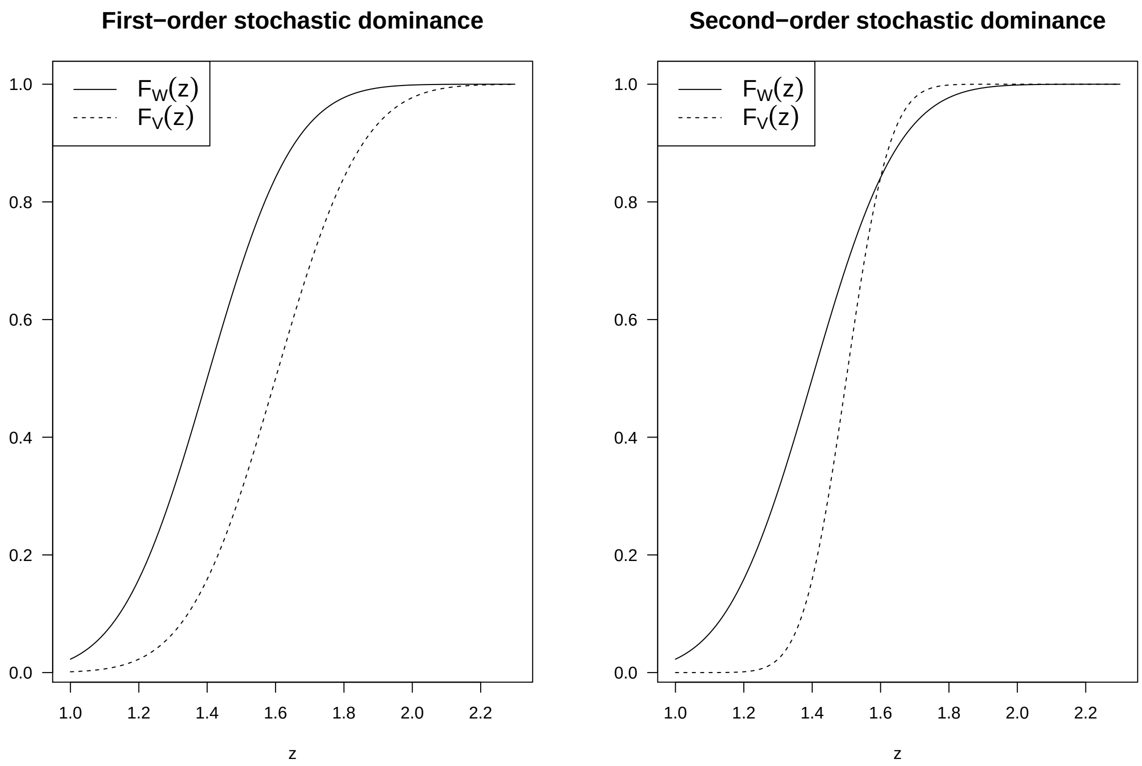

Figure 3.3: Examples of first- and second-order stochastic dominance

For taking these concepts to the data, we rely on statistical measures and the bootstrapping procedure proposed by Barrett and Donald (2003). We suppose to have two samples of observations for V (v_i with $i = 1, \ldots, n_V$) and W (w_i with $i = 1, \ldots, n_W$). In our specific application these are, for example, the efficiency measures of the cars pertaining to different groups. Testing FSD requires estimating the CDFs by their corresponding empirical distribution functions (EDF)

$$\widehat{F}_V(z) = n_V^{-1} \sum_{i=1}^{n_V} I(v_i \leq z) \quad \text{and} \quad \widehat{F}_W(z) = n_W^{-1} \sum_{i=1}^{n_W} I(w_i \leq z) \tag{3.3}$$

with $I(.)$ denoting the usual indicator function and using the Kolmogorov-Smirnov-type test statistic

$$\widehat{D}_{\text{FSD}} = \max_z \{\widehat{F}_V(z) - \widehat{F}_W(z)\} \tag{3.4}$$

[19] In a productivity context Delgado et al. (2002) as well as Fariñas and Ruano (2005) use first-order stochastic dominance to compare productivity distributions of Spanish manufacturing firms. Second-order stochastic dominance is not applied in that context as far as we are aware of.

to test the null hypothesis $H_0 : V \succeq_{\text{FSD}} W$ against the alternative $H_1 : W \succ_{\text{FSD}} V$. In our implementation this statistic is evaluated over an equally spaced grid of 100 points for z spanning the whole range of all observations. The null hypothesis is rejected in favor of the alternative for large values of the test statistic. We compute the p-values from the finite sample distribution approximated by $B = 2000$ bootstrap replications.[20] The application of the bootstrap here amounts to resample B times with replacement from the joint sample $(v_1, \ldots, v_{n_V}, w_1, \ldots, w_{n_W})$ and then to compute the test statistic repeatedly, resulting in $D^*_{\text{FSD},1}, \ldots, D^*_{\text{FSD},B}$. Resampling from the joint sample simulates the test statistics under the equality of the distributions. As Abadie (2002, p. 287) notes, this mimics the least favorable case for the null hypothesis stated above. The p-value of the test is subsequently computed as the fraction of the bootstrap statistics exceeding the statistic $\widehat{D}_{\text{FSD}}$ computed from the original samples, i.e. by the formula

$$p = \frac{1}{B} \sum_{b=1}^{B} I(D^*_{\text{FSD},b} > \widehat{D}_{\text{FSD}}). \tag{3.5}$$

Note that not rejecting the null hypothesis does not reveal whether the distributions are equal or $V \succ_{\text{FSD}} W$. Therefore, one has to test whether the converse hypothesis is rejected. If this is the case then $V \succ_{\text{FSD}} W$. Otherwise, $F_V(z) = F_W(z)$. Moreover, rejecting the null hypothesis does not allow to conclude $W \succ_{\text{FSD}} V$ because when the distributions cross the null hypothesis would be also rejected. Only if the converse null hypothesis is not rejected we can conclude $W \succ_{\text{FSD}} V$. If this test also rejects the null hypothesis then the distributions cross and a test for second-order stochastic dominance needs to be applied. Figure 3.4 presents an overview of the testing procedure and possible results of the test for first-order stochastic dominance.

Testing SSD affords the computation of the empirical analogs of the integrals appearing in the definition. Subjecting these integrals to partial integration we get

$$\int_{-\infty}^{t} F_V(z)dz = \int_{-\infty}^{t} (t-z)dF_V(z) \text{ and } \int_{-\infty}^{t} F_W(z)dz = \int_{-\infty}^{t} (t-z)dF_W(z) \tag{3.6}$$

with the empirical analogs

$$\widehat{G}_V(z) = n_V^{-1} \sum_{i=1}^{n_V} (z - v_i) I(v_i \leq z) \text{ and } \hat{G}_W(z) = n_W^{-1} \sum_{i=1}^{n_W} (z - w_i) I(w_i \leq z). \tag{3.7}$$

As a test statistic for testing the null hypothesis $H_0 : V \succeq_{SSD} W$ against the alternative $H_1 : W \succ_{SSD} V$ we now compute

$$\widehat{D}_{SSD} = \max_z \{\widehat{G}_V(z) - \widehat{G}_W(z)\} \tag{3.8}$$

and reject the null hypothesis for large values of the test statistic. Again 2000 bootstrap replications are computed according the procedure outlined above and the p-value is computed analogous to (3.5).

[20] For applied references on bootstrapping see Davison and Hinkley (1997) or Efron and Tibshirani (1993).

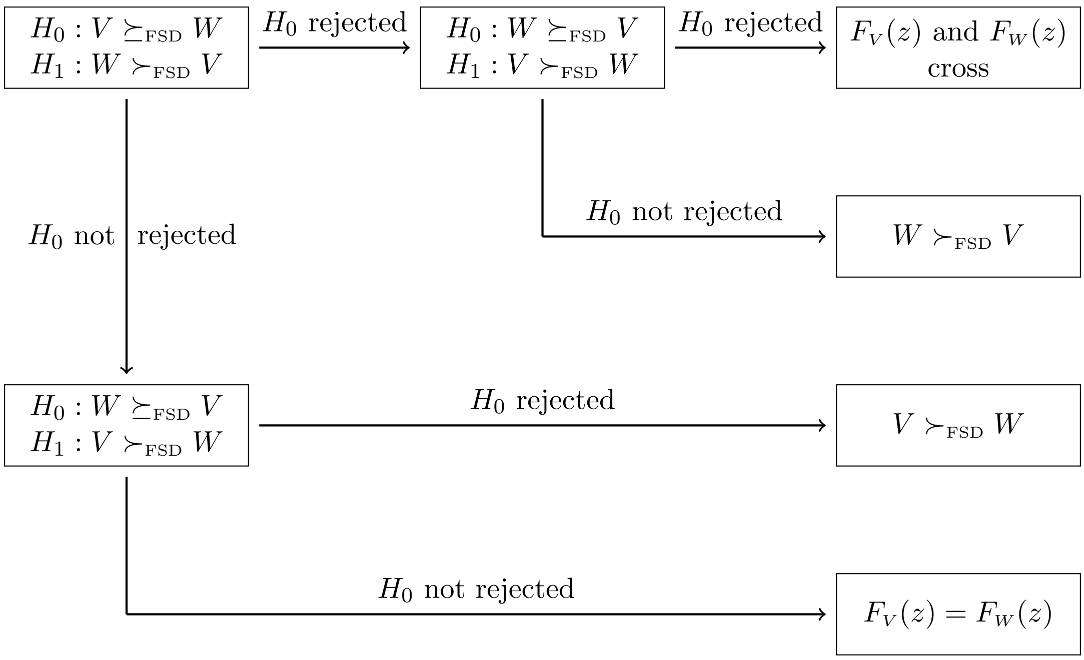

Figure 3.4: Structure of the test by Barrett and Donald (2003)

3.2.3 Bootstrapping the efficiency scores

By construction the technology set estimated by nonparametric methods is a subset of the true but unknown technology. Therefore, the obtained output-oriented efficiency measures are biased downwards (see Simar and Wilson (2000) for an overview of this problem). Simar and Wilson (1998) propose a bootstrap algorithm to estimate the bias and to correct the efficiency results. This approach simulates the data generating process (DGP) by first using the distance function estimates to place all observations on the frontier. Random deviations from the frontier are then used to generate bootstrap samples. For each observation in the original sample the distance to the original frontier and to the frontiers of each of the bootstrap samples is calculated. The average difference in the distance to the bootstrapped frontiers and to the original frontier is used as an estimation of the bias in the efficiency score. In the following we present this approach in more detail. Our description of the algorithm for the directional distance function results is geared to the presentation in Simar and Wilson (2008, p. 462). It can be analogously applied to the Farrell output measure of technical efficiency. The algorithm can be summarized by the following steps:

1. Estimate the directional distance functions using the original data set $\mathcal{X}_n^{\text{Env}}$ resulting in $\widehat{\beta}_i = \widehat{\beta}(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i; \boldsymbol{g}) \geq 0$ for $i = 1, \ldots, n$.
2. Select the bandwidth parameter h by cross validation (i.e. by the function `eff.bw()` from the *FEAR* package for R applied to $\widehat{\beta}_i + 1$, $i = 1, \ldots, n$).[21]
3. Draw $\delta_1^*, \ldots, \delta_n^*$ with replacement from the set $\left\{\widehat{\beta}_1, \ldots, \widehat{\beta}_n, -\widehat{\beta}_1, \ldots, -\widehat{\beta}_n\right\}$ imposing the

[21] See Wilson (2008) for the documentation of the package.

reflection around the boundary of zero here.

4. Draw $\varepsilon_1^*, ..., \varepsilon_n^*$ as independent standard normal variates and compute $\delta_i^{**} = \delta_i^* + h\varepsilon_i^*$ for $i = 1, \dots, n$.
5. Compute $\delta_i^{***} = \bar{\delta}^* + (1 + h^2\widehat{\sigma}_\delta^{-2})^{-1/2} \cdot (\delta_i^{**} - \bar{\delta}^*)$ with $\bar{\delta}^* = n^{-1}\sum_{i=1}^n \delta_i^*$ and $\widehat{\sigma}_\delta^2 = (n-1)^{-1}\sum_{i=1}^n(\delta_i^* - \bar{\delta}^*)^2$. Then compute $\beta_i^* = |\delta_i^{***}|$ for $i = 1, \dots, n$ to remove the reflection.
6. Create the bootstrap sample $\mathcal{X}_n^{\text{Env}^*} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i^*, \boldsymbol{u}_i^*), i = 1, \dots, n\}$ with $\boldsymbol{y}_i^* = (1+\widehat{\beta}_i \boldsymbol{g_y})/(1+\beta_i^* \boldsymbol{g_y}) \cdot \boldsymbol{y}_i$ and $\boldsymbol{u}_i^* = (1 - \widehat{\beta}_i \boldsymbol{g_u})/(1 - \beta_i^* \boldsymbol{g_u}) \cdot \boldsymbol{u}_i$ for $i = 1, \dots, n$.
7. Compute the DDF efficiency measures $\widehat{\beta}_i^* = \widehat{\beta}_i^*(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i; \boldsymbol{g})$ for the data point $(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i)$, $i = 1, \dots, n$, and the reference set $\mathcal{X}_n^{\text{Env}^*}$.
8. Repeat steps 3 to 7 B times to obtain the bootstrap estimates $\left\{\widehat{\beta}_{bi}^*, b = 1, \dots, B\right\}$.

These bootstrap estimates can be used to calculate the bias of the efficiency estimate for an observation i as

$$\widehat{\text{bias}}_B\left(\widehat{\beta}_i\right) = \frac{1}{B}\sum_{b=1}^{B}\left[\widehat{\beta}_{bi}^* - \widehat{\beta}_i\right]. \tag{3.9}$$

The bias corrected estimate then reads as

$$\widehat{\beta}_i^{\text{BC}} = \widehat{\beta}_i - \widehat{\text{bias}}_B\left(\widehat{\beta}_i\right) \tag{3.10}$$

However, since the bias correction leads to additional noise we correct for it only if

$$\frac{\left|\widehat{\text{bias}}_B\left(\widehat{\beta}_i\right)\right|}{\widehat{\sigma}_{\widehat{\beta}_{bi}^*}} > \frac{1}{\sqrt{3}} \tag{3.11}$$

holds where $\widehat{\sigma}_{\widehat{\beta}_{bi}^*}$ is the estimated standard deviation of the bootstrap estimates of the directional distance function.[22] Furthermore, when applying the frontier separation approach we correct the results for the managerial efficiency only if the bias corrected managerial efficiency result is lower than the overall efficiency results. Otherwise, we would obtain a program efficiency smaller than one and thus the theoretically implausible result that the group frontier would be located above the overall frontier.

In the following section we present the efficiency analysis of the automobiles. After discussing the data used in the study and the applied efficiency measures we will present the results for both the overall sample as well as the different groups analyzed.

3.3 Efficiency analysis of the automobiles

3.3.1 Data of the automobiles

The data for the automobiles including the emissions of carbon dioxide are collected from the "ADAC Autokatalog 2010" (ADAC (2010)) which is an continuously updated database published

[22] See Efron and Tibshirani (1993) for a discussion of the additional noise problem and the threshold $\frac{1}{\sqrt{3}}$.

online by the German automobile club ADAC. It contains data for 55 car producers with 403 product lines and 9256 model variants that are sold in Germany.[23] Since many of these variants only differ in aspects that are not relevant for our analysis (e.g. optional equipment like airbags) we eliminate these duplets resulting in 3961 remaining observations. These cars have been divided into 7 vehicle classes according to some of their characteristics (e.g size, price etc., see ADAC (2009)).[24] Since some of these classes only contain very few observations we aggregate them to 3 main classes:

- Compact class (combining "Microwagen", "Kleinstwagen" and "Kleinwagen")
- Middle class (combining "Untere Mittelklasse" and "Mittelklasse")
- Upper class (combining "Obere Mittelklasse" and "Oberklasse").

These three classes refer to the general structure of the automobile market and we will use the frontier separation approach to compare the performance of cars belonging to the different classes. We expect the compact cars to be more efficient as the middle and upper class vehicles because we assume that the additional luxury equipment lowers the technical and environmental performance hence resulting in a trade-off.

Beside the car classes we also analyze the following groups:[25]

- Engine type (differentiating cars with gasoline, diesel and natural gas engines)
- Sport utility vehicles (SUVs).

The analysis of engine types is interesting because diesel and natural gas engines cars are often found to produce less emissions compared to gasoline cars (see Eberle (2008)). In our analysis we will check whether this leads to increased efficiency if multiple inputs and outputs are accounted for. We do not analyze hybrid cars separately because there is only a very small number of observations (14) present in our dataset (see Choi and Oh (2010) for an analysis of hybrid cars). The SUVs are of special interest since their market share in Europe increased during the last years (see Zervas (2010)) while their environmental performance is questionable (see Plotkin (2004)).

The literature on nonparametric efficiency analysis of automobiles is very heterogeneous in terms of the analyzed datasets and the included variables. To determine which inputs and outputs we use in our analysis we consider a highly stylized conception of a car which orients at Papahristodoulou (1997). We assume that the car is bought (input variable: price) to transport a load (output variable: payload) which in combination with the engine power (output variable: engine power) and the weight (input variable: net weight) leads to the acceleration (input variable: acceleration) and the final top speed (output: top speed) of the car.[26] This transportation process requires fuel (input variable: fuel consumption) and produces carbon dioxide emissions

[23] In this study we use the same terminology as in Cantner et al. (2012) (see table A.1 in appendix A for an example).
[24] These classes are: Microwagen, Kleinstwagen, Kleinwagen, Untere Mittelklasse, Mittelklasse, Obere Mittelklasse and Oberklasse.
[25] Note that the groups are not disjoint, e.g. cars with gasoline engines also belong to one of the car classes.
[26] Acceleration is included as an input because it is measured in seconds needed to accelerate the car from 0 to 100 km/h. Hence, a lower value of the variable is associated with a larger acceleration.

(undesirable output variable: CO_2). We limit our research focus on this technical view of an automobile as plainly providing transportation services and ignore variables like luxury equipment (e.g air conditioner) because the incorporation of qualitative variables in DEA or DDF analyses is problematic (see e.g. Dyson et al. (2001)) and the input variable price captures these features. Table 3.1 contains some descriptive statistics of the variables used in our study.

Table 3.1: Descriptive statistics of the automobile data

	Min.	1. Qu.	Median	Mean	3. Qu.	Max.	SD
Inputs							
Price [€]	6990	20890	28860	36871	39990	523838	36812.72
Fuel consumption [l/km]	3.3	5.9	7.1	7.53	8.6	21.3	2.40
Net weight [kg]	825	1355	1545	1562	1730	2855	312.56
Acceleration [s]	3.2	8.1	10.2	10.29	12.2	23.6	3.14
Outputs							
Engine power [PS]	52	109	145	172.4	200	670	95.59
Top speed [km/h]	135	180	200	204	226	340	33.39
Payload [kg]	115	425	484	496.4	547	1160	127.60
Undesirable output							
CO_2 [g/km]	87	147	172	183	205	495	53.74

It is evident from the descriptive statistics that our dataset covers a wide range of different car types comprising low-budget cars and luxury vehicles (price varying from 6990€ (Dacia Sandero) to 523838€ (Maybach)) as well as small city cars and SUVs (payload varying from 115 kg (Daihatsu Copen) to 1160 kg (Ford Ranger)). Interestingly, less than 25% of the cars achieve the emission target set by the European Union which limits CO_2 emissions to 130 g/km.

3.3.2 Measures to evaluate the efficiency of automobiles

In our subsequent analysis we will evaluate the efficiency of automobiles applying three different measures. To compare our results to the previous literature we start by conducting an analysis ignoring the carbon dioxide emissions. In this case we will use the Farrell output measure of technical efficiency ($\phi(\boldsymbol{x}, \boldsymbol{y})$) to estimate the efficiency. We apply an output-oriented measure to make the results comparable to those obtained using a directional distance function when emissions are included as the single bad (weak disposable) output. We apply the directional vector

$$\boldsymbol{g}_{\mathrm{O}} = \begin{bmatrix} \boldsymbol{y} \\ 0 \end{bmatrix} \tag{3.12}$$

to measure the efficiency in the presence of the additional constraint of the undesirable output. Comparing the results to $\phi(\boldsymbol{x}, \boldsymbol{y})$ allows to measure the influence of the undesirable outputs on the efficiency results. Finally, a second directional vector

$$\boldsymbol{g}_{\mathrm{UO}} = \begin{bmatrix} \boldsymbol{y} \\ u \end{bmatrix} \tag{3.13}$$

is used to measure the efficiency of cars if the reduction of emissions is regarded as an equally important target as the increase of conventional outputs. This measure can be compared to the results obtained with the vector $\boldsymbol{g}_{\mathrm{O}}$ to evaluate whether a trade-off between technical and environmental targets exists.

3.3.3 Results for the overall sample of automobiles

In this section we present and discuss the results for the three efficiency measures presented for the overall sample of 3961 automobiles. Descriptive statistics for the results can be found in table 3.2 and density plots which provide a graphical visualization of the results can be found in figure 3.5. To shorten the notation in the tables and figures we have abbreviated the bias corrected distance estimates as $\widehat{\rho}$ for the Farrell output measure of technical efficiency, $\widehat{\beta}_{\mathrm{O}}$ for the directional distance function using vector $\boldsymbol{g}_{\mathrm{O}}$ and by $\widehat{\beta}_{\mathrm{UO}}$ for the directional distance function using vector $\boldsymbol{g}_{\mathrm{UO}}$. To make the results for the Farrell measure and the directional distance function comparable we present the latter with an added value of one (see the chapter on general concepts for the relationship between Farrell measures, Shephard distance functions and the directional distance function).

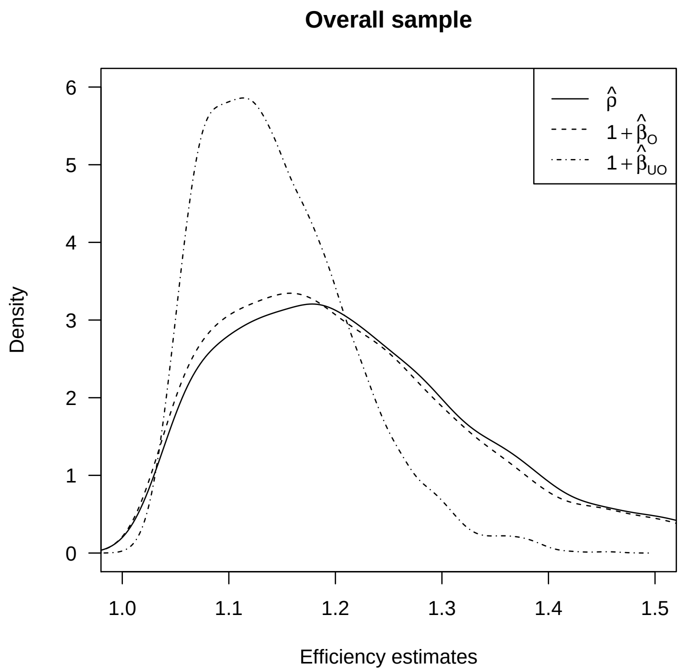

Figure 3.5: Density functions of the efficiency measures (overall sample)

Table 3.2: Descriptive statistics of the efficiency measures (overall sample)

Efficiency measure	Min.	1. Qu.	Median	Mean	3. Qu.	Max.	SD
$\widehat{\rho}$	1.0093	1.1294	1.2081	1.2354	1.3052	2.1201	0.1473
$1+\widehat{\beta}_{\mathrm{O}}$	1.0135	1.1207	1.1959	1.2240	1.2921	2.0129	0.1411
$1+\widehat{\beta}_{\mathrm{UO}}$	1.0136	1.0918	1.1345	1.1449	1.1873	1.4589	0.0690

The results for the analysis excluding the carbon dioxide emissions show that the inefficiency of cars ranges from 0.93% (VW Polo 1.2 TSI Comfortline) to 112.01% (Nissan Patrol 3.0 Di SE Automatik) with the average inefficiency being 23.54%. Hence, the average car has to increase all its outputs by 23.54% holding inputs constant to become efficient. Due to the bias correction of the efficiency scores no car is indicated as fully efficient ($\widehat{\rho} = 1$). The bias corrected efficiency estimates are on the average 2.9% larger than the non-corrected efficiency scores. This shows that not accounting for the bias leads to a non-neglible underestimation of the inefficiency of the cars. However, comparing the bias with the descriptive statistics we find that while the magnitude is non-neglible the results are not crucially driven by the bias.

Comparing our findings with the results obtained in previous studies is somewhat difficult since the literature mentioned in the introduction of this study is very heterogeneous in terms of analyzed datasets, selected variables and applied efficiency measures. The two most related studies are Papahristodoulou (1997) and Cantner et al. (2012). Both studies analyze cars using German data. Papahristodoulou (1997) uses data obtained from the automobile magazine “Auto, Motor und Sport” applying an input-oriented DEA analysis while Cantner et al. (2012) focus on compact class cars using the order-m method to analyze data from the “ADAC”, the same data source as in our analysis. Comparing our results with the results of these two studies we find substantially larger inefficiencies.[27] Papahristodoulou (1997) finds that the minimal efficiency in his analysis is 85%. Hence, the maximal inefficiency is 15%. In our analysis the inefficiency is at the average 23.54%. Cantner et al. (2012) obtain a mean technical inefficiency of 6.6% for the year 2005. Comparing these differences with the magnitude of the bias in our study shows that the bias correction alone can not explain the large differences. Moreover, the bias correction is only a difference to the approach by Papahristodoulou (1997) since the order-m approach applied in Cantner et al. (2012) does not lead to biased results. To compare whether our results are due to a different model specification we also reestimated the efficiency using the specification in Cantner et al. (2012). The inefficiencies obtained in this analysis were even larger than those obtained in our model.[28] To explain the differences in the efficiency results we emphasize the significant larger number of observations included in our dataset. This follows for two reasons. Firstly, Cantner et al. (2012) focus on compact class cars and exclude all other types of cars. However, as we will show below focusing on compact class cars also reveals large inefficiencies. Moreover, the group efficiency of the compact class cars is very high. Therefore,

[27] Note that this also holds if we compare our results with those of less comparable studies like Oh et al. (2010).

[28] We also estimated our specification using the order-m method and also found larger inefficiencies. Moreover, Cantner et al. (2012) find that the inefficiency of cars in the German market is very stable over different time periods. Therefore, the different year for which the analysis was conducted can not explain the large differences.

the additionally covered car classes do not largely effect the results. Secondly, we only exclude dupletts that do not differ with respect to variables not included in the analysis. In Cantner et al. (2012) the cars are aggregated on the "car model" level. Hence, efficiency differences due to differences in model variants are not included in this analysis and the amount of observations in the compact class is about two thirds of our number. The large difference in the inefficiency shows that this aggregation of the data leads to an analysis that does not reveal all possibilities to increase the efficiency of the cars.

Table 3.3: Tests of first-order stochastic dominance (overall dataset)

H_0	p-value
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.999
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} \widehat{\rho}$	0.003
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.824
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} \widehat{\rho}$	0.000
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.647
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.000

The effect of including carbon dioxide emissions on the efficiency results for the automobiles can be analyzed by comparing the results for $\widehat{\rho}$ and $1 + \widehat{\beta}_{\text{O}}$. Table 3.3 shows the related p-values of the tests for stochastic dominance between the efficiency measures for the overall dataset.

We find that including the emissions lowers the inefficiency of the cars. This follows naturally from the estimation of the efficiency scores. Since the optimal value of the linear programming problem can not increase if an additional restriction is incorporated the results for $1 + \widehat{\beta}_{\text{O}}$ have to be lower or equal to the results for $\widehat{\rho}$. Using the test for stochastic dominance we find that the results for $\widehat{\rho}$ dominate stochastically those of $1 + \widehat{\beta}_{\text{O}}$ and therefore the inefficiency measured by the latter is lower. However, comparing the magnitude of the inefficiency we find that these significant differences are rather small. The inefficiency measured by $1 + \widehat{\beta}_{\text{O}}$ ranges from 1.35% to 101.29% and the average value being 22.4% which is approximately 1% less than the analysis excluding them. While the most efficient car changes if emissions are included (Audi A5 Coupe 2.0 TFSI) the most inefficient car remains the same (Nissan Patrol 3.0 Di SE).

To explain this relatively small difference note that we have incorporated the fuel consumption as an input in the car model. Moreover, we find that the fuel consumption and the carbon dioxide emissions are highly correlated (correlation coefficient: 0.95).[29] This follows because of the close technical relationship of fuel consumption and carbon dioxide emissions (see Schäfer and Basshuysen (1995)). The small difference in the efficiency results indicates that the factors beside the fuel consumption that influence carbon dioxide emissions (see e.g. the overview by Zervas and Diamantopoulos (2009)) that are incorporated in the analysis by adding the amount of produced emissions do not largely influence the results of the efficiency analysis. Hence, we can conclude that ignoring carbon dioxide leads to an overestimation of the inefficiency of automobiles but only to a small extent if the fuel consumption is included as a variable.

[29] Note that a high correlation among variables itself does not indicate whether the inclusion of a variable leads to significant changes in the efficiency results or not (see Nunamaker (1985)).

Measuring efficiency by simultaneously increasing the good outputs while emissions are reduced leads to results that are both statistically and economically significant lower than those obtained by the other measures. The reduced inefficiencies, ranging from 1.36% to 45% with an average of 14.49%, show that the cars are more inefficient with regard to the technical output characteristics than they are with regard to the production of emissions. Moreover, the reduced inefficiency can be also interpreted as reduced possibilities to increase technical features of the cars. Therefore, the analysis shows the trade-off between environmental and technical performance. However, the results may be driven by the correlation between fuel consumption and carbon dioxide emissions. For example, a car may be compared to a reference car that uses the same amount of all inputs except the fuel consumption and uses less fuel than the DMU under evaluation leading to a slack with regard to this input. Moreover, since it uses less fuel it produces less emissions. In this case the efficiency measure would indicate a reduction possibility for carbon dioxide emissions that is due to decreased fuel consumption. In the given analysis we would therefore indirectly measure fuel consumption in the efficiency analysis. To check wether the results are driven by this possibility we reestimated the results introducing an equality constraint for the fuel consumption input in the linear programming problem. The results of this analysis do not show economically significant differences from the results assuming strong disposability of inputs. Therefore, the results are not driven by the correlation of fuel consumption and carbon dioxide emissions.

The results in this section are based on a very large dataset including various types of automobiles. Other studies like Kortelainen and Kuosmanen (2007) and Cantner et al. (2012) focus solely on a particular group of cars (SUVs, compact class cars etc.). In the following section we will present the analysis of different groups of cars and we will compare their relative performance. The decomposition into managerial and program efficiency is of particular interest since studies that focus solely on one group eliminate the effect of program inefficiencies from their analysis.

3.3.4 Results for different groups of automobiles

We start our discussion with a comparison of the efficiency of different engine types (gasoline, diesel and natural gas). The following figure 3.6 presents the boxplots of the group results. To visualize whether the groups show different efficiency patterns than the overall dataset we have also included the boxplots for the analysis of the total sample in this figure. A boxplot can be read as follows. The box shows the interquartile range of the efficiency estimates with the lower end indicating the first quartile and the upper end indicating the third quartile. The bold horizontal line within the box shows the median value. The whiskers span to the most extreme observations that lie within 1.5 times the interquartile range. The remaining observations are indicated by a circle.

The descriptive statistics of the results for each engine and efficiency type can be found in table 3.4. Table A.2 in appendix A presents the p-values for the tests comparing the efficiency scores between different engine types. Moreover, the p-values for the tests of stochastic dominance comparing different efficiency measures for each engine type can be found in table A.3.

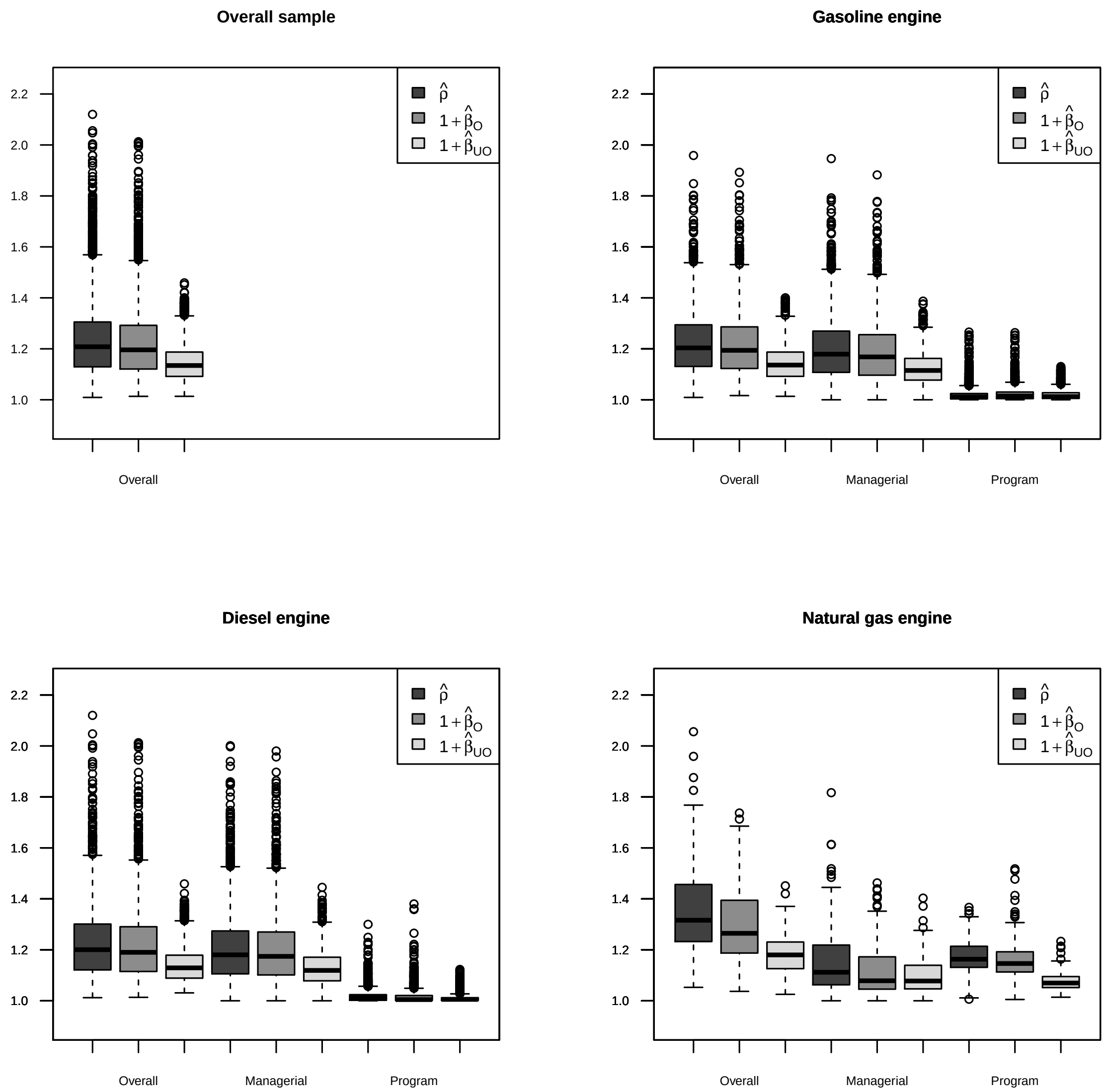

Figure 3.6: Boxplots of efficiency results (overall dataset and engine types)

The results for the overall efficiency of the gasoline cars are very similar to the results obtained for the overall dataset. With average inefficiencies of 22.6%, 21.7% and 14.5% for the three efficiency measures $\widehat{\rho}$, $1+\widehat{\beta}_{\mathrm{O}}$ and $1+\widehat{\beta}_{\mathrm{UO}}$ the difference to the total sample is within 1%. Comparing the median results the difference is even smaller. Moreover, we again find only small effects of the incorporation of emissions on the efficiency results while the accounting for the reduction of emissions has a larger effect.

The decomposition of the overall inefficiencies into car specific (managerial) and group specific (program) inefficiencies shows that the results are largely due to managerial inefficiencies for all three efficiency measures. This indicates that the group frontier of the gasoline cars is located close to the overall frontier. This result is of particular interest because car analyses that focus on a one-dimensional comparison of carbon dioxide efficiency with regard to a particular car characteristic (see e.g. Sullivan et al. (2004) with regard to the weight of the car) find that diesel

Table 3.4: Descriptive statistics of the efficiency measures (engine types)

Efficiency measure	Engine type	Efficiency type	Min.	1. Qu.	Median	Mean	3. Qu.	Max.	SD
$\widehat{\rho}$	Gasoline	Overall	1.009	1.131	1.204	1.226	1.294	1.959	0.130
		Managerial	1.000	1.108	1.179	1.201	1.270	1.946	0.126
		Program	1.000	1.004	1.010	1.021	1.025	1.266	0.030
	Diesel	Overall	1.012	1.121	1.200	1.233	1.301	2.120	0.156
		Managerial	1.000	1.106	1.180	1.210	1.274	2.001	0.148
		Program	1.000	1.002	1.012	1.019	1.024	1.300	0.027
	Natural gas	Overall	1.053	1.232	1.316	1.356	1.456	2.056	0.181
		Managerial	1.000	1.063	1.112	1.157	1.219	1.817	0.129
		Program	1.006	1.131	1.163	1.170	1.214	1.367	0.069
$1+\widehat{\beta}_{\mathrm{O}}$	Gasoline	Overall	1.017	1.123	1.194	1.217	1.286	1.893	0.128
		Managerial	1.000	1.097	1.168	1.189	1.256	1.882	0.123
		Program	1.000	1.005	1.015	1.024	1.030	1.263	0.030
	Diesel	Overall	1.013	1.115	1.190	1.224	1.291	2.013	0.153
		Managerial	1.000	1.101	1.174	1.204	1.270	1.980	0.146
		Program	1.000	1.002	1.005	1.017	1.021	1.380	0.031
	Natural gas	Overall	1.037	1.187	1.265	1.296	1.394	1.737	0.154
		Managerial	1.000	1.046	1.079	1.118	1.172	1.463	0.099
		Program	1.005	1.113	1.146	1.159	1.192	1.518	0.083
$1+\widehat{\beta}_{\mathrm{UO}}$	Gasoline	Overall	1.014	1.092	1.136	1.145	1.187	1.400	0.068
		Managerial	1.000	1.077	1.115	1.123	1.162	1.387	0.063
		Program	1.000	1.005	1.013	1.019	1.028	1.131	0.020
	Diesel	Overall	1.031	1.089	1.129	1.141	1.179	1.459	0.068
		Managerial	1.000	1.078	1.119	1.131	1.171	1.445	0.069
		Program	1.000	1.002	1.005	1.009	1.012	1.123	0.013
	Natural gas	Overall	1.025	1.126	1.180	1.182	1.231	1.451	0.076
		Managerial	1.000	1.047	1.078	1.099	1.139	1.403	0.072
		Program	1.014	1.052	1.070	1.076	1.095	1.233	0.036

cars are considerably more efficient than gasoline cars. In our analysis accounting for multiple inputs and outputs such a dominance would imply that the overall frontier is constructed solely by diesel cars and hence the program inefficiencies should be large for the gasoline cars. However, our analysis does not show large group inefficiencies for the gasoline cars. Therefore, the group of gasoline cars can not be regarded as a generally very less efficient group of cars even if the efficiency evaluation accounts for carbon dioxide emissions.

For the diesel cars we find results that are very similar to those of the gasoline cars and therefore to the overall dataset. Again, the boxplots show that the incorporation of emissions has no large effect. Indeed, in this case the test for stochastic dominance can not reject the hypothesis of equal distributions for the results for $\widehat{\rho}$ and $1 + \widehat{\beta}_{\mathrm{O}}$ for both the overall and the managerial efficiency. Moreover, we find that the overall inefficiencies are based on managerial inefficiencies while the program efficiency is quite high.

Comparing the group efficiency between the diesel and the gasoline cars (see the p-values comparing the program efficiency in table A.2) we find that gasoline cars dominate diesel cars for all three efficiency measures. Since domination implies larger inefficiency this indicates that the frontier of diesel cars is located closer to the overall frontier than the group frontier of the gasoline cars. However, while this difference is statistically significant the magnitude of the difference is very small ranging from 0.2% for the averages of $\widehat{\rho}$ to 1% for the averages of $1 + \widehat{\beta}_{\mathrm{UO}}$. Furthermore, as shown above the effect of the program inefficiency on the overall inefficiency is very small. Therefore, the statistical significant difference in the program efficiency does not lead to a significant difference in the overall efficiency for the measures $\widehat{\rho}$ and $1 + \widehat{\beta}_{\mathrm{O}}$. With regard to the results of $1 + \widehat{\beta}_{\mathrm{UO}}$ we find that the overall efficiency differs significantly between the groups. However, the magnitude of the difference (at the average 0.4%) is again very small. This result again shows that a general favorization of the diesel over the gasoline cars can not be supported in our analysis, neither if emissions are incorporated in the efficiency analysis nor if they are excluded.

In contrast to the very homogeneous results for the gasoline and diesel cars we find that cars with natural gas engines (NGEs) show very different results. Compared to the findings for the analysis of the total sample the overall inefficiency of the NGEs is far higher. The estimated average inefficiencies obtained by the three efficiency measures $\widehat{\rho}$, $1+\widehat{\beta}_{\mathrm{O}}$ and $1+\widehat{\beta}_{\mathrm{UO}}$ are 35.6%, 29.6% and 18.2%. However, we observe that the differences in the average results for the NGEs and the overall sample decline significantly in the chosen efficiency measure. With regard to $\widehat{\rho}$ the difference is approximately 12% while it lowers to approximately 7% using $1 + \widehat{\beta}_{\mathrm{O}}$ and to approximately 4% using $1 + \widehat{\beta}_{\mathrm{UO}}$. This indicates that accounting for the production and reduction of carbon dioxide enhances the relative performance of the NGEs compared to the total sample. In particular the increased efficiency if emissions are incorporated in the analysis is clearly visible from the boxplots for the overall efficiency. While the boxes for $\widehat{\rho}$ and $1+\widehat{\beta}_{\mathrm{O}}$ are nearly the same for the overall dataset as well as the gasoline and diesel cars, we observe that the box for $1+\widehat{\beta}_{\mathrm{O}}$ is located considerably lower than the box for $\widehat{\rho}$ for the NGEs. The difference in these measures is therefore both economically and statistically (see the p-values in table A.3) significant. Hence, combining these findings with the results for the gasoline and diesel cars shows that while the efficiency of cars with conventional engines accounting for the emissions of carbon dioxide can be estimated relatively precisely by incorporating the fuel consumption of the cars we observe that this is not the case for cars with natural gas engines.

Another difference to the cars with conventional engines can be found with regard to the sources of the inefficiencies. The decomposition shows that group specific inefficiencies contribute more to the overall inefficiencies than the individual specific inefficiencies for the measures $\widehat{\rho}$ and $1+\widehat{\beta}_{\mathrm{O}}$. In contrast, for the measure $1+\widehat{\beta}_{\mathrm{UO}}$ we find that the individual effects contribute slightly more than the group effects. This shows that the group inefficiency reduces if the reduction of emissions is incorporated in the analysis. But comparing the overall efficiency scores for the gasoline and diesel cars with those of the NGEs (see table A.2 for the p-values) we find that gasoline and diesel cars show a statistically and economically larger efficiency than the NGEs irrespective of which of the three efficiency measures is applied in the analysis. The source for this larger efficiency can be clearly identified as group specific aspects while the managerial efficiency of the NGEs is higher than that of the gasoline and diesel cars. Combining both findings indicates that NGEs are located more closely to their group frontier than diesel or gasoline cars are to theirs. In contrast, the group frontier of the NGEs is located further away from the overall frontier. This latter effect dominates and therefore the overall efficiency of the NGEs is lower than the overall efficiency of the cars with conventional engines. The differences in the group frontiers become obvious when comparing the minimal program inefficiencies for the different engine types. For both, diesel and gasoline cars, we find for all efficiency measures minimal scores of one which indicates that at least one car of the respective group is located on the overall frontier. In contrast, the minimal program efficiency score for the NGEs is for all efficiency measure larger than one indicating that no car of this group belongs to the overall frontier.

These findings again underline the importance of accounting for multiple inputs and outputs when analyzing the efficiency of automobiles. In contrast to other studies which do not incorporate these inputs and outputs (see e.g. MacLean and Lave (2000)) we find that natural gas cars show significantly larger inefficiencies. Including the reduction of emissions in the analysis enhances their relative performance, but they are still found to be less efficient than cars with conventional engines.

We now turn to the results for the car classes and the SUVs. The following figure 3.7 presents the boxplots of the efficiency results for both groups. The related summary statistics can be found in tables 3.5 and 3.6. Tables A.4, A.5 and A.6 in appendix A present the results of the tests for stochastic dominance.

The boxplots for the different classes of automobiles show that the patterns we observed for the conventional engine types can be found again when analyzing compact, middle and upper class cars. The overall efficiency results for all three classes are largely influenced by individual inefficiencies rather than by group specific inefficiencies. This result is of importance because it shows that analyzing a larger dataset than in the previous literature capturing various types of automobiles does not necessarily lead to results only driven by a special group of automobiles. Therefore, the restriction of analyses to relatively small samples seems questionable given that they fail to reveal all inefficiencies as shown above. The results also show that focusing solely

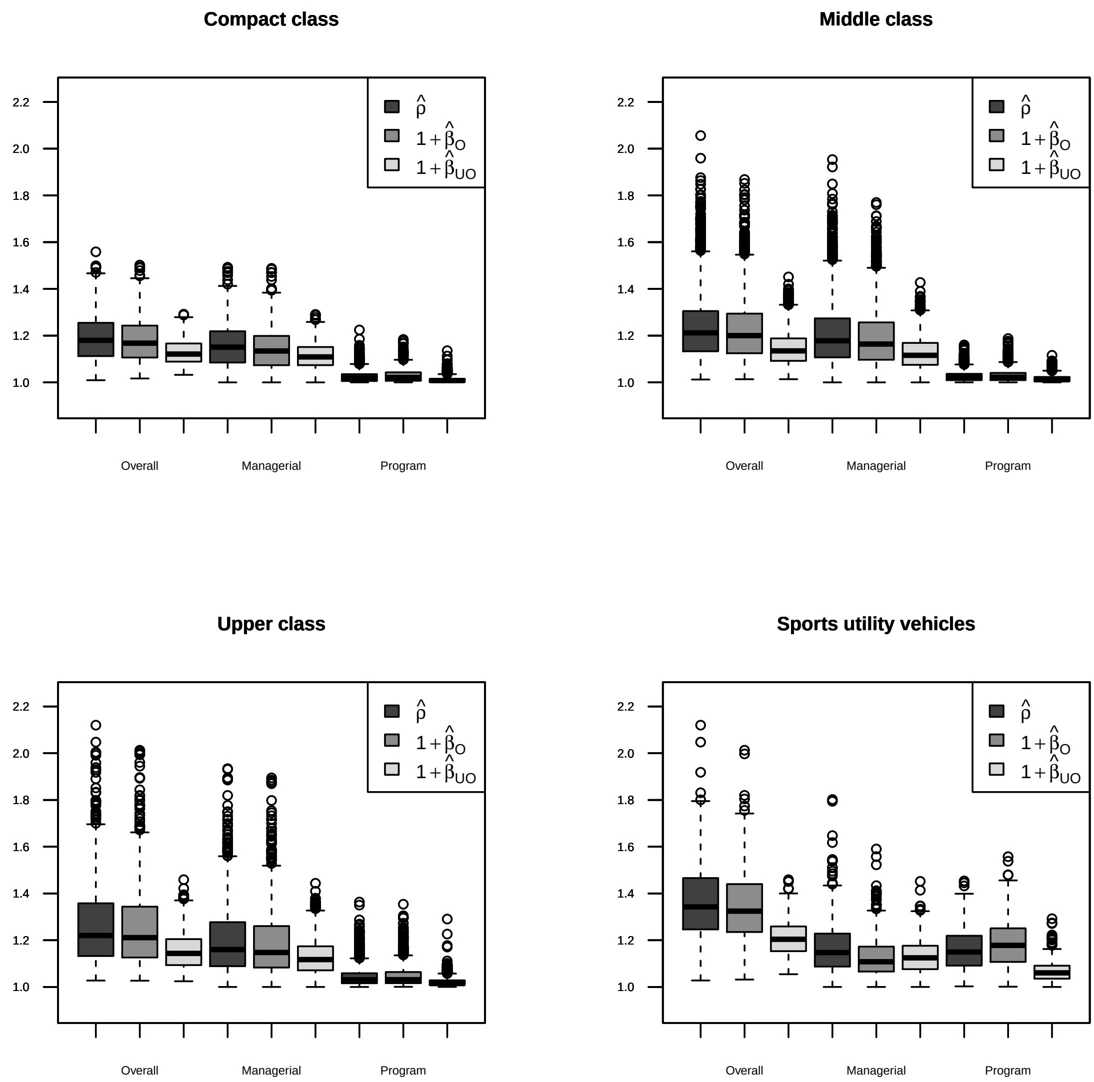

Figure 3.7: Boxplots of efficiency results (car classes and SUVs)

on one particular class of automobiles like it is done e.g. in Oh et al. (2010) and Cantner et al. (2012) leads to an underestimation of the inefficiency because the group inefficiencies are not captured. However, since we find only small effects of group inefficiencies we can conclude that the underestimation is not a severe problem.

Analyzing the results for the car classes we find again that the inclusion of carbon dioxide does only lead to small changes in the efficiency results. These changes are not significant for the upper class cars and only weakly significant for the compact class cars (see the results of the test for stochastic dominance in table A.4). But we again find a large decrease in the inefficiency if the reduction of emissions is included in the efficiency analysis indicating that the cars are more efficient with regard to the emission of CO_2 than with regard to the conventional outputs. This holds for all three car classes.

Comparing the results between the car classes (see the p-values in table A.6) we observe that

Table 3.5: Descriptive statistics of the efficiency measures (car classes)

Efficiency measure	Car class	Efficiency type	Min.	1. Qu.	Median	Mean	3. Qu.	Max.	SD
$\widehat{\rho}$	Compact	Overall	1.009	1.112	1.180	1.189	1.255	1.559	0.101
		Managerial	1.000	1.085	1.151	1.158	1.219	1.493	0.098
		Program	1.000	1.006	1.019	1.027	1.035	1.224	0.029
	Middle	Overall	1.012	1.133	1.212	1.235	1.305	2.056	0.140
		Managerial	1.000	1.107	1.178	1.204	1.274	1.953	0.135
		Program	1.000	1.010	1.021	1.026	1.037	1.160	0.023
	Upper	Overall	1.027	1.132	1.220	1.266	1.358	2.120	0.180
		Managerial	1.000	1.089	1.160	1.207	1.277	1.934	0.161
		Program	1.000	1.016	1.031	1.049	1.059	1.363	0.050
$1+\widehat{\beta}_{\mathrm{O}}$	Compact	Overall	1.017	1.106	1.168	1.180	1.243	1.502	0.096
		Managerial	1.000	1.074	1.134	1.145	1.199	1.488	0.094
		Program	1.000	1.007	1.021	1.030	1.043	1.184	0.031
	Middle	Overall	1.013	1.125	1.200	1.223	1.294	1.868	0.132
		Managerial	1.000	1.097	1.164	1.189	1.257	1.770	0.126
		Program	1.000	1.010	1.023	1.029	1.041	1.188	0.026
	Upper	Overall	1.027	1.126	1.211	1.256	1.343	2.013	0.175
		Managerial	1.000	1.083	1.148	1.196	1.260	1.895	0.155
		Program	1.001	1.016	1.031	1.051	1.064	1.354	0.052
$1+\widehat{\beta}_{\mathrm{UO}}$	Compact	Overall	1.032	1.089	1.121	1.129	1.166	1.292	0.052
		Managerial	1.000	1.074	1.109	1.115	1.151	1.291	0.055
		Program	1.000	1.002	1.008	1.013	1.015	1.136	0.017
	Middle	Overall	1.014	1.092	1.135	1.145	1.188	1.451	0.069
		Managerial	1.000	1.075	1.116	1.126	1.169	1.427	0.068
		Program	1.000	1.005	1.011	1.017	1.023	1.116	0.016
	Upper	Overall	1.024	1.093	1.144	1.156	1.205	1.459	0.078
		Managerial	1.000	1.071	1.117	1.130	1.174	1.444	0.077
		Program	1.000	1.009	1.015	1.023	1.028	1.291	0.024

with regard to the overall efficiency the compact class cars are more efficient than both middle and upper class cars for all three efficiency measures. The differences in the class efficiency are also economically significant (see the descriptive statistics in table 3.5). The differences to the middle class cars are largely due to differences in the managerial efficiency. Disparities in the group frontiers are not statistically significant for $\widehat{\rho}$ and significant but with a small magnitude for $1+\widehat{\beta}_{\mathrm{O}}$ and $1+\widehat{\beta}_{\mathrm{UO}}$ (see the descriptive statistics). Compared to upper class cars we find that compact class cars are statistically significantly more efficient for all efficiency measures and types. However, while the differences in group specific inefficiencies are larger than those found from a comparison of the compact cars to middle class cars the magnitude of the differences in managerial inefficiencies is still much larger. This result is quite surprising. Given that we do not account for the additional luxury equipment included in upper class cars we would have expected very large group inefficiencies because the price of the car is a variable in the analysis. Obviously, this is not the case in our analysis indicating that differences in the luxury equipment do only account for small efficiency differences if multiple technical characteristics are incorporated.

The last group to be analyzed are the sport utility vehicles (SUVs). The boxplots of the results can be found in figure 3.7 while the descriptive statistics are presented in table 3.6. The p-values of the test for stochastic dominance comparing the results for different efficiency measures can be found in table A.5 in the appendix.

Table 3.6: Descriptive statistics of the efficiency measures (SUVs)

Efficiency measure	Efficiency type	Min.	1. Qu.	Median	Mean	3. Qu.	Max.	SD
	Overall	1.028	1.247	1.342	1.362	1.466	2.120	0.170
$\widehat{\rho}$	Managerial	1.000	1.088	1.147	1.175	1.228	1.802	0.122
	Program	1.002	1.092	1.150	1.160	1.219	1.454	0.090
	Overall	1.032	1.235	1.324	1.341	1.440	2.013	0.160
$1+\widehat{\beta}_{\mathrm{O}}$	Managerial	1.000	1.066	1.108	1.133	1.172	1.590	0.095
	Program	1.001	1.108	1.178	1.184	1.251	1.558	0.102
	Overall	1.054	1.153	1.204	1.208	1.259	1.459	0.077
$1+\widehat{\beta}_{\mathrm{UO}}$	Managerial	1.000	1.076	1.125	1.132	1.176	1.452	0.075
	Program	1.000	1.035	1.060	1.068	1.091	1.292	0.047

Regarding the overall efficiency results we find that with average inefficiencies of 36.2%, 34.1% and 20.8% for the three measures SUVs are very inefficient compared to the total sample. Comparing the maximal inefficiencies for the overall sample (see table 3.2) and the SUVs we observe that the most inefficient car in the dataset belongs to the group of SUVs irrespective of which efficiency measure is used. These findings are interesting because they show that SUVs are very inefficient compared to other cars even if an analysis accounts for features of cars (e.g. engine power) that favor SUVs.

The decomposition of the overall efficiency (see the related boxplots in figure 3.7) ignoring carbon dioxide emissions shows that individual and group specific inefficiencies contribute nearly equally to the overall efficiency results. Therefore, the distance between the SUVs to their group frontier and the distance of the group frontier to the overall frontier are of comparable magnitude. This result changes if carbon dioxide emissions are incorporated in the analysis. From the tests of stochastic dominance (see table A.5) we find that there is no statistically significant effect of the inclusion of emissions on the overall efficiency. However, from the boxplots and the descriptive statistics we observe that the decomposition of the inefficiencies changes. In contrast to the analysis ignoring emissions, where the effects of managerial and program inefficiencies were equally large, the inclusion of emissions lowers the managerial inefficiency and increases the program inefficiency. This indicates that incorporating the emissions shifts the group frontier of the SUVs farther away from the overall frontier. Since this new group frontier lies closer to the SUV observations the managerial efficiency increases. However, since the relative position of the SUVs to the overall frontier remains unchanged the overall efficiency results are not affected by the incorporation of emissions. Comparing this result with the findings for other groups we observe again that including the fuel consumption leads to an efficiency measurement that also accounts for environmental aspects. But our results for the SUVs using the frontier separation approach indicate that the decomposition into group and individual effects may change significantly if emissions are additionally included in the analysis.

3.4 Summary

In this study we have presented a nonparametric analysis of automobiles sold in Germany in 2010. Focusing on the role of carbon dioxide emissions we found that the mere inclusion of the emissions does have a statistically significant yet not economically large effect on the estimated efficiency. This result is due to the simultaneous inclusion of the fuel consumption of cars which is highly correlated with the carbon dioxide emissions. This shows that existing studies on the efficiency of automobiles implicitly account for environmental aspects when incorporating the fuel consumption as a variable. Nonetheless, including the reduction of emissions in the efficiency analysis we still find significant reduction potentials even if the fuel consumption remains unchanged. In general our findings showed larger inefficiencies compared to previous studies which results because of the use of less aggregated data. These data were capable to reveal larger possibilities to enhance the performance of the cars. Our analysis of the performance of different groups of automobiles using a frontier separation approach showed that in an analysis incorporating multiple inputs and outputs some clear dominance relations obtained in previous studies (e.g. the significantly better performance of diesel compared to gasoline cars) can not be uphold. Only for SUVs and cars with natural gas engine we find economically and statistically significant group inefficiencies. While these findings strengthen previous results on the inefficiency of SUVs they also question previously conducted analyses that find large efficiency advantages of cars with natural gas engines. We want to emphasize, that this result does not automatically imply the bold conclusion that the technology of natural gas engines is inferior to standard engines but that it also could be due to the fact that the cars that are currently soled are in an early stage of the technological development and many potential efficiency improvements appear to be unexploited yet.

4 Separating environmental efficiency into production and abatement efficiency

4.1 Motivation

In the chapter on general concepts we have presented an environmental DEA technology ($\widehat{T}^{\mathrm{Env}}$) to estimate the efficiency of DMUs. In this technology environmental factors are modeled as weak disposable outputs. As we have noted there, the literature provides various other theoretical approaches to model emissions like the incorporation of the inverse of the emissions (see Lovell et al. (1995)) or the translation approach where the undesirable outputs are subtracted from a sufficient large positive number (see Seiford and Zhu (2002)).

These models have in common that they treat the DMUs as "black boxes" (see figure 4.1) which use inputs ($\boldsymbol{x}$) and produce desirable ($\boldsymbol{y}$) as well as undesirable outputs ($\boldsymbol{u}$) (e.g. power plants using coal and producing electricity and SO_2 emissions). Environmental efficiency is analyzed without taking into account that the DMUs produce desirable outputs and try to abate undesirable outputs in different stages, which is the basic idea behind classic end-of-pipe abatement technologies (e.g. scrubber technologies).[30] Moreover, these approaches have in common that they neither formulate an explicit production nor an abatement process of the undesirable outputs (see Førsund (2009) for critical remarks).

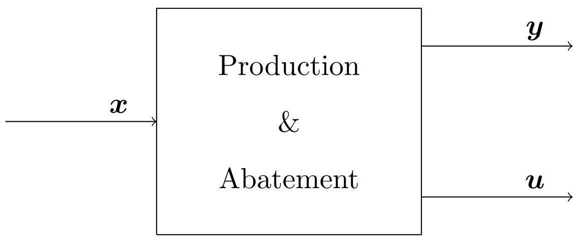

Figure 4.1: Environmental "black box"

As a result, little research has been conducted to reveal sources of possible inefficiencies with regard to the emission of pollutants. For instance a DMU might be inefficient because it uses too much of a polluting input in the production stage or the amount of emissions which are abated using an abatement technology is too low.

The existing literature which presents more detailed analyses of environmental efficiency also has some drawbacks. Hua and Bian (2008) propose a network based efficiency measure incorporating undesirable outputs but no production theoretical background of this measure is presented. Yang et al. (2008) propose an analysis with both a production and an abatement stage, but they do not separate the efficiencies of these stages. Coelli et al. (2007) suggest an approach where technical and environmental allocative efficiency is analyzed. However, abatement efficiency is only shortly noted and not included in their analysis. Färe et al. (forthcoming) and Murty et al. (2012) provide models which account for abatement activities in a network setting. As

[30] See Zotter (2004) for an applied analysis of an end-of-pipe technology.

discussed in more detail below the model by Färe et al. (forthcoming) is very inflexible in the way efficiency is measured while Murty et al. (2012) rely on implausible assumptions regarding the abatement technology.

In the following study we propose a new model to evaluate the environmental efficiency of DMUs and to separate it into production and abatement efficiency. In contrast to the existing literature, we explicitly formulate a production theoretical model of production and pollution abatement activities. Furthermore, an environmental efficiency measure is proposed which can be decomposed into the effects of production and abatement inefficiencies on environmental efficiency and we show how nonparametric methods can be used to estimate this measure and its components. To show the empirical applicability of our new approach, we analyze the environmental efficiency of U.S. coal-fired power plants in the year 2009 with regard to sulfur dioxide emissions.

4.2 Modeling the production and abatement technologies

In our model we assume that the production process of a DMU can be divided into two stages, the production stage with technology T_1 and the abatement stage with technology T_2. In the first stage, the DMUs use inputs $\boldsymbol{x}_1^F$, $\boldsymbol{x}^P$ and $\boldsymbol{x}_1^S$ to produce outputs $\boldsymbol{y}$. $\boldsymbol{x}_1^F$ denotes pollution free (or non-polluting) inputs, which means that the use of these inputs does not lead to any pollution (e.g. labor), while pollution containing (or polluting) inputs are denoted by $\boldsymbol{x}^P$ (e.g. coal).[31] $\boldsymbol{x}_1^S$ denotes the amount of shared inputs $\boldsymbol{x}^S$ used in the production stage. In contrast to the other inputs the shared inputs are used by both stages. Hence, it is possible to reallocate these inputs between the two stages to increase efficiency. Our approach to shared inputs follows Färe et al. (1997) but without the assumption that the overall amount of shared inputs is fixed. Alternative approaches to shared inputs like Cook et al. (2000) could be also included in the analysis. The desirable outputs of the DMUs consist of final outputs $\boldsymbol{y}^f$ and intermediate inputs $\boldsymbol{y}^2$ (see Färe and Grosskopf (1996b)) which are inputs of the abatement stage. The use of the pollution containing inputs $\boldsymbol{x}^P$ to produce outputs $\boldsymbol{y} = \boldsymbol{y}^f + \boldsymbol{y}^2$ leads to a production of undesirable outputs $\boldsymbol{u}'$ (e.g. carbon dioxide emissions).[32] These undesirable outputs are inputs of the abatement process (e.g. scrubbers) where they are reduced to the final amount of undesirable outputs $\boldsymbol{u}''$ which are emitted to the environment by using non-polluting inputs $\boldsymbol{x}_2^F$, the amount $\boldsymbol{x}_2^S$ of the shared inputs and the intermediate inputs $\boldsymbol{y}^2$. The structure of this production process is depicted in figure 4.2.

[31] Since we assume that polluting inputs are only used in the first stage, they do not have subscripts.

[32] Note that we assume that neither $\boldsymbol{y}^2$ nor $\boldsymbol{x}^S$ contain any pollution, so that no additional pollution can be created at the abatement stage.

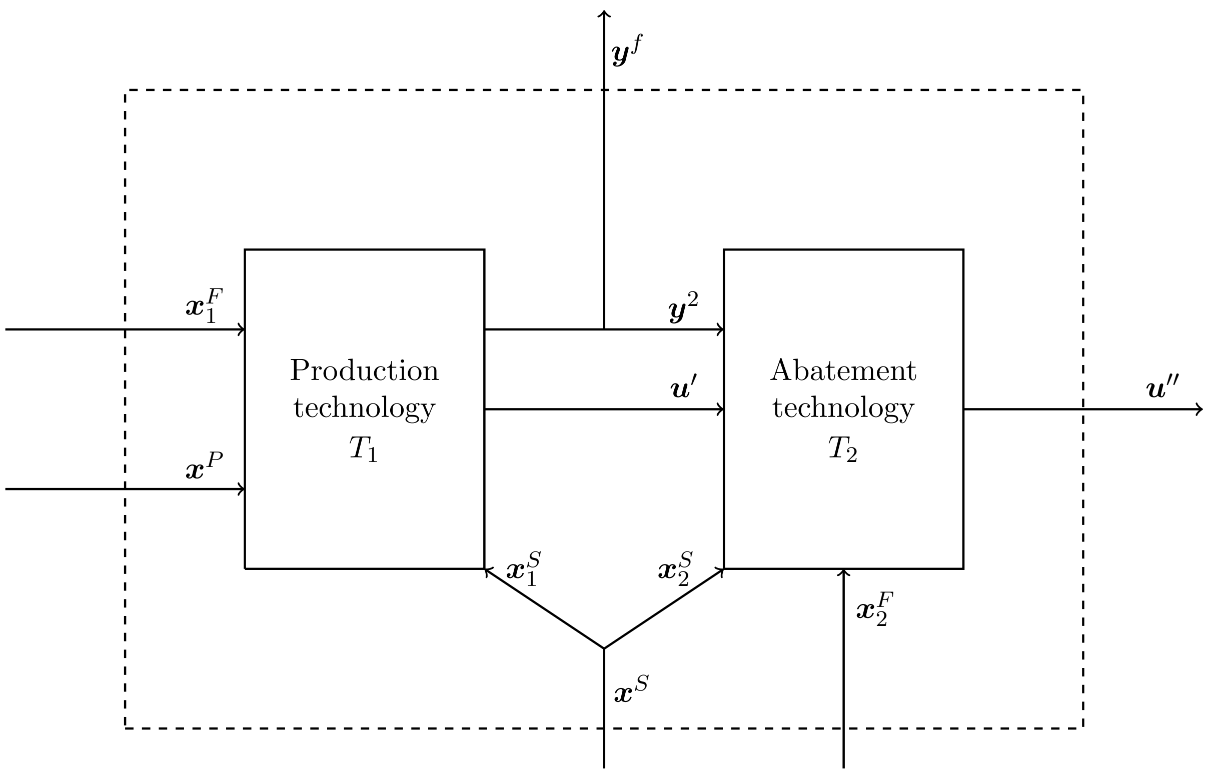

Figure 4.2: Structure of the two-stage production process

To formally define the technology consider n DMUs that are using m inputs $\boldsymbol{x} \in \mathbb{R}_+^m$ which can be split into m^F pollution free, m^P pollution containing and m^S shared inputs to produce s desirable outputs $\boldsymbol{y} \in \mathbb{R}_+^s$. m_1^F non-polluting inputs are used in the production stage and m_2^F are used in the abatement stage. As a result of the use of polluting inputs to produce $\boldsymbol{y}$ in the first stage, r undesirable outputs $\boldsymbol{u}' \in \mathbb{R}_+^r$ are produced. They are reduced to $\boldsymbol{u}'' \in \mathbb{R}_+^r$ in the abatement process. We further define $\boldsymbol{l} = \boldsymbol{u}' - \boldsymbol{u}'' \in \mathbb{R}_+^r$ as the amount of abated pollution. Given the definitions above the first stage of the overall technology, the production technology T_1, can be defined as

$$T_1 = \left\{ \left(\boldsymbol{x}_1^F, \boldsymbol{x}^P, \boldsymbol{x}_1^S, \boldsymbol{y}, \boldsymbol{u}'\right) \in \mathbb{R}_+^{m_1^F + m^P + m^S + s + r} : \left(\boldsymbol{x}_1^F, \boldsymbol{x}^P, \boldsymbol{x}_1^S\right) \text{ can produce } \left(\boldsymbol{y}, \boldsymbol{u}'\right) \right\}. \quad (4.1)$$

In our model we treat the pollutants $\boldsymbol{u}'$ as the residuals of the production stage.[33] A practical problem arises from the fact that in general $\boldsymbol{u}'$ is not observable for the researcher. To overcome this problem we assume that no abatement activities are conducted in the production stage. Therefore, the amount of undesirable outputs which are the inputs of the abatement stage can be derived from the materials balance condition (MBC).[34] This concept, introduced by Ayers and Kneese (1969), can be simplified as “what goes in must come out”. It is based on fundamental physical laws and states that the amount of materials bound in the inputs must be equal to those that are bound in the desirable and undesirable outputs.[35] Therefore, we can estimate

[33] See Pethig (2006) for microeconomic foundations of the residual generation in production processes.
[34] Of course, the MBC also holds if abatement activities are introduced in the production stage. But in this case the amount of abatement would have to be considered in the MBC resulting in an equivalent data problem.
[35] See Lauwers (2009) for a discussion and Coelli et al. (2007) for an application of the MBC in nonparametric

the amount of pollutants resulting from the first stage by the equality

$$\boldsymbol{u}' = \boldsymbol{\Pi}\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}^f \tag{4.2}$$

where $\boldsymbol{\Pi}$ is a $r \times m^P$ matrix of factors that indicate the amount of undesirable outputs bound in the polluting inputs and $\boldsymbol{\Psi}$ is a $r \times s$ matrix of factors that indicate the amount of pollutants bound in the final outputs.[36] Note that the matrices $\boldsymbol{\Pi}$ and $\boldsymbol{\Psi}$ can differ among the DMUs since the inputs and outputs might not be completely homogeneous e.g. the quality of coal and hence the sulfur content of it may differ among power plants.

Since we assume that the amount of $\boldsymbol{u}'$ follows this equality as a residual of the production we can split the production technology in two parts.[37] The production of the desirable outputs $\boldsymbol{y}$

$$T_1^{\boldsymbol{y}} = \left\{ \left(\boldsymbol{x}_1^F, \boldsymbol{x}^P, \boldsymbol{x}_1^S, \boldsymbol{y}\right) \in \mathbb{R}_+^{m_1^F + m^P + m^S + s} : \left(\boldsymbol{x}_1^F, \boldsymbol{x}^P, \boldsymbol{x}_1^S\right) \text{ can produce } \boldsymbol{y} \right\} \tag{4.3}$$

and the resulting pollution according to the MBC

$$T_1^{\boldsymbol{u}'} = \left\{ \left(\boldsymbol{x}^P, \boldsymbol{y}^f, \boldsymbol{u}'\right) \in \mathbb{R}_+^{m^P + s + r} : \boldsymbol{u}' = \boldsymbol{\Pi}\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}^f \right\}. \tag{4.4}$$

We assume that the technology $T_1^{\boldsymbol{y}}$ satisfies the standard axioms for nonparametric technology sets (e.g. strong disposability of in- and outputs, convexity) as proposed by Shephard (1970) (see the discussion in the chapter "General concepts").

Given the observed combinations $\left(\boldsymbol{x}_{1i}^F, \boldsymbol{x}_i^P, \boldsymbol{x}_{1i}^S, \boldsymbol{y}_i^f, \boldsymbol{y}_i^2\right)$ for $i = 1, \dots, n$ the DEA estimation of the technology $T_1^{\boldsymbol{y}}$ reads as

$$\begin{aligned} \widehat{T}_1^{\boldsymbol{y}} = \Big\{ \left(\boldsymbol{x}_1^F, \boldsymbol{x}^P, \boldsymbol{x}_1^S, \boldsymbol{y}\right) \in \mathbb{R}_+^{m_1^F + m^P + m^S + s} : \; & \boldsymbol{x}_1^F \geqq \boldsymbol{X}_1^F \boldsymbol{\lambda}, \boldsymbol{x}^P \geqq \boldsymbol{X}^P \boldsymbol{\lambda}, \boldsymbol{x}_1^S \geqq \boldsymbol{X}_1^S \boldsymbol{\lambda}, \\ & \boldsymbol{y}^f + \boldsymbol{y}^2 \leqq \left(\boldsymbol{Y}^f + \boldsymbol{Y}^2\right) \boldsymbol{\lambda}, \boldsymbol{\lambda} \geqq 0 \Big\} \end{aligned} \tag{4.5}$$

where $\boldsymbol{X}_1^F$ represents the $m_1^F \times n$ matrix of non-polluting inputs, $\boldsymbol{X}^P$ represents the $m^P \times n$ matrix of polluting inputs, $\boldsymbol{X}_1^S$ denotes the $m^S \times n$ matrix of the amount of shared inputs in the production stage. The outputs consist of a $s \times n$ matrix $\boldsymbol{Y}^f$ of final outputs and a $s \times n$ matrix $\boldsymbol{Y}^2$ of intermediate inputs. Note that only for notational convenience the matrices $\boldsymbol{Y}^f$ and $\boldsymbol{Y}^2$ have the same dimensions. If not all final outputs are also intermediate inputs, the matrix $\boldsymbol{Y}^2$ contains rows of zeros for these outputs. $\boldsymbol{\lambda}$ denotes a $n \times 1$ vector of weight factors, with $\boldsymbol{\lambda} \geqq 0$ indicating constant returns to scale. However, the technology can also be modified to exhibit variable returns to scale by adding the additional restriction $\boldsymbol{1}^T\boldsymbol{\lambda} = 1$.

In the second stage an abatement technology is used to reduce the undesirable outputs which are residuals of the use of polluting inputs and the production of final outputs according to $T_1^{\boldsymbol{u}'}$.

efficiency analysis.

[36] By assumption, the intermediate inputs do not contain any pollution.

[37] For a detailed discussion of the splitting of a technology into the production of good output and the residual generation see Murty et al. (2012).

The abatement technology T_2 can be defined as

$$T_2 = \left\{ \left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}', \boldsymbol{l}\right) \in \mathbb{R}_+^{m_2^F + m^S + s + 2r} : \left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}'\right) \text{ can produce } \boldsymbol{l} \right\} \tag{4.6}$$

where $\boldsymbol{l} = \boldsymbol{u}' - \boldsymbol{u}''$ and hence has the same dimension as the r undesirable outputs. We use $\boldsymbol{l}$ as the output of the abatement stage since in contrast to $\boldsymbol{u}''$ it is a desirable output (see Coelli et al. (2007, p. 9)). Like the production technology this technology is assumed to satisfy the Shephard (1970) axioms. In particular:

- Strong disposability of inputs:
  If $\left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}', \boldsymbol{l}\right) \in T_2$ and $\left(\tilde{\boldsymbol{x}}_2^F, \tilde{\boldsymbol{x}}_2^S, \tilde{\boldsymbol{y}}^2, \tilde{\boldsymbol{u}}'\right) \geqq \left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}'\right)$ then $\left(\tilde{\boldsymbol{x}}_2^F, \tilde{\boldsymbol{x}}_2^S, \tilde{\boldsymbol{y}}^2, \tilde{\boldsymbol{u}}', \boldsymbol{l}\right) \in T_2$.
  Strong disposability of $\boldsymbol{u}'$ means that an increase of the emissions that are an input to the abatement stage results in an equal increase in $\boldsymbol{u}''$, such that none of the additional emissions are abated.

- Strong disposability of outputs:
  If $\left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}', \boldsymbol{l}\right) \in T_2$ and $\tilde{\boldsymbol{l}} \leqq \boldsymbol{l}$ then $\left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}', \tilde{\boldsymbol{l}}\right) \in T_2$.
  It is possible to increase the amount of emissions $\boldsymbol{u}''$ until they are equal to $\boldsymbol{u}'$. This boundary follows from the non-negativity of $\boldsymbol{l}$.

These standard assumptions are explicitly mentioned because they allow to differentiate our model from existing approaches which rely on non-standard assumptions to model abatement processes and will be discussed below.

The DEA estimation of this technology is created using observations of $\left(\boldsymbol{x}_{2i}^F, \boldsymbol{x}_{2i}^S, \boldsymbol{y}_i^2, \boldsymbol{u}_i''\right)$ and estimations of $\boldsymbol{u}_i'$ and $\boldsymbol{l}_i$ by the MBC for $i = 1, \dots, n$ and reads as

$$\begin{aligned} \widehat{T}_2 = \Big\{ \left(\boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{y}^2, \boldsymbol{u}', \boldsymbol{l}\right) \in \mathbb{R}_+^{m_2^F + m^S + s + 2r} : &\ \boldsymbol{x}_2^F \geqq \boldsymbol{X}_2^F \boldsymbol{z}, \boldsymbol{x}_2^S \geqq \boldsymbol{X}_2^S \boldsymbol{z}, \boldsymbol{y}^2 \geqq \boldsymbol{Y}^2 \boldsymbol{z}, \\ &\ \boldsymbol{u}' \geqq \boldsymbol{U}' \boldsymbol{z}, \boldsymbol{l} \leqq \boldsymbol{L} \boldsymbol{z}, \boldsymbol{z} \geqq \boldsymbol{0} \Big\} \end{aligned} \tag{4.7}$$

where $\boldsymbol{X}_2^F$ denotes the $m_2^F \times n$ matrix of non-polluting inputs, $\boldsymbol{X}_2^S$ represents the $m^S \times n$ matrix of shared inputs used in the abatement stage, $\boldsymbol{U}'$ denotes the $r \times n$ matrix of undesirable outputs and $\boldsymbol{L}$ represents the $r \times n$ matrix of abated undesirable outputs, hence $\boldsymbol{L} = \boldsymbol{U}' - \boldsymbol{U}''$. $\boldsymbol{z}$ denotes the $n \times 1$ vector of weight factors. Note that these weight factors do not have to equal the $\boldsymbol{\lambda}$-values of the first stage. Hence, the reference observations may differ. Again, the technology may be modified to exhibit variable returns to scale by adding the constraint $\boldsymbol{1}^T \boldsymbol{z} = 1$.

The overall technology of the two-stage production process T_N is constructed by combining all three subtechnologies $\left(T_1^{\boldsymbol{y}}, T_1^{\boldsymbol{u}'} \text{and } T_2\right)$ to one network technology and can be estimated as

$$
\begin{aligned}
\widehat{T}_N = \Big\{ \left(\boldsymbol{x}^P, \boldsymbol{x}_1^F, \boldsymbol{x}_1^S, \boldsymbol{x}_2^F, \boldsymbol{x}_2^S, \boldsymbol{x}^S, \boldsymbol{y}^f, \boldsymbol{y}^2, \boldsymbol{u}', \boldsymbol{u}''\right) \in \mathbb{R}^{m^P+m_1^F+m^S+m_2^F+2m^S+2s+2r} :& \\
\boldsymbol{x}^P &\geqq \boldsymbol{X}^P\boldsymbol{\lambda} \\
\boldsymbol{x}_1^F &\geqq \boldsymbol{X}_1^F\boldsymbol{\lambda} \\
\boldsymbol{x}_1^S &\geqq \boldsymbol{X}_1^S\boldsymbol{\lambda} \\
\boldsymbol{y}^f + \boldsymbol{y}^2 &\leqq \left(\boldsymbol{Y}^f + \boldsymbol{Y}^2\right)\boldsymbol{\lambda} \\
\boldsymbol{\lambda} &\geqq \boldsymbol{0} \\
\boldsymbol{u}' &= \boldsymbol{\Pi}\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}^f \\
\boldsymbol{y}^2 &\geqq \boldsymbol{Y}^2\boldsymbol{z} \\
\boldsymbol{x}_2^F &\geqq \boldsymbol{X}_2^F\boldsymbol{z} \\
\boldsymbol{x}_2^S &\geqq \boldsymbol{X}_2^S\boldsymbol{z} \\
\boldsymbol{u}' &\geqq \boldsymbol{U}'\boldsymbol{z} \\
\boldsymbol{u}' - \boldsymbol{u}'' &\leqq \left(\boldsymbol{U}' - \boldsymbol{U}''\right)\boldsymbol{z} \\
\boldsymbol{z} &\geqq \boldsymbol{0} \\
\boldsymbol{x}_1^S + \boldsymbol{x}_2^S &\leqq \boldsymbol{x}^S \quad \Big\}.
\end{aligned}
\tag{4.8}
$$

In addition to the three subtechnologies the last inequality is included which states that the sum of shared inputs used in both stages can not exceed an exogenous total amount of shared inputs. Our technology is similar to the one presented in Färe and Grosskopf (1996a) which also contains shared and intermediate inputs but is not constructed for the analysis of environmental efficiency.

4.3 Measuring and separating environmental efficiency

4.3.1 Analysis under constant returns to scale

In this section we present a new possibility to evaluate the environmental efficiency of the DMUs given the network technology defined in the last section and to decompose it into production and abatement efficiency. Moreover, we show how the stage and the network efficiency can be separated providing more detailed insights in the measurement of environmental efficiency. As described above, we assume that the production process has a two-stage structure with the production of desirable outputs in the first stage and the reduction of the undesirable outputs, which are the residuals of the output production, in the second stage.

In the literature of environmental economics different measures for environmental efficiency have been proposed (see e.g. Tyteca (1996) for an overview). For a nonparametric analysis of environmental efficiency incorporating undesirable outputs as weak disposable outputs Färe et al. (2004) have developed an index that is based on the ratio of good to bad outputs. An index which is defined as the value added over the weighted amount of emissions has been proposed by Kuosmanen and Kortelainen (2005). In their model the weights are endogenously determined by using DEA in the multiplier form. For a network application Murty et al. (2012) have developed an efficiency measure that is based on the Russell measure of technical efficiency (see Färe and Lovell (1978)). However, with exception of the measure proposed by Murty et al.

(2012) these measures are not able to separate production from abatement inefficiencies when evaluating environmental efficiency. Moreover, the measure by Murty et al. (2012) relies on very restrictive assumption with regard to the structure of the technologies to be able to differentiate production and abatement efficiency.[38] Therefore, we propose a new environmental efficiency measure (EEM) which is defined as

$$\frac{\boldsymbol{\omega}^T \boldsymbol{u}''^*}{\boldsymbol{\omega}^T \boldsymbol{u}''} = \underbrace{\frac{\boldsymbol{\omega}^T \boldsymbol{u}'^*}{\boldsymbol{\omega}^T \boldsymbol{u}'}}_{\text{PE}} \cdot \underbrace{\frac{\boldsymbol{\omega}^T \boldsymbol{u}''^*}{\boldsymbol{\omega}^T \boldsymbol{u}'^*} \cdot \left[\frac{\boldsymbol{\omega}^T \boldsymbol{u}''}{\boldsymbol{\omega}^T \boldsymbol{u}'}\right]^{-1}}_{\text{AE}}. \tag{4.9}$$

This measure is given by the ratio of the weighted minimal amount of emissions $\left(\boldsymbol{\omega}^T \boldsymbol{u}''^*\right)$ released to the environment to the equally weighted actual observed amount of emissions $\left(\boldsymbol{\omega}^T \boldsymbol{u}''\right)$ of the DMU with a value less than one indicating environmental inefficiency. $\boldsymbol{\omega}^T$ denotes the transpose of the $r \times 1$ vector of aggregation weights for the emissions.[39] In case of a single pollutant ω can be set to one.

The environmental efficiency measure can be decomposed into the product of the efficiencies of two sources. The first term $\left(\boldsymbol{\omega}^T \boldsymbol{u}'^* / \boldsymbol{\omega}^T \boldsymbol{u}'\right)$ captures the effect of production efficiency (PE) on the environmental efficiency. It measures how much the weighted optimal amount of produced emissions $\boldsymbol{\omega}^T \boldsymbol{u}'^*$ differs from the weighted actual amount of produced emissions $\boldsymbol{\omega}^T \boldsymbol{u}'$. The second term measures the effect of abatement efficiency (AE) on environmental efficiency and is the quotient of two ratios: the first measuring the ratio of weighted minimal emissions released to the environment $\left(\boldsymbol{\omega}^T \boldsymbol{u}''^*\right)$ to the weighted optimal amount of produced emissions $\left(\boldsymbol{\omega}^T \boldsymbol{u}'^*\right)$ while the second measures the initial ratio of weighted emission released to the environment $\left(\boldsymbol{\omega}^T \boldsymbol{u}''\right)$ to the weighted produced emissions $\left(\boldsymbol{\omega}^T \boldsymbol{u}'\right)$. For an intuition of this measure assume that only one pollutant exists. In this case the denominator measures how much of a pollutant is emitted to the environment per unit of produced pollutant. If the abatement process is operated efficiently then the same rate of emissions should be emitted even if the amount of pollution produced in the production stage decreases. Hence, if u' is decreased to u'^* and the rate of pollution released to the environment is kept constant, then u'' is reduced to u''^* and the measure of abatement efficiency equals one. If there are no abatement activities present the environmental efficiency measure equals the measure of the production efficiency.

Both parts of the environmental efficiency measure can be further decomposed to differentiate stage from network inefficiencies. The measure for production efficiency can be decomposed into

$$\frac{\boldsymbol{\omega}^T \boldsymbol{u}'^*}{\boldsymbol{\omega}^T \boldsymbol{u}'} = \frac{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{ Prod}}}{\boldsymbol{\omega}^T \boldsymbol{u}'} \cdot \frac{\boldsymbol{\omega}^T \boldsymbol{u}'^*}{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{ Prod}}} \tag{4.10}$$

where $\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Prod}}$ indicates the amount of weighted emissions if the efficiency of the production stage is evaluated using only technology T_1, hence ignoring the network structure. $\boldsymbol{\omega}^T \boldsymbol{u}'^* / \boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Prod}}$ represents the effect of the network analysis. This second term is an environmental equivalent to the "black box" bias measure which Fukuyama and Matousek (2011) have proposed for a

[38] This point will be presented in more detail when our new model is compared to this approach.
[39] We will discuss the role of these weights in more detail below.

network analysis of cost efficiency.[40]

This measure is important for our analysis because it allows to explore several aspects of environmental efficiency which only result in the network approach. For example, assume that the efficiency of the productions stage is measured input-oriented and the two stages are connected by $\boldsymbol{u}'$ and $\boldsymbol{x}^S$. Even if the stage analysis indicates that the production stage is efficient this may not hold in the network analysis. This follows from the possibility to reallocate of $\boldsymbol{x}^S$. Given substitution possibilities of $\boldsymbol{x}^P$ and $\boldsymbol{x}^S$ it may be environmentally efficient to increase $\boldsymbol{x}^P$ (and therefore $\boldsymbol{u}'$) and decrease $\boldsymbol{x}_1^S$ to reallocate the shared input to the abatement stage where it is used to decrease the final emissions. Hence, while the stage efficiency measure always takes values less or equal to one the "black box" bias measure can also exhibit values larger than one. Whether the measure of production efficiency in the network $\boldsymbol{\omega}^T\boldsymbol{u}'^*/\boldsymbol{\omega}^T\boldsymbol{u}'$ takes a value less or larger than one depends on which of the two effects dominates. In the discussion of variable returns to scale the measure of the bias becomes even more important because it allows to analyze situations in which an inefficient production stage may be environmentally efficient.

Similar to the measure for the production efficiency the measure for abatement efficiency also can be decomposed into a stage efficiency and a network efficiency component:

$$\frac{\boldsymbol{\omega}^T\boldsymbol{u}''^*}{\boldsymbol{\omega}^T\boldsymbol{u}'^*}\cdot\left[\frac{\boldsymbol{\omega}^T\boldsymbol{u}''}{\boldsymbol{\omega}^T\boldsymbol{u}'}\right]^{-1}=\frac{\boldsymbol{\omega}^T\boldsymbol{u}''^{\text{Abat}}}{\boldsymbol{\omega}^T\boldsymbol{u}'}\cdot\left[\frac{\boldsymbol{\omega}^T\boldsymbol{u}''}{\boldsymbol{\omega}^T\boldsymbol{u}'}\right]^{-1}\cdot\frac{\boldsymbol{\omega}^T\boldsymbol{u}''^*}{\boldsymbol{\omega}^T\boldsymbol{u}'^*}\cdot\left[\frac{\boldsymbol{\omega}^T\boldsymbol{u}''^{\text{Abat}}}{\boldsymbol{\omega}^T\boldsymbol{u}'}\right]^{-1} \tag{4.11}$$

where $\boldsymbol{\omega}^T\boldsymbol{u}''^{\text{Abat}}$ is obtained by analyzing the abatement efficiency using only technology T_2 and therefore ignoring the network structure. As in the case of the production stage the network measure may be larger than one. This would indicate that in a network analysis it is efficient to release more emissions to the environment per unit of emissions generated at the production stage compared to the minimal emission rate of an efficient abatement stage. This follows because the reduction of $\boldsymbol{\omega}^T\boldsymbol{u}'$ through reallocating the shared inputs to the production stage can overcompensate this effect.

In our definition of the environmental efficiency measure we have not specified an orientation of the measurement beside the minimization of $\boldsymbol{u}''$. Therefore, the production stage can be measured either in input or output orientation. Thus, the environmental efficiency of DMUs is evaluated given a fixed amount of outputs or inputs. To show how our model is connected with the previous literature on environmental efficiency we will focus our discussion on an input-oriented analysis of the production stage. The direction of the measurement of the abatement efficiency is predetermined by the minimization of $\boldsymbol{u}''$ which is a part of the output of that stage.

Assuming an input-oriented evaluation of the efficiency of the production stage the environmental efficiency measure can be obtained by solving the following linear programming problem:

[40] Kao (2009) also presents models to separate efficiency in network models using the multiplier form of DEA. However, the models rely on measuring the efficiency of all stages in the same orientation which is not the case in our model.

$$
\begin{array}{rrcl}
\min\limits_{\boldsymbol{x}^P,\boldsymbol{x}_1^F,\boldsymbol{x}_1^S,\boldsymbol{x}_2^S,\boldsymbol{u}',\boldsymbol{u}'',\boldsymbol{\lambda},\boldsymbol{z}} & & & \boldsymbol{\omega}^T\boldsymbol{u}'' \\
\text{s.t.} & \boldsymbol{x}^P & \geqq & \boldsymbol{X}^P\boldsymbol{\lambda} \\
& \boldsymbol{x}_1^F & \geqq & \boldsymbol{X}_1^F\boldsymbol{\lambda} \\
& \boldsymbol{x}_1^S & \geqq & \boldsymbol{X}_1^S\boldsymbol{\lambda} \\
& \boldsymbol{y}_i^f+\boldsymbol{y}_i^2 & \leqq & \left(\boldsymbol{Y}^f+\boldsymbol{Y}^2\right)\boldsymbol{\lambda} \\
& \boldsymbol{\lambda} & \geqq & \mathbf{0} \\
& \boldsymbol{u}' & = & \boldsymbol{\Pi}\boldsymbol{x}^P-\boldsymbol{\Psi}\boldsymbol{y}_i^f \\
& \boldsymbol{y}_i^2 & \geqq & \boldsymbol{Y}^2\boldsymbol{z} \\
& \boldsymbol{x}_{2i}^F & \geqq & \boldsymbol{X}_2^F\boldsymbol{z} \\
& \boldsymbol{x}_2^S & \geqq & \boldsymbol{X}_2^S\boldsymbol{z} \\
& \boldsymbol{u}' & \geqq & \boldsymbol{U}'\boldsymbol{z} \\
& \boldsymbol{u}'-\boldsymbol{u}'' & \leqq & \left(\boldsymbol{U}'-\boldsymbol{U}''\right)\boldsymbol{z} \\
& \boldsymbol{u}'-\boldsymbol{u}'' & \geqq & \mathbf{0} \\
& \boldsymbol{u}'' & \geqq & \mathbf{0} \\
& \boldsymbol{z} & \geqq & \mathbf{0} \\
& \boldsymbol{x}_1^S+\boldsymbol{x}_2^S & \leqq & \boldsymbol{x}_i^S.
\end{array}
\tag{4.12}
$$

In this program the weighted emissions released to the environment are minimized over the inputs of the production stage.[41] Since the pollution containing inputs are linked to $\boldsymbol{u}'$ by the materials balance condition the optimization also includes $\boldsymbol{u}'$. Furthermore, the shared inputs allowing for a reallocation between the two stages and the intensity variables $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ are included in the optimization. Since we measure the production efficiency input-oriented the intermediate inputs $\boldsymbol{y}^2$ remain unchanged. In addition to the network technology we have included two constraints ($\boldsymbol{u}'-\boldsymbol{u}''\geqq\mathbf{0}$ and $\boldsymbol{u}''\geqq\mathbf{0}$) which restrict the estimated optimal values of $\boldsymbol{l}$ and $\boldsymbol{u}''$ to be non-negative. The estimated minimal amount of weighted emissions released to the environment $\left(\boldsymbol{\omega}^T\widehat{\boldsymbol{u}}''^*\right)$ is used together with the estimated minimal amount of emissions generated in the production stage $\left(\boldsymbol{\omega}^T\widehat{\boldsymbol{u}}'^*\right)$ to evaluate the environmental efficiency as well as to decompose it into the components of production and abatement efficiency as described above.

To further differentiate the two efficiency components into stage and network inefficiencies the efficiency of both the production and the abatement stage need to be calculated ignoring the network structure of the overall technology. Therefore, the stage efficiency of the production stage $\left(\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Prod}}/\boldsymbol{\omega}^T\boldsymbol{u}'\right)$ is estimated by minimizing $\boldsymbol{\omega}^T\boldsymbol{u}'$ using the technology T_1. In case of an input-oriented measurement of the stage efficiency the corresponding linear program reads as

[41] In this formulation we assume that all inputs of the production stage are variable. However, if some inputs are fixed the approach by Banker and Morey (1986) can also be applied.

$$\begin{array}{rrcl}
\min\limits_{\boldsymbol{x}^P,\boldsymbol{x}_1^F,\boldsymbol{u}',\boldsymbol{\lambda},} & & & \boldsymbol{\omega}^T\boldsymbol{u}' \\
\text{s.t.} & \boldsymbol{x}^P & \geqq & \boldsymbol{X}^P\boldsymbol{\lambda} \\
& \boldsymbol{x}_1^F & \geqq & \boldsymbol{X}_1^F\boldsymbol{\lambda} \\
& \boldsymbol{x}_{1i}^S & \geqq & \boldsymbol{X}_1^S\boldsymbol{\lambda} \\
& \boldsymbol{y}_i^f + \boldsymbol{y}_i^2 & \leqq & \left(\boldsymbol{Y}^f + \boldsymbol{Y}^2\right)\boldsymbol{\lambda} \\
& \boldsymbol{u}' & = & \boldsymbol{\Pi}\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}_i^f \\
& \boldsymbol{\lambda} & \geqq & \boldsymbol{0}.
\end{array} \tag{4.13}$$

Note that since the network structure is ignored in this analysis the efficiency is evaluated without optimizing over the shared inputs.

Our approach to the efficiency of the production stage ignoring the network technology is similar to the approach by Coelli et al. (2007) which presents an analysis of environmental efficiency without incorporating abatement activities. The proposed model minimizes the amount of a single pollutant by applying the MBC. Their environmental efficiency measure and its decomposition read as[42]

$$\frac{\boldsymbol{\Pi}\boldsymbol{x}^{P*}}{\boldsymbol{\Pi}\boldsymbol{x}^{P}} = \frac{\boldsymbol{\Pi}\boldsymbol{x}^{P,\text{Tech}}}{\boldsymbol{\Pi}\boldsymbol{x}^{P}} \cdot \frac{\boldsymbol{\Pi}\boldsymbol{x}^{P*}}{\boldsymbol{\Pi}\boldsymbol{x}^{P,\text{Tech}}} \tag{4.14}$$

where $\boldsymbol{x}^{P*}$ is the environmentally efficient amount of inputs given a fixed amount of outputs and $\boldsymbol{x}^{P,\text{Tech}}$ is the technical efficient amount. The latter is obtained by applying a radial Farrell input measure of technical efficiency. Therefore, the above expression can be reformulated as

$$\frac{\boldsymbol{\Pi}\boldsymbol{x}^{P*}}{\boldsymbol{\Pi}\boldsymbol{x}^{P}} = \frac{\boldsymbol{\Pi}\theta\left(\boldsymbol{x},\boldsymbol{y}\right)\boldsymbol{x}^{P}}{\boldsymbol{\Pi}\boldsymbol{x}^{P}} \cdot \frac{\boldsymbol{\Pi}\boldsymbol{x}^{P*}}{\boldsymbol{\Pi}\theta\left(\boldsymbol{x},\boldsymbol{y}\right)\boldsymbol{x}^{P}} = \theta\left(\boldsymbol{x},\boldsymbol{y}\right) \cdot \frac{\boldsymbol{\Pi}\boldsymbol{x}^{P*}}{\boldsymbol{\Pi}\theta\left(\boldsymbol{x},\boldsymbol{y}\right)\boldsymbol{x}^{P}} \tag{4.15}$$

where $\theta\left(\boldsymbol{x},\boldsymbol{y}\right)$ denotes the Farrell input measure of technical efficiency. The second term represents a measure of environmental allocative efficiency.

Our measure of the stage efficiency can be seen as a generalization of this approach to a situation with more than one pollutant. As mentioned above we denote by $\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Prod}}$ the amount of produced emissions given the production stage operates environmentally efficient. Moreover, let $\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Tech}}$ denote the technical efficient amount of produced emissions.[43] Hence, our measure of the stage efficiency can be decomposed similar to Coelli et al. (2007) as:

$$\frac{\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Prod}}}{\boldsymbol{\omega}^T\boldsymbol{u}'} = \frac{\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Tech}}}{\boldsymbol{\omega}^T\boldsymbol{u}'} \cdot \frac{\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Prod}}}{\boldsymbol{\omega}^T\boldsymbol{u}'^{\text{Tech}}} \tag{4.16}$$

where the first term on the right hand side represents the generalized measure of technical efficiency while the second term presents a generalized measure of allocative environmental efficiency. Note that the decomposition into a radial input efficiency measure and an allocative

[42] To avoid notational confusion we use our notation instead of the notation of Coelli et al. (2007).
[43] See table B.1 in appendix B for an overview of the used measures, their abbreviations and the formal definitions.

efficiency term is only possible if $\boldsymbol{\Psi} = \mathbf{0}$. Otherwise,

$$\frac{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Tech}}}{\boldsymbol{\omega}^T \boldsymbol{u}'} = \frac{\boldsymbol{\omega}^T \left(\boldsymbol{\Pi}\theta\left(\boldsymbol{x},\boldsymbol{y}\right)\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}^f\right)}{\boldsymbol{\omega}^T\left(\boldsymbol{\Pi}\boldsymbol{x}^P - \boldsymbol{\Psi}\boldsymbol{y}^f\right)} \neq \theta\left(\boldsymbol{x},\boldsymbol{y}\right). \tag{4.17}$$

However, in this case it is still possible to differentiate the technical and the allocative efficiency but the technical efficiency can not be represented by a radial input efficiency measure.

The stage efficiency of the abatement stage can be calculated by minimizing $\boldsymbol{\omega}^T\boldsymbol{u}''$ using only T_2. Hence, by again ignoring the network structure of the production process. The corresponding linear programming problem reads as

$$\begin{array}{lrcl}
\min\limits_{\boldsymbol{u}'',\boldsymbol{z}} & & \boldsymbol{\omega}^T\boldsymbol{u}'' & \\
\text{s.t.} & \boldsymbol{y}_i^2 & \geqq & \boldsymbol{Y}^2\boldsymbol{z} \\
 & \boldsymbol{x}_{2i}^F & \geqq & \boldsymbol{X}_2^F\boldsymbol{z} \\
 & \boldsymbol{x}_{2i}^S & \geqq & \boldsymbol{X}_2^S\boldsymbol{z} \\
 & \boldsymbol{u}_i' & \geqq & \boldsymbol{U}'\boldsymbol{z} \\
 & \boldsymbol{u}_i' - \boldsymbol{u}'' & \leqq & \left(\boldsymbol{U}' - \boldsymbol{U}''\right)\boldsymbol{z} \\
 & \boldsymbol{u}_i' - \boldsymbol{u}'' & \geqq & \mathbf{0} \\
 & \boldsymbol{u}'' & \geqq & \mathbf{0} \\
 & \boldsymbol{z} & \geqq & \mathbf{0}.
\end{array} \tag{4.18}$$

When evaluating solely the efficiency of the abatement stage $\boldsymbol{u}'$ is not an optimization variable since it is given from the production stage. Moreover, analogously to the analysis of the production stage $\boldsymbol{x}_2^S$ is not an optimization variable because of ignoring the network structure. The obtained estimation $\boldsymbol{\omega}^T\widehat{\boldsymbol{u}}''^{\text{Abat}}$ can be used to estimate the stage efficiency. The stage efficiency can be used together with the abatement efficiency obtained from an analysis using the network technology to estimate the "black box" bias measure.

In the following we present a numerical example to show how our model works. As described above we measure the efficiency of the production stage input-oriented and the efficiency of the abatement stage with regard to minimal emissions released to the environment. Consider the following input-output structure of the 3 DMUs A, B and C which is graphically presented in figure 4.3.

Table 4.1: Input-output structure of 3 DMUs

DMU	x^P	y^f	x_2^F	l
A	1	1	1	1
B	4	4	1	4
C	4	2	1	1

In this graph the upper right quadrant shows the production stage where the DMUs use a single polluting input x^P to produce a single desirable output y. The production frontier is constructed by DMUs A and B which therefore are production efficient. In contrast, C is production inefficient since it lies in the interior of the production technology. In the lower

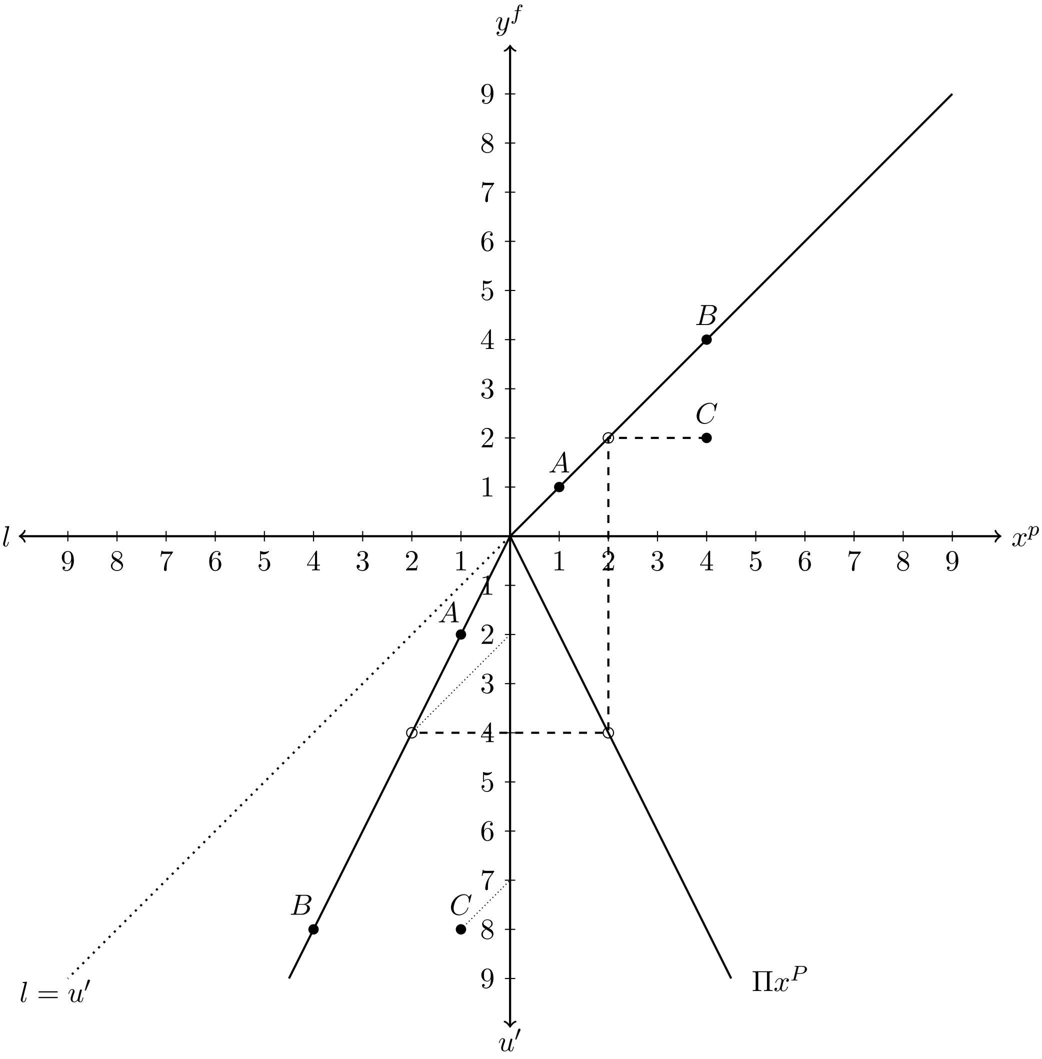

Figure 4.3: Example of the environmental efficiency analysis (CRS)

right quadrant the amount of the single pollutant u' given x^P is obtained by applying the materials balance condition. We assume that no emissions are bound in the output, hence $\Psi = 0$ and $u' = \Pi x^P$. In our example $\Pi = 2$ for all DMUs. The lower left quadrant depicts the abatement technology where l is "produced" using u' and a single non-polluting input x_2^F. For the graphical representation it is assumed that $x_2^F = 1$ for all DMUs. The dotted line in this quadrant represents all points for which $l = u'$ and hence $u'' = 0$ holds. Since we assume that $u'' \geqq 0$ no point left of this line can exist. Again, A and B are classified efficient, while C is not.

The environmental efficiency analysis of C can be visualized by following the dashed line in the plot. In the production stage C has to reduce its use of x^P from 4 to 2 units holding y constant to become efficient. From the MBC it follows that this equals 4 units of emissions. Hence, the production efficiency is estimated as

$$\frac{\widehat{u}_C'^*}{u_C'} = \frac{4}{8} = 0.5. \tag{4.19}$$

The amount of emissions released to the environment (u'') associated with a point of interest in the abatement quadrant can be obtained by drawing a line with slope equal to one through this point. The intersection of this line with the u'-axis denotes the associated amount of u''. For DMU C we therefore obtain an initial amount of emissions released to the environment of $u''_C = 7$. Given $\widehat{u}'^*_C = 4$ the maximal amount of abatement is $\widehat{l}^*_C = 2$ and thus $\widehat{u}''^*_C = 2$. Therefore, the abatement efficiency of DMU C in the above presented example is estimated as

$$\frac{\widehat{u}''^*_C}{\widehat{u}'^*_C} \cdot \left[\frac{u''_C}{u'_C}\right]^{-1} = \frac{2}{4} \cdot \left[\frac{7}{8}\right]^{-1} \approx 0.571. \tag{4.20}$$

Since the environmental efficiency is the product of the production and the abatement efficiency it is calculated as $0.5 \cdot 0.571 \approx 0.286$. This indicates that DMU C could reduce its amount of emissions released to the environment to 28.6% given that the production and the abatement stage would operate efficiently.

Note that this numerical example is kept as simple as possible to be able to present the efficiency analysis in a graphical visualization. Due to this limitation not all details of the proposed model can be included in this example. Most important, assuming that no shared or intermediate inputs exist leads to a special case of the analysis where the network environmental efficiency measure is equal the product of the stage efficiencies. Therefore, the “black box” bias measure equals one for both stages. Moreover, since we assume that only one input is used in the production stage the environmental allocative efficiency of the production stage is equal to one for each DMU and the whole inefficiency of the production stage is due to technical inefficiencies.

However, in contrast to this limited graphical example our application of the model to power plants will include all the presented details of this new model. Furthermore, in the case of variable returns to scale it is possible that even without shared or intermediate inputs the network analysis can lead to different results than the analysis ignoring the network structure. This will be discussed in the next section.

4.3.2 Analysis under variable returns to scale

As mentioned when presenting the modeling of the technologies the efficiency analysis can be extended to the case of variable returns to scale by adding the constraints $\mathbf{1}^T\boldsymbol{\lambda} = 1$ and $\mathbf{1}^T\boldsymbol{z} = 1$ to the nonparametric technology estimations.

To show how the assumption of variable returns to scale may lead to different results of the stage and the network analysis even without the assumption of shared and intermediate inputs we present a slightly modified version of our numerical example under constant returns to scale. To visualize our argument figure 4.4 represents the analysis of the abatement stage under variable returns to scale.

In this graph the observations A, B and C remain unchanged but we add an additional observation D with $u'_D = 3$ and $l_D = 2.5$. From the graph we can see that D is located on the frontier of the abatement stage. To keep the example as simple as possible we assume that D is also production efficient. However, the inclusion of D has an effect on the efficiency analysis of A.

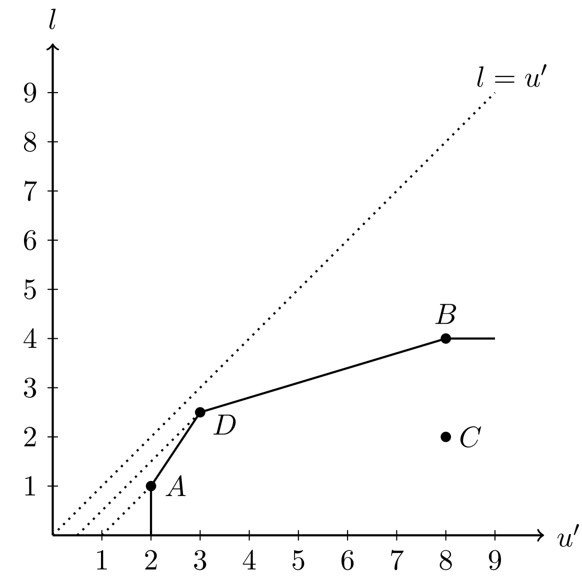

Figure 4.4: Efficiency analysis of the abatement stage (VRS)

Measuring the efficiency of A separately for each stage would render A efficient in both cases because it is located on both frontiers.[44] But in the network analysis this result does not prevail. From figure 4.4 it appears that D uses more u' but produces less u'' than A. Hence, from an environmental network perspective it would be more efficient if A moves to point D. Since we measure the production stage input-oriented, this implies that A should increase its use of x^P to produce more u' holding all other factors constant. Therefore, the network approach reveals that from an environmental view an inefficient production stage of A could be more environmentally efficient than an efficient stage.

Given the network analysis under variable returns to scale the environmental efficiency of A is thus estimated as

$$\frac{\widehat{u}''^{*}_{A}}{u''_{A}} = \frac{u''_{D}}{u''_{A}} = \frac{0.5}{1} = 0.5. \tag{4.21}$$

The decomposition of the production efficiency

$$\frac{\widehat{u}'^{*}_{A}}{u'_{A}} = \frac{\widehat{u}'^{\text{Prod}}_{A}}{u'_{A}} \cdot \frac{\widehat{u}'^{*}_{A}}{\widehat{u}'^{\text{Prod}}_{A}} = \frac{2}{2} \cdot \frac{3}{2} = 1.5 \tag{4.22}$$

and the abatement efficiency

$$\frac{\widehat{u}''^{*}_{A}}{\widehat{u}'^{*}_{A}} \cdot \left[\frac{u''_{A}}{u'_{A}}\right]^{-1} = \frac{\widehat{u}''^{\text{Abat}}_{A}}{u'_{A}} \cdot \left[\frac{u''_{A}}{u'_{A}}\right]^{-1} \cdot \frac{\widehat{u}''^{*}_{A}}{\widehat{u}'^{*}_{A}} \cdot \left[\frac{\widehat{u}''^{\text{Abat}}_{A}}{u'_{A}}\right]^{-1} = \frac{1}{2} \cdot \left[\frac{1}{2}\right]^{-1} \cdot \frac{0.5}{3} \cdot \left[\frac{1}{2}\right]^{-1} = \frac{1}{3} \tag{4.23}$$

into the stage efficiency and the “black box” bias measure shows that the whole inefficiency of DMU A obtained in the network analysis can be explained by network effects. This follows because in both decompositions the stage efficiency term equals one.

[44] Since A uses the smallest amount of x^P in the production stage it is under VRS production efficient irrespective of how much output it produces.

Note that the above presented result occurs because the slope of the frontier connecting A and D is larger than one. Since the frontier under CRS has to exhibit a slope smaller than one (it lies below the $l = u'$ line), increasing returns to scale of the abatement process are a necessary but not a sufficient condition to observe such a situation.[45]

4.3.3 The role of the aggregation weights

In the above presented numerical example we have assumed that only a single pollutant results from the production of the good output. However, for more general applications our model allows to incorporate multiple pollutants. In this case our measure of environmental efficiency is based on the use of different weights which are exogenously given to aggregate emissions in the index. We therefore differentiate from approaches which weight different pollutants equally (see e.g. Pasurka (2006)). The use of weighted emissions leads to several advantages.

It allows the index to be modified to analyze efficiency to a specific issue, e.g. global warming, for which possible choices of the weights are given by global warming potentials to convert different emissions into CO_2 equivalents. For example, the global warming potential of methane (CH_4) is 25 (see Forster et al. (2007)). This means, if the DMU emits 1 ton of CO_2 and 1 ton of CH_4 our efficiency measure aggregates the pollutants to 26 tons of CO_2. As another example it is also possible to measure efficiency in a monetary setting. In this case environmental taxes or emission allowance prices can be used as the aggregation weights. Olsthoorn et al. (2001) summarize that the view on environmental indicators depends on the user of the indicator. For example, the question how important different emissions are and hence how much they should influence a measure of environmental efficiency may be answered differently from economists, managers or national authorities. In this sense, the introduction of different weights allows for a very flexible approach to environmental efficiency. Furthermore, not weighting all pollutants equally incorporates the possibility to improve efficiency not only by reducing pollutants but also by substituting among different pollutants. For example, regulated emissions which are taxed may be substituted through non-taxed emissions (see Färe et al. (2012) for an analysis of the substitution among different pollutants by U.S. power plants).

Beside these advantages, the use of exogenous weights also leads to several drawbacks. As Kuosmanen and Kortelainen (2005) point out, the possibility to choose from different weights leads to a source of subjectivity. For example, Toffel and Marshall (2004) analyze how different weighting schemes influence the results of comparing the impact of different chemicals contained in the Toxic Release Inventory by the U.S. Environmental Protection Agency. They find that no optimal weighting scheme can be identified but that the appropriate scheme depends on the target of the research, e.g. the impact on human health or the environment may be analyzed using different weighting schemes. To overcome the subjectivity, Kuosmanen and Kortelainen (2005) propose to use a multiplier form of DEA to estimate the weights of different pollutants. But they note that this may lead to implausible results for the weights. Therefore they suggest to use an assurance region approach to limit possible choices of the weights (see Pedraja-Chaparro

[45] For empirical analyses on whether increasing returns to scale are present in abatement activities see Andreoni and Levinson (2001) and Managi (2006).

et al. (1997) for an overview of weight restriction in DEA models). In our model such bounds may be needed if the aggregation weights are not known with certainty. E.g. the climate change effects of different pollutants may not be known precisely (see Harvey (1993)). Kuosmanen and Post (2001) and Camanho and Dyson (2005) present models to analyze cost efficiency if the price data are uncertain. These models could be modified for an analysis of environmental efficiency and incorporated in our model.

4.3.4 Comparison to existing network models

In the following we want to compare our model to two existing models which have a similar structure and discuss some points which we consider as advantages of our approach.

Murty et al. (2012) and Färe et al. (forthcoming) both present models which aim at analyzing the efficiency of DMUs with regard to environmental pollution in a more detailed technology modeling than the standard “black box” models. Murty et al. (2012) discuss microeconomic concepts and implications of modeling undesirable outputs as byproducts of the desirable outputs and formulate an overall technology as the union of a production and an residual generation technology which also allows to account for abatement activities. Their overall technology ($\widehat{T}^{\mathrm{MRL}}$) reads as[46]

$$\widehat{T}^{\mathrm{MRL}} = \Big\{ \Big(\boldsymbol{x}^P, \boldsymbol{x}_1^F, \boldsymbol{y}^f, \boldsymbol{u}''\Big) \in \mathbb{R}^{m^P+m^F+k+s} : \boldsymbol{x}^P \geqq \boldsymbol{X}^P\boldsymbol{\lambda}, \boldsymbol{x}_1^F \geqq \boldsymbol{X}_1^F\boldsymbol{\lambda}, \boldsymbol{y}^f \leqq \boldsymbol{Y}^f\boldsymbol{\lambda}, \\ \boldsymbol{\lambda} \geqq \boldsymbol{0}, \boldsymbol{x}^P \leqq \boldsymbol{X}^P\boldsymbol{z}, \boldsymbol{u}'' \geqq \boldsymbol{U}''\boldsymbol{z}, \boldsymbol{z} \geqq \boldsymbol{0}\Big\}. \tag{4.24}$$

To account for abatement activities the authors propose to incorporate an abatement output which is an output of the first stage and an input of the second stage. This abatement output therefore is the same approach as our incorporation of intermediate inputs. In this modeling of the residual generation process the authors do not model an explicit relation of the polluting input and the emissions like the MBC but introduce the assumption of “costly disposability” of $\boldsymbol{x}^P$ and $\boldsymbol{u}''$. This assumption is represented by the restrictions $\boldsymbol{x}^P \leqq \boldsymbol{X}^P\boldsymbol{z}$ and $\boldsymbol{u}'' \geqq \boldsymbol{U}''\boldsymbol{z}$ in the technology $\widehat{T}^{\mathrm{MR}}$.[47] Hence, modeling “costly disposability” leads to a formulation of the residual generation where undesirable outputs are treated as inputs and polluting inputs are treated as outputs. This formulation leads to a model that exhibits the technical possibility to produce an infinite amount of emissions with a given amount of polluting inputs. This is physically impossible and one of the arguments against modeling emissions as inputs in a “black box” model (see the chapter on general concepts). In our model this problem is not present. Furthermore, the costly disposability of polluting inputs is problematic if one considers the two subtechnologies separately. In our model all subtechnologies $T_1^{\boldsymbol{y}}$, $T_1^{\boldsymbol{u}'}$ and T_2 present plausible models even if they are not combined in a network. In the above described model this is not the case since the costly disposability of $\boldsymbol{x}^P$ allows to produce positive amounts of emissions without using any polluting input if the residual generation is analyzed separately. Again, this problem is absent in our approach.

[46] As before, our notation is applied.

[47] In contrast, our model relies on standard disposability assumptions for the production and the abatement stage.

To measure the efficiency of DMUs given the above defined technology, Murty et al. (2012) propose three different approaches which are all based on scaling both the desirable outputs and the undesirable outputs. They apply the hyperbolic efficiency measure by Färe et al. (1989) and the directional (output) distance function in an environmental setting. Moreover, they propose an index which is based on the Russell measure of technical efficiency by Färe and Lovell (1978). But the proposed decomposition of this index into a production and an emission efficiency index relies on the independence of the abatement stage from the production of good outputs and hence does not allow for a network application with multiple dependencies. In contrast, our measure allows for a decomposition also in a general network setting.

The second model which we want to compare with our approach is the environmental network DEA model by Färe et al. (forthcoming). Similar to ours the model consists of a production and an abatement subtechnology. In this network emissions are modeled as a weak disposable output of the first stage which is a strong disposable input in the second stage and undesirable outputs released to the environment are a weak disposable output of the abatement stage. Färe et al. (forthcoming) also include shared and intermediate inputs in their network model. To measure technical efficiency a directional (output) distance function with directional vector $\boldsymbol{g} = (1, 1)$ is applied.[48] The linear programming problem to obtain the measure is given by

$$\begin{array}{llcl}
\max\limits_{\beta,\boldsymbol{x}_1^S,\boldsymbol{x}_2^S,y^2,u',\boldsymbol{\lambda},\boldsymbol{z}} & \beta & & \\
s.t. & \boldsymbol{x}_1^S & \geqq & \boldsymbol{X}_1^S\boldsymbol{\lambda} \\
& y^2 + (y_i^f + \beta) & \leqq & \boldsymbol{Y}\boldsymbol{\lambda} \\
& u' & = & \boldsymbol{U}'\boldsymbol{\lambda} \\
& \boldsymbol{\lambda} & \geqq & \boldsymbol{0} \\
& & & \\
& y^2 & \geqq & \boldsymbol{Y}^2\boldsymbol{z} \\
& u' & \geqq & \boldsymbol{U}'\boldsymbol{z} \\
& \boldsymbol{x}_2^S & \geqq & \boldsymbol{X}_2^S\boldsymbol{z} \\
& u_i'' - \beta & = & \boldsymbol{U}''\boldsymbol{z} \\
& \boldsymbol{z} & \geqq & \boldsymbol{0} \\
& & & \\
& \boldsymbol{x}_1^S + \boldsymbol{x}_2^S & \leqq & \boldsymbol{x}_i^S
\end{array} \tag{4.25}$$

In this model Färe et al. (forthcoming) do not differentiate between polluting and non-polluting inputs and all inputs are part of a common source. This assumption would imply in our model that all inputs are shared between the stages.

As in case of the model of Murty et al. (2012) this approach is less flexible than ours regarding the measurement of efficiency. To see this, consider the case that the DMUs are evaluated without changing the production stage and emissions should be minimized by improving only the efficiency of the abatement process. In our model this would imply an optimization over u'' and $\boldsymbol{z}$ and in the model by Färe et al. (forthcoming) the efficiency measure would only scale u_i''. Comparing the models we find that this leads to very different results. In the model by Färe

[48] Note that in the paper by Färe et al. (forthcoming) only a single pollutant and a single good output is modeled in the network setting.

et al. (forthcoming) we obtain that $\widehat{\boldsymbol{z}} = \mathbf{0}$ is optimal for all DMUs. Hence, a positive output production would lead to an optimal amount of zero undesirable output in the case of a single pollutant. In contrast, in our model $\widehat{\boldsymbol{z}} = \mathbf{0}$ can not be optimal. This follows because evaluating only the abatement stage remains $u_i' > 0$ unchanged. Therefore, $\widehat{\boldsymbol{z}} = \mathbf{0}$ would imply $\widehat{l_i} = 0$ and $\widehat{u}_i'' = u_i'$. Thus, $\widehat{\boldsymbol{z}} \geq \mathbf{0}$ and assuming that no observation abates all its emissions this implies that $\widehat{u}_i''^{*} > 0$.

With regard to the above presented findings from a comparison of network models we argue that our model which does not rely on non-standard disposability assumptions and leads to plausible results even if the stages are evaluated separately provides clear advantages compared to the existing approaches.

4.4 Application to U.S. power plants

For an empirical illustration we apply our model to an efficiency analysis of U.S. coal-fired power plants in the year 2009. These plants have been addressed by several previous studies (e.g. Färe et al. (2005), Färe et al. (2007), and Sueyoshi et al. (2010)) analyzing their environmental efficiency as described in the motivation of this study. The amount of available data, especially for abatement activities, has significantly increased in the last years enabling us to conduct a detailed analysis of the potential sources of environmental inefficiency.

We analyze the environmental efficiency with regard to the sulfur dioxide (SO_2) emissions of the power plants. The reason for choosing SO_2 emissions in our analysis is two folded. Firstly, coal fired power plants contribute 73% of all SO_2 emissions in the United States (EPA (2012)) and therefore their efficiency has a significant influence on the total generation of SO_2 emissions in the United States. Secondly, the abatement of these emissions by flue gas desulfurization units (FGDs) exists as an end-of-pipe technology and hence can be analyzed with our network model (see Srivastava and Josewicz (2001) for a description of FGDs).

4.4.1 Constructing the dataset of U.S. power plants

The sources of the data used in our study are the files EIA-767, EIA-860 and EIA-923 of the U.S. Energy Information Administration (EIA), where EIA-923 provides detailed information on the inputs and outputs of the production stage of the power plants, whereas EIA-767 and EIA-860 contain data on their abatement activities. These files present the data on boiler and generator level (EIA-923), respectively on FGD unit level (EIA-767 and EIA-860). In addition to the data from EIA we use the Clean Air Markets data from the U.S. Environmental Protection Agency (EPA). These data provide information on the amount of SO_2 emissions released to the environment by each boiler of the plants. Finally, we include plant-level data on the structure costs and labor input from Form 1 of the Federal Energy Regulatory Commission (FERC).

Since the data are reported on different levels we have to aggregate them to estimate plant-level efficiency. This aggregation is done as follows: In the first step, we exclude all observations (boilers, generators and FGDs) with missing data (see the paragraph below for a description of the used inputs and outputs). We also exclude boilers for which coal contributes to less than 95% of the used heat content of the fuels and those for which fuels other than coal, oil or gas contribute to more than 0.0001% of the used heat content.[49] FGDs are excluded if they are either non-operating or if the generators linked with these units have an installed capacity that is lower than 100 megawatts. The last exclusion is due to previous studies (see Eastern Research Group (2009)) which find that while medium (100 – 500 MW) and large (> 500 MW) FGDs are comparable e.g. with regard to capital cost per capacity, smaller (< 100 MW) FGDs show significant differences to large and medium FGDs. To avoid this comparison we exclude those observations. The remaining observations are checked whether all linked units are still included in the data set (e.g. if all boilers that are linked to one generator are still part of the data). If the data are complete we sum the single parts up to estimate the data on power plant level. Otherwise, we do not include the data. As a result of this procedure, not necessarily all generators, boilers and FGDs of a power plant are included in our analysis. However, we prefer this method to simply summing up all boilers, generators and FGDs to one plant since this would lead to more serious problems. For example, our approach avoids comparing the environmental efficiency of boilers and generators without FGDs with those that are equipped with FGDs using the same estimated technology set.

For our efficiency analysis we use the following input and output variables. In the production stage we include the sum of the heat content (measured by British thermal units (BTUs)) of the fuels used by the power plants (coal, oil and natural gas) as the pollution-containing input.[50] We do not include the different fuels separately.[51] Since our final sample only consists of 23 power plants we want to reduce the number of inputs. Therefore, we aggregate all used fuels by multiplying the physical quantities of the fuels with their heat content (reported in the file EIA-923) and sum up the results to one input (total heat content). The non-polluting input of the production stage is given by the capital stock of the power plant which we estimate following Pasurka (2006). The FERC Form 1 contains data on the cost of structures and equipment of the power plants. Since no data on investment is available we assume that changes in these costs represent net investments (NI). Using the Handy-Whitman index (HW) by Whitman, Requardt & Associates (2009) to present it in 1970$ we estimate the capital stock (CS) by:

$$\mathrm{CS}_T = \sum_{t=1}^{T} \frac{\mathrm{NI}_t}{\mathrm{HW}_t}. \tag{4.26}$$

Since the FERC data are publicly available only back to 1994 we assume that all previous investments were done in the year the last unit (generator) of the plant was installed. Moreover, since we do not include the whole plants due to missing data we estimate the partial capital stock by multiplying the overall capital stock with the ratio of the generating capacity of the included generators of the plant to the generating capacity of the overall power plant. The output of the production stage, the produced amount of electricity, can be split up into two parts. The first part, the net generation of electricity, is the amount (measured in gigawatt hours (GWh))

[49] This step is done following the definition of a power plant as coal-fired by Färe et al. (2007).
[50] A BTU is the amount of thermal energy needed to raise the temperature of one pound of water at $68°F$ by $1°F$ (Çengel (2008, p. 9)).
[51] In our analysis coal consists of anthracite and bituminous (BIT), lignite (LIG) and subbituminous (SUB) coal. Oil consists of distillate (DFO) and residual fuel (RFO) oil.

of electricity produced by the plant excluding the amount of electricity used by the plant. The second part is the amount of electricity used by the FGD units to abate SO_2 emissions. Since this electricity is both an output of the production stage and an input of the abatement stage, it can be viewed as an intermediate input. To estimate the amount of SO_2 emissions (in tons) which are generated in the production stage we multiply the physical quantities of the used fuels by their sulfur content and by the uncontrolled emission factors that are reported in appendix A of the Electric Power Annual Report (EIA (2011)).

Beside these emissions and the electricity used by the FGD units, the capital stock of the installed FGD structures is included as an input of the abatement stage. The capital stock is obtained by again applying the methodology of Pasurka (2006) using the historical data on costs of structures and equipment of the FGD units collected in the EIA file 767. They are reported back to 1985 and we assume that all prior investments were done in the year the FGD unit went into operation. The single output of the abatement stage is the amount of abated SO_2 emissions which we obtain by subtracting the amount of released SO_2 emissions (given by the EPA data) from the estimated amount of SO_2 emissions produced in the first stage.

In addition to the inputs described above which are only used by either the production or the abatement stage we also include one shared input, the number of employed workers. Two problems arise from the fact that FERC data only report the overall number of workers employed at the power plant. Firstly, our dataset does not necessarily cover the whole plants as explained above. Therefore, we assume that the number of workers of the plants is proportional to the plants capacity and hence the total amount of labor in our dataset is estimated by

$$\text{Number of workers (total)} = \text{Number of workers (FERC)} \cdot \frac{\mathrm{MW}_{\mathrm{Data}}}{\mathrm{MW}_{\mathrm{Total}}} \tag{4.27}$$

where $\mathrm{MW}_{\mathrm{Data}}$ is the capacity of the plant in our dataset and $\mathrm{MW}_{\mathrm{Total}}$ is the total capacity of the plant. Secondly, we have to attribute the total amount of workers to the production and the abatement stage. To estimate the number of workers operating the FGD units we use two power laws which were developed by Srivastava (2000) to estimate the cost of operating labor for FGD units. For wet scrubbers the amount is estimated by

$$\text{Number of workers (FGD)} = 41.69041 \cdot \mathrm{MW}^{-0.322307} \cdot \frac{\mathrm{MW}}{100} \tag{4.28}$$

and for dry scrubbers by

$$\text{Number of workers (FGD)} = (18.25 - 2.278 \cdot \ln(\mathrm{MW})) \cdot \frac{\mathrm{MW}}{100} \tag{4.29}$$

where MW denotes the capacity of the generators linked to the FGD unit. The difference of the estimated total number of workers and the estimated FGD operating labor is attributed to the production stage. Table 4.2 presents the descriptive statistics for the data used in the production and the abatement stage, respectively. The relatively small sample size (23 power plants) is largely driven by the FERC data since for many power plants no labor and/or cost data were available.

Table 4.2: Descriptive statistics of the power plant data

$n = 23$ Power plants	Min	Mean	Median	Max	SD
Production stage					
Inputs					
Total heat content (Bio. BTUs)	9049.93	54947.11	43108.64	127135.37	34595.67
Capital stock (1000$, 1970)	114038.54	321271.60	283597.19	773356.78	160345.37
Output					
Net generation (GWh)	1108.39	5774.96	4278.05	14664.33	3819.71
Abatement stage					
Inputs					
SO_2	3878.54	101553.51	54928.84	304364.04	96448.70
Capital stock (1000$, 1970)	10103.63	46316.15	31674.81	141685.67	34907.40
Output					
Abated SO_2 emissions (tons)	2102.27	94133.96	49150.88	280130.00	91300.74
Intermediate input					
FGD electricity (GWh)	5.79	98.94	57.54	371.91	98.23
Shared input					
Labor total (worker)	57.00	155.61	129.00	398.00	78.08
Labor production (worker)	25.00	102.26	88.00	330.00	65.65
Labor FGD (worker)	12.00	40.70	33.00	77.00	20.59

4.4.2 Results of the analysis of U.S. power plants

We estimate the environmental efficiency of the power plants by solving the linear programming problem (4.12).[52] In this network analysis we minimize the amount of final emissions ($\boldsymbol{u}''$) over the inputs of the first stage ($\boldsymbol{x}_1^F$, $\boldsymbol{x}_1^P$) as well as the shared inputs ($\boldsymbol{x}_1^S$, $\boldsymbol{x}_2^S$) and the emission input of the second stage ($\boldsymbol{u}'$). Since we aim at estimating the efficiency of the production stage input-oriented the final output as well as the intermediate input remain unchanged. Matrix $\boldsymbol{\Pi}$ is estimated for each plant by dividing the initial amount of sulfur dioxide emissions ($\boldsymbol{u}'$) by the total heat content. Since the final output (net generation of electricity) does not contain any pollution the analysis of power plants allows for a detailed decomposition of the production stage efficiency into technical efficiency (measured by a radial Farrell input measure of technical efficiency) and environmental allocative efficiency as proposed by Coelli et al. (2007) (see table B.1 in appendix B for a summary of the decompositions). The analysis is conducted for both constant and variable returns to scale. Detailed results of the environmental efficiency as well as the decomposition into the different components for each plant can be found in tables B.2 and B.3 in appendix B.

[52] To solve this programming problem we use the package “lpSolve” for the statistical software R.

Table 4.3: Descriptive statistics of the power plant efficiency results (CRS)

Efficiency type	Min.	Mean	Median	Max.	SD	# Efficient
Network						
Environmental	0.0200	0.2739	0.0861	0.9936	0.3089	0
Production	0.7719	0.8776	0.8758	1.0000	0.0677	1
Abatement	0.0223	0.3054	0.1017	1.0000	0.3374	1
Production stage						
Stage efficiency	0.7705	0.8625	0.8478	1.0000	0.0629	1
Technical	0.8031	0.9012	0.8841	1.0000	0.0665	3
Allocative	0.8524	0.9575	0.9587	1.0000	0.0248	1
"Black box" bias	1.0000	1.0173	1.0169	1.0427	0.0142	5
Abatement stage						
Stage efficiency	0.0223	0.3624	0.1017	1.0000	0.3999	5
"Black box" bias	0.5071	0.9282	1.0000	1.0000	0.1552	17

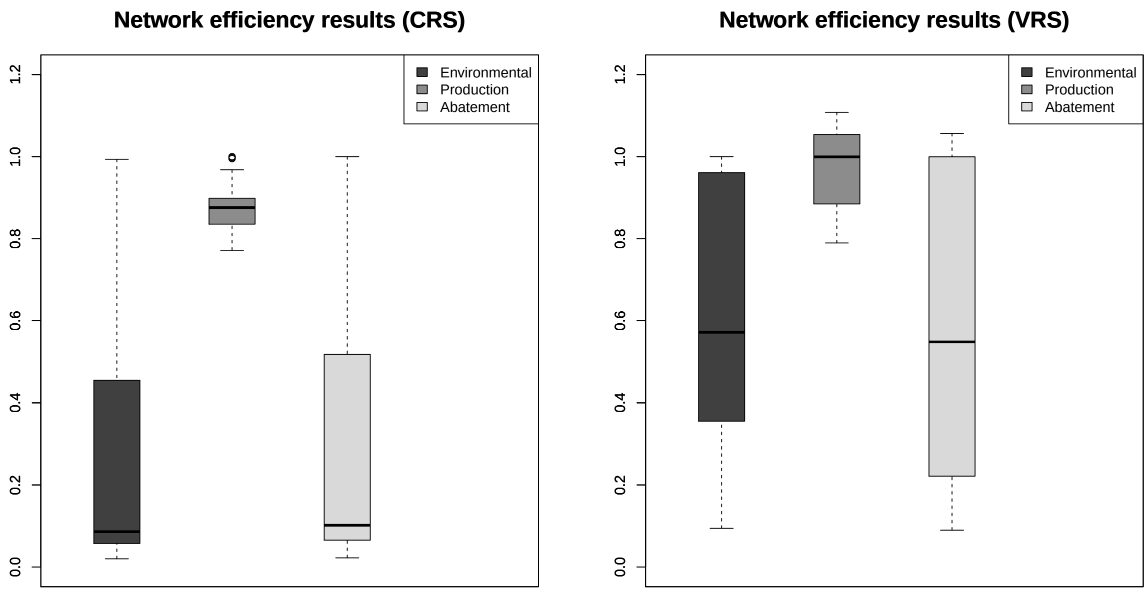

Figure 4.5: Boxplots of the power plant efficiency results

We start our discussion by presenting the results for the analysis under constant returns to scale. Table 4.3 contains the descriptive statistics of the results and all decompositions while the left graph in figure 4.5 contains boxplots for the results of the network analysis.

The results show that the power plants exhibit significant potentials to reduce their SO_2 emissions. The average environmental efficiency is 0.2739 which indicates that at the average the power plants could reduce their emissions by about 73 % holding their electricity output constant. This result becomes even more obvious when inspecting the median which shows that 50% of the power plants could reduce their emissions to less than 10% of the current level. One plant of our dataset (Valmont) is production efficient, another one (East Bend) is abatement efficient. Since no plant is both production and abatement efficient no plant can be identified as environmentally efficient given the analysis under constant returns to scale.

The decomposition shows that the environmental inefficiency is largely driven by abatement inefficiencies while the production efficiency is relatively high. This can be seen from both the boxplots as well as the descriptive statistics which show that the average production efficiency is 0.8776 while the average abatement efficiency is 0.3054. This indicates that at the average the amount of emissions resulting in the production stage could be reduced by about 12% if the power plants would operate their production stage efficiently. Moreover, the power plants could on average reduce the amount of emissions released to the environment per unit of produced emissions by approximately 69% if the abatement process would be conducted efficiently.

The results for the production stage ignoring the network structure show that stage inefficiencies can be attributed to technical inefficiency while the allocative efficiency is quite high. This indicates that the reduction of the emissions is largely obtained by proportionally scaling down both the pollution-free as well as the polluting input. Not exploited substitution possibilities between the two inputs do only contribute to a smaller extend to the overall production stage efficiency. However, while three power plants are technical efficient only a single one is allocative efficient. Therefore, only one plant (Valmont) is production stage efficient. The results for the “black box” bias measure of the production stage show that the results for the production efficiency of five plants are not changed if the analysis is conducted using the network technology. The remaining observations exhibit values for the bias that are larger than one. These plants substitute polluting inputs for shared inputs. Therefore, they do not have to reduce their emissions as much as possible given the stage analysis and instead have to reallocate the shared input from the production to the abatement stage to reduce the amount of emissions released to the environment.

Regarding the efficiency of the abatement stage we find that five power plants are stage efficient. The remaining observations are highly inefficient. The most stage inefficient plant (Valmont) could reduce its amount of emissions released to the environment per unit of produced emissions by approximately 98%. Since this plant is found to be production efficient the whole environmental inefficiency is due to the large inefficiency of the abatement stage. For the other abatement inefficient plants we also clearly observe from the boxplots that this inefficiency is the main driver of the overall environmental inefficiency. The majority of 17 plants exhibits a value for the “black box” bias measure equal to one showing that the abatement inefficiency is largely due to stage inefficiencies and not to inefficiencies caused by network effects.

To evaluate whether these large inefficiencies obtained from the analysis assuming constant returns to scale are due to scale inefficiencies we have also estimated the environmental efficiency measure and its decompositions assuming variable returns to scale. Table 4.4 contains the descriptive statistics of the results while the right graph in figure 4.5 presents the boxplots.

Table 4.4: Descriptive statistics of the power plant efficiency results (VRS)

Efficiency type	Min.	Mean	Median	Max.	SD	# Efficient
Network						
Environmental	0.0942	0.5965	0.5721	1.0000	0.3252	4
Production	0.7896	1.1886	0.9994	2.4456	0.5257	5
Abatement	0.0897	0.5854	0.5485	1.0566	0.3712	4
Production stage						
Stage	0.7896	0.8922	0.8835	1.0000	0.0637	2
Technical	0.8065	0.9212	0.9103	1.0000	0.0660	6
Allocative	0.8461	0.9699	0.9898	1.0000	0.0491	5
“Black box” bias	1.0000	1.3423	1.0000	2.7128	0.6207	13
Abatement stage						
Stage	0.1238	0.7227	1.0000	1.0000	0.3377	13
“Black Box” bias	0.3503	0.7609	0.8130	1.0566	0.2463	5

The results for the analysis under VRS show that the environmental efficiency increases largely compared to the results assuming constant returns to scale. The average inefficiency decreases from 73% to 40%. The results for the median are even larger with the inefficiency decreasing from 91% to 43%. Moreover, we find that in addition to the plant Valmont three more plants (Cope, Reid Gardner and Trimble County) are both production and abatement efficient and hence environmentally efficient. The decomposition of the environmental efficiency shows that in contrast to the analysis under CRS some plants exhibit a value for the production efficiency larger than one indicating that in a network analysis these plants have to increase the emission of pollutants in the production stage to become efficient. As discussed before, this follows because shared inputs are reallocated to the abatement stage and, furthermore, increasing returns to scale at the abatement stage can be exploited. The abatement efficiency shows significantly larger values than under CRS indicating that the results are largely influenced by scale effects.

With regard to the stage analysis of the production stage we only find minor differences under VRS compared to the results under CRS. Interestingly, the “black box” bias results show that a majority of 13 plants does not exhibit different results for the production stage if the network structure is taken into account under VRS. However, the remaining plants show very large differences. This indicates that for a minor group of plants scale effects have a crucial influence on the efficiency if they are evaluated using a network approach instead of a “black box” model. The results for the abatement stage show that the inefficiency decreases largely if we exclude scale inefficiencies as it is done by assuming variable returns to scale. For example, the average inefficiency decreases from 64% to 38% while the median inefficiency decreases from 90% to 0%. However, we find that accounting for the network structure renders less plants efficient than under CRS. But the stage effect dominates and hence the overall abatement efficiency increases largely. This indicates that the large environmental inefficiencies found in the analysis under CRS are driven by scale inefficiencies of the abatement stage. But we have to emphasize that our dataset only covers a small sample of plants. Moreover, due to the above presented data limitations some plants are not completely included. Hence, the scale inefficiencies may be influenced by the construction of the dataset. However, our findings of scale inefficiencies are in line with findings in previous studies on the efficiency of power plants in the United States (e.g. Yaisawarng and Klein (1994) and Zeitsch and Lawrence (1996)). However, our results suggest that these inefficiencies are more important for the abatement process than for the production process.

4.5 Summary

In this chapter we have presented a new approach to evaluate the environmental efficiency of decision making units. Our model presents a network technology which consists of a production and an abatement stage with the environmental efficiency being determined by the efficiency of both stages. The proposed modeling of the two stages does not need special assumptions on the disposability of in- and outputs and allows for an efficiency analysis of both stages ignoring the network structure. Therefore, it addresses some of the shortcomings of previous approaches. Building on this two-stage network we have proposed a new measure of environmental efficiency which can be decomposed into the effects of production and abatement inefficiencies. Further decompositions allow for an analysis of stage and network effects. The new measure and all its components can be evaluated assuming both constant and variable returns to scale.

Our application of the model to a sample of U.S. power plants shows that there are significant potentials to reduce SO_2 emissions. While we found only minor inefficiencies with regard to the production stage, large inefficiencies were found for the abatement activities. Since usual "black box" models do not or only implicitly account for abatement processes, this finding highlights the importance of a detailed analysis to reveal all environmental inefficiencies of the DMUs. However, we have to emphasize again that the sample of our analysis covers only 23 power plants and is thus merely illustrative. Therefore, the results should not be overstressed as providing information on the efficiency of the total power plant infrastructure in the United States.

5 Optimal directions for directional distance functions[‡]

5.1 Motivation

In the literature on nonparametric efficiency analysis of decision making units the directional distance function which has been introduced in the chapter on general concepts is a frequently used approach to measure efficiency in the presence of undesirable outputs. A microeconomic analysis applying this function was conducted in chapter 3 of this dissertation. Examples of macroeconomic applications of the DDF accounting for undesirable outputs are, among others, Arcelus and Arocena (2005), Färe et al. (2001), Lozano and Gutiérrez (2008) and Picazo-Tadeo et al. (2005). The main advantage of the DDF is the possibility to define a different direction of measurement for each input or output. Therefore, it is possible to analyze the efficiency by increasing good outputs while simultaneously decreasing bad outputs. However, the variety of possible directions allows for a large extend of subjectivity regarding the importance of the production of good and the abatement of bad outputs. Färe et al. (2011) propose a method to endogenously determine the directions for a slacks-based directional measure. In this study we modify their approach to the analysis using the environmental directional distance function by Chung et al. (1997) and propose an alternative method to obtain optimal directions in a dynamic setting.

We apply both methods to an efficiency analysis of the 62 countries which together produced 90% of the average total global emissions of greenhouse gases (GHG) of the years 2000 and 2005. Using the results of the analysis we calculate by how much the GHG emissions could be reduced if countries would produce efficiently. This question is of great importance given that anthropogenic GHG emissions are regarded as the main driver of climate change (see the introduction to this dissertation). The literature about the economic effects of climate change is summarized e.g. in the Stern Review (Stern (2007)) and subsequently by Tol (2009) and Aldy et al. (2010). Aldy et al. (2010) argue that stabilizing global warming over pre-industrial levels at 2.9°C (2.1°C) requires a stabilization of CO_2 equivalent greenhouse gas emissions concentration in the atmosphere at 550 ppm (450 ppm).[53] According to Stern (2007) this amounts to a reduction of CO_2 equivalent greenhouse gas emissions of 25-30 percent (70 percent) until 2050 relative to 2005. Likewise, the European Commission (2011) announces the commitment of member states to a reduction of greenhouse gas emissions by 20 percent until 2020.

In this chapter we want to shed some light on the feasibility of these percentage reduction targets from a production-economic perspective. In our analysis we compare the reduction potentials obtained by the efficiency analysis with the necessary reductions to limit the effects of climate change. We examine whether the targets can be achieved by reducing inefficiencies in the production processes of the countries. In applying nonparametric methods to assess emissions reduction targets our study is closely linked with Färe et al. (2012) who present results of the optimal timing of greenhouse gas emissions reductions.

[‡] This chapter is based on Hampf and Krüger (2012).

[53] CO_2 equivalent greenhouse gas emissions are the sum of CO_2 emissions and several greenhouse effective gases, denominated in equivalent tons of CO_2. See the data description below. The abbreviation ppm stands for parts per million.

In the following section we discuss different approaches for computing optimal directional vectors for the environmental directional distance function. We start by applying the model by Färe et al. (2011) to the analysis of environmental efficiency. We then propose a novel approach based on an extension to a dynamic nonparametric analysis.

5.2 Deriving optimal directions

5.2.1 Static approaches to optimal directions

In recent literature the question of how to obtain optimal directional vectors for directional distance functions has received some attention. Peyrache and Daraio (2012) present an empirical approach to investigate this question which is merely a robustness assessment while Färe et al. (2011) present a theoretical model to calculate the directions endogenously. Their model estimates the optimal directions by maximizing the inefficiency of the DMU under evaluation over the directional vector. We follow Färe et al. (2011) and apply their model to an environmental efficiency analysis. The original paper presents the model for the case of two DMUs and two outputs and applies the slacks-based directional measure by Färe and Grosskopf (2010). Extending their analysis we consider the case of multiple DMUs using multiple in- and outputs in an environmental setting. We start the discussion by applying a distance function that scales all outputs in weighted proportions. Then we will show how this approach is related to the slacks-based measure applied by Färe et al. (2011).[54]

The first distance function has the advantage that it allows to connect the ideas of Färe et al. (2011) with the literature of dynamic efficiency analysis as presented in the next section. It can be calculated by solving the nonlinear programming problem

$$
\begin{array}{lrcl}
\max\limits_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} & \beta & & \\
\text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\
& \boldsymbol{y}_i + \beta\boldsymbol{\alpha} \odot \boldsymbol{y}_i & \leqq & \boldsymbol{Y\lambda} \\
& \boldsymbol{u}_i - \beta\boldsymbol{\delta} \odot \boldsymbol{u}_i & = & \boldsymbol{U\lambda} \\
& \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} & = & 1 \\
& \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} & \geqq & \mathbf{0}.
\end{array}
\tag{5.1}
$$

The elements of the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ represent the different weights of the good and bad outputs, while $\odot$ denotes the Hadamard (or direct) product of two vectors. The fourth constraint is a normalization constraint. The non-negativity assumption for $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ implies that only directions which do not increase bad outputs or decrease good outputs are selected. This model maximizes the distance function and hence the inefficiency of a DMU by endogenously selecting the optimal directional vector. This vector is optimal in the sense that it is directed to the furthest feasible point on the frontier compared to the DMU under evaluation. The resulting elements of the $\widehat{\boldsymbol{\lambda}}$ vector identify the reference observations for the analyzed DMU. The resulting

[54] In most applications slacks-based measures are compared to radial distance functions like the Shephard (1970) output distance function. However, since we consider a weighted scaling as well as increasing good and decreasing bad outputs the term “radial” is not appropriate in this analysis.

efficiency measure of this program can be denoted as $\widehat{\beta}\left(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i\right)$. For notational simplification we abbreviate this measure by $\widehat{\beta}$. The nonlinear programming problem can be transformed into a linear one by dividing all constraints by β and introducing the new variables $\gamma = 1/\beta$ and $\boldsymbol{\mu} = \boldsymbol{\lambda}/\beta$.[55] The linear model then reads as

$$\begin{array}{llcl}
\min\limits_{\gamma,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\mu}} & \gamma & & \\
\text{s.t.} & \gamma \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\mu} \\
 & \gamma \boldsymbol{y}_i + \boldsymbol{\alpha} \odot \boldsymbol{y}_i & \leqq & \boldsymbol{Y}\boldsymbol{\mu} \\
 & \gamma \boldsymbol{u}_i - \boldsymbol{\delta} \odot \boldsymbol{u}_i & = & \boldsymbol{U}\boldsymbol{\mu} \\
 & \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} & = & 1 \\
 & \gamma, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\mu} & \geqq & \mathbf{0}.
\end{array} \tag{5.2}$$

This program has no feasible solution if the DMU under evaluation lies on the strong efficient part of the frontier because in this case $\widehat{\beta} = 0$ and hence $1/\widehat{\beta}$ is not defined. But this does not lead to a problem for obtaining optimal vectors because Färe et al. (2011) argue that the vector for efficient DMUs may be chosen arbitrarily. In this case they propose to use equal weights for all outputs which in our application would imply that all elements of $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ are set equal to $1/(s+r)$.

As mentioned above, the original approach by Färe et al. (2011) is based on a slacks-based distance measure. In the present setting the program to obtain this measure can be stated as

$$\begin{array}{llcl}
\max\limits_{\boldsymbol{\beta_y},\boldsymbol{\beta_u},\boldsymbol{\lambda}} & \mathbf{1}^T\boldsymbol{\beta_y} + \mathbf{1}^T\boldsymbol{\beta_u} & & \\
\text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\lambda} \\
 & \boldsymbol{y}_i + \boldsymbol{\beta_y} \odot \boldsymbol{e_y} & \leqq & \boldsymbol{Y}\boldsymbol{\lambda} \\
 & \boldsymbol{u}_i - \boldsymbol{\beta_u} \odot \boldsymbol{e_u} & = & \boldsymbol{U}\boldsymbol{\lambda} \\
 & \boldsymbol{\beta_y}, \boldsymbol{\beta_u}, \boldsymbol{\lambda} & \geqq & \mathbf{0}
\end{array} \tag{5.3}$$

where $\boldsymbol{e_y}$ ($\boldsymbol{e_u}$) denotes a vector containing one unit of each good (bad) output to render $\boldsymbol{\beta_y}$ ($\boldsymbol{\beta_u}$) a vector of dimensionless measures that can be summed up. Modifying this model by dividing each restriction on the good (bad) outputs by the amount of the associated good (bad) output of DMU i the problem reads as

$$\begin{array}{llcl}
\max\limits_{\boldsymbol{\beta_y},\boldsymbol{\beta_u},\boldsymbol{\lambda}} & \mathbf{1}^T\boldsymbol{\beta_y} + \mathbf{1}^T\boldsymbol{\beta_u} & & \\
\text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\lambda} \\
 & \frac{\boldsymbol{y}_i}{\boldsymbol{y}_i} + \frac{\boldsymbol{\beta_y}}{\boldsymbol{y}_i} & \leqq & \frac{\boldsymbol{Y}\boldsymbol{\lambda}}{\boldsymbol{y}_i} \\
 & \frac{\boldsymbol{u}_i}{\boldsymbol{u}_i} - \frac{\boldsymbol{\beta_u}}{\boldsymbol{u}_i} & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}}{\boldsymbol{u}_i} \\
 & \boldsymbol{\beta_y}, \boldsymbol{\beta_u}, \boldsymbol{\lambda} & \geqq & \mathbf{0}.
\end{array} \tag{5.4}$$

In this presentation of the model we slightly abuse the matrix notation. The fraction of two vectors refers to an element-wise division (similar to the Hadamard product) and the fraction of a matrix and a vector refers to an element-wise division of each column of the matrix by the

[55] For a further discussion of linear and nonlinear environmental DEA models see Zhou et al. (2008b).

vector. In the modified program the vectors $\boldsymbol{e_y}$ and $\boldsymbol{e_u}$ are replaced by $\frac{\mathbf{1}}{\boldsymbol{y}_i}$ and $\frac{\mathbf{1}}{\boldsymbol{u}_i}$. Therefore, $\boldsymbol{\beta_y}$ and $\boldsymbol{\beta_u}$ are again dimensionless. Denoting $\frac{\boldsymbol{\beta_y}}{\boldsymbol{y}_i} = \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}$ and $\frac{\boldsymbol{\beta_u}}{\boldsymbol{u}_i} = \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}$ the model becomes

$$
\begin{aligned}
\max_{\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}, \boldsymbol{\lambda}} \quad & \mathbf{1}^T \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} + \mathbf{1}^T \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} \\
\text{s.t.} \quad & \boldsymbol{x}_i \geqq \boldsymbol{X\lambda} \\
& \mathbf{1} + \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} \leqq \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\
& \mathbf{1} - \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} = \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\
& \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}, \boldsymbol{\lambda} \geqq \mathbf{0}.
\end{aligned}
\tag{5.5}
$$

This programming problem is linear, hence optimal values $\widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{y}}$, $\widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{u}}$ and $\widehat{\boldsymbol{\lambda}}$ can be calculated without transformation using the conventional simplex algorithm. However, the optimal values of (5.3) and (5.5) are not equal because (5.3) is not independent of the units in which the inputs and outputs are measured and hence the transformation leading to (5.5) changes the results. In contrast, the optimal values of the objective functions of programs (5.1) and (5.5) can be shown to be equal. A proof of this equality can be found in appendix C. Therefore, $\widehat{\beta} = \mathbf{1}^T \widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{y}} + \mathbf{1}^T \widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{u}}$ for a DMU under evaluation.

In the empirical part of this study we apply model (5.1) to an analysis of greenhouse gas emissions and calculate potential reductions given the directional vectors obtained from it.

5.2.2 A dynamic approach to optimal directions

In the previous section we have presented an application of the static model by Färe et al. (2011). Now we propose an extension of this model to a dynamic analysis. We propose that an optimal vector can be derived as the direction in which innovating DMUs have shifted the frontier of a technology between two periods. This dynamic approach can be summarized by three steps:

1. Calculate the direction of movement between periods t and $t+1$ for all DMUs.
2. Evaluate which of the DMUs is an innovator given the directions obtained in the first step.
3. Identify the reference innovator for each non-innovating DMU. Estimate the efficiency of the DMUs by using the directional vector of the reference innovator.

This model is based on dynamic nonparametric productivity measures which we have introduced in the chapter on general concepts. In particular, the Malmquist-Luenberger index proposed by Chung et al. (1997) who estimate this index by assuming that the vector of the efficiency analysis is given by equal weights for the good and bad outputs which in our model is the case $\boldsymbol{\alpha} = \boldsymbol{\delta} = 1/(r+s)$.

Other studies like Jeon and Sickles (2004) or Kumar (2006) follow this approach and also treat the reduction of bad outputs and the increase of good outputs as equally important. This is a very restrictive assumption and we propose to extend the above discussed approach of endogenous directional vectors to a dynamic analysis. Our approach is based on the idea of estimating the direction of shifts in the frontier. Färe et al. (2001) propose conditions to identify observations that shift the frontier and hence can be regarded as “innovators” of the technology. An innovating DMU can be identified by checking whether it fulfills the following conditions:

$$\widehat{\text{MLTech}}_i^{t,t+1} > 1 \tag{5.6}$$

$$\widehat{\beta}^t \left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i}\right) < 0 \tag{5.7}$$

$$\widehat{\beta}^{t+1} \left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i}\right) = 0. \tag{5.8}$$

The first condition states that technical progress must have occurred between two periods. The second states that the input-output combination of DMU i in $t+1$ must lie outside the technology in t and the third condition states that DMU i must be part of the frontier in $t+1$. If all conditions are met simultaneously then i is identified as an innovator or frontier-shifting DMU between the periods t and $t+1$. In the previous literature the above stated conditions are evaluated using the directional vectors specified earlier which assign all good and bad outputs the same weight. This direction of the analysis may not be the direction of the movement of the innovating DMUs. Hence, the direction of the shift of the frontier and the direction of the measurement of technical change as well as efficiency change may be different. For a graphical illustration of this argument consider figure 5.1 that depicts two DMUs (A and B) and the technological frontier for two periods t and $t+1$.

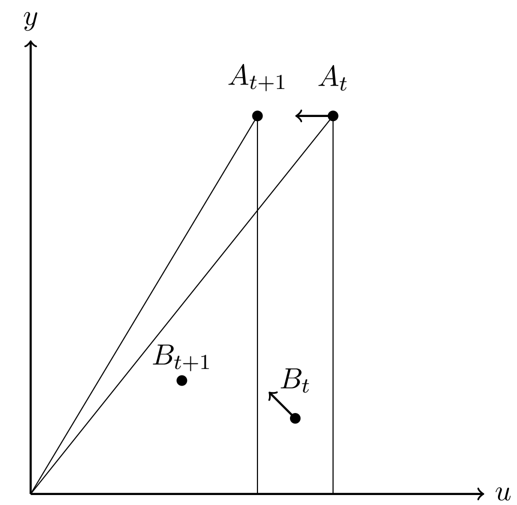

Figure 5.1: Example of a frontier shift

In this example DMUs A and B are supposed to use one unit of an input x in period t and $t+1$ to produce a single good output y and a single bad output u with the quantities of both outputs being indicated by the filled circles in the graph. Technical progress occurs between the two periods because DMU A is able to produce less bad output in period $t+1$ compared to the quantity in t holding input and good output constant. Hence, the direction which captures this movement and therefore would provide a plausible direction of the measurement of technical progress is given by ($\alpha = 0, \delta = 1$) or using the notation without weights $(0, -u)$ as indicated in figure 5.1 by the arrow associated with point A_t. However, in the standard Malmquist-Luenberger index the direction is defined as ($\alpha = 0.5, \delta = 0.5$) or $(y, -u)$ as indicated by the arrow for observation B_t.

To overcome this problem we propose to first determine the directional vectors of the DMUs in a dataset by measuring the direction of their movement between two periods. More precisely, the optimal directions are obtained by calculating changes in the output structure of the DMUs. A problem arises in this analysis because in contrast to the figure presented above where it is assumed that $x_t = x_{t+1}$ it is likely that the DMUs change their output as well as their input quantity between two periods. In the existing literature different approaches to this problem have been proposed. Färe and Grosskopf (2012) address technical change in a nonparametric setting by using the idea of a technical change matrix developed by Simon (1951). To calculate this matrix the technology matrices in t and $t+1$ need to be constructed. The technology matrix in t contains the input-output structure of all DMUs and can be written as:[56]

$$\boldsymbol{T}^t = \begin{bmatrix} -x_{t,11} & \cdots & -x_{t,1n} \\ \vdots & \vdots & \vdots \\ -x_{t,m1} & \cdots & -x_{t,mn} \\ y_{t,11} & \cdots & y_{t,1n} \\ \vdots & \vdots & \vdots \\ y_{t,s1} & \cdots & y_{t,sn} \\ -u_{t,11} & \cdots & -u_{t,1n} \\ \vdots & \vdots & \vdots \\ -u_{t,r1} & \cdots & -u_{t,rn} \end{bmatrix} \tag{5.9}$$

with each column referring to one DMU. Analogously the technology matrix in $t+1$, $\boldsymbol{T}^{t+1}$, can be constructed by collecting the input-outputs structure of all DMUs in $t+1$. Assuming that the analyzed DMUs are the same in each period (e.g. no DMU shuts down operations between the periods), the technological change matrix $\Delta\boldsymbol{T}^{t,t+1}$ can be calculated as

$$\Delta\boldsymbol{T}^{t,t+1} = \boldsymbol{T}^{t+1} - \boldsymbol{T}^t. \tag{5.10}$$

Inputs and undesirable outputs are included with a negative sign in the technology matrices so that the technological change matrix contains positive elements for inputs and bad outputs if they are reduced and for good outputs if they are increased between two periods. Färe and Grosskopf (2012) propose to use only those DMUs as reference observations which exhibit non-negative elements in their respective column of $\Delta\boldsymbol{T}^{t,t+1}$. This is a very restrictive assumption. For example, consider an observation that has reduced its input use and increased all but one output which it has decreased between the periods. Given the above stated assumption this DMU is excluded although it may have increased its productivity and hence may be an innovator. Moreover, this assumption may lead to situations where no DMU can be identified as a reference observation because none exhibits only non-negative elements.

In a different approach Otsuki (2012) proposes to measure the effect of different directional vectors on technical change by fixing the input vector over the analyzed periods. However, this

[56] In the original works by Simon (1951) and Färe and Grosskopf (2012) undesirable outputs are not incorporated. To show the similarity to our approach we include them.

vector is arbitrarily chosen and hence may not be related to the actual data situation.

In our model we build upon these ideas by proposing a dynamic approach that conducts an output-oriented analysis and obtains appropriate directional vectors by analyzing changes in the output structure of the innovating DMUs. The first step consists of identifying the optimal directions for each DMU. In contrast to Otsuki (2012) we analyze changes in the output structure by fixing the input vector of each DMU to the quantities actually used in period t. Hence, we first derive the hypothetical output quantities of the DMU under evaluation in period $t+1$ given the input vector of period t. This can be done by solving the following linear programming problem:

$$\begin{array}{llcl} \max\limits_{\lambda,\boldsymbol{y},\boldsymbol{u}} & \lambda & & \\ \text{s.t.} & \boldsymbol{x}_{t,i} & \geqq & \boldsymbol{x}_{t+1}\lambda \\ & \boldsymbol{y} & \leqq & \boldsymbol{y}_{t+1}\lambda \\ & \boldsymbol{u} & = & \boldsymbol{u}_{t+1}\lambda \\ & \lambda,\boldsymbol{y},\boldsymbol{u} & \geqq & \boldsymbol{0}. \end{array} \tag{5.11}$$

The right hand side of the input output structure shows that the boundary of this technology is given by the input-output combination of the DMU under evaluation in $t+1$ and all input-output combinations that result from proportionally scaling the vectors by the scalar λ. This last part follows because the technology is assumed to exhibit constant returns to scale. Maximizing λ leads to $\widehat{\boldsymbol{y}}$ and $\widehat{\boldsymbol{u}}$ values that are associated with at least one binding input constraint. Using this result the above stated linear programming problem can be easily solved for a particular DMU i by finding $\widehat{\lambda}$ such that

$$\widehat{\lambda} = \min_j \left\{ \frac{x_{t,ji}}{x_{t+1,ji}} \right\} \quad j = 1,\ldots,m. \tag{5.12}$$

The quantities of good and bad outputs in period $t+1$ given $\boldsymbol{x}_{t,i}$ can then be calculated as $\widehat{\boldsymbol{y}} = \boldsymbol{y}_{t+1,i}\widehat{\lambda}$ and $\widehat{\boldsymbol{u}} = \boldsymbol{u}_{t+1,i}\widehat{\lambda}$. These output quantities are used to identify observations that increased at least one good and/or decreased at least one bad output and hence are possible innovators. This variant of choosing reference observations is less restrictive than Färe and Grosskopf (2012) because it does not assume that all good and bad outputs have to change in an appropriate direction. Moreover, since we are correcting for changes in the input structure we may also consider observations which use more inputs in period $t+1$ compared to t. The optimal directional vectors $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\delta}}$ can be obtained by first setting the $\widehat{\alpha}$ and $\widehat{\delta}$ values for all good outputs which have been decreased and all bad outputs which have been increased to zero. The directions for the remaining good and bad outputs are then calculated by solving the nonlinear programming problem

$$\begin{array}{llcl} \min\limits_{\beta,\boldsymbol{\alpha}_e,\boldsymbol{\delta}_e,\boldsymbol{\lambda}} & & \beta & \\ \text{s.t.} & \boldsymbol{y}_{t,ie} + \beta\boldsymbol{\alpha}_e \odot \boldsymbol{y}_{t,ie} & = & \widehat{\boldsymbol{y}}_e \\ & \boldsymbol{u}_{t,ie} - \beta\boldsymbol{\delta}_e \odot \boldsymbol{u}_{t,ie} & = & \widehat{\boldsymbol{u}}_e \\ & \boldsymbol{1}^T\boldsymbol{\alpha}_e + \boldsymbol{1}^T\boldsymbol{\delta}_e & = & 1 \\ & \beta,\boldsymbol{\alpha}_e,\boldsymbol{\delta}_e & \geqq & \boldsymbol{0} \end{array} \tag{5.13}$$

which can be transformed into a linear programming problem similar to the model by Färe et al. (2011). The subscript e indicates that this program calculates weights only for those good and bad outputs which have changed between period t and $t+1$ with an appropriate direction. The weights are optimal in the sense that they lead to a minimal distance between the values of outputs obtained in t and those obtained in $t+1$ using the input vector of period t. We use these vectors to estimate the distance functions and the Malmquist-Luenberger index for all observations which are included in the above discussed programming problem to identify the innovators and hence the "innovating" directions which are the reference directions for all non-innovating DMUs. To choose among this set of vectors we calculate the euclidean distance between the innovating and the non-innovating DMUs. The closest innovator provides the appropriate directional vector used to estimate the efficiency of a non-innovating DMU. Hence, the chosen directional vector for a non-innovating DMU is the one for the closest innovating DMU.

5.3 Analysis of major GHG emitting countries

5.3.1 Data of the analyzed countries

In this section we apply the methods described above to a sample of major emitting countries. The data are obtained from two sources. World Bank (2011) provides data for total greenhouse gas (GHG) emissions of the countries (measured in thousand metric tons) which are computed as the sum of carbon dioxide (CO_2) emissions and the CO_2 equivalents of methane (CH_4) emissions, nitrous oxide (N_2O) emissions and other greenhouse gas emissions (i.e. fluorinated gases like hydrofluorocarbons, perfluorocarbons, and sulfur hexafluoride). We take the average of the data for the years 2000 and 2005 which are available for most countries of the world.

Regarding the other data, we use the Penn World Table (PWT) (Heston et al. (2011)) which provides national accounts data for the period 1950-2009 to compute real GDP as the desired output, the number of workers as labor input and cumulated investment using the perpetual inventory method as capital input. Actually used are the series real GDP per capita (rgdpl), real GDP per worker (rgdpwok), population (pop) and the investment share (ki). From these GDP is computed as rgdpl $\cdot$ pop, labor input as rgdpl $\cdot$ pop/rgdpwok and capital input from real investment data (ki/100) $\cdot$ rgdpl $\cdot$ pop by the perpetual inventory method.[57] Analogous to the GHG emissions we average the annual values over the period 2000-2005. Descriptive statistics of the data can be found in table 5.1.

We restrict the sample to those countries which are the largest emitters and together represent 90 percent of total world GHG emissions (in the average of 2000 and 2005). This leaves us with a sample of 62 countries. These countries are listed in appendix A in the order of emission volume.

[57] For the perpetual inventory method the initial capital stock is calculated by the formula $K_0 = I_0 \cdot (1+g)/(g+\delta)$ (see Park (1995)) where g is the average growth rate of investment over the first ten years for which investment data are available (or five years if $(g+\delta) < 0$) and δ is the depreciation rate fixed at 0.05. Subsequent capital stocks are calculated by the recursion $K_t = K_{t-1} \cdot (1-\delta) + I_t$ with $t = 1, 2, ...$

Table 5.1: Descriptive statics of the data (62 major emitting countries)

	Min	Median	Mean	Max	SD
Labor (1000 workers)	1115.39	11948.86	40249.36	742462.65	106171.75
Capital stock (bio. \$, 2005)	16.02	689.25	2200.93	28969.82	4560.95
GDP (bio. \$, 2005)	7.27	246.04	785.36	11591.17	1678.75
GHG (mio. tons of CO_2 eqv.)	64.79	181.99	548.36	7175.31	1210.00

To check whether our sample contains outliers we applied the method by Wilson (1995) to detect influential observations. The results indicated that for directions associated with large weights to the decrease of emissions Great Britain and Sweden are identified as influential observations. To check whether these observations are outliers we estimated the value for the directional distance function using direction $\delta = 1$ for the total sample and for the sample excluding these two observations. Histograms of the results are presented in figure 5.2. We find that a non-negligible share of countries exhibits only small inefficiencies ($\widehat{\beta} \leq 0.1$). Since this is observed irrespective of whether the two observations are included or not, we do not consider Great Britain and Sweden to be outliers. Hence, we include them into the subsequent analysis.

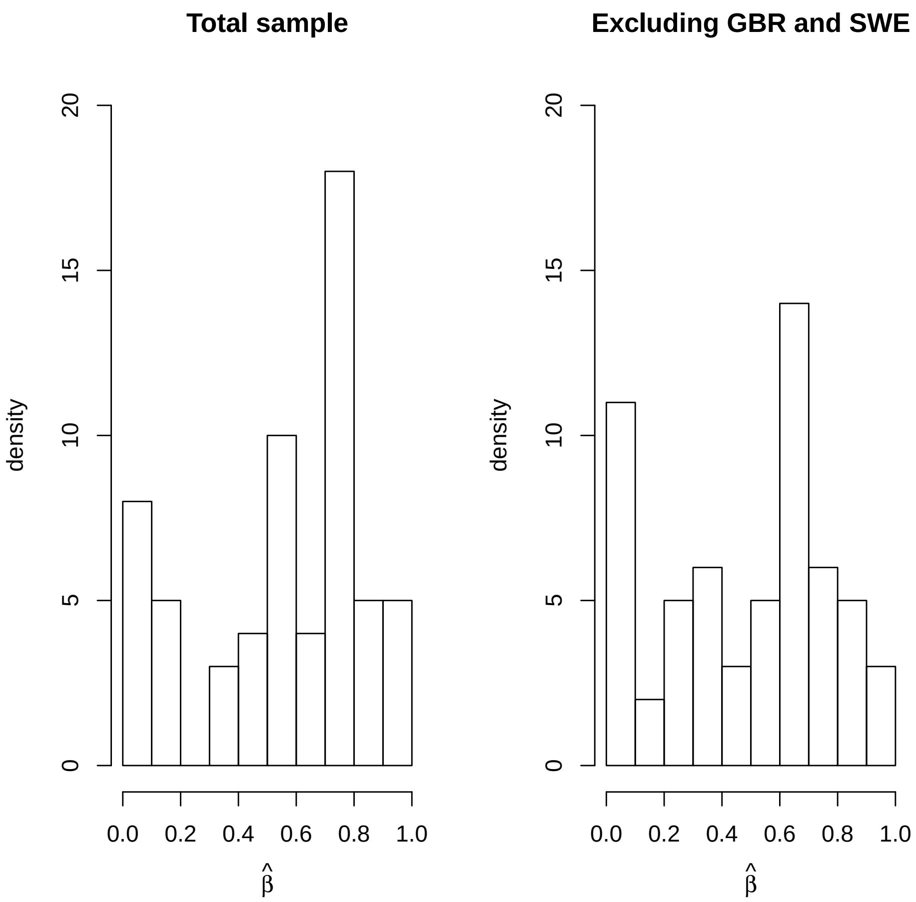

Figure 5.2: Effects of influential observations

5.3.2 Results of the analysis using fixed directions

In the following we present the results of our efficiency analysis for 62 major greenhouse gas emitting countries. We start by discussing the results for an analysis using a fixed grid of directional vectors. Afterwards we compare the results for the different optimization approaches presented in the last section.

In this analysis we estimate the efficiency of the countries using a grid of 11 different weights $\delta = 0, 0.1, \ldots, 1$ for the reduction of the bad output (GHG).[58] The corresponding weights for the enhancement of the single good output (GDP) are then $\alpha = 1 - \delta$. The reduction potentials of GHG as well as the potentials to increase GDP associated with each weight obtained by an analysis of the whole sample of countries can be found in columns two to five in the upper part of table 5.2. The first column of the table shows the weight of GHG used for the analysis. The second column represents the estimated absolute changes (in billions of international dollars of the year 2005, the currency of the Penn World tables) of GDP given that all countries remove their inefficiency, while column three shows the change relative to the current level of GDP. Columns four and five present the absolute (measured in million tons of CO_2 equivalents) and relative reduction of greenhouse gas emissions associated with the chosen weight δ.

The polar cases of the efficiency analysis are given by the weights $\delta = 0$ and $\delta = 1$. In the first case, efficiency is measured purely in terms of possible increases of GDP while the second case measures efficiency only in terms of reductions of GHG. Given that efficiency is measured with regard exclusively to increases of GDP the results show that the total GDP of the sample countries could increase by about 12600 billions of international dollars. This is approximately the GDP of the United States in the year 2005. In the opposite case, the GHG emissions could be reduced by nearly 17 billion tons if the countries increase their efficiency by focusing exclusively on the reduction of emissions. This amount of CO_2 equivalents exceeds the combined production of the two largest producers of carbon dioxide emissions, the United States and China. This result shows that a significant reduction of greenhouse gas emissions can be achieved without the invention of new technologies by just focusing on the reduction of inefficiencies in the abatement of bad outputs oriented at the efficient peers and adopting their practice.

Comparing the relative results for the cases $\delta = 0$ and $\delta = 1$ we find that the relative increase in GDP ($\approx 26\%$) is much lower than the relative decrease of emissions ($\approx 50\%$). This indicates that the inefficiency in the direction of the reduction of emissions is much higher than the inefficiency in the direction of the production of good outputs. However, these values do not provide information on whether this result holds likewise for all countries or whether it is driven by the efficiency of the largest countries. Therefore, table C.1 in appendix C contains the efficiency results for each country for the analysis with weights $\delta = 0$ and $\delta = 1$. The columns show that the inefficiency is larger for the majority of countries if the reduction of emissions is addressed ($\delta = 1$) compared to the case where efficiency is measured exclusively by potential increases in GDP ($\delta = 0$). Hence, we observe that the structure of the results for the larger countries is quite similar to the results of the smaller countries and we find consistent evidence of larger inefficiencies with regard to the abatement of bad outputs.

[58] The results in the rows “static” and “dynamic” will be discussed later on.

Supposing that the increase of good and the reduction of bad outputs are regarded as equally important targets ($\delta = 0.5$) we observe that GDP could be increased by more than 17% while emissions are reduced by more than 23%. Note that the percentage changes are not the same for GDP and GHG. This follows, because we present the aggregate changes and countries may exhibit different values of inefficiency. Hence, while for each country the relative changes in good and bad outputs are the same, on the aggregate level (which is also influenced by the size of the countries) they do not have to be equal.

Admittedly, the above presented results have to be interpreted with caution because our focus on the 62 major emitting countries leads to a very heterogeneous group of countries which is compared using a constant returns to scale technology. Therefore we have tested for constant returns to scale by the test procedure suggested by Simar and Wilson (2002). Regardless of which weight we used in the test (optimized and non-optimized) we could not reject the null hypothesis of constant returns to scale. The lowest p-value (0.163) was obtained by the test using the weights obtained by applying the method by Färe et al. (2011). Another issue is that the results may be driven by the heterogeneity of the countries. To examine the influence of heterogeneity we repeat the analysis but divide the countries into three groups. In the efficiency measurement a country in a specific group is only compared to peers out of this group. The intention is that peers determined in this way are more similar and thus more relevant for the country under evaluation.

To obtain groups of similar countries we separate them regarding to income per capita as well as regarding to the level of development measured by the human development index (HDI). The HDI is a composite index used for ranking countries and also reflects dimensions of human well-being beyond just income per capita. It is published in the annual Human Development Reports of the United Nations Development Programme (UNDP) and can be accessed by the website `http://hdr.undp.org/en/statistics/`. For both indicators we divide the countries in three groups pertaining to the lower, middle and upper terciles of the indicator. The results for the analysis using income terciles as well as the results using terciles of the HDI can also be found in table 5.2.

Comparing the results for the two methods to obtain the groups we observe that they do not differ significantly. In both cases the maximum reduction of emissions is about 12 bio. tons and the maximum increase in GDP is about 9 billions of international dollars. In contrast, comparing the results to the overall sample analysis we find that the numbers differ. The absolute values of reduction potentials of greenhouse gas emissions as well as potential increase in GDP are larger in the overall analysis indicating that the heterogeneity of countries influences the results. Peers are more similar to the countries in the respective groups analysis and this drives the smaller inefficiencies. However, even if we account for the heterogeneity by using the group analysis, the reduction potential for GHG emissions remains striking. Moreover, the difference between the maximum increase of GDP and the maximum decrease of GHG is also visible in the group analysis confirming larger inefficiencies with regard to the abatement of GHG.

To show in more detail how the choice of peers influences the results of the efficiency analysis and the implications for politics the results of the efficiency analysis using a grid of weights for

Table 5.2: Potentials to reduce GHG and increase GDP

	Overall				Income groups				HDI groups			
Weight	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$
$\delta = 0$	12604.02	25.89	0.00	0.00	8559.84	17.58	0.00	0.00	7811.11	18.48	0.00	0.00
$\delta = 0.1$	12300.29	25.26	1341.13	3.94	7987.44	16.40	837.53	2.46	7561.95	17.73	789.38	2.32
$\delta = 0.2$	11601.43	23.83	2812.14	8.27	7319.11	15.03	1713.00	5.04	7187.44	16.70	1688.12	4.97
$\delta = 0.3$	10689.08	21.95	4388.46	12.91	6640.74	13.64	2624.42	7.72	6782.97	15.62	2711.94	7.98
$\delta = 0.4$	9657.28	19.83	6085.77	17.90	5950.55	12.22	3589.83	10.56	6328.09	14.40	3898.93	11.47
$\delta = 0.5$	8398.01	17.25	7859.64	23.12	5145.55	10.57	4591.90	13.51	5623.49	12.67	5152.73	15.16
$\delta = 0.6$	6987.85	14.35	9772.44	28.74	4292.03	8.81	5730.08	16.85	4550.05	10.16	6157.18	18.11
$\delta = 0.7$	5463.21	11.22	11674.47	34.34	3361.34	6.90	6762.23	19.89	3501.95	7.70	7120.07	20.94
$\delta = 0.8$	3744.98	7.69	13669.67	40.21	2299.81	4.72	7777.58	22.88	2361.05	5.20	8041.17	23.65
$\delta = 0.9$	1896.52	3.89	15513.30	45.63	1274.91	2.62	10043.17	29.54	1300.79	2.78	10266.29	30.20
$\delta = 1$	0.00	0.00	16698.04	49.11	0.00	0.00	11429.08	33.62	0.00	0.00	11429.71	33.62
static	5807.70	11.93	13518.73	39.76	6224.22	12.78	6528.15	19.20	6111.11	12.14	7662.78	22.54
dynamic	5792.40	11.90	10980.64	32.30	3243.48	6.66	6786.82	19.96	3417.34	11.70	7053.50	20.75

Germany (as an example of a highly developed country) are presented in table 5.3. For Peru, a less developed country, the results are presented in table 5.4. For the analysis using the whole sample of 62 countries we observe that the inefficiency of Germany increases with the weight associated with the reduction of emissions. Therefore, Germany is less efficient with regard to the abatement of bad output than it is with the production of good outputs. Regarding the peers that are used to evaluate the efficiency of Germany we find that they change with the direction of measurement. Given that the efficiency is measured only with regard to the increase of GDP ($\delta = 0$) we find that peers for Germany are the United States, Great Britain and Austria.[59] If the reduction of emissions is assigned a large weight the United States and Great Britain are no longer peers for Germany and also the importance of Austria declines. In contrast, Sweden becomes a peer and thus if Germany aims at reducing its inefficiency with regard to the abatement of emissions it should focus on Sweden's technology. Since the reference countries for Germany are very similar in terms of per capita income and HDI the peers and the results for Germany do not change if the analysis is restricted to groups of similar countries.

The opposite is the case for Peru. Given the analysis using the whole sample of countries we find that the inefficiency of Peru is high regardless of which direction of the measurement is chosen. The peers for Peru are the United States and Sweden. Both countries are neither in terms of per capita income nor in terms of the HDI similar to Peru. Hence, the results of the efficiency analysis change largely if the reference group is restricted to more similar countries. Comparing Peru to countries which are similar with regard to the HDI we find that instead of the United States and Sweden, Turkey and Portugal are reference observations and compared with these more homogeneous peers Peru is found to be more efficient than in the analysis using the whole sample of countries. Even more striking, we find that given a comparison with countries that are similar in terms of per capita income Peru is classified as efficient and hence $\widehat{\lambda}$-PER (which due to space limitations is not included in table 5.4) is equal to one for all directions.

To gain more insights in the structure of the inefficiencies the next section presents the results for the analysis with directions computed with the different methods which have been explained in the previous section.

5.3.3 Results of the analysis using optimal directions

We first look at the δ values calculated with the two approaches outlined above. Histograms of the weights can be found in figure 5.3. For each of the reference groups (the total sample, the income groups and the HDI groups) a histogram of the weights (referred to as "Static") obtained by an application of the method by Färe et al. (2011) is presented. The histogram entitled "Dynamic" refers to the weights obtained by our novel dynamic approach. Note that the histograms do not include the results for the DMUs classified as efficient because as explained above the weights for these DMUs can not be uniquely determined and have been arbitrarily set equal to 0.5.

[59] Note that since the size of the peers differ, the λ-values can not be compared in terms of more or less important peers.

Table 5.3: Results of the efficiency analysis (Germany)

	Overall					Income groups					HDI groups				
Weight	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-GBR	$\widehat{\lambda}$-AUT	$\widehat{\lambda}$-SWE	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-GBR	$\widehat{\lambda}$-AUT	$\widehat{\lambda}$-SWE	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-GBR	$\widehat{\lambda}$-AUT	$\widehat{\lambda}$-SWE
$\delta = 0$	0.1101	0.0394	0.0886	7.9325	0.0000	0.1101	0.0394	0.0886	7.9325	0.0000	0.1101	0.0394	0.0886	7.9325	0.0000
$\delta = 0.1$	0.1206	0.0365	0.0922	8.0152	0.0000	0.1206	0.0365	0.0922	8.0152	0.0000	0.1206	0.0365	0.0922	8.0152	0.0000
$\delta = 0.2$	0.1333	0.0329	0.0965	8.1152	0.0000	0.1333	0.0329	0.0965	8.1152	0.0000	0.1333	0.0329	0.0965	8.1152	0.0000
$\delta = 0.3$	0.1489	0.0285	0.1019	8.2388	0.0000	0.1489	0.0285	0.1019	8.2388	0.0000	0.1489	0.0285	0.1019	8.2388	0.0000
$\delta = 0.4$	0.1687	0.0229	0.1088	8.3951	0.0000	0.1687	0.0229	0.1088	8.3951	0.0000	0.1687	0.0229	0.1088	8.3951	0.0000
$\delta = 0.5$	0.1946	0.0156	0.1177	8.5995	0.0000	0.1946	0.0156	0.1177	8.5995	0.0000	0.1946	0.0156	0.1177	8.5995	0.0000
$\delta = 0.6$	0.2299	0.0057	0.1298	8.8780	0.0000	0.2299	0.0057	0.1298	8.8780	0.0000	0.2299	0.0057	0.1298	8.8780	0.0000
$\delta = 0.7$	0.2796	0.0000	0.0308	8.6627	1.0205	0.2796	0.0000	0.0308	8.6627	1.0205	0.2796	0.0000	0.0308	8.6627	1.0205
$\delta = 0.8$	0.3154	0.0000	0.0000	6.6816	2.9655	0.3154	0.0000	0.0000	6.6816	2.9655	0.3154	0.0000	0.0000	6.6816	2.9655
$\delta = 0.9$	0.3544	0.0000	0.0000	4.0032	5.3236	0.3544	0.0000	0.0000	4.0032	5.3236	0.3544	0.0000	0.0000	4.0032	5.3236
$\delta = 1$	0.4044	0.0000	0.0000	0.5695	8.3465	0.4044	0.0000	0.0000	0.5695	8.3465	0.4044	0.0000	0.0000	0.5695	8.3465

Table 5.4: Results of the efficiency analysis (Peru)

	Overall					Income groups					HDI groups				
Weight	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-TUR	$\widehat{\lambda}$-PRT	$\widehat{\lambda}$-SWE	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-TUR	$\widehat{\lambda}$-PRT	$\widehat{\lambda}$-SWE	$\widehat{\beta}$	$\widehat{\lambda}$-USA	$\widehat{\lambda}$-TUR	$\widehat{\lambda}$-PRT	$\widehat{\lambda}$-SWE
$\delta = 0$	0.5036	0.1042	0.0000	0.0000	0.0452	0.0000	0.0000	0.0000	0.0000	0.0000	0.1603	0.0000	0.1314	0.4160	0.0000
$\delta = 0.1$	0.5261	0.0919	0.0000	0.0000	0.1117	0.0000	0.0000	0.0000	0.0000	0.0000	0.1635	0.0000	0.1243	0.4279	0.0000
$\delta = 0.2$	0.5508	0.0785	0.0000	0.0000	0.1844	0.0000	0.0000	0.0000	0.0000	0.0000	0.1669	0.0000	0.1169	0.4402	0.0000
$\delta = 0.3$	0.5778	0.0638	0.0000	0.0000	0.2643	0.0000	0.0000	0.0000	0.0000	0.0000	0.1705	0.0000	0.1092	0.4531	0.0000
$\delta = 0.4$	0.6077	0.0476	0.0000	0.0000	0.3524	0.0000	0.0000	0.0000	0.0000	0.0000	0.1741	0.0000	0.1012	0.4665	0.0000
$\delta = 0.5$	0.6408	0.0295	0.0000	0.0000	0.4501	0.0000	0.0000	0.0000	0.0000	0.0000	0.1780	0.0000	0.0929	0.4806	0.0000
$\delta = 0.6$	0.6778	0.0094	0.0000	0.0000	0.5591	0.0000	0.0000	0.0000	0.0000	0.0000	0.1820	0.0000	0.0841	0.4952	0.0000
$\delta = 0.7$	0.6688	0.0000	0.0000	0.0000	0.5872	0.0000	0.0000	0.0000	0.0000	0.0000	0.1862	0.0000	0.0750	0.5106	0.0000
$\delta = 0.8$	0.6269	0.0000	0.0000	0.0000	0.5504	0.0000	0.0000	0.0000	0.0000	0.0000	0.1906	0.0000	0.0654	0.5266	0.0000
$\delta = 0.9$	0.5899	0.0000	0.0000	0.0000	0.5179	0.0000	0.0000	0.0000	0.0000	0.0000	0.1952	0.0000	0.0554	0.5435	0.0000
$\delta = 1.0$	0.5570	0.0000	0.0000	0.0000	0.4891	0.0000	0.0000	0.0000	0.0000	0.0000	0.2001	0.0000	0.0448	0.5611	0.0000

The histograms for the static approach show that independent of which group is used as a reference most countries are assigned a weight for the reduction of bad outputs (δ) that is larger than 0.5 with an obvious peak for weights in the interval 0.9 to 1.[60] This confirms our findings from the analysis using a grid of weights. Given that we optimize the directions to maximize the inefficiency of the countries we find that most countries are assigned a direction that gives the reduction of greenhouse gas emissions a higher weight than the increase of GDP. Therefore, most countries show significant larger inefficiencies with regard to the reduction of emissions. However, for the analysis using the HDI groups we also find a smaller peak of weights lying between 0 and 0.1. Hence, for a minority of countries we find that accounting for differences in development may lower the inefficiency with regard to the abatement of emissions.

The results obtained by our dynamic model are shown in the upper right graph of figure 5.3. Note that the directions of the innovators have only been calculated for the analysis of the whole sample of countries. This has been done because innovators are assumed to shift the overall frontier. Shifts of the group frontier may not be due to the innovation of a country in this group but to shifts of the overall frontier. The intervals with density larger than zero indicate that the innovators have shifted the frontier in directions that lead to weights between 0 and 0.1 as well as 0.6 and 1. However, the graph shows that the vast majority of countries get assigned weights that are either near to 0 or near to 1. This indicates that the innovators which are more similar to the majority of non-innovating countries have predominantly focused on technical progress for either the reduction of bad or the increase of good outputs. This follows because in the dynamic approach the nearest innovator is chosen to calculate the direction of the efficiency measurement. The small number of countries that are assigned a direction that combines enhancement with regard to both the production of good and the abatement of bad outputs indicate that the countries that innovated in this direction are rather different compared to the remaining countries in the sample.

The aggregated potentials for reducing greenhouse gas emissions as well as increasing the production of GDP associated with the optimized weights can be found in the lower two lines of table 5.2. For visualizing the differences in the results of the efficiency measurement given the different directions, figure 5.4 shows the potentials changes in GDP and emissions for each of the chosen weights as well as for each type of reference groups. The effect of the optimization by maximizing the inefficiency for the directional distance function (the “static” approach) is clearly visible in figure 5.4 independent of the chosen reference group. Compared to the results for the grid analysis with fixed values of δ which are the same for all countries, the combination of GDP increase and decrease of GHG is located further to the right. This indicates that the optimization finds larger potentials to enhance efficiency than the non-optimizing approaches. Given the analysis of the whole sample of countries (depicted in the left graph of figure 5.4) we observe the static approach leads to a far larger decrease of emissions compared to the increase of GDP. This again confirms our previous findings that the largest inefficiencies are associated with the abatement of emissions. However, the results change if we account for the heterogeneity by restricting the reference groups. The results in table 5.2 show that the potentials to increase

[60] Note that the mean for the weights obtained by the “static” approach using the overall sample is 0.683 and the median is 0.714.

GDP are very constant about 12% and not influenced by the choosen reference group of countries. This is also visible from the plots in figure 5.4.[61] In contrast, we find large differences in the potentials to reduce emissions. The analysis using the whole sample of countries find the largest potentials of 39%, this number is lowered to 19% if countries are compared with regard to their per capita income and to 23% in the HDI group analysis. This shows that the heterogeneity of countries does not affect the efficiency measurement in terms of GDP enhancement but exerts significant effects on the reduction potentials of GHG emissions. The finding that the emission efficiency depends on the income group is in line with findings by Taskin and Zaim (2000). Note that this result is not driven by significant changes in the weights for the directional vectors as indicated by the relatively small differences in the histograms for the three groups. Combining this finding with the result that changing the reference group has a large effect on the potential decrease of emissions and nearly no effect on the potential decrease of emissions shows that the group specific frontier differs largely from the overall frontier with regard to the abatement of emissions. The maximal production of good output given inputs is not affected by the change of the frontier.

The results for the dynamic analysis show that in contrast to the static approach no maximization of the inefficiency is targeted. From figure 5.4 we find the combination of GDP increase and GHG decrease is for all groups in line with the results of the grid analysis using $\delta = 0.7$. Similar to the results of the static approach we find (see the last row of table 5.2) that the potentials to decrease emissions vary largely with the reference group used in the analysis. Moreover, for the analysis using HDI groups we observe that the potential change is close to the result obtained by the analysis of the overall sample. A difference can be found by comparing the results for the income groups. In this case the potential to decrease emissions lowers to 19% like in the static analysis but the potential to increase GDP also lowers to 7% which nearly half the potential obtained from the analysis using the whole sample.

Combining the results of large inefficiencies with regard to the emissions and smaller inefficiencies with regard to the production of good outputs with the direction of the movement of the innovators leads to an interesting conclusion. For most non-innovating countries the technical progress of the most similar (in terms of input and outputs quantities) innovator was oriented either only at the production of good outputs or at the abatement of bad outputs. Therefore the large difference in the efficiency results for good and bad outputs may indicate that the countries were more capable to follow technical progress with regard to the production of good outputs than with regard to the abatement of bad outputs. This can be interpreted as supporting the importance for technology transfers between countries in order to reduce the generation of emissions.

[61] Note that the scaling of the axis of each graph differ.

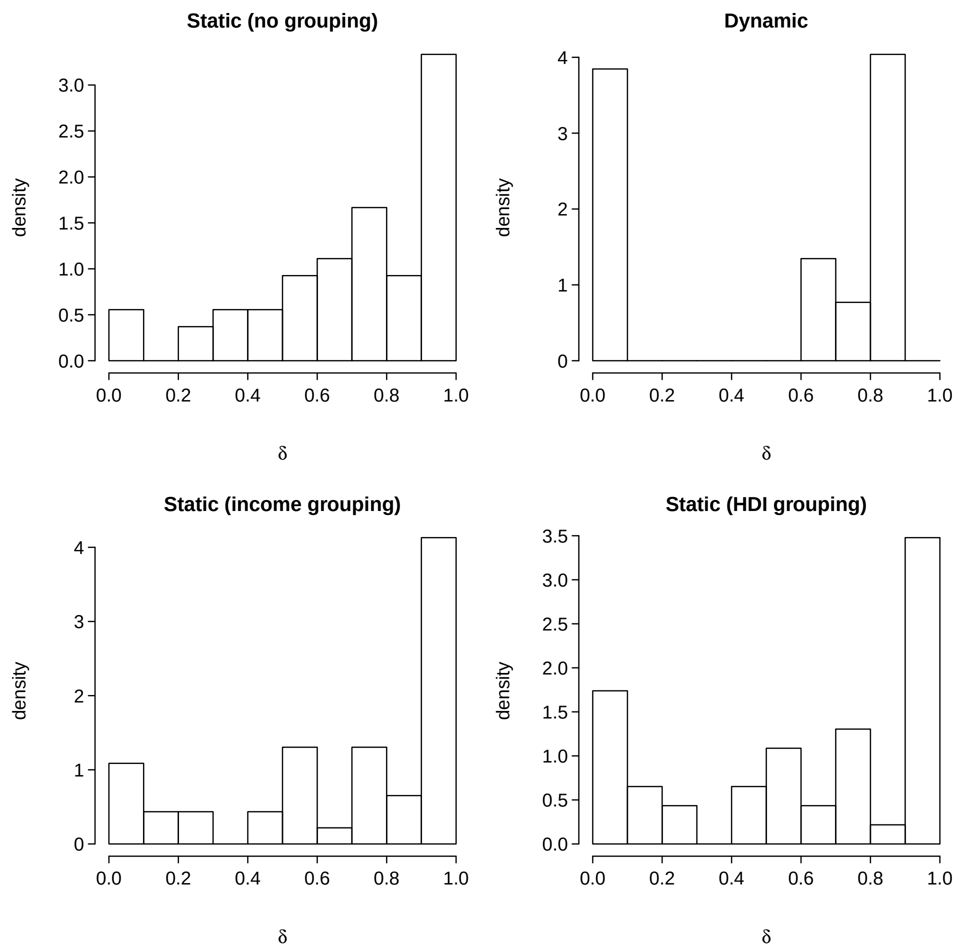

Figure 5.3: Histograms of optimal weights

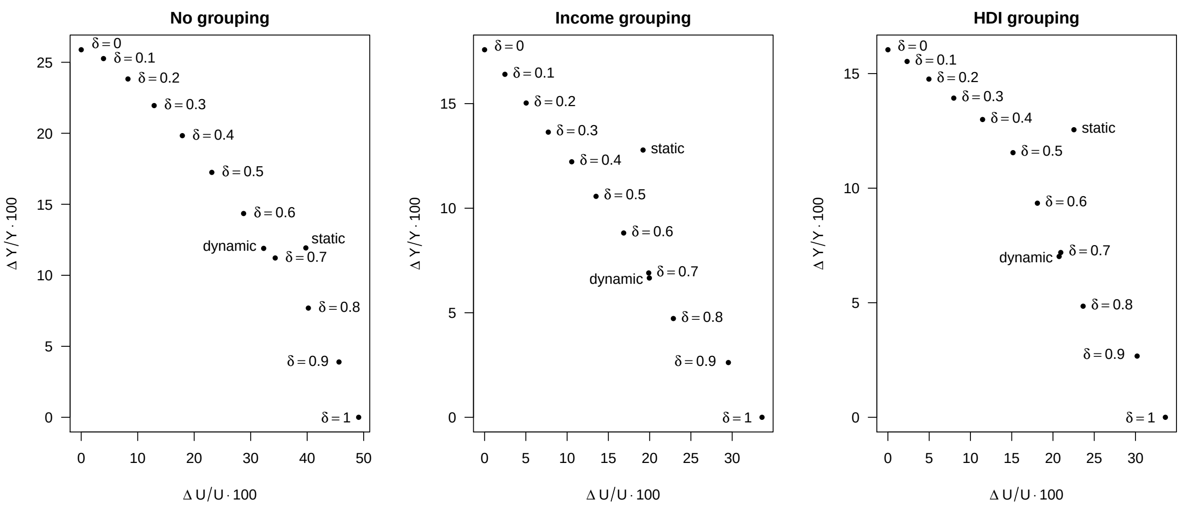

Figure 5.4: Comparison of results using optimized and non-optimized weights

The above discussed findings show sizable potentials for emission reduction expressed in absolute or in percentage terms. Permitting the selection of efficient peers from the whole sample these are nearly 40% for the static approach to determine the optimal directions and about 32% for the dynamic approach. When restricting the choice of efficient peers to the respective income or HDI groups potential emission reductions amount to roughly 20%.

How can these numbers be put into perspective? According to the results summarized in Aldy et al. (2010) a stabilization of the temperature increase of pre-industrial levels at about 2°C (resp. 3°C) requires a stabilization of CO_2 equivalents at a concentration of 450 ppm (resp. 550 ppm). According to the Stern (2007) review (see table 8.2) it would require emission reductions of about 70% of global emissions of the year 2005 until 2050 to reach the 2°C target and emission reductions of $25 - 30\%$ to reach the 3°C target. The latter target is thus not too far away from the reduction potentials we have calculated in this study. It is also a compromise since part of the inefficiency may be simultaneously realized in the form of increasing output by more than 10%.

The central question is how this numbers are to be assessed. They are, of course, rough estimates that are associated with measurement error. Furthermore, they are also biased estimates, although it is not a priori clear in which direction. On the one hand, the numbers are an underestimation of the reduction potentials since the distances measured with the directional distance function are downward biased estimates of the true but unknown values. Moreover, it is assumed that no emission reduction is possible for the frontier countries which are defining the best-practices (e.g. the US or Russia for most choices of δ). It is of course unrealistic that no emission reduction at all is feasible in these countries. As a consequence we are also faced with an underestimation for all other countries which are compared with these best practices. On the other hand, the numbers can be viewed as an overestimation of the reduction potentials since it is debatable whether the indicated best-practices can be adopted in reality. This is surely not reachable in the short run, but may also not be easily achieved in the longer run.

Nevertheless, the current analysis offers some quantitative orientation about potential emission reductions which could be realized by adopting best-practices and varying degrees of foregoing possible output enhancements.

5.4 Summary

In this study we have addressed the problem of endogenously determining the directions for directional distance functions. Modifying the approach by Färe et al. (2011) we demonstrated how optimal directions can be obtained in a static analysis of environmental efficiency. Moreover, we have proposed a new method to derive the directions in a dynamic setting. The directions obtained by this model estimate the movement of the frontier, hence the direction of technical change. With these methods we provide a solution to the practical issue that efficiency results depend on the directions chosen by the researcher. Therefore, endogenizing the directions eliminates this source of subjectivity.

That different directions indeed have a significant influence on the efficiency estimates was shown for an analysis of 62 major emitting countries. Using a grid of directions we found that the efficiency increases if the reduction of emissions is assigned a large weight. Moreover, applying the optimization approaches to calculate directions we found that large potentials to reduce greenhouse gases exist. While these potentials decrease if we account for the heterogeneity among the countries in our sample, they nonetheless provide an important possibility to limit climate change.

6 Macroeconomic productivity and the Kyoto Protocol

6.1 Motivation

In 1997 the Kyoto Protocol, the first international treaty setting legally binding reduction goals on greenhouse gas emissions (GHG), was signed. In a debate which followed the signing economists discussed whether the protocol would lead to significant macroeconomic costs associated with the reduction of emissions. Studies like Barrett (1998) and Nordhaus and Boyer (1999) which were conducted shortly after the Kyoto Protocol was signed argue that the protocol is likely to fail due to large costs which it burdens on countries which have to decrease their emissions significantly. In contrast, Böhringer and Vogt (2003, 2004) argue that the protocol and especially several changes in it (e.g. increased allowability of sinks) allow a "business as usual" strategy and therefore are not likely to induce significant costs. Macroeconomic costs arise if countries need to reallocate inputs from the production of good outputs (like GDP) to the abatement of bad outputs in order to decrease their emissions. Therefore, the amounts of produced desirable outputs are reduced.

In the literature on nonparametric analysis two different methods to account for the effect of environmental regulation have been applied. Färe et al. (1989) analyze efficiency of Swedish paper mills incorporating bad outputs alternatively as weak or strong disposable outputs. The difference in the efficiency results are interpreted as effects of regulation. This approach is criticized by Førsund (2009) for using an implausible assumption on bad outputs (strong disposability) to derive the measure. Alternatively, in a dynamic setting the effect of regulation can be analyzed by comparing the Malmquist and the Malmquist-Luenberger index which have been presented in the previous chapter on general concepts. Since shifts of inputs to abatement activities are not captured by the conventional Malmquist index but incorporated in the Malmquist-Luenberger index, comparing those indices allows to analyze whether significant reallocations have occurred (see e.g. Weber and Domazlicky (2001) and Jeon and Sickles (2004) for previous studies that compare these indices).

An empirical problem arises from the question whether potential differences in the productivity patterns are due to environmental regulations. Previous cross-country studies analyzing effects of the Kyoto Protocol (e.g. Yörük and Zaim (2005) and Kumar (2006)) use panel data regression models and compare regulated and unregulated countries. But this approach is problematic because it separates very heterogeneous groups of countries.[62] For example, Kumar (2006) compares Annex-I and Non-Annex-I countries of the Kyoto Protocol to account for regulations. Annex-I of the Kyoto Protocol covers those countries which have to reduce their production of carbon dioxide emissions. This approach is similar to separating industrialized and non-industrialized countries and therefore the results of the regressions are questionable given that only few control variables are included and the R^2-values are often low. Moreover, Simar and Wilson (2007) have shown that the application of simple regression models to explain efficiency and productivity differences leads to a severe problem of endogeneity and thus to invalid results.

[62] For example, in the last chapter we have shown how heterogeneity influences the results of macroeconomic efficiency analyses.

In contrast to previous studies we use an approach which is analogous to the difference-in-differences methods in regression analysis (see e.g. Wooldridge (2010)). We compare productivity changes between two groups of countries before and after regulations were implemented. More precisely, we focus on the EU15 countries which under the Kyoto Protocol have to reduce their emissions by 8% during the commitment period 2008-2012 compared to their emissions in 1990 (UNFCCC (2008)). In 1998 these countries have introduced the EU Burden Sharing Agreement (EU-BSA) which reallocated the overall reduction target among the countries. We use this program to analyze two groups of countries, those who are supposed to actually reduce their emissions and those who are allowed to increase emissions or hold them constant. Since the EU15 is a more homogenous group of countries to be compared than the groups exploited in the previous literature we argue that this is a more convincing way of estimating whether the Kyoto Protocol had an effect on macroeconomic productivity.

In the following section we present the theoretical background of the analysis of European countries. This section extends the introduction of dynamic methods in the chapter on general concepts.

6.2 Theoretical concepts for the analysis of European countries

6.2.1 Contemporaneous and sequential output sets

In the discussion of the general concepts we presented the DEA estimation of output sets in a static setting. Estimating the distance functions for the Malmquist and the Malmquist-Luenberger index makes it necessary to estimate these output sets given several time periods. In most dynamic analyses the output set for a period t ($t = 1, \ldots, T$) is constructed using only observations of period t. These output sets are called “contemporaneous” (see Shestalova (2003)). Given a sample of observed input-output combinations $\left(\boldsymbol{x}_{t,i}, \boldsymbol{y}_{t,i}, \boldsymbol{u}_{t,i}\right)$ with $i = 1, \ldots, n$ the DEA estimation of the contemporaneous environmental output sets of period t reads as

$$\widehat{P}_t^{\text{Env}}\left(\boldsymbol{x}_t\right) = \left\{\left(\boldsymbol{y}_t, \boldsymbol{u}_t\right) \in \mathbb{R}_+^{s+r} : \boldsymbol{x}_t \geqq \boldsymbol{X}_t\boldsymbol{\lambda}_t, \boldsymbol{y}_t \leqq \boldsymbol{Y}_t\boldsymbol{\lambda}_t, \boldsymbol{u}_t = \boldsymbol{U}_t\boldsymbol{\lambda}_t, \boldsymbol{\lambda}_t \geqq \boldsymbol{0}\right\} \tag{6.1}$$

where $\boldsymbol{X}_t$ denotes the $m \times n$ matrix of inputs, $\boldsymbol{Y}_t$ denotes the $s \times n$ matrix of good outputs and $\boldsymbol{U}_t$ denotes the $r \times n$ matrix of undesirable outputs of period t. $\boldsymbol{\lambda}_t$ is a vector of weight factors with $\boldsymbol{\lambda}_t \geqq 0$ indicating constant returns to scale. The output sets not accounting for the production of bad outputs $\left(\widehat{P}_t\left(\boldsymbol{x}_t\right)\right)$ can be constructed by eliminating the constraint for $\boldsymbol{u}_t$.

In contrast, “sequential” output sets for t are constructed by using observations from period t and all previous time periods (see Tulkens and Vanden Eeckaut (1995)). Formally, the sequential output sets are given by:[63]

$$\tilde{P}_t^{\text{Env}}\left(\boldsymbol{x}_t\right) = \text{convex}\left\{P_1^{\text{Env}}\left(\boldsymbol{x}_t\right) \cup P_2^{\text{Env}}\left(\boldsymbol{x}_t\right) \cup \cdots \cup P_t^{\text{Env}}\left(\boldsymbol{x}_t\right)\right\} \tag{6.2}$$

Hence, the sequential output sets in t are the convex unions of all contemporaneous output sets

[63] The operator “convex” means that the union set also contains all convex combinations of the sets. This operator is needed because the union of convex sets is not necessarily a convex set.

from period 1 up to period t given $\boldsymbol{x}_t$.
The DEA estimation of the sequential environmental output sets reads as

$$\widehat{\tilde{P}}_t^{\text{Env}}(\boldsymbol{x}_t) = \left\{(\boldsymbol{y}_t, \boldsymbol{u}_t) \in \mathrm{R}_+^{s+r} : \boldsymbol{x}_t \geqq \tilde{\boldsymbol{X}}_t\tilde{\boldsymbol{\lambda}}_t, \boldsymbol{y}_t \leqq \tilde{\boldsymbol{Y}}_t\tilde{\boldsymbol{\lambda}}_t, \boldsymbol{u}_t = \tilde{\boldsymbol{U}}_t\tilde{\boldsymbol{\lambda}}_t, \tilde{\boldsymbol{\lambda}}_t \geqq \boldsymbol{0}\right\}. \tag{6.3}$$

Here, $\tilde{\boldsymbol{X}}_t$ denotes the $m \times (n \cdot t)$ matrix of inputs, $\tilde{\boldsymbol{Y}}_t$ represents the $s \times (n \cdot t)$ matrix of good outputs and $\tilde{\boldsymbol{U}}_t$ denotes the $r \times (n \cdot t)$ matrix of bad outputs from period 1 to t. $\tilde{\boldsymbol{\lambda}}_t$ represents the $(n \cdot t) \times 1$ vector of weight factors. To simplify the notation we assume that the dataset is a balanced panel. Hence, in each period n columns are attached to the matrices. As in the case of contemporaneous output sets the output sets ignoring pollution can be constructed by eliminating the constraint for $\boldsymbol{u}_t$.

To visualize the difference of contemporaneous and sequential output sets the following figure 6.1 presents a contemporaneous output set for periods t and $t+1$ as well as the resulting sequential output set for period $t+1$.

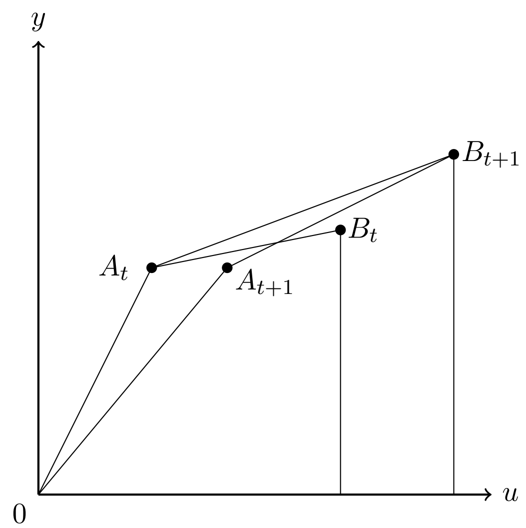

Figure 6.1: Example of contemporaneous and sequential output sets

The frontier of the contemporaneous output set for period t $(t+1)$ is given by $\overline{0A_tB_t}$ $\left(\overline{0A_{t+1}B_{t+1}}\right)$ and the vertical extension to B_t (B_{t+1}). The frontier of the sequential output set for period $t+1$ is constructed by $\overline{0A_tB_{t+1}}$ and the vertical extension to B_{t+1}.

In the following analysis we will apply the Malmquist and the Malmquist-Luenberger index introduced in the chapter on general concepts. Shestalova (2003) has shown that the Malmquist index captures effects of business cycles on productivity in the technical change component, e.g. a recession is indicated as technical regress, if it is computed based on contemporaneous output sets. Using sequential output sets this effect is captured in the efficiency change component. To exclude the effect of business cycles on productivity measurement we calculate pure technical change by applying the decomposition proposed by Färe et al. (1992) and the sequential output sets.[64] In contrast to this, we use contemporaneous output sets to calculate pure efficiency

[64] The use of sequential output sets excludes the possibility of technological regress by construction. While this may be problematic for an industry-level analysis (e.g. Shestalova (2003) provides the example of the

change. The productivity change adjusted for business cycles is then obtained by multiplying pure technical and pure efficiency change.[65] Excluding the effects of business cycles prevents the comparison of different groups of countries from being biased by different industry structures of the countries. For example, energy-intensive industries react very sensitive to business cycles (see Moomaw (1996)). When including emissions into the analysis the same methodology is used for the Malmquist-Luenberger index for which a sequential model has been developed by Oh and Heshmati (2010).

6.2.2 Bootstrapping the indices

The Malmquist index as well as the Malmquist-Luenberger index are constructed using ratios of distance functions. Since the nonparametric technology estimation and resulting the distance functions are biased estimators the indices are biased, too. Simar and Wilson (1999) provide a method for correcting the bias of the Malmquist index. They take into account the intertemporal dependencies between the distance functions and hence create bootstrap samples simultaneously for two periods. This method has also been extended to the analysis of the Malmquist-Luenberger index (see e.g. Jeon and Sickles (2004)). The algorithm to obtain bootstrap samples to correct the bias of the indices can be summarized by the following steps:[66]

1. Estimate the directional distance functions $\widehat{\beta}^t\left(\boldsymbol{x}_{t,i}, \boldsymbol{y}_{t,i}, \boldsymbol{u}_{t,i}; \boldsymbol{g}_t\right) = \widehat{\beta}_i^t\left(t\right)$ and $\widehat{\beta}^{t+1}\left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i}; \boldsymbol{g}_{t+1}\right) = \widehat{\beta}_i^{t+1}\left(t+1\right)$ for $i = 1, \ldots, n$. Denote

$$\boldsymbol{A} = \left[1 + \widehat{\beta}_1^t\left(t\right), \ldots, 1 + \widehat{\beta}_n^t\left(t\right)\right]^T \tag{6.4}$$

and

$$\boldsymbol{B} = \left[1 + \widehat{\beta}_1^{t+1}\left(t+1\right), \ldots, 1 + \widehat{\beta}_n^{t+1}\left(t+1\right)\right]^T \tag{6.5}$$

with the elements of $\boldsymbol{A}$ and $\boldsymbol{B}$ being bounded from below at unity.

2. Calculate the matrix of reflected values as

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{2} - \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{2} - \boldsymbol{A} & \boldsymbol{2} - \boldsymbol{B} \\ \boldsymbol{A} & \boldsymbol{2} - \boldsymbol{B} \end{bmatrix} \tag{6.6}$$

where $\boldsymbol{2}$ denotes a $n \times 1$ vector of twos.

mining industry where regress may be observable) technological regress seems unlikely to be observed on a macroeconomic level of industrialized countries.

[65] Note that in general the results of the index are not the same using sequential and contemporaneous outputs sets. Hence, it is not possible to uniquely calculate the effect of business cycles. However, it is possible to exclude these effects as explained above.

[66] In this presentation we follow Jeon and Sickles (2004, pp. 588 - 589) but we are using our own definition of the variables. Moreover, we adapt the approach to the Malmquist-Luenberger index while Jeon and Sickles (2004) present it for the Malmquist index.

3. Estimate the covariance matrix $\widehat{\boldsymbol{V}}$ of the columns of $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$ and the covariance matrix $\widehat{\boldsymbol{V}}_R$ of the columns of $\begin{bmatrix}\boldsymbol{2}-\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$.[67]
4. Draw with replacement n rows from $\boldsymbol{\Delta}$ and denote the resulting $n\times 2$ matrix $\boldsymbol{\Delta}^*$. Moreover, denote the elements of $\boldsymbol{\Delta}^*$ as δ_{ij} with $i=1,\ldots,n$ and $j=1,2$. The mean values of the columns of $\boldsymbol{\Delta}^*$ are given by $\bar{\delta}_j=\frac{1}{n}\sum_{i=1}^{n}\delta_{ij}$.
5. Calculate the lower triangular matrix of a cholesky decomposition of the covariance matrix $\widehat{\boldsymbol{V}}$ and denote the result

$$\boldsymbol{L}=\begin{bmatrix} l_1 & 0 \\ l_2 & l_3 \end{bmatrix}. \tag{6.7}$$

6. Draw $2n$ times form a standard normal distribution and form n pairs (z_1,z_2).
7. Construct the $n\times 2$ matrix $\boldsymbol{E}$ which ith row consists of $(l_1z_1, l_2z_1+l_3z_2)$ if the corresponding row of the $n\times 2$ matrix $\boldsymbol{\Delta}^*$ was drawn from $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$ or $\begin{bmatrix}\boldsymbol{2}-\boldsymbol{A} & \boldsymbol{2}-\boldsymbol{B}\end{bmatrix}$ and of $(l_1z_1,-l_2z_1+l_3z_2)$ if the corresponding row of $\boldsymbol{\Delta}^*$ was drawn from $\begin{bmatrix}\boldsymbol{2}-\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$ or $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{2}-\boldsymbol{B}\end{bmatrix}$. The rows of $\boldsymbol{E}$ simulate draws from bivariate normal distributions $N\left(\boldsymbol{0},\widehat{\boldsymbol{V}}\right)$ and $N\left(\boldsymbol{0},\widehat{\boldsymbol{V}}_R\right)$.
8. Calculate the $n\times 2$ matrix $\boldsymbol{\Gamma}$ as

$$\boldsymbol{\Gamma}=\left(1+h^2\right)^{-1/2}\left(\boldsymbol{\Delta}^*+h\boldsymbol{E}-\boldsymbol{C}\begin{bmatrix}\bar{\delta}_1 & 0\\ 0 & \bar{\delta}_2\end{bmatrix}\right)+\boldsymbol{C}\begin{bmatrix}\bar{\delta}_1 & 0\\ 0 & \bar{\delta}_2\end{bmatrix} \tag{6.8}$$

where $\boldsymbol{C}$ denotes a $n\times 2$ matrix of ones and the bandwidth h is equal to $\left(\frac{4}{5n}\right)^{1/6}$. This bandwidth is chosen by Simar and Wilson (1999) following suggestions by Silverman (1986).
9. To remove the reflection around unity denote the elements of $\boldsymbol{\Gamma}$ as γ_{ij} and set $\gamma_{ij}^*=\gamma_{ij}$ if $\gamma_{ij}\geqq 1$ and $\gamma_{ij}^*=2-\gamma_{ij}$ otherwise. Denote the simulated deviations from the frontier as $\beta_i^{t*}(t)=\gamma_{i1}^*-1$ and $\beta_i^{t+1*}(t+1)=\gamma_{i2}^*-1$.[68]
10. Use the simulated distance functions to construct a bootstrap sample $\left(\boldsymbol{x}_{t,i}^*,\boldsymbol{y}_{t,i}^*,\boldsymbol{u}_{t,i}^*\right)$ with $i=1,\ldots,n$ as

$$\boldsymbol{x}_{t,i}^*=\boldsymbol{x}_{t,i} \tag{6.9}$$
$$\boldsymbol{y}_{t,i}^*=\left(1+\widehat{\beta}_i^t(t)\right)/\left(1+\beta_i^{t*}(t)\right)\cdot\boldsymbol{y}_{t,i} \tag{6.10}$$
$$\boldsymbol{u}_{t,i}^*=\left(1-\widehat{\beta}_i^t(t)\right)/\left(1-\beta_i^{t*}(t)\right)\cdot\boldsymbol{u}_{t,i} \tag{6.11}$$

[67] Note that the covariance matrix $\widehat{\boldsymbol{V}}$ of the columns of $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\end{bmatrix}$ and the covariance matrix of the columns of $\begin{bmatrix}\boldsymbol{2}-\boldsymbol{A} & \boldsymbol{2}-\boldsymbol{B}\end{bmatrix}$ are equal. The same holds for $\widehat{\boldsymbol{V}}_R$ and $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{2}-\boldsymbol{B}\end{bmatrix}$.

[68] Note that γ_{ij} is not bounded above by 2. Hence, the resulting simulated distance function may become larger than one resulting in negative values of the simulated amounts of undesirable outputs. This problem is not addressed in the previous literature but occurs in our application. To overcome this problem and to avoid a spurious mass of observations at 2 which would result if all values larger than 2 are set equal to 2 we replace these values by draws from a $U(\min(\gamma_{i1}),2)$ or $U(\min(\gamma_{i2}),2)$ distribution. The choice of the distribution depends on whether the value that should be replaced is part of the first or the second column of $\boldsymbol{\Gamma}$.

and for period $t+1$ as

$$\boldsymbol{x}^{*}_{t+1,i} = \boldsymbol{x}_{t+1,i} \tag{6.12}$$

$$\boldsymbol{y}^{*}_{t+1,i} = \left(1+\widehat{\beta}_i^{t+1}(t+1)\right) / \left(1+\beta_i^{t+1*}(t+1)\right) \cdot \boldsymbol{y}_{t+1,i} \tag{6.13}$$

$$\boldsymbol{u}^{*}_{t+1,i} = \left(1-\widehat{\beta}_i^{t+1}(t+1)\right) / \left(1-\beta_i^{t+1*}(t+1)\right) \cdot \boldsymbol{u}_{t+1,i} \tag{6.14}$$

11. Repeat steps 4 to 10 B times to generate B bootstrap samples. Use these bootstrap samples to estimate the bootstrap Malmquist-Luenberger indices $\widehat{\mathrm{ML}}_{bi}^{t,t+1}$ and their components $\widehat{\mathrm{MLEff}}_{bi}^{t,t+1}$ and $\widehat{\mathrm{MLTech}}_{bi}^{t,t+1}$.

The obtained values can be used to estimate the bias of the Malmquist-Luenberger index as[69]

$$\widehat{\mathrm{Bias}}_B\left(\widehat{\mathrm{ML}}_i^{t,t+1}\right) = \frac{1}{B}\sum_{b=1}^{B}\left[\widehat{\mathrm{ML}}_{bi}^{t,t+1} - \widehat{\mathrm{ML}}_i^{t,t+1}\right] \tag{6.15}$$

The index is corrected for the bias as

$$\widehat{\mathrm{ML}}_i^{t,t+1,\mathrm{BC}} = \widehat{\mathrm{ML}}_i^{t,t+1} - \widehat{\mathrm{Bias}}_B\left(\widehat{\mathrm{ML}}_i^{t,t+1}\right) \tag{6.16}$$

if

$$\frac{\left|\widehat{\mathrm{Bias}}_B\left(\widehat{\mathrm{ML}}_i^{t,t+1}\right)\right|}{\sigma_{\widehat{\mathrm{ML}}_{bi}^{t,t+1}}} \geq \frac{1}{\sqrt{3}} \tag{6.17}$$

where $\sigma_{\widehat{\mathrm{ML}}_{bi}^{t,t+1}}$ denotes the standard deviation of the bootstrapped values of the index.[70]

In our analysis we correct the bias using $B = 2000$ bootstrap replications. Note that in the case of the technical change component of the sequential index we only correct for the bias if the above condition holds and if the resulting bias corrected technical change component is not smaller than unity to exclude technical regress as it is the case in the theoretical construction of this index.

6.2.3 A statistical test for the comparison of productivity results

Our following analysis of European countries evaluates whether there is an effect of the Kyoto Protocol on the productivity by comparing two groups of countries before and after the protocol was signed. This leads to a difficulty regarding an appropriate test to compare the groups. Compared to microeconomic applications (see e.g. the chapter on an efficiency analysis of automobiles) the number of observations is considerably smaller. This makes it necessary to find a test that performs well even in relatively small samples. A test accounting for this problem is given by the test by Li (1996) which has been adapted to the context of nonparametric

[69] For the components of the indices the method can be applied analogously.

[70] This threshold is introduced because the bias corrected indices may have a larger mean square error than the original indices. Therefore, Simar and Wilson (1999) follow Efron and Tibshirani (1993) and correct only if the bias is large compared to the standard deviation of the bootstrap indices.

estimates by Simar and Zelenyuk (2006).[71] However, the application of this tests also leads to a problem because while other tests analyze shifts in the distributions (like the Wilcoxon ranksum test) or dominance relations (see the test for stochastic dominance applied in chapter 3 of this dissertation) this test does only analyze whether two distributions are equal or not. Hence, to derive more detailed results we analyze density plots in combination with the test results. In the following we describe this test and its application to the analysis of productivity changes.

Consider two random variables V and W for which random samples (v_i with $i = 1, \dots, n_V$ and w_i with $i = 1, \dots, n_W$) are available. In our analysis these samples are the productivity (or efficiency and technical) changes of two different groups of European countries. Let f_l with $l = V, W$ denote the density functions of the random variables, then the null and the alternative hypothesis of the test are given by

$$\begin{aligned} H_0 &: f_V(z) = f_W(z) \\ H_1 &: f_V(z) \neq f_W(z). \end{aligned} \tag{6.18}$$

These hypotheses are tested by considering the integrated squared difference of the density functions

$$\begin{aligned} T &= \int (f_V(z) - f_W(z))^2 \, dz = \int f_V^2(z)dz + \int f_W^2(z)dz - 2\int f_V(z)f_W(z)dz \\ &= \int f_V(z)dF_V(z) + \int f_W(z)dF_W(z) - \int f_V(z)dF_W(z) - \int f_W(z)dF_V(z). \end{aligned} \tag{6.19}$$

Li (1996) showed that a consistent estimation of this statistic can be obtained by replacing the unknown distribution functions by the empirical distribution functions and the density functions by kernel density estimators. The estimation of the test statistic then reads as

$$\begin{aligned} \widehat{T} = \Bigg\{ & \frac{1}{hn_V(n_V-1)} \sum_{j=1}^{n_V} \sum_{k\neq j, k=1}^{n_V} K\left(\frac{v_j - v_k}{h}\right) \\ & + \frac{1}{hn_W(n_W-1)} \sum_{j=1}^{n_W} \sum_{k\neq j, k=1}^{n_W} K\left(\frac{w_j - w_k}{h}\right) \\ & - \frac{1}{hn_V(n_W-1)} \sum_{j=1}^{n_W} \sum_{k\neq j, k=1}^{n_V} K\left(\frac{w_j - v_k}{h}\right) \\ & - \frac{1}{hn_W(n_V-1)} \sum_{j=1}^{n_V} \sum_{k\neq j, k=1}^{n_W} K\left(\frac{v_j - w_k}{h}\right) \Bigg\} \end{aligned} \tag{6.20}$$

where $K(.)$ denotes the kernel function and h the bandwidth of the kernel estimation. In our application we follow Simar and Zelenyuk (2006) and use a Gaussian kernel function. Moreover, the rule-of-thumb by Silverman (1986) is used to estimate the optimal bandwidth.[72] In the test statistic the center term ($k = j$) is removed because Li (1996) showed that this estimation

[71] For macroeconomic applications of the test see e.g. Kumar and Russell (2002) and Henderson and Zelenyuk (2007).

[72] We apply this method because the Monte-Carlo simulations of the test's performance are based on it as well.

performs better using Monte-Carlo simulations. Moreover, he proved that

$$\widehat{J} = \frac{n_V h^{1/2} \widehat{T}}{\widehat{\sigma}} \xrightarrow{d} N(0,1) \tag{6.21}$$

Here,

$$\begin{aligned}\widehat{\sigma}_T^2 = 2\Bigg\{ & \frac{1}{hn_V^2} \sum_{j=1}^{n_V} \sum_{k=1}^{n_V} K\left(\frac{v_j - v_k}{h}\right) + \frac{\lambda^2}{hn_W^2} \sum_{k=1}^{n_W} \sum_{j=1}^{n_W} K\left(\frac{w_j - w_k}{h}\right) \\ & - \frac{\lambda}{hn_V n_W} \sum_{j=1}^{n_W} \sum_{k=1}^{n_V} K\left(\frac{w_j - v_k}{h}\right) - \frac{\lambda}{hn_V n_W} \sum_{j=1}^{n_V} \sum_{k=1}^{n_W} K\left(\frac{v_j - w_k}{h}\right) \Bigg\} \\ & \times \int K^2(u) du, \end{aligned} \tag{6.22}$$

with $\lambda = \frac{n_V}{n_W}$.

In our application we use a gaussian kernel function to estimate the density functions. Therefore, the last term in the expression for $\widehat{\sigma}_T^2$, $\int K^2(u)du$, is given by $\frac{1}{2\sqrt{\pi}} = 0.2821$ (see Pagan and Ullah (1999, p. 73)).

Simar and Zelenyuk (2006) have proposed a bootstrap algorithm based on Li (1999) to estimate the p-value of the above defined test. This bootstrapping is necessary because the standard normal distribution of the test statistic is only an asymptotic result with little relevance for practical implementations. Since our study aims at comparing the bias corrected productivity, efficiency and technical changes we adapt their algorithm to our analysis. In the following we present the different steps to bootstrap the p-value of the test for a comparison of two groups with regard to their technical change. The same procedure can be applied to the productivity and efficiency change. However, in this case step 1b.) is not needed because it addresses the problem of a boundary at unity which does only occur for the technical change component. The steps of the bootstrapping procedure are:

1.) For each observation ($i = 1, \ldots, n$) estimate the technical change. Denote the results as $\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}}$.

1b.) Smooth the estimated values of $\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}}$ according to the following procedure

$$\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}*} = \begin{cases} \widehat{\text{MLTech}}_i^{t,t+1,\text{BC}} + \epsilon_i & \text{if } \widehat{\text{MLTech}}_i^{t,t+1,\text{BC}} = 1 \\ \widehat{\text{MLTech}}_i^{t,t+1,\text{BC}} & \text{otherwise} \end{cases} \tag{6.23}$$

with $\epsilon_i \sim U(0, a-1)$ where a is the 5%-quantile of the empirical distribution function of all observations exhibiting $\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}} > 1$. In Simar and Zelenyuk (2006) this procedure is called “algorithm 2”. It accounts for the problem of a spurious mass of observations at the boundary of one. Another approach, called “algorithm 1”, simply removes all observations with $\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}} = 1$ from the sample.[73]

[73] In our application we estimate p-values for both methods.

2.) Form subsamples V and W of the $\widehat{\text{MLTech}}_i^{t,t+1,\text{BC}*}$ values. These are the groups which shall be compared with regard to the equality of their distributions.

3.) Estimate the statistic $\widehat{J}$ using V and W and the bandwidth $h = \min(h_V, h_W)$. The bandwidth for both groups is obtained by using the rule-of-thumb by Silverman (1986).

4.) Draw n_V and n_W observations with replacement from the larger one of the two subsamples V and W and denote the new subsamples V^* and W^*.[74]

5.) Estimate the statistic $\widehat{J}^*$ using the new subsamples. The bandwidth is given by $h^* = \min(h_{V^*}, h_{W^*})$ using the same rule as in step 3.

6.) Repeat steps 4.) and 5.) B times to obtain $\widehat{J}_b^*$ $(b = 1, \ldots, B)$ estimates.

7.) Calculate the bootstrapped p-value as

$$p = \frac{1}{B} \sum_{b=1}^{B} I(\widehat{J}_b^* > \widehat{J}) \tag{6.24}$$

where I denotes the indicator function which takes the value 1 if $\widehat{J}_b^* > \widehat{J}$ and 0 otherwise.

In our study we estimate the p-values using $B = 2000$ replications. The calculation of the test is done using our own programming for the statistical software R.

6.3 Productivity analysis of European countries

6.3.1 Data of the analysis

Our sample consists of 30 OECD countries and covers the years 1990 to 2007. Although the focus lies on the EU15 countries we analyze them as a part of the joint sample of OECD countries. Since DEA estimations of the technology suffer from the "curse of dimensionality" (see e.g. Simar (2007)) the increased number of observations allows for more precise estimations. As a rule-of-thumb the number of observations should be larger than three times the number of inputs and outputs (see e.g. Fernandes and Pacheco (2002)). In our analysis we include 3 inputs and 2 outputs, hence more than 15 observations should be included. To meet this requirement we estimate the productivity using the sample of 30 OECD countries.

We include in our analysis three inputs. To estimate the amount of labor and the capital stock data from the Penn World Table (PWT) by Heston et al. (2009) are used. The capital stock is estimated by applying the perpetual inventory method (PIM) where the initial capital stock K_0 and the following capital stocks K_{t+1} are calculated by the following equations (see Park (1995)):

$$K_0 = I_0 \cdot \frac{(1+g)}{g+\delta} \tag{6.25}$$

[74] This procedure mimics a draw under the null hypothesis. It is also possible to mimic the null hypothesis by drawing from the joint sample, but Li (1999) shows that the resampling from one of the subsamples performs better in Monte-Carlo simulations.

$$K_{t+1} = I_{t+1} + K_t \cdot (1 - \delta) \tag{6.26}$$

Here, I_0 denotes investments in the initial period while I_{t+1} denotes investments in period $t+1$. g represents the growth rate of investments and δ symbolizes the rate of depreciation. In our analysis we assume a depreciation rate of $\delta = 0.05$ which is in line with common assumptions in empirical analyses of the OECD countries (see e.g. Nonneman and Vanhoudt (1996)).[75] The third input, obtained from the Worldbank database, is given by the energy consumption of the countries. The Worldbank also provides data on the single bad output CO_2. For the years 1990 and 1991 the Worldbank is lacking data for the Slovak Republic and for the year 1990 the amount of emissions is missing for Germany. For these observations we use data from the European Environment Agency (EEA) instead. The single good output GDP is obtained from the PWT. Table 6.1 presents some descriptive statistics of our sample.

Table 6.1: Descriptive statistics of the country data (1990-2007)

	Min	Median	Mean	Max	SD
Inputs					
Capital (bio. $, 2005)	28.01	1025.31	3356.96	42112.60	6252.15
Labor (1000 employees)	128.26	5175.12	17572.69	153197.49	27260.07
Energy (kt. oil eq.)	2034.00	51324.00	168139.75	2336546.00	387219.53
Output					
GDP (bio. $, 2005)	6.40	274.05	952.69	12921.00	1917.12
Undesirable output					
CO_2 (kt.)	1892.17	119404.85	408786.20	5595357.96	956947.33

Our dataset covers a broad range of countries with the United States being the largest producer of CO_2 (in 2005) and the country with the highest GDP (in 2007). The observation with the lowest emissions and GDP is Iceland in 1992. Regarding our inputs and outputs we find a high correlation (0.9978) between the energy consumption and the carbon dioxide emissions which is in line with the results of previous studies (see Oh and Heshmati (2010)).

Figure 6.2 shows the time series of the aggregate CO_2 emissions and the GDP for the EU15 countries as well as the OECD countries which are not part of the EU15. The graphs show that GDP has grown for both groups over the sample period. In contrast, while the emissions of the EU15 have slightly decreased (from 3.31 billion tons to 3.28 billion tons) the emissions of the remaining sample have increased (from 8.19 billion tons to 9.64 billion tons). But we do not observe a clear change in the trend after the signing of the Kyoto Protocol in 1997 in this graph (indicated by the vertical line).

[75] We have also conducted the analysis using depreciation rates of 3% and 8%. However, these depreciation rates lead to similar results.

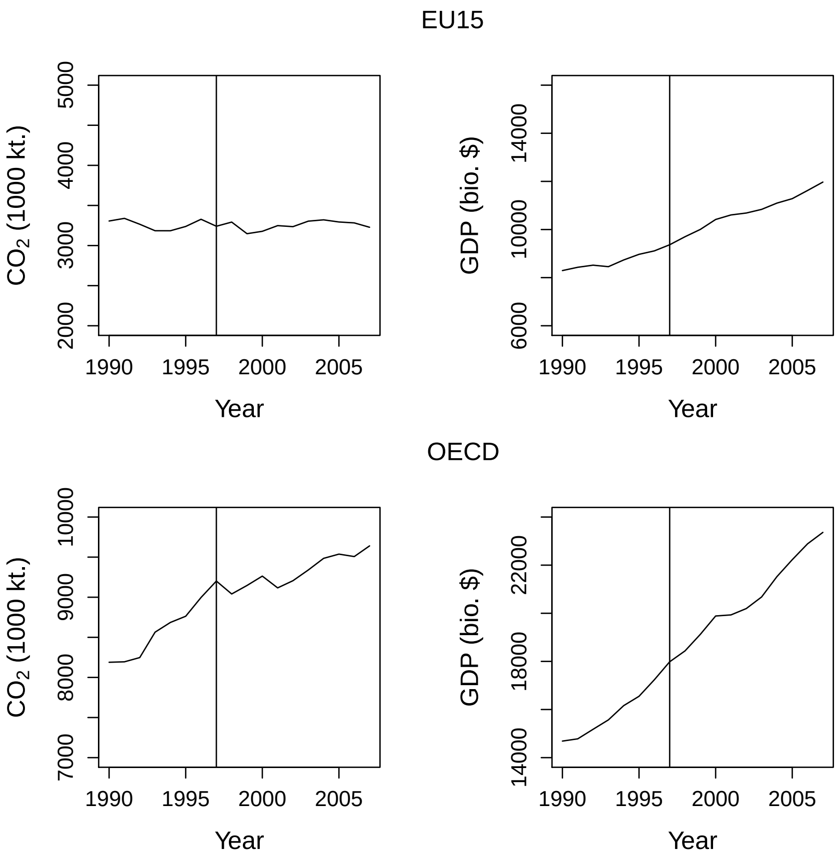

Figure 6.2: Trends in CO_2 and GDP

6.3.2 Results of the analysis of European countries

The median results of the bias-corrected Malmquist and Malmquist-Luenberger index as well as its decomposition can be found for each of the EU15 countries in the tables D.1 and D.2 in appendix D. As explained in the chapter on general concepts a value larger than one for the indices and/or its components indicates progress while a value less than one indicates regress. Moreover, $100 \cdot \left(\widehat{\mathrm{Malm}}_i^{t,t+1,\mathrm{BC}} - 1 \right)$ can be interpreted as the percentage change in productivity.[76]

As a concrete example for the interpretation of the results consider the case of Austria for which results are presented in the first row of tables D.1 and D.2. According to the Malmquist index the median productivity increase was 1.48% per year. This increase resulted from a combination of a slight decrease in efficiency (−0.47%) and a larger effect of technical progress (2.35%). The Malmquist-Luenberger index indicates a smaller increase in productivity if the reduction of carbon dioxide emissions is accounted for. In this case the median increase of the productivity was 0.59%. While the efficiency of Austria remained unchanged the driver of this increase was technical progress (0.59%).

A result which is particular different compared to the findings for the remaining countries can

[76] Note that in case the ML index was not computable we included neither the Malmquist nor the ML index result for an observation. This is done to avoid comparing different observations in the subsequent analysis.

be observed for Luxembourg. While the Malmquist index shows an increase of productivity by 2.96% the Malmquist-Luenberger index indicates an increase by 10.1% which is far above the results for the other countries. Hence, the results for the very small country Luxembourg may be outliers detering the results for our analysis of the Kyoto Protocol. Therefore, we have also conducted the following analysis excluding the results for Luxembourg (see appendix D for the graphs and test results). However, we decided not to exclude Luxembourg from our sample. Firstly, the related literature on nonparametric productivity analysis also contains this observation (see e.g. Kumar (2006)). Secondly, we applied the testing procedure by Wilson (1995) to detect influential observations in our dataset. The results showed that Luxembourg is indeed influential but very similar results were obtained for Ireland and the United Kingdom, two countries which do not show unusual patterns in their productivity results. Hence, we do not consider Luxembourg as an outlier.

For an overview of the development of the productivity in the EU15 countries, figure 6.3 presents the time series of the median productivity change and its decomposition into efficiency and technical change as measured by both the Malmquist index and the Malmquist-Luenberger (ML) index. More detailed density plots of the results will be presented in the subsequent group analysis.

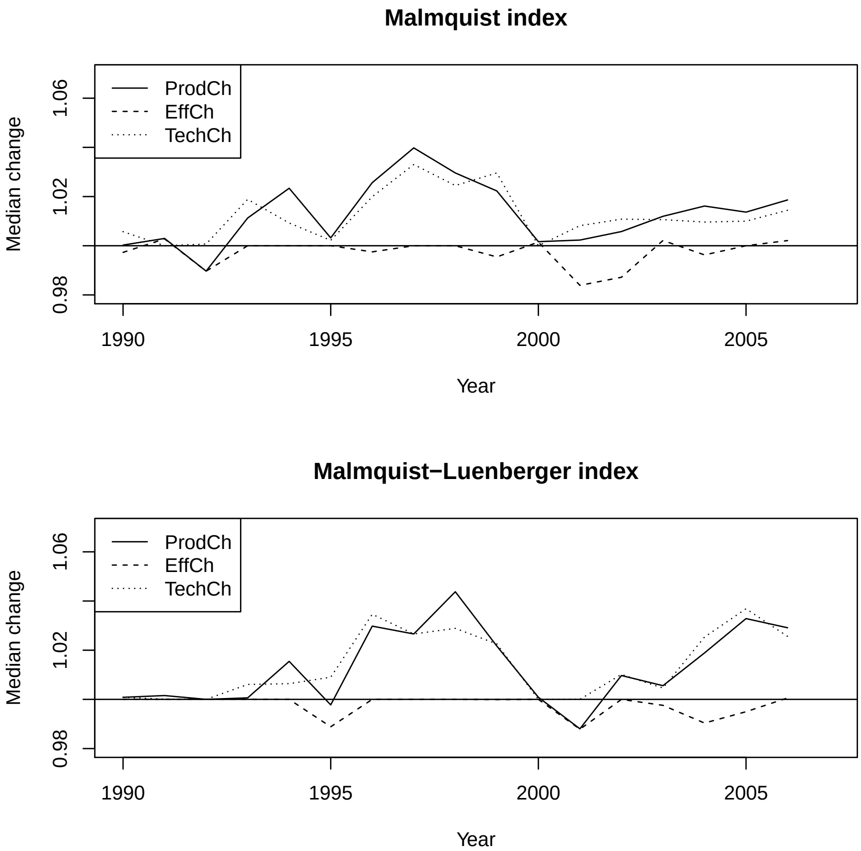

Figure 6.3: Median productivity, efficiency and technical change for the EU15

For both indices we find that the median productivity has increased in the majority of years. This is largely due to technical progress while efficiency improvements are small and in some years we even find efficiency regress. Note that due to the use of sequential output sets technical regress is excluded by construction. Moreover, we find that productivity growth significantly slowed down in Europe after the year 2000 which is in line with the findings of the literature that measures productivity not using nonparametric frontier methods (see van Ark et al. (2008)).

To analyze which countries have induced the technical progress which has driven productivity growth tables D.3 and D.4 in the appendix contain the innovating countries for each period. An "innovator" is a country which fulfills the following conditions (see the previous chapter for a further discussion of innovators in nonparametric models):

$$\widehat{\mathrm{MLTech}}_i^{t,t+1} > 1 \tag{6.27}$$

$$\widehat{\beta}^t\left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i}; \boldsymbol{g}_{t+1}\right) < 0 \tag{6.28}$$

$$\widehat{\beta}^{t+1}\left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i}; \boldsymbol{g}_{t+1}\right) = 0. \tag{6.29}$$

The results show that members of the EU15 were constantly among those countries which shifted the frontier irrespective of whether the production of carbon dioxide emissions is included in the analysis. In contrast, non-European countries (e.g. the United States) are among the innovators only in a few periods. This analysis also reveals why Ireland, Luxembourg and the United Kingdom are indicated as influential observations. The results show that these countries were among those who shifted the frontier for many years. With regard to differences between the Malmquist and the Malmquist-Luenberger index we find that when accounting for the reduction of emissions more countries can be identified as innovators. In particular, Sweden, Norway and Switzerland become important observations with this regard.

In the following we analyze whether the hypothesis by Böhringer and Vogt (2003, 2004) that the Kyoto Protocol has not led to significant changes in the countries efforts to reduce emissions and therefore did not induce significant macroeconomic costs on them is justified. To evaluate the hypothesis for the EU15 countries we divide the countries into those which according to the EU-BSA have to decrease their emissions and those which are allowed to increase their emissions or hold them constant. EU15 countries which have to decrease emissions (referred to as DC countries) are Austria, Belgium, Denmark, Germany, Italy, Luxembourg, the Netherlands and the United Kingdom while Finland, France, Greece, Ireland, Portugal, Spain, and Sweden have not to decrease emissions (referred to as NDC countries). Table 6.2 contains the emission targets of all EU15 countries for the commitment period 2008-2012 compared to the emission level in the base year 1990 (see European Environment Agency (2011a)). The targets show that the EU-BSA aims at shifting the overall reduction target of the Kyoto Protocol from the poorer cohesion countries like Spain and Portugal to the richer European countries like Germany or Denmark.

Table 6.2: Emission targets (EU15)

Country	Emission target
Austria	-13%
Belgium	-7.5%
Denmark	-21%
Finland	0%
France	0%
Germany	-21%
Greece	25%
Ireland	13%
Italy	-6.5%
Luxembourg	-28%
Netherlands	-6%
Portugal	27%
Spain	15%
Sweden	4%
United Kingdom	-8%

Figures 6.4, 6.5 and 6.6 show the density plots of the productivity changes and their decompositions for the two groups of EU countries before and after the EU Burden Sharing Agreement was implemented. As noted before, the results excluding Luxembourg can be found in the appendix. The density plots of the technical changes were obtained using the function `show.dens()` from the package "FEAR" by Wilson (2008) for the statistical software R. This function estimates the density of estimates bounded at unity using the reflection method by Boneva et al. (1971) discussed in Silverman (1986). Table 6.3 presents the bootstrapped p-values of the Li (1996) test adapted by Simar and Zelenyuk (2006) which tests the null hypothesis of equal distributions for both groups. The row "Technical change 1" refers to algorithm 1 in the paper by Simar and Zelenyuk (2006) while "Technical change 2" refers to algorithm 2 (see the description of the test above for more details on the algorithms).

Table 6.3: p-values of the Li test comparing DC and NDC countries

	Malmquist index		ML index	
	Before EU-BSA	After EU-BSA	Before EU-BSA	After EU-BSA
Productivity change	0.472	0.945	0.558	0.148
Efficiency change	0.011	0.004	0.018	0.006
Technical change 1	0.003	0.949	0.035	0.217
Technical change 2	0.060	0.636	0.002	0.418

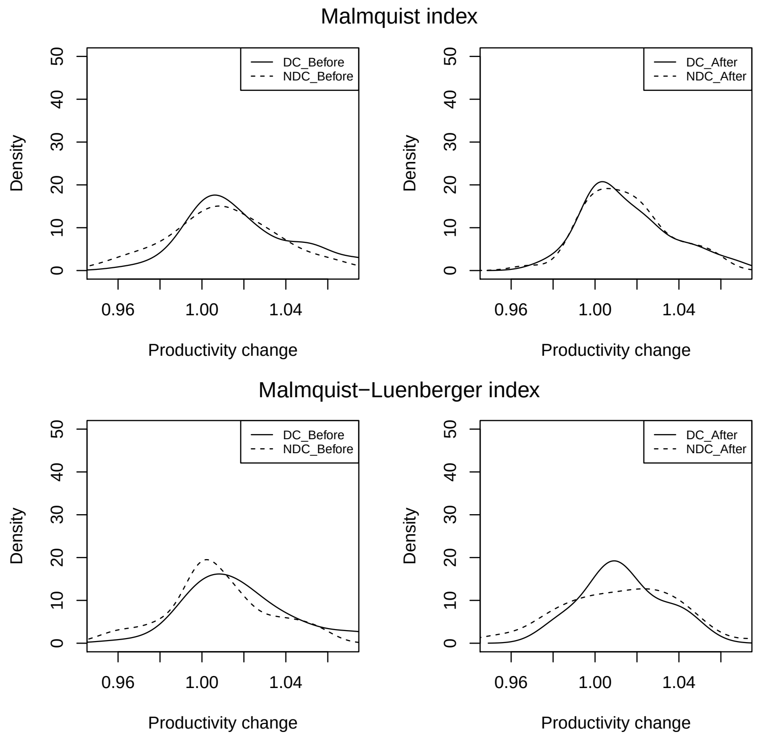

Figure 6.4: Density plots of productivity change of EU15 countries

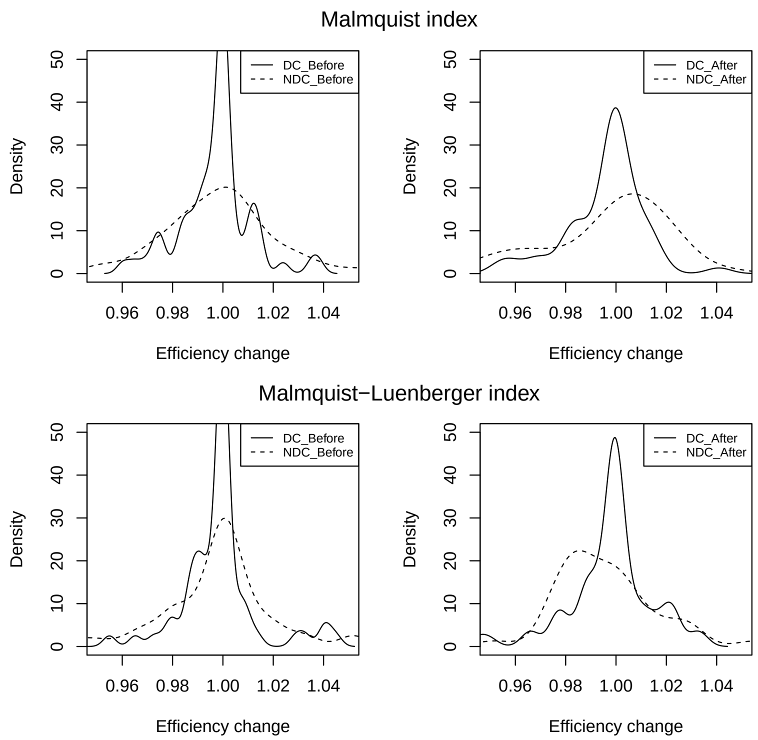

Figure 6.5: Density plots of efficiency change of EU15 countries

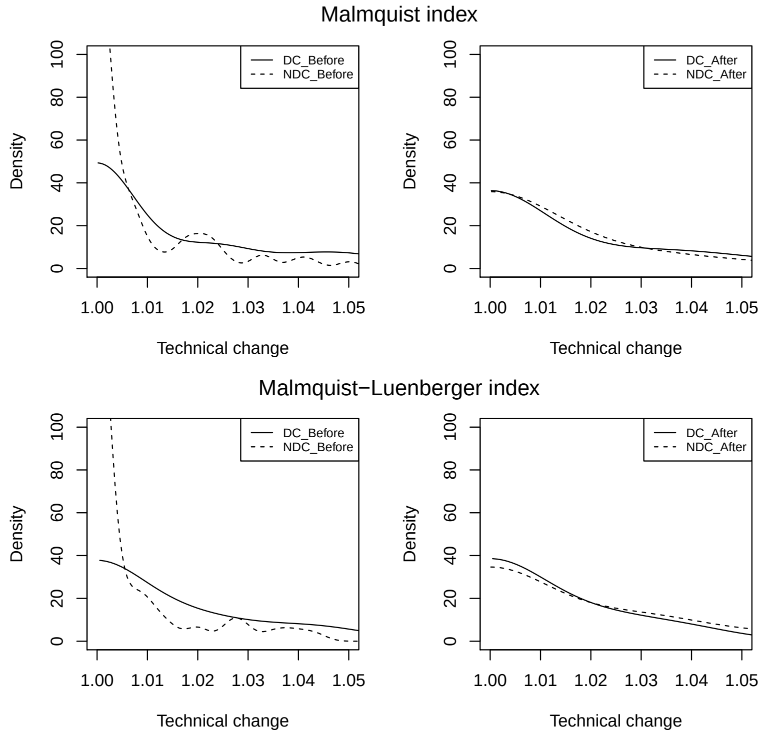

Figure 6.6: Density plots of technical change of EU15 countries

We start our discussion with the results for the period before the implementation of the EU-BSA (1990-1998). The related density plots can be found in the left columns of the figures. For the productivity change measured by the the Malmquist index the density plots show only minor differences between the groups indicating that both groups showed homogeneous patterns in their productivity development. This result is also statistically confirmed by the Li test which does not reject the null hypothesis (p-value: 0.472). Accounting for the reduction of emissions when evaluating the productivity of countries as it is done by the Malmquist-Luenberger index leads to similar results. Although the graph for the ML results exhibits slightly larger differences for the groups before the EU-BSA the test does not reject that the distributions are equal (p-value: 0.558). Hence, like in the case of the Malmquist index we find homogeneous patterns in the productivity results. These findings do not change if Luxembourg is excluded from the analysis (see figure D.1 and table D.5 in appendix D).

Regarding the components of the productivity change we find statistically significant differences in the efficiency results before the EU-BSA. However, the differences are only minor as can be seen from the plots and the difference is only weakly significant if Luxembourg is excluded from the comparison. For technical change we find slightly larger results for the DC countries before the EU-BSA which is also statistically significant. But as shown above these differences do not affect the productivity comparison which finds homogeneous patterns for both groups.

Analyzing the groups after the implementation of the EU-BSA we again do not find significant differences in the productivity patterns. The results of the Li test indicate that the productivity between the groups differs significantly neither for the Malmquist nor for the Malmquist-

Luenberger index which is also visible from the density plots. We still find significant differences in the efficiency change results but even the small differences in technical change have vanished. Therefore, we find that after the EU-BSA was signed the very homogeneous patterns in productivity change prevail and that again differences in the sources of productivity change do not influence the overall productivity comparison.

The results obtained by this analysis can be used to address the question of a "business-as-usual" strategy of the EU15 countries. Given that we compare productivity change measured by the Malmquist index for two homogeneous groups we would expect that countries which have reallocated inputs to the reduction of the emissions and therefore face significant macroeconomic costs to meet their reduction targets should actually show a shift to the left in the density plots compared to the other group of countries. This is to be expected since the Malmquist index does not capture increasing abatement activities which lead to reallocations of inputs from the production of good outputs to the abatement of bad outputs. Hence, the productivity change estimated by the Malmquist index should be lower. In the above presented analysis we do not observe such a shift between the DC and the NDC countries. This indicates that the DC countries did not reallocate significant more inputs to abatement after the implementation of the EU-BSA.

However, this interpretation of the Malmquist index results implicitly assumes that the DC countries facing reduction targets did not shift inputs to the abatement activities and at the same time increased their productivity with regard to the good outputs to a larger extent than the NDC countries. In this case, the latter effect would cancel out the first and we would also obtain no significant differences comparing the productivity results between the groups. To exclude this possibility the results for the ML index become important. If the DC countries would have reallocated inputs and increased their productivity compared to the NDC countries, the productivity changes measured by the ML index should be significantly larger for the DC countries compared to the results for the NDC countries. This follows because the ML index accounts for both the increased production of good outputs as well as the decreased emission of pollutants. This is not the case in our analysis where the results for both indices do not differ statistically significant between the groups before and after the EU-BSA was signed.

Our results obtained for the European countries therefore raise doubts on whether significant amounts of inputs have been shifted to the abatement of emissions after the Kyoto Protocol was signed. Hence, our findings support the hypothesis by Böhringer and Vogt (2003, 2004) that the Kyoto Protocol enables the European countries to follow a "business-as-usual" strategy to achieve the Kyoto targets.

This finding of our study raises the question why the DC countries did not reallocate significant amounts of inputs. One particularly important country with this regard is Germany which according to the EU-BSA faces a large reduction target. With this agreement the initial Kyoto target (-8%) was tightened to -21%. To explain this change the development of emissions in Germany after the reunification of East and West Germany is important. As Schleich et al. (2001) analyze, Germanys emissions dropped by more than 18% after the reunification due to structural change in East Germany. Hence, even before the Kyoto Protocol was signed Germany had more than fulfilled its initial Kyoto target. The EU-BSA therefore has reallocated these "wall-fall profits" (Schleich et al. (2001, p. 364)) to allow several other countries to increase their emissions. Moreover, taking into account the already achieved reductions, the EU-BSA target for Germany is only slightly ambitous. Hence, Germany may not require large changes in the input allocation. However, this finding does not explain why the remaining DC countries have not shifted additional inputs to abatement activities. A possible explanation for this can be given by our findings for the innovating countries and the source of productivity growth. As we have shown using the ML index European countries were those who shifted the frontier if emissions are included in the efficiency analysis. Moreover, we found that the efficiency changes for the countries are rather small while technical progress spurred the increased productivity. Thus, the non-innovating countries were able to follow the technical changes induced by the innovating countries. Hence, European countries were able to produce more good outputs while reducing emissions holding inputs constant. Therefore, a tentative answer to the question why the countries did not reallocate inputs after the signing of the Kyoto Protocol can be given by technical change having reduced the necessity to reallocate inputs to abatement activities and thus preventing the countries from large macroeconomic costs.

6.4 Summary

In this study we have investigated whether the Kyoto Protocol and its European implementation, the EU Burden Sharing Agreement, had a significant influence on the macroeconomic productivity of European countries. The analysis of the EU15 showed that their productivity has increased over time. We found that this increase was driven by technical progress mostly induced by European countries which have been identified as the innovators among 30 OECD countries.

Our analysis of two groups of European countries aimed at evaluating whether countries which have to reduce emissions according to the EU-BSA have faced significant costs compared to those who had not to decrease their emissions. Our comparison of the conventional Malmquist and the Malmquist-Luenberger index between the countries did not find any evidence that the Kyoto Protocol has changed "business-as-usual" politics by the European countries, supporting the hypothesis by Böhringer and Vogt (2003, 2004). Therefore, our analysis questions whether these countries have faced significant costs in the following of the signing of the Kyoto Protocol.

7 Conclusion

In this dissertation we have addressed several questions of current research in nonparametric efficiency analysis incorporating environmental factors. In the following we summarize the main findings of the studies, show connections between them and highlight potentials for future research.

Chapter 3 of this dissertation presented an analysis on whether the incorporation of carbon dioxide emissions influences the results of an efficiency analysis of automobiles. We found that the mere inclusion of CO_2 has no large effect if fuel consumption is included simultaneously. By accounting for the reduction of CO_2 we obtained significant potentials to decrease these emissions. Our analysis also showed that these potentials are not associated with an reduction of fuel consumption. Thus, we could identify car groups which are more efficient with regard to simultaneously increasing the good outputs and decreasing the bad output. However, this result is based on an analysis that due to data limitations accounts only for a single pollutant. In chapter four we have introduced the materials balance condition which states that the amount of materials bound in the inputs has to equal the amount of materials bound in the outputs. Applying this function to the results of the automobile study shows that decreasing the CO_2 emissions while holding the fuel consumption constant has to result in an increase of other pollutants (e.g. CO). Therefore, while some cars may be more efficient with regard to CO_2 they may be indicated as less efficient if we would account for other pollutants. This problem faced by economists when conducting environmental analyses is already being discussed in the macroeconomic literature e.g. with regard to the environmental Kuznets curve (see Dasgupta et al. (2002)). In contrast, microeconomic analyses have not yet accounted for this issue. Therefore, future research should focus on accounting for results from natural science when conducting environmental analyses.[77] Especially, it is important to clarify which different pollutants can result from the conversion of particular inputs into outputs so that missing variables in a dataset are not mistaken as inefficiency.

In our analysis of automobiles we also compared our results for different car groups with those obtained in the previous literature. We have shown that this literature ignores the influence of groups specific inefficiencies and thus tends to underestimate the inefficiency of automobiles when focusing solely on a particular group (e.g. compact class cars). Our study is the first to actually compare results with previous analyses which are similar to a certain extend. The literature on nonparametric efficiency analysis of automobiles is very heterogeneous with regard to the applied methods, which are often novel ones, and the analyzed dataset. Therefore, it is difficult to explain whether differences in the results are due to methodological or data differences. To overcome this problem future research may provide a comparison of the proposed models and efficiency measures using a single dataset. This could improve the understanding of whether the broad range of methodological approaches indeed leads to very different results when evaluating the efficiency of automobiles.

In chapter four we presented a new possibility to measure the environmental efficiency of DMUs

[77] See Fischer et al. (2012) for an overview of interdisciplinary research previously conducted by economics and natural science.

in a network model consisting of a production and an end-of-pipe abatement stage. Furthermore, we have shown how the new environmental efficiency measure can be used to decompose environmental efficiency into production and abatement efficiency as well as stage and network effects. In our model we focus on end-of-pipe abatement processes because these are predominantly evaluated in economic analyses. However, the large attention that end-of-pipe technologies receive in economics is not necessarily supported by empirical data. For example, Frondel et al. (2007) find for OECD countries that the majority of firms invests more in process-integrated than in end-of-pipe abatement activities. Hence, modifying the presented model to incorporate process-integrated abatement of pollution could lead to an approach suitable for a larger number of empirical applications.

In this study we have demonstrated the decomposition of the new environmental efficiency measure into the different components assuming constant returns to scale. The results and influences of the assumption of variable returns to scale were shortly discussed. Nevertheless, it is possible to extend the proposed decompositions into technical and scale effects using CRS and VRS technologies (see e.g. Lozano (2011) for scale effects in network technologies). However, this proposal comes with the caveat that further decomposing the efficiency components may lead to a situation similar to the Malmquist index for which various decompositions have been proposed (see e.g. Lovell (2003) for an overview) but are not often applied to empirical data presumably because each further decomposition weakens the clarity of the results.

From an empirical point of view the literature on network models in nonparametric efficiency analysis is more homogeneous compared to the literature on nonparametric evaluations of automobiles. The data sources used in our study, the U.S. Environmental Protection Agency, the U.S. Energy Information Administration and the Federal Energy Regulatory Commission, are also exploited by several other studies. As in our analysis these studies, especially Färe et al. (forthcoming), are faced with very small numbers of observations. Moreover, this major research interest for U.S. power plants can only be partly explained by the importance of their environmental efficiency. It is also worth noting that to our knowledge these data sources are the only publicly available ones that provide detailed microlevel information on abatement activities for a group of observations which environmental efficiency has a large influence e.g. with regard to regulatory actions. The remaining datasets (e.g. exploited in Coelli et al. (2007) or Yang et al. (2008)) provide information for very specialized agricultural applications. Hence, while the theoretical models become more sophisticated the data to evaluate them remain very limited. Therefore, future research should focus on collecting and compiling further environmental datasets.

The fifth chapter of this dissertation focused on different methods to calculate optimal directional vectors for directional distance functions. This distance function had been applied in chapter three for the analysis of automobiles but there we followed the previous literature by exogenously determining the vector. In this study we showed how an existing model to endogenize the directions can be adapted to the analysis of environmental efficiency and proposed an extension to a dynamic analysis. While the static approach results in a vector that projects the DMU under evaluation on the furthest feasible point on the boundary of the technology the dynamic approach estimates the direction by calculating the movements of the innovators and hence of the frontier.

Both methods are used to obtain optimal vectors in terms of weights for the increase of good and the decrease of bad outputs. In the dynamic approach the direction of the efficiency measurement for a non-innovating DMU is given by the innovating direction of the closest located innovator. To identify the closest innovator the euclidean distance is used. However, the movement of the frontier segment a non-innovating DMU is compared to is likely to be only partly captured by the movement of the closest innovator. For example, if this frontier segment is constructed by a linear combination of two innovators with different innovating directions than the DMU should be assigned a direction that captures this combination to measure more precisely the direction of the frontier change, hence the direction of technical change. Therefore, an extension of the proposed dynamic model could include the closest linear combination of innovators to identify the direction of technical change. To obtain this combination approaches like Aparicio et al. (2007) could be modified and applied in this model. A further extension may address the endogenization of the orientation of the measurement. In the presented study we assumed output orientation and hence only endogenized the weights of the outputs. A generalization could include an endogenized decision of whether an input, output or perhaps mixed orientation of the efficiency analysis should be applied. Such an endogenization would be also very useful for an application to the new environmental efficiency measure proposed in chapter four for which the orientation with regard to the production efficiency is not predetermined.

Applying the theoretical approaches to optimal directions to a macroeconomic efficiency analysis of the major greenhouse gas emitting countries we found significant potentials to reduce these emissions even when accounting for heterogeneity in these countries. However, while the results indicate that by comparing inefficient countries with the best-practice countries we observe reduction potentials they do not readily provide information on how to achieve these reduction potentials in practice. For example, if these potentials are due to differences in the industry structure they may not be easily exploited in the short run. This problem of benchmarking with DEA has been addressed on the microeconomic level by combining DEA with other benchmarking methods (see e.g. Talluri (2000)) to provide more practical information on efficiency improvement possibilities. Future research may apply these methods to macroeconomic analyses to provide a more detailed view on how much of the reduction potentials found in our analysis can be realized.

The final study in this dissertation, presented in chapter six, provided a dynamic analysis of the productivity of European countries. The research aimed at analyzing whether international environmental regulation, in particular the Kyoto Protocol in form of the EU Burden Sharing Agreement, have lead to a significant redistribution of inputs from the production of good outputs to the abatement of bad outputs. Therefore, we compared the productivity changes for two groups of European countries. Our results did not indicate that significant amounts of inputs have been reallocated by countries which have to reduce their emissions compared to countries which are allowed to increase their emissions. Thus, we could not find a significant influence of the Kyoto Protocol on the macroeconomic productivity of these countries. In our analysis we applied the Malmquist and the Malmquist-Luenberger index. With regard to the latter we faced the problem that it was not possible to calculate the index for each country in each period due to the infeasibility of the mixed-period distance function. Hence, the number of observations used in this study is smaller than the number of countries times the number of periods. This problem could be addressed in future research by applying the global Malmquist and Malmquist-Luenberger indices (see Pastor and Lovell (2005) as well as Oh (2010)). Combining these approaches with the analysis of sequential and contemporaneous output sets applied in this dissertation allows to furthermore exclude the effects of business cycles in these indices.

For this study of the influence of the Kyoto Protocol we have used a dataset that covers the years before and after the signing of the Protocol but does not include the actual commitment period. Therefore, to account for the possibility that the European countries have waited until the commitment period to begin with reallocating inputs data on this period should be included in the analysis when available.

The discussion of the effects of the Kyoto Protocol which was established to decrease the amount of greenhouse gas emissions on a global level and thus to limit the increase in global temperature leads back to the introduction of this dissertation. There we have discussed the challenges associated with climate change and the necessity of economic analyses accounting for the production of pollutants. In our presented research we addressed this issue in multiple ways. Microeconomically and macroeconomically by analyzing objects ranging from single automobiles to entire countries. Implicitly and explicitly by measuring productivity changes including emissions and estimating reduction potentials for specific pollutants. Empirically and theoretically by analyzing power plants applying a newly developed production model. In general, the results of this dissertation have shown that economically significant reduction potentials for pollutants exist highlighting the relevance of the conducted research. Nonetheless, since the combination of environmental economics and nonparametric efficiency analysis is still a young field of research many open questions, as demonstrated above, remain to be answered in the future.

References

Abadie, A. (2002). “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models”. In: *Journal of the American Statistical Association* 97, pp. 284–292.

ADAC (2009). “Der ADAC-Autotest”. In: *Informationen aus der Fahrzeugtechnik* 03/09, pp. 1–9.

— (2010). *ADAC-Autokatalog*. [Retrieved February 03, 2010]. URL: `http://www.adac.de/infotestrat/autodatenbank/default.aspx`.

Aigner, D. and S. F. Chu (1968). “On Estimating the Industry Production Function”. In: *American Economic Review* 58, pp. 826–839.

Aigner, D., C. A. K. Lovell, and P. Schmidt (1977). “Formulation and Estimation of Stochastic Frontier Production Function Models”. In: *Journal of Econometrics* 6, pp. 21–37.

Aldy, J. E., A. J. Krupnick, R. G. Newell, I. W. H. Parry, and W. A. Pizer (2010). “Designing Climate Mitigation Policy”. In: *Journal of Economic Literature* 48, pp. 903–934.

Andor, M. and F. Hesse (2012). “The StoNED Age: The Departure into a New Era of Efficiency Analysis? A MC Study Comparing StoNED and the Oldies (SFA and DEA)”. In: *CAWM Discussion Paper* No. 60. Centrum für Angewandte Wirtschaftsforschung Münster.

Andreoni, J. and A. Levinson (2001). “The Simple Analytics of the Environmental Kuznets Curve”. In: *Journal of Public Economics* 80, pp. 269–286.

Aparicio, J., J. L. Ruiz, and I. Sirvent (2007). “Closest Targets and Minimum Distance to the Pareto-Efficient Frontier in DEA”. In: *Journal of Productivity Analysis* 28, pp. 209–218.

Aragon, Y., A. Daouia, and C. Thomas-Agnan (2005). “Nonparametric Frontier Estimation: A Conditional Quantile-Based Approach”. In: *Econometric Theory* 21, pp. 358–389.

Arcelus, F. J. and P. Arocena (2005). “Productivity Differences Across OECD Countries in the Presence of Environmental Constraints”. In: *Journal of the Operational Research Society* 56, pp. 1352–1362.

Ayers, R. U. and A. V. Kneese (1969). “Production, Consumption, and Externalities”. In: *The American Economic Review* 59, pp. 282–297.

Banker, R. D., A. Charnes, and W. W. Cooper (1984). “Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis”. In: *Management Science* 30, pp. 1078–1092.

Banker, R. D. and R. C. Morey (1986). “Efficiency Analysis for Exogenously Fixed Inputs and Outputs”. In: *Operations Research* 34, pp. 513–521.

Barrett, G. F. and S. G. Donald (2003). “Consistent Tests for Stochastic Dominance”. In: *Econometrica* 71, pp. 71–104.

Barrett, S. (1998). “Political Economy of the Kyoto Protocol”. In: *Oxford Review of Economic Policy* 14, pp. 20–39.

Böhringer, C. and C. Vogt (2003). “Economic and Environmental Impacts of the Kyoto Protocol”. In: *Canadian Journal of Economics* 36, pp. 475–494.

— (2004). “The Dismantling of a Breakthrough: The Kyoto Protocol as Symbolic Policy”. In: *European Journal of Political Economy* 20, pp. 597–617.

Boden, T. A., G. Maryland, and R. J. Andres (2010). “Global, Regional and National Fossil-Fueled CO_2 Emissions”. Carbon Dioxide Information Analysis Center.

Boneva, L. I., D. Kendall, and I. Stefanov (1971). "Spline Transformations: Three New Diagnostic Aids for the Statistical Data-Analyst". In: *Journal of the Royal Statistical Society* 33, pp. 1–71.

Botzen, W. J. W. and J. C. J. M. van den Bergh (2012). "How Sensitive is Nordhaus to Weitzman? Climate Policy in DICE with an Alternative Damage Function". In: *Economics Letters* 117, pp. 372–374.

Brekke, K. A. and O. Johansson-Stenman (2008). "The Behavioural Economics of Climate Change". In: *Oxford Review of Economic Policy* 24, pp. 280–297.

Briec, W. (2000). "An Extended Färe-Lovell Technical Efficiency Measure". In: *International Journal of Production Economics* 65, pp. 191–199.

Briec, W. and K. Kerstens (2009). "Infeasibility and Directional Distance Functions with Application to the Determinateness of the Luenberger Productivity Indicator". In: *Journal of Optimization Theory and Applications* 141, pp. 55–73.

Brockett, P. L. and B. Golany (1996). "Using Rank Statistics for Determining Efficiency Differences in Data Envelopment Analysis". In: *Management Science* 42, pp. 466–472.

Camanho, A. S. and R. G. Dyson (2005). "Cost Efficiency Measurement with Price Uncertainty: A DEA Application to Bank Branch Assessments". In: *European Journal of Operational Research* 161, pp. 432–446.

Canadell, J. G., C. Le Quére, M. R. Raupach, C. B. Field, E. T. Buitenhuis, P. Ciais, T. J. Conway, N. P. Gillett, R. A. Houghton, and G. Marland (2007). "Contributions to Accelerating Atmospheric CO_2 from Economic Activity, Carbon Intensity and Efficiency of Natural Sinks". In: *Proceedings of the National Academy of Sciences* 104, pp. 18866–18870.

Cantner, U., J. J. Krüger, and R. Söllner (2012). "Product Quality, Product Price, and Share Dynamics in the German Compact Car Market". In: *Industrial and Corporate Change* 21, pp. 1 –31.

Caves, D. W., L. R. Christensen, and W. E. Diewert (1982). "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity". In: *Econometrica* 50, pp. 1393–1414.

Cazals, C., J.-P. Florens, and L. Simar (2002). "Nonparametric Frontier Estimation: A Robust Approach". In: *Journal of Econometrics* 106, pp. 1–25.

Çengel, Y. A. (2008). *Introduction to Thermodynamics and Heat Transfer*. 2nd ed. New York: McGraw-Hill.

Chambers, R. G., Y. Chung, and R. Färe (1996). "Benefit and Distance Functions". In: *Journal of Economic Theory* 70, pp. 407–419.

Charnes, A. and W. W. Cooper (1962). "Programming with Linear Fractional Functionals". In: *Naval Research Logistics Quarterly* 9, pp. 181–186.

Charnes, A., W. W. Cooper, and E. Rhodes (1978). "Measuring the Efficiency of Decision Making Units". In: *European Journal of Operational Research* 2, pp. 429–444.

— (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through". In: *Management Science* 27, pp. 668–697.

Choi, H. and I. Oh (2010). "Analysis of Product Efficiency of Hybrid Vehicles and Promotion Policies". In: *Energy Policy* 38, pp. 2262–2271.

Chung, Y. H., R. Färe, and S. Grosskopf (1997). "Productivity and Undesirable Outputs: A Directional Distance Function Approach". In: *Journal of Environmental Management* 51, pp. 229–240.

Cline, W. R. (2007). *Global Warming and Agriculture: Impact Estimates by Country*. Massachusetts: Center for Global Devlopment and Perter Institute for International Economics.

Coelli, T., L. Lauwers, and G. Van Huylenbroeck (2007). "Environmental Efficiency Measurement and the Materials Balance Condition". In: *Journal of Productivity Analysis* 28, pp. 3–12.

Coelli, T. and S. Perelman (1999). "A Comparison of Parametric and Non-Parametric Distance Functions: With Application to European Railways". In: *European Journal of Operational Research* 117, pp. 326–339.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell, and G. E. Battese (2005). *An Introduction to Efficiency and Productivity Analysis*. New York: Springer.

Commission of the European Communities (2007). "Proposal for a Regulation of the European Parliament and of the Council - Setting Emission Performance Standards for New Passenger Cars as Part of the Community's Integrated Approach to Reduce CO_2 Emissions from Light-Duty Vehicles". Dossier COD/2007/0297.

Cook, W. D., M. Hababou, and H. J. H. Tuenter (2000). "Multicomponent Efficiency Measurement and Shared Inputs in Data Envelopment Analysis: An Application to Sales and Service Performance in Bank Branches". In: *Journal of Productivity Analysis* 14, pp. 209–224.

Cook, W. D. and L. M. Seiford (2009). "Data Envelopment Analysis (DEA) – Thirty Years On". In: *European Journal of Operational Research* 192, pp. 1–17.

Cropper, M. L. and W. E. Oates (1992). "Environmental Economics: A Survey". In: *Journal of Economic Literature* 30, pp. 675–740.

Crowley, T. J. (2000). "Causes of Climate Change over the Past 1000 Years". In: *Science* 289, pp. 270–277.

Dai, A. (2011). "Drought under Global Warming: A Review". In: *Wiley Interdisciplinary Reviews: Climate Change* 2, pp. 45–65.

Daly, H. E. (1997). "Georgescu-Roegen versus Solow/Stiglitz". In: *Ecological Economics* 22, pp. 261–266.

Daraio, D. and L. Simar (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. New York: Springer.

Dasgupta, S., B. Laplante, H. Wang, and D. Wheeler (2002). "Confronting the Environmental Kuznets Curve". In: *Journal of Economic Perspectives* 16, pp. 147–168.

Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and Their Application*. Cambridge (Mass.): Cambridge University Press.

Davison, J. (2007). "Performance and Costs of Power Plants with Capture and Storage of CO_2". In: *Energy* 32, pp. 1163–1176.

Debreu, G. (1951). "The Coefficient of Resource Utilization". In: *Econometrica* 19, pp. 273–292.

Delgado, M. A., J. C. Fariñas, and S. Ruano (2002). "Firm Productivity and Export Markets: A Non-parametric Approach". In: *Journal of International Economics* 57, pp. 397–422.

Deprins, D., L. Simar, and H. Tulkens (1984). “Measuring Labor-Efficiency in Post Offices”. In: *The Performance of Public Enterprises: Concepts and Measurement*. Ed. by M. Marchand, P. Prestieau, and H. Tulkens. Amsterdam: New-Holland, pp. 243–268.

Dietz, S. and N. Stern (2008). “Why Economic Analysis Supports Strong Action on Climate Change: A Response to the Stern Review’s Critics”. In: *Review of Environmental Economics and Policy* 2, pp. 94–113.

Dyson, R. G., R. Allen, A. S. Camanho, V. V. Podinovski, C. S. Carrico, and E. A. Shale (2001). “Pitfalls and Protocols in DEA”. In: *European Journal of Operational Research* 132, pp. 245–259.

Eastern Research Group (2009). *Electricity Framework 5 Years Review - Control Technologies Review*.

Eberle, M. K. (2008). “Efficiency in Oil Use and Alternatives to Oil”. In: *Climate Change and Energy Pathways for the Mediterranean: Workshop Proceedings, Cyprus*. Ed. by E. J. Moniz. Dordrecht: Springer, pp. 127–144.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.

EIA (2011). *Electric Power Annual 2009*. Report DOE/EIA-0348(2009).

Emrouznejad, A., B. R. Parker, and G. Tavares (2008). “Evaluation of Research in Efficiency and Productivity: A Survey and Analysis of the First 30 Years of Scholarly Literature in DEA”. In: *Socio-Economic Planning Sciences* 42, pp. 151–157.

Endres, A. (2011). *Environmental Economics: Theory and Policy*. Cambridge (Mass.): Cambridge University Press.

EPA (2012). *Sulfur Dioxide*. [Retrieved May 23, 2012]. URL: http://www.epa.gov/oaqps001/sulfurdioxide/.

European Commission (2011). *A Roadmap for Moving to a Competitive Low Carbon Economy in 2050*. URL: http://ec.europa.eu/clima/policies/roadmap/index_en.htm.

European Environment Agency (2011a). *Annual European Union Greenhouse Gas Inventory 1990-2009 and Inventory Report 2011*.

— (2011b). *Most Carmakers must Further Improve Carbon Efficiency by 2015*. [Retrieved September 28, 2012]. URL: http://www.eea.europa.eu/highlights/most-carmakers-must-further-improve/.

Färe, R. (2010). “Directional Distance Functions and Public Transportation: A Comment”. In: *Transportation Research Part D* 15, pp. 108–109.

Färe, R., R. Grabowski, S. Grosskopf, and S. Kraft (1997). “Efficiency of a Fixed but Allocatable Input: A Non-Parametric Approach”. In: *Economics Letters* 56, pp. 187–193.

Färe, R. and S. Grosskopf (1983). “Measuring Output Efficiency”. In: *European Journal of Operational Research* 13, pp. 173–179.

— (1996a). *Intertemporal Production Frontiers: With Dynamic DEA*. Norwell, Massachusetts (U.S.): Kluwer Academic Publishers.

— (1996b). “Productivity and Intermediate Products: A Frontier Approach”. In: *Economics Letters* 50, pp. 65–70.

Färe, R. and S. Grosskopf (2003). "Nonparametric Productivity Analysis with Undesirable Outputs: Comment". In: *American Journal of Agricultural Economics* 85, pp. 1070–1074.

— (2004). *New Directions: Efficiency and Productivity.* Boston/London/Dordrecht: Kluwer Academic Publishers.

— (2009). "A Comment on Weak Disposability in Nonparametric Production Analysis". In: *American Journal of Agricultural Economics* 91, pp. 535–538.

— (2010). "Directional Distance Functions and Slacks-Based Measures of Efficiency". In: *European Journal of Operational Research* 200, pp. 320–322.

— (2012). "DEA and Technical Change: A New Look". In: *Proceedings of DEA Symposium 2012*, pp. 1–7.

Färe, R., S. Grosskopf, and F. Hernandez-Sancho (2004). "Environmental Performance: An Index Number Approach". In: *Resource and Energy Economics* 26, pp. 343–352.

Färe, R., S. Grosskopf, B. Lindgren, and P. Roos (1992). "Productivity Changes in Swedish Pharmacies 1980-1989: A Non-Parametric Malmquist Approach". In: *Journal of Productivity Analysis* 3, pp. 85–101.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985). *The Measurement of Efficiency of Production.* Dordrecht: Kluwer Academic Publishers.

Färe, R., S. Grosskopf, C. A. K. Lovell, and C. Pasurka (1989). "Multilateral Productivity Comparisons when some Outputs are Undesirable: A Nonparametric Approach". In: *Review of Economics and Statistics* 71, pp. 90–98.

Färe, R., S. Grosskopf, T. Lundgren, P.-O. Marklund, and W. Zhou (2012). "Productivity: Should we Include Bads?" In: *CERE Working Paper* 2012:13.

Färe, R., S. Grosskopf, D. Magaritis, and W. L. Weber (2012). "Technological Change and Timing Reductions in Greenhouse Gas Emissions". In: *Journal of Productivity Analysis* 37, pp. 205–216.

Färe, R., S. Grosskopf, D.-W. Noh, and W. Weber (2005). "Characteristics of a Polluting Technology: Theory and Practice". In: *Journal of Econometrics* 126, pp. 469–492.

Färe, R., S. Grosskopf, C. Parsurka, and W. Weber (2012). "Substitutability Among Undesirable Outputs". In: *Applied Economics* 44, pp. 39–47.

Färe, R., S. Grosskopf, and C. Pasurka (2001). "Accounting for Air Pollution Measures of State Manufacturing Productivity Growth". In: *Journal of Regional Science* 41, pp. 381–409.

— (2007). "Pollution Abatement Activities and Traditional Productivity". In: *Ecological Economics* 62, pp. 673–682.

— (forthcoming). "Joint Production of Good and Bad Outputs with a Network Application". In: *Encyclopedia of Energy, Natural Resources and Environmental Economics.* Elsevier.

Färe, R., S. Grosskopf, and G. Whittaker (2011). "Directional Distance Functions: With Endogenous Directions". Paper presented at the XII European Workshop on Efficiency and Productivity Analysis, Verona.

Färe, R. and C. A. K. Lovell (1978). "Measuring the Technical Efficiency of Production". In: *Journal of Economic Theory* 19, pp. 150–162.

Färe, R. and D. Primont (1995). *Multi-Output Production and Duality: Theory and Applications.* Boston/London/Dordrecht: Kluwer Academic Publishers.

Fariñas, J. C. and S. Ruano (2005). "Firm Productivity, Heterogeneity, Sunk Costs and Market Selection". In: *International Journal of Industrial Organization* 23, pp. 505–534.

Farrell, M. J. (1957). "The Measurement of Productive Efficiency". In: *Journal of the Royal Statistic Society* 120, pp. 253–281.

Fernandes, E. and R. R. Pacheco (2002). "Efficient Use of Airport Capacity". In: *Transportation Reseach Part A* 36, pp. 225–238.

Fernandez-Castro, A. S. and P. C. Smith (2002). "Lancaster's Characteristics Approach Revisited: Product Selection Using Non-Parametric Methods". In: *Managerial and Decision Economics* 23, pp. 83–91.

Fischer, A. R. H., H. Tobi, and A. Ronteltap (2012). "When Natural met Social: A Review of Collaboration between the Natural and Social Sciences". In: *Interdisciplinary Science Reviews* 36, pp. 341–358.

Fisher, I. (1922). *The Making of Index Numbers – A Study of Their Varieties, Tests, and Reliability*. Boston: Houghton Mifflin.

Forster, P. V., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D. W. Fahey, J. Haywood, J. Lean, D. C. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz, and R. Van Dorland (2007). "IPCC Guidelines for National Greenhouse Gas Inventories". In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller. Cambridge: Cambridge University Press, pp. 129–234.

Førsund, F. R. (2009). "Good Modelling of Bad Outputs: Pollution and Multiple-Output Production". In: *International Review of Environmental and Resource Economics* 3, pp. 1–38.

Førsund, F. R. and N. Sarafoglu (2002). "On the Origins of Data Envelopment Analysis". In: *Journal of Productivity Analysis* 17, pp. 23–40.

Fried, H. O., C. A. K. Lovell, and S. S. Schmidt (2008). "Efficiency and Productivity". In: *The Measurement of Productive Efficiency and Productivity Growth*. Ed. by H. O. Fried, C. A. K. Lovell, and S. S. Schmidt. Oxford University Press, pp. 3–91.

Frondel, M., J. Horbach, and K. Rennings (2007). "End-of-Pipe or Cleaner Production? An Empirical Comparison of Environmental Innovation Decisions Across OECD Countries". In: *Business Strategy and the Environment* 16, pp. 571–584.

Fukuyama, H. and R. Matousek (2011). "Efficiency of Turkish Banking: Two-Stage Network System. Variable Returns to Scale Model". In: *Journal of International Fincanial Markets, Institutions & Money* 21, pp. 75–91.

González, E., J. Ventura, and A. Cárcaba (2012). "Product Efficiency in the Spanish Automobile Market". In: *Investigaciones Europeas de Dirección y Economía de la Empresa*. DOI: 10.1016/j.iedee.2012.07.003.

Greene, W. H. (2008). *Econometric Analysis*. 6th ed. New Jersey: Pearson Education.

Grosskopf, S. (2003). "Some Remarks on Productivity and its Decompositions". In: *Journal of Productivity Analysis* 20, pp. 459–474.

Hackman, S. T. (2008). *Production Economics - Integrating the Microeconomic and Engineering Perspectives*. Berlin: Springer.

Hailu, A. and T. S. Veeman (2001). “Non-Parametric Productivity Analysis with Undesirable Outputs: An Application to the Canadian Pulp and Paper Industry”. In: *American Journal of Agricultural Economics* 83, pp. 605–616.

Hampf, B. and J. J. Krüger (2010). “Technical Efficiency of Automobiles – A Nonparametric Approach Incorporating Carbon Dioxide Emissions”. In: *Darmstadt Discussion Papers in Economics* 198.

— (2012). “Optimal Directions for Directional Distance Functions: An Exploration of Potential Reductions of Greenhouse Gases”. Unpublished Manuscript, Darmstadt University of Technology.

Harvey, L. D. D. (1993). “A Guide to Global Warming Potentials (GWPs)”. In: *Energy Policy* 21, pp. 24–34.

Henderson, D. J. and V. Zelenyuk (2007). “Testing for (Efficiency) Catching-Up”. In: *Southern Economic Journal* 73, pp. 1003–1019.

Heston, Alan, Robert Summers, and Bettina Aten (2009). *Penn World Table Version 6.3*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, August 2009.

— (2011). *Penn World Table Version 7.0*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, June 2011.

Hua, Z. and Y. Bian (2008). “Performance Measurement for Network DEA with Undesirable Factors”. In: *International Journal of Management and Decision Making* 9, pp. 141–153.

Jaffe, A. B., R. G. Newell, and R. N. Stavins (2003). “Technological Change and the Environment”. In: *Handbook of Environmental Economics*. Ed. by K-G. Mahler and J. R. Vincent. Vol. 1. Amsterdam: New-Holland, pp. 461–516.

Jamasb, T. and M. Pollitt (2001). “Benchmarking and Regulation: International Electricity Experience”. In: *Utilities Policy* 9, pp. 107–130.

Jeon, B. M. and R. C. Sickles (2004). “The Role of Environmental Factors in Growth Accounting”. In: *Journal of Applied Econometrics* 19, pp. 567–591.

Kao, C. (2009). “Efficiency Decomposition in Network Data Envelopment Analysis: A Relational Model”. In: *European Journal of Operational Research* 192, pp. 949–962.

Karl, T. R. and K. E. Trenberth (2003). “Modern Global Climate Change”. In: *Science* 302, pp. 1719–17.

Kemeny, J. G., O. Morgenstern, and G. L. Thompson (1956). “A Generalization of the von Neumann Model of an Expanding Economy”. In: *Econometrica* 24, pp. 115–135.

Kneip, A., B. U. Park, and L. Simar (1998). “A Note on the Convergence of Nonparametric DEA Estimators for Production Efficiency Scores”. In: *Econometric Theory* 14, pp. 783–793.

Koopmans, T. C. (1951). “Analysis of Production as an Efficient Combination of Activities”. In: *Activity Analysis of Production and Allocation*. Ed. by T. C. Koopmans. New York: John Wiley & Sons, pp. 33–97.

Kortelainen, M. and T. Kuosmanen (2007). “Eco-Efficiency Analysis of Consumer Durables Using Absolute Shadow Prices”. In: *Journal of Productivity Analysis* 28, pp. 57–69.

Krautzberger, L. and H. Wetzel (2012). "Transport and CO_2: Productivity Growth and Carbon Dioxide Emissions in the European Commercial Transport Industry". In: *Environmental and Resource Economics* 53, pp. 435–454.

Krüger, J. J. (2012). "A Monte Carlo Study of Old and New Frontier Methods for Efficiency Measurement". In: *European Journal of Operational Research* 222, pp. 137–148.

Kumar, S. (2006). "Environmentally Sensitive Productivity Growth: A Global Analysis using Malmquist-Luenberger Index". In: *Ecological Economics* 56, pp. 280–293.

Kumar, S. and R. R. Russell (2002). "Technological Change, Technological Catch-Up, and Capital Depening: Relative Contributions to Growth and Convergence". In: *American Economic Review* 92, pp. 527–548.

Kumbhakar, S. C. and C. A. K. Lovell (2000). *Stochastic Frontier Analysis*. Cambridge (Mass.): Cambridge University Press.

Kuosmanen, T. and M. Kortelainen (2005). "Measuring Eco-Efficiency of Production with Data Envelopment Analysis". In: *Journal of Industrial Ecology* 9, pp. 59–72.

— (2007). "Valuing Environmental Factors in Cost-Benefit Analysis Using Data Envelopment Analysis". In: *Ecological Economics* 62, pp. 56–65.

— (2012). "Stochastic Non-Smooth Envelopment of Data: Semi-Parametric Frontier Estimation Subject to Shape Constraints". In: *Journal of Productivity Analysis* 38, pp. 11–28.

Kuosmanen, T. and T. Post (2001). "Measuring Economic Efficiency with Incomplete Price Information: With an Application to European Commercial Banks". In: *European Journal of Operational Research* 134, pp. 43–58.

Lauwers, L. (2009). "Justifying the Incorporation of the Materials Balance Principle into Frontier-Based Eco-Efficiency Models". In: *Ecological Economics* 68, pp. 1605–1614.

Levy, H. (1992). "Stochastic Dominance and Expected Utility: Survey and Analysis". In: *Management Science* 38, pp. 555–593.

Li, Q. (1996). "Nonparametric Testing of Closeness between Two Unknown Distribution Functions". In: *Econometric Reviews* 15, pp. 261–274.

— (1999). "Nonparametric Testing the Similarity of Two Unknown Density Functions: Local Power and Bootstrap Analysis". In: *Journal of Nonparametric Statistics* 11, pp. 189–213.

Liu, J. S., L. Y. Y. Lu, W.-M. Lu, and B. J. Y. Lin (2013). "Data Envelopment Analysis 1978-2010: A Citation-Based Literature Survey". In: *Omega* 41, pp. 3–15.

Liu, W. B., W. Meng, X. X. Li, and D. Q. Zhang (2010). "DEA Models with Undesirable Outputs". In: *Annals of Operations Research* 173, pp. 177–194.

Los, B. and B. Verspagen (2009). "Localized Innovation, Localized Diffusion, and the Environment: An Analysis of Reductions of CO_2 Emissions by Passenger Cars". In: *Journal of Evolutionary Economics* 19, pp. 507–526.

Lovell, C. A. K. (2003). "The Decomposition of Malmquist Productivity Indexes". In: *Journal of Productiocvity Analysis* 20, pp. 437–458.

Lovell, C. A. K., J. T. Pastor, and J. A. Turner (1995). "Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries". In: *European Journal of Operational Research* 87, pp. 507–518.

Lozano, S. (2011). "Scale and Cost Efficiency Analysis of Networks of Processes". In: *Expert Systems with Applications* 38, pp. 612–6617.

Lozano, S. and E. Gutiérrez (2008). "Non-Parametric Frontier Approach to Modelling the Relationships among Population, GDP, Energy Consumption and CO_2 Emissions". In: *Ecological Economics* 66, pp. 687–699.

Luenberger, D. and Y. Ye (2008). *Linear and Nonlinear Programming.* 3. ed. New York: Springer.

Luenberger, D. G. (1992). "Benefit Functions and Duality". In: *Journal of Mathematical Economics* 21, pp. 461–481.

MacLean, H. L. and L. B. Lave (2000). "Environmental Implications of Alternative-Fueled Automobiles: Air Quality and Greenhouse Gas Tradeoffs". In: *Environmental Science & Technology* 34, pp. 225–231.

Malmquist, S. (1953). "Index Numbers and Indifference Surfaces". In: *Trabajos de Estadística y de Investigacíon Operativa* 4, pp. 209–242.

Managi, S. (2006). "Are there Increasing Returns to Pollution Abatement? Empirical Analytics of the Environmental Kuznets Curve in Pesticides". In: *Ecological Economics* 58, pp. 617–636.

McMichael, A. J., R. E. Woodruff, and S. Hales (2006). "Climate Change and Human Health: Present and Future Risks". In: *Lancet* 367, pp. 859–869.

Meeusen, W. and J. van den Broeck (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error". In: *International Economic Review* 18, pp. 435–444.

Mendelsohn, R. O. (2006). "A Critique of the Stern Report". In: *Regulation* 29, pp. 42–46.

Moomaw, W. R. (1996). "Industrial Emissions of Greenhouse Gases". In: *Energy Policy* 24, pp. 951–968.

Murillo-Zamorano, L. R. (2004). "Economic Efficiency and Frontier Techniques". In: *Journal of Economic Surveys* 18, pp. 33–77.

Murty, S., R. R. Russell, and S. B. Levkoff (2012). "On Modeling Pollution-Generating Technologies". In: *Journal of Environmental Economics and Management* 64, pp. 117–135.

Nicholls, R. J. and A. Cazenave (2010). "Sea-Level Rise and Its Impact on Coastal Zones". In: *Science* 328, pp. 1517–1520.

Nonneman, W. and P. Vanhoudt (1996). "A Further Augmentation of the Solow Model and the Empirics of Economic Growth for OECD Countries". In: *Quarterly Journal of Economics* 111, pp. 943–953.

Nordhaus, W. D. (2007). "A Review of the 'Stern Review on the Economics of Climate Change'". In: *Journal of Economic Literature* 45, pp. 686–702.

Nordhaus, W. D. and J. G. Boyer (1999). "Requiem for Kyoto: An Economic Analysis of the Kyoto Protocol". In: *Energy Journal* 20 (Special Issue I), pp. 93–130.

Nunamaker, T. R. (1985). "Using Data Envelopment Analysis to Measure the Efficiency of Non-Profit Organizations: A Cricital Evaluation". In: *Managerial and Decision Economics* 6, pp. 50–58.

Odeck, J. and L. Hjalmarsson (1996). "The Performance of Trucks – An Evaluation using Data Envelopment Analysis". In: *Transportation Planning and Technology* 20, pp. 49–66.

Oh, D.-H. (2010). "A Global Malmquist-Luenberger Productivity Index". In: *Journal of Productivity Analysis* 34, pp. 183–197.

Oh, D.-H. and A. Heshmati (2010). "A Sequential Malmquist-Luenberger Productivity Index: Environmentally Sensitive Productivity Growth Considering the Progressive Nature of Technology". In: *Energy Economics* 32, pp. 1345–1355.

Oh, I., J.-D. Lee, S. Hwang, and A. Heshmati (2010). "Analysis of Product Efficiency in the Korean Automobile Market from a Consumers Perspective". In: *Empirical Economics* 38, pp. 119–137.

Olsthoorn, X., D. Tyteca, W. Wehrmayer, and M. Wagner (2001). "Environmental Indicators for Business: A Review of the Literature and Standardisation Methods". In: *Journal of Cleaner Production* 9, pp. 453–463.

O'Neill, B. C. (2000). "The Jury ist Still Out on Global Warming Potentials". In: *Climate Change* 44, pp. 427–443.

Oreskes, N. (2004). "The Scientific Consensus on Climate Change". In: *Science* 306, p. 1686.

Otsuki, T. (2012). "Nonparametric Measurement of the Overall Shift in the Technology Frontier: An Application to Multiple-Output Agricultural Production Data in the Brazilian Amazon". In: *Empirical Economics*. DOI: 10.1007/s00181-012-0582-4.

Pagan, A. and A. Ullah (1999). *Nonparametric Econometrics*. Cambridge University Press.

Papahristodoulou, C. (1997). "A DEA Model to Evaluate Car Efficiency". In: *Applied Economics* 29, pp. 1493–1508.

Pareto, V. (1909). *Manuel d'Economie Politique*. Paris: Giars & Briere.

Park, B. U., S.-O. Jeong, and L. Simar (2010). "Asymptotic Distribution of Conical-Hull Estimators of Directional Edges". In: *Annals of Statistics* 38, pp. 1320–1340.

Park, W. G. (1995). "International R&D Spillovers and OECD Economic Growth". In: *Economic Inquiry* 33, pp. 571–591.

Pastor, J. T. and C. A. K. Lovell (2005). "A Global Malmquist Productivity Index". In: *Economics Letters* 88, pp. 266–271.

Pasurka, C. (2006). "Decomposing Electric Power Plant Emissions within a Joint Production Framework". In: *Energy Economics* 28, pp. 26–43.

Pedraja-Chaparro, F., J. Salinas-Jiminez, and P. Smith (1997). "On the Role of Weight Restrictions in Data Envelopment Analysis". In: *Journal of Productivity Analysis* 8, pp. 215–230.

Pethig, R. (2006). "Non-Linear Production, Abatement, Pollution and Materials Balance Reconsidered". In: *Journal of Environmental Economics and Management* 51, pp. 185–204.

Peyrache, A. and C. Daraio (2012). "Empirical Tools to Assess the Sensitivity of Directional Distance Function to Direction Selection". In: *Applied Economics* 44, pp. 933–943.

Picazo-Tadeo, A. J., E. Reig-Martínez, and F. Hernández-Sancho (2005). "Directional Distance Functions and Environmental Regulation". In: *Resource and Energy Economics* 27, pp. 131–142.

Pielke Jr., R. A. (2005). "Misdefining 'climate change': Consequences for Science and Action". In: *Envrionmental Science & Policy* 8, pp. 548–561.

Plotkin, S. (2004). “Is Bigger Better?: Moving Toward a Dispassionate View of SUVs”. In: *Environment: Science and Policy for Sustainable Development* 46, pp. 8–21.

Popp, D., R. G. Newell, and A. B. Jaffe (2010). “Energy, the Environment, and Technological Change”. In: *Handbook of the Economics of Innovation*. Ed. by B. H. Hall and N. Rosenberg. Vol. 2. Amsterdam: New-Holland, pp. 873–937.

Portela, M. C. A. S. and E. Thanassoulis (2001). “Decomposing School and School-Type Efficiency”. In: *European Journal of Operational Research* 132, pp. 357–373.

Quiggin, J. (2008). “Stern and his Critics on Discounting and Climate Change: An Editorial Essay”. In: *Climate Change* 89, pp. 195–205.

Raven, J., K. Caldeira, H. Elderfield, O. Hoegh-Guldberg, P. Liss, U. Riebesell, J. Shepherd, C. Turley, A. Watson, R. Heap, R. Banes, and R. Quinn (2005). “Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide”. Policy document 12/05. London: The Royal Society.

Reinhard, S., C. A. K. Lovell, and G. J. Thijssen (2000). “Environmental Efficiency with Multiple Environmentally Detrimental Variables; Estimated with SFA and DEA”. In: *European Journal of Operational Research* 121, pp. 287–303.

Scheel, H. (2001). “Undesirable Outputs in Efficiency Valuations”. In: *European Journal of Operational Research* 13, pp. 400–410.

Schäfer, F. and R. van Basshuysen (1995). *Reduced Emissions and Fuel Consumption in Automobile Engines*. Wien: Springer.

Schleich, J., W. Eichhammer, U. Boede, F. Gagelmann, E. Jochem, B. Schlomann, and H.-J. Ziesing (2001). “Greenhouse Gas Reductions in Germany – Lucky Strike or Hard Work?” In: *Climate Policy* 1, pp. 363–380.

Seiford, L. M. and J. Zhu (2002). “Modeling Undesirable Factors in Efficiency Evaluation”. In: *European Journal of Operational Research* 142, pp. 16–20.

Shephard, R. W. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.

Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures*. New York: Chapman & Hall.

Shestalova, V. (2003). “Sequential Malmquist Indices of Productivity Growth: An Application to OECD Industrial Activities”. In: *Journal of Productivity Analysis* 19, pp. 211–226.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Simar, L. (2003). “Detecting Outliers in Frontier Models: A Simple Approach”. In: *Journal of Productivity Analysis* 20, pp. 391–424.

— (2007). “How to Improve the Performances of DEA/FDH Estimators in the Presence of Noise?” In: *Journal of Productivity Analysis* 28, pp. 183–201.

Simar, L. and P. W. Wilson (1998). “Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models”. In: *Management Science* 44, pp. 49–61.

— (1999). “Estimating and Bootstrapping Malmquist Indices”. In: *European Journal of Operational Research* 115, pp. 459–471.

Simar, L. and P. W. Wilson (2000). "A General Methodology for Bootstrapping in Nonparametric Frontier Models". In: *Journal of Applied Statistics* 27, pp. 779–802.

— (2002). "Non-Parametric Tests of Returns to Scale". In: *European Journal of Operational Research* 139, pp. 115–132.

— (2007). "Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes". In: *Journal of Econometrics* 136, pp. 31–64.

— (2008). "Statistical Inference in Nonparametric Frontier Models". In: *The Measurement of Productive Efficiency and Productivity Growth.* Ed. by H. O. Fried, C. A. K. Lovell, and S. S. Schmidt. Oxford University Press, pp. 421–521.

Simar, L. and V. Zelenyuk (2006). "On Testing Equality of Distributions of Technical Efficiency Scores". In: *Econometric Reviews* 25, pp. 497–522.

Simon, H. A. (1951). "Effects of Technological Change in a Linear Model". In: *Activity Analysis of Production and Allocation.* Ed. by T. C. Koopmans. New York: John Wiley & Sons, pp. 260–281.

Simpson, G. (2005). "Programmatic Efficiency Comparisons between Unequally Sized Groups of DMUs in DEA". In: *Journal of the Royal Operational Research Society* 56, pp. 1431–1438.

— (2007). "A Cautionary Note on Methods of Comparing Programmatic Efficiency between Two or More Groups of DMUs in Data Envelopment Analysis". In: *Journal of Productivity Analysis* 28, pp. 141–147.

Solow, R. M. (1997). "Reply - Georgescu-Roegen versus Solow/Stiglitz". In: *Ecological Economics* 22, pp. 267–268.

Srivastava, R. K. (2000). "Contolling SO_2 Emissions: A Review of Technologies". In: *U.S. Environmental Protection Agency, Office of Research and Development, EPA/600/R-00/093.*

Srivastava, R. K. and W. Josewicz (2001). "Flue Gas Desulfurization: The State of the Art". In: *Journal of the Air & Waste Management Association* 51, pp. 1676–1688.

Staat, M., H. H. Bauer, and M. Hammerschmidt (2002). "Structuring Product-Markets: An Approach Based on Consumer Value". In: *Marketing Theory and Applications* 13, pp. 205–212.

Stern, N. (2007). *The Economics of Climate Change - The Stern Review.* Cambridge (Mass.): Cambridge University Press.

Stokes, J. R. and P. W. Claar (2004). "Evaluating Agricultural Tractor Performance: A Data Envelopment Analysis Approach". In: *SAE Technical Paper* 2004-01-2705. DOI: 10.4271/2004-01-2705.

Sueyoshi, T., M. Goto, and T. Ueno (2010). "Performance Analysis of US Coal-Fired Power Plants by Measuring Three DEA Efficiencies". In: *Energy Policy* 38, pp. 1675–1688.

Sullivan, J. L., R. E. Baker, B. A. Boyer, R. H. Hammerle, T. E. Kenney, L. Muniz, and T. J. Wallington (2004). "CO_2 Emission Benefit of Diesel (Versus Gasoline) Powered Vehicles". In: *Environmental Science & Technology* 38, pp. 3217–3223.

Talluri, S. (2000). "A Benchmarking Method for Business-Process Reengineering and Improvement". In: *International Journal of Flexible Manufacturing Systems* 12, pp. 291–304.

Taskin, F. and O. Zaim (2000). "Searching for a Kuznets Curve in Environmental Efficiency Using Kernel Estimation". In: *Economics Letters* 68, pp. 217–223.

ten Raa, T. (2008). "Debreus's Coefficient of Resource Utilization, the Solow Residual, and TFP: The Connection by Leontief Preferences". In: *Journal of Productivity Analysis* 30, pp. 191–199.

Thanassoulis, E., M. C. A. S. Portela, and O. Despić (2008). "Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis". In: *The Measurement of Productive Efficiency and Productivity Growth*. Ed. by H. O. Fried, C. A. K. Lovell, and S. S. Schmidt. Oxford University Press, pp. 251–420.

Toffel, M. W. and J. D. Marshall (2004). "Improving Environmental Performance Assessment - A Comparative Analysis of Weighting Methods Used to Evaluate Chemical Release Inventories". In: *Journal of Industrial Ecology* 8, pp. 143–172.

Tol, R. S. J. (2003). "Is the Uncertainty about Climate Change Too Large for Expected Cost-Benefit Analysis?" In: *Climate Change* 56, pp. 265–289.

— (2006). "The Stern Review of the Economics of Climate Change: A Comment". In: *Energy & Environment* 17, pp. 977–981.

— (2009). "The Economic Effects of Climate Change". In: *Journal of Economic Perspectives* 23, pp. 29–51.

Tol, R. S. J., T. E. Downing, O. J. Kuik, and J. B. Smith (2004). "Distributional Aspects of Climate Change Impacts". In: *Global Environmental Change* 14, pp. 259–272.

Tone, K. (2001). "A Slacks-Based Measure of Efficiency in Data Envelopment Analysis". In: *European Journal of Operational Research* 130, pp. 498–509.

Tulkens, H. and P. Vanden Eeckaut (1995). "Non-Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects". In: *European Journal of Operational Research* 80, pp. 474–499.

Tyteca, D. (1996). "On the Measurement of the Environmental Performance of Firms - A Literature Review and a Productive Efficiency Perspective". In: *Journal of Environmental Management* 46, pp. 281–308.

UNFCCC (2008). *Kyoto Protocol Reference Manual - On Accounting of Emissions and Assigned Amount*.

van Ark, B., M. O'Mahony, and M. P. Timmer (2008). "The Productivity Gap Between Europe and the United States: Trends and Causes". In: *Journal of Economic Perspectives* 22, pp. 25–44.

Wagner, U. J. (2002). "The Desing of Stable International Environmental Agreements: Economic Theory and Political Economy". In: *Journal of Economic Surveys* 15, pp. 377–411.

Ward, P. L. (2009). "Sulfur Dioxide Initiates Global Climate Change in Four Ways". In: *Thin Solid Films* 517, pp. 3188–3203.

Weber, W. L. and B. Domazlicky (2001). "Productivity Growth and Pollution in State Manufacturing". In: *Review of Economics and Statistics* 83, pp. 195–199.

Whitman, Requardt & Associates (2009). *The Handy-Whitman Index of Public Utility Construction Costs: Trends of Construction Costs*.

Wilson, P. W. (1993). "Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs". In: *Journal of Business and Economic Statistics* 11, pp. 319–323.

Wilson, P. W. (1995). "Detecting Influential Observations in Data Envelopment Analysis". In: *Journal of Productivity Analysis* 6, pp. 27–45.

— (2008). "FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R". In: *Socio-Economic Planning Sciences* 42, pp. 247–254.

Winsten, C. B. (1957). "Discussion on Mr. Farrell's Paper". In: *Journal of the Royal Statistical Society* 120, pp. 282–284.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd Edition. Cambridge (Mass.): The MIT Press.

World Bank (2011). *The World at a Glance - Key Development Indicators from the World Bank*. [Retrieved September 03, 2012]. URL: http://data.worldbank.org/topic/environment.

Yaisawarng, S. and J. D. Klein (1994). "The Effects of Sulfur Dioxide Controls on Productivity Changes in the U.S. Electric Power Industry". In: *Review of Economics and Statistics* 76, pp. 447–460.

Yang, C-C., C.-K. Hsaio, and M.-M. Yu (2008). "Technical Efficiency and Impact of Environmental Regulations in Farrow-to-Finish Swine Production in Taiwan". In: *Agricultural Economics* 39, pp. 51–61.

Yörük, B. K. and O. Zaim (2005). "Productivity Growth in OECD Countries: A Comparison with Malmquist Indices". In: *Journal of Comparative Economics* 33, pp. 401–420.

Zeitsch, J. and D. Lawrence (1996). "Decomposing Economic Inefficiency in Base-Load Power Plants". In: *Journal of Productivity Analysis* 7, pp. 359–378.

Zervas, E. (2009). "Analysis of CO_2 Emissions and Other Characteristics of New German Passenger Cars". In: *Energy & Fuels* 23, pp. 244–252.

— (2010). "Analysis of the CO_2 Emissions and of the other Characteristics of the European Market of New Passenger Cars. 2. Segment Analysis". In: *Energy Policy* 38, pp. 5426–5441.

Zervas, E. and E. Diamantopoulos (2009). "Relationships Between CO_2 Emitted from New Passenger Cars in European Union and their Engine and Vehicle Characteristics". In: *SAE Technical Paper* 2009-01-1825. DOI: 10.4271/2009-01-1825.

Zhou, P., B. W. Ang, and K. L. Poh (2008a). "A Survey of Data Envelopment Analysis in Energy and Environmental Studies". In: *European Journal of Operational Research* 189, pp. 1–18.

— (2008b). "Measuring Environmental Performance under Different Environmental DEA Technologies". In: *Energy Economics* 30, pp. 1–14.

Zotter, K. A. (2004). "'End-of-Pipe' versus 'Process-Integrated' Water Conservation Solutions: A Comparison of Planning, Implementation and Operating Phases". In: *Journal of Cleaner Production* 12, pp. 685–695.

Appendix A

Table A.1: Automobile terminologies

Category	Example
Brand	VW
Product line	Golf
Model	Golf 1.6
Model variant	Golf 1.6 Trendline

Source: Cantner et al. (2012. p. 8)

Table A.2: p-values of tests for first-order stochastic dominance (between engine types)

H_0	Efficiency type	$\widehat{\rho}$	$1+\widehat{\beta}_{\mathrm{O}}$	$1+\widehat{\beta}_{\mathrm{UO}}$
Gasoline $\succeq_{\mathrm{FSD}}$ Diesel	Overall	0.122	0.147	0.163
	Managerial	0.207	0.054	0.024
	Program	0.185	0.852	0.993
Gasoline $\succeq_{\mathrm{FSD}}$ Natural gas	Overall	0.000	0.000	0.000
	Managerial	0.962	0.961	0.938
	Program	0.000	0.000	0.000
Diesel $\succeq_{\mathrm{FSD}}$ Gasoline	Overall	0.121	0.164	0.006
	Managerial	0.334	0.885	0.996
	Program	0.000	0.000	0.000
Diesel $\succeq_{\mathrm{FSD}}$ Natural gas	Overall	0.000	0.000	0.000
	Managerial	0.829	0.933	0.978
	Program	0.000	0.000	0.000
Natural gas $\succeq_{\mathrm{FSD}}$ Gasoline	Overall	0.998	0.987	1.000
	Managerial	0.000	0.000	0.000
	Program	0.990	0.993	0.996
Natural gas $\succeq_{\mathrm{FSD}}$ Diesel	Overall	0.997	0.892	0.979
	Managerial	0.000	0.000	0.000
	Program	0.989	0.987	0.994

Table A.3: p-values of tests for first-order stochastic dominance (within engine types)

H_0	Gasoline			Diesel			Natural Gas		
	Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\widehat{\rho} \succeq_{\text{FSD}} 1+\widehat{\beta}_{\text{O}}$	0.997	0.998	0.000	0.961	0.988	0.609	0.974	0.980	0.807
$1+\widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} \widehat{\rho}$	0.029	0.006	0.824	0.168	0.296	0.000	0.001	0.000	0.008
$\widehat{\rho} \succeq_{\text{FSD}} 1+\widehat{\beta}_{\text{UO}}$	0.997	0.989	0.000[a)]	0.334	0.142	0.988	0.987	0.985	0.934
$1+\widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} \widehat{\rho}$	0.000	0.000	0.022[a)]	0.000	0.000	0.000	0.000	0.000	0.000
$1+\widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} 1+\widehat{\beta}_{\text{UO}}$	0.973	0.959	0.000[b)]	0.265	0.083[c)]	0.003[d)]	0.989	0.674	0.951
$1+\widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} 1+\widehat{\beta}_{\text{O}}$	0.000	0.000	0.000[b)]	0.000	0.000[c)]	0.000[d)]	0.000	0.034	0.000

a) Tests for SSD find $1+\widehat{\beta}_{\text{UO}} \succ_{\text{SSD}} \widehat{\rho}$.
b) Tests for SSD find $1+\widehat{\beta}_{\text{O}} \succ_{\text{SSD}} 1+\widehat{\beta}_{\text{UO}}$.
c) Tests for SSD find $1+\widehat{\beta}_{\text{O}} \succ_{\text{SSD}} 1+\widehat{\beta}_{\text{UO}}$.
d) Tests for SSD find $1+\widehat{\beta}_{\text{O}} \succ_{\text{SSD}} 1+\widehat{\beta}_{\text{UO}}$.

Table A.4: p-values of tests for first-order stochastic dominance (within car classes)

H_0	Compact cars			Middle class			Upper class		
	Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.769	0.995	0.010	0.994	0.924	0.006	0.998	0.996	0.429
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} \widehat{\rho}$	0.074	0.005	0.756	0.030	0.001	0.821	0.174	0.090	0.758
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.182	0.038	0.990	0.988	0.922	0.995	0.992	0.994	0.990
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} \widehat{\rho}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.376	0.014[e)]	0.971	0.929	0.579	0.994	0.944	0.970	0.992
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.000	0.000[e)]	0.000	0.000	0.000	0.000	0.000	0.000	0.000

e) Tests for SSD find $1 + \beta_{\text{O}} \succ_{\text{SSD}} 1 + \beta_{\text{UO}}$.

Table A.5: p-values of tests for first-order stochastic dominance (within SUVs)

H_0	SUVs		
	Overall	Managerial	Program
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.936	0.864	0.000
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} \widehat{\rho}$	0.255	0.000	0.969
$\widehat{\rho} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.907	0.975	0.990
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} \widehat{\rho}$	0.000	0.000	0.000
$1 + \widehat{\beta}_{\text{O}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{UO}}$	0.943	0.008	0.994
$1 + \widehat{\beta}_{\text{UO}} \succeq_{\text{FSD}} 1 + \widehat{\beta}_{\text{O}}$	0.000	0.381	0.000

Table A.6: p-values of tests for first-order stochastic dominance (between car classes)

H_0	Efficiency type	$\widehat{\rho}$	$1+\widehat{\beta}_{\mathrm{O}}$	$1+\widehat{\beta}_{\mathrm{UO}}$
Compact $\succeq_{\mathrm{FSD}}$ Middle	Overall	0.000	0.000	0.000
	Managerial	0.000	0.000	0.001
	Program	0.006[f)]	0.010	0.000
Compact $\succeq_{\mathrm{FSD}}$ Upper	Overall	0.000	0.000	0.000
	Managerial	0.000	0.000	0.000
	Program	0.000	0.000	0.000
Middle $\succeq_{\mathrm{FSD}}$ Compact	Overall	0.999	0.995	0.496
	Managerial	0.998	0.998	0.538
	Program	0.061[f)]	0.128	0.846
Middle $\succeq_{\mathrm{FSD}}$ Upper	Overall	0.000	0.000	0.002
	Managerial	0.045[g)]	0.011[h)]	0.090
	Program	0.000	0.000	0.000
Upper $\succeq_{\mathrm{FSD}}$ Compact	Overall	0.999	0.998	0.995
	Managerial	0.996	0.998	0.145
	Program	0.988	0.992	0.989
Upper $\succeq_{\mathrm{FSD}}$ Middle	Overall	0.701	0.977	0.999
	Managerial	0.000[g)]	0.000[h)]	0.135
	Program	0.996	0.996	0.996

[f)] Tests for SSD find no dominance.

[g)] Tests for SSD find Middle $\succ_{\mathrm{SSD}}$ Upper.

[h)] Tests for SSD find Upper $\succ_{\mathrm{SSD}}$ Middle.

Appendix B

Table B.1: Efficiency types, abbreviations and formal definitions

Efficiency type	Abbreviation	Formal definition
Network		
Environmental	EE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''^*}{\boldsymbol{\omega}^T \boldsymbol{u}''}$
Production	PE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}'^*}{\boldsymbol{\omega}^T \boldsymbol{u}'}$
Abatement	AE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''^*}{\boldsymbol{\omega}^T \boldsymbol{u}'^*} \cdot \left[\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''}{\boldsymbol{\omega}^T \boldsymbol{u}'}\right]^{-1}$
Production stage		
Stage	SE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Prod}}}{\boldsymbol{\omega}^T \boldsymbol{u}'}$
Technical	TE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Tech}}}{\boldsymbol{\omega}^T \boldsymbol{u}'}$
Allocative	EAE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Prod}}}{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Tech}}}$
"Black box" bias	Bias	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}'^*}{\boldsymbol{\omega}^T \boldsymbol{u}'^{\text{Prod}}}$
Abatement stage		
Stage	SE	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''^{\text{Abat}}}{\boldsymbol{\omega}^T \boldsymbol{u}'} \cdot \left[\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''}{\boldsymbol{\omega}^T \boldsymbol{u}'}\right]^{-1}$
"Black box" bias	Bias	$\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''^*}{\boldsymbol{\omega}^T \boldsymbol{u}'^*} \cdot \left[\dfrac{\boldsymbol{\omega}^T \boldsymbol{u}''^{\text{Abat}}}{\boldsymbol{\omega}^T \boldsymbol{u}'}\right]^{-1}$

Table B.2: Power plant results of the efficiency analysis (CRS)

		Network			Production				Abatement	
Plant ID	Plant name	EE	PE	AE	SE	TE	EAE	Bias	SE	Bias
6137	A B Brown	0.1824	0.7719	0.2363	0.7719	0.8031	0.9612	1.0000	0.3713	0.6365
994	AES Petersburg	0.0575	0.8758	0.0657	0.8560	0.8944	0.9570	1.0232	0.0657	1.0000
1915	Allen S King	0.0676	0.9630	0.0702	0.9526	0.9905	0.9618	1.0108	0.0702	1.0000
8042	Belews Creek	0.3649	0.9952	0.3667	0.9570	1.0000	0.9570	1.0399	0.3667	1.0000
645	Big Bend	0.6968	0.8478	0.8219	0.8478	0.8841	0.9590	1.0000	1.0000	0.8219
1363	Cane Run	0.0861	0.8467	0.1017	0.8415	0.8665	0.9711	1.0062	0.1017	1.0000
1001	Cayuga	0.7113	0.8972	0.7928	0.8698	0.9074	0.9585	1.0316	0.7928	1.0000
469	Cherokee	0.0200	0.7982	0.0250	0.7982	0.8285	0.9635	1.0000	0.0250	1.0000
113	Cholla	0.0572	0.8797	0.0651	0.8651	0.9019	0.9592	1.0169	0.0651	1.0000
1893	Clay Boswell	0.0701	0.8322	0.0842	0.8210	0.8579	0.9571	1.0136	0.0842	1.0000
470	Comanche	0.0408	0.8388	0.0486	0.8312	0.8670	0.9587	1.0091	0.0486	1.0000
7210	Cope	0.1759	0.9679	0.1817	0.9482	0.9891	0.9587	1.0207	0.2118	0.8581
51	Dolet Hills	0.0301	0.7877	0.0383	0.7705	0.8032	0.9593	1.0223	0.0383	1.0000
6018	East Bend	0.8709	0.8709	1.0000	0.8447	0.8811	0.9586	1.0311	1.0000	1.0000
1356	Ghent	0.4637	0.8762	0.5292	0.8434	0.8810	0.9574	1.0389	1.0000	0.5292
8069	Huntington	0.0711	0.8918	0.0797	0.8721	0.9103	0.9580	1.0226	0.0797	1.0000
2727	Marshall	0.9936	0.9978	0.9958	0.9570	0.9988	0.9581	1.0427	1.0000	0.9958
1364	Mill Creek	0.4469	0.8814	0.5071	0.8524	1.0000	0.8524	1.0340	1.0000	0.5071
4162	Naughton	0.0525	0.8389	0.0625	0.8276	0.8621	0.9600	1.0136	0.0625	1.0000
6085	R M Schahfer	0.0677	0.8046	0.0842	0.7891	0.8236	0.9581	1.0196	0.0842	1.0000
2324	Reid Gardner	0.0983	0.8215	0.1197	0.8215	0.8397	0.9784	1.0000	0.1197	1.0000
6071	Trimble County	0.6530	0.9002	0.7255	0.8990	0.9376	0.9589	1.0013	0.7255	1.0000
477	Valmont	0.0223	1.0000	0.0223	1.0000	1.0000	1.0000	1.0000	0.0223	1.0000

Table B.3: Power plant results of the efficiency analysis (VRS)

		Network			Production				Abatement	
Plant ID	Plant name	EE	PE	AE	SE	TE	EAE	Bias	SE	Bias
6137	A B Brown	0.4331	0.7896	0.5485	0.7896	0.8886	0.8886	1.0000	1.0000	0.5485
994	AES Petersburg	0.1875	0.8934	0.2098	0.8934	0.9103	0.9815	1.0000	0.3307	0.6345
1915	Allen S King	0.9973	0.9845	1.0129	0.9845	0.9949	0.9896	1.0000	1.0000	1.0129
8042	Belews Creek	0.3834	1.0000	0.3834	1.0000	1.0000	1.0000	1.0000	0.3834	1.0000
645	Big Bend	0.7812	0.8835	0.8842	0.8835	0.8845	0.9989	1.0000	1.0000	0.8842
1363	Cane Run	0.1442	0.8858	0.1627	0.8685	0.8685	1.0000	1.0199	0.4000	0.4069
1001	Cayuga	0.7266	0.9031	0.8046	0.9031	0.9288	0.9723	1.0000	1.0000	0.8046
469	Cherokee	0.2684	2.2342	0.1201	0.8236	0.8321	0.9898	2.7128	0.2852	0.4212
113	Cholla	0.1529	1.7051	0.0897	0.9008	0.9028	0.9978	1.8929	0.1547	0.5796
1893	Clay Boswell	0.5721	1.9655	0.2910	0.8493	1.0000	0.8493	2.3142	0.5970	0.4875
470	Comanche	0.5313	2.2795	0.2331	0.8620	0.8707	0.9901	2.6444	0.5650	0.4125
7210	Cope	1.0000	1.0000	1.0000	0.9743	1.0000	0.9743	1.0264	1.0000	1.0000
51	Dolet Hills	0.3853	0.7987	0.4824	0.7987	0.8065	0.9902	1.0000	1.0000	0.4824
6018	East Bend	0.9247	0.8752	1.0566	0.8752	0.8855	0.9884	1.0000	1.0000	1.0566
1356	Ghent	0.7159	0.8805	0.8130	0.8805	0.8813	0.9990	1.0000	1.0000	0.8130
8069	Huntington	0.3706	2.4456	0.1515	0.9075	0.9124	0.9946	2.6948	0.4326	0.3503
2727	Marshall	0.9984	0.9994	0.9990	0.9994	0.9994	1.0000	1.0000	1.0000	0.9990
1364	Mill Creek	0.7112	0.8894	0.7997	0.8894	1.0000	0.8894	1.0000	1.0000	0.7997
4162	Naughton	0.3403	1.1080	0.3071	0.8461	1.0000	0.8461	1.3096	0.3506	0.8760
6085	R M Schahfer	0.0942	0.8173	0.1153	0.8173	0.8283	0.9867	1.0000	0.1238	0.9313
2324	Reid Gardner	1.0000	1.0000	1.0000	0.8471	0.8471	1.0000	1.1805	1.0000	1.0000
6071	Trimble County	1.0000	1.0000	1.0000	0.9273	0.9451	0.9812	1.0784	1.0000	1.0000
477	Valmont	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Appendix C

Sample countries (in order of emisson volume (ISO three-letter country codes in parentheses)):

United States (USA), China (CHN), India (IND), Russia (RUS), Japan (JPN), Brazil (BRA), Germany (GER), Canada (CAN), United Kingdom (GBR), Mexico (MEX), Indonesia (IDN), Australia (AUS), Italy (ITA), Iran (IRN), France (FRA), South Korea (KOR), South Africa (ZAF), Spain (ESP), Saudi Arabia (SAU), Poland (POL), Thailand (THA), Argentina (ARG), Pakistan (PAK), Turkey (TUR), Venezuela (VEN), Egypt (EGY), Nigeria (NGA), Netherlands (NLD), Malaysia (MYS), Kazakhstan (KAZ), Vietnam (VNM), Uzbekistan (UZB), Algeria (DZA), Bangladesh (BGD), United Arab Emirates (ARE), Czech Republic (CZE), Colombia (COL), Philippines (PHL), Belgium (BEL), Sudan (SDN), Greece (GRC), Ethiopia (ETH), Chile (CHL), Republic of Congo (COG), New Zealand (NZL), Syria (SYR), Austria (AUT), Hungary (HUN), Portugal (PRT), Angola (AGO), Peru (PER), Tanzania (TZA), Morocco (MAR), Finland (FIN), Singapore (SGP), Sweden (SWE), Bolivia (BOL), Israel (ISR), Turkmenistan (TKM), Libya (LBY), Norway (NOR), Denmark (DNK), Ireland (IRL)

Proof of $\widehat{\beta} = \mathbf{1}^T\widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{y}} + \mathbf{1}^T\widehat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{u}}$

In the following we will proof the equality of optimal values of the objective functions of programs (5.1) and (5.5).

To start, consider the slack-based approach of program (5.5) in its general formulation with m inputs, s good outputs and r bad outputs

$$\begin{array}{lrcl}\max\limits_{\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}},\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}},\boldsymbol{\lambda}} & \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} + \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \mathbf{1} + \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} & \leqq & \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\ & \mathbf{1} - \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} & = & \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\ & \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}, \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{5.5}$$

In the optimum the restrictions on both the good and bad outputs hold with equality. To see this, consider the case that for a DMU i under evaluation the obtained values are given by $\tilde{\boldsymbol{\beta}}^*_{\boldsymbol{y}}$, $\tilde{\boldsymbol{\beta}}^*_{\boldsymbol{u}}$ and $\boldsymbol{\lambda}^*$. Moreover, assume that the constraint for the jth output of i is non-binding, hence it reads $1 + \tilde{\beta}^*_{\boldsymbol{y},j} < \frac{\boldsymbol{Y}_{j\cdot}}{y_{ji}}\boldsymbol{\lambda}^*$ with $\boldsymbol{Y}_{j\cdot}$ denoting the jth row of $\boldsymbol{Y}$. In this case there exists a $\tilde{\beta}^{**}_{\boldsymbol{y},j} > \tilde{\beta}^*_{\boldsymbol{y},j}$ for which the constraint holds with equality and thus the vector $\tilde{\boldsymbol{\beta}}^*_{\boldsymbol{y}}$ cannot be an optimal solution to (5.5).

Therefore, we can rearrange the constraints as

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} = \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \tag{C.1}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} = \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} - \mathbf{1}. \tag{C.2}$$

Substituting these conditions into the objective function of (5.5) leads to the transformed linear program

$$\begin{array}{lrcl} \max\limits_{\boldsymbol{\lambda}} & \mathbf{1}^T \left(\frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \right) + \mathbf{1}^T \left(\frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} - \mathbf{1} \right) & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{C.3}$$

Analogously, consider program (5.1) which finds the maximal value of the measure β by optimizing the weights assigned to the increase of good and the reduction of bad outputs.

$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \boldsymbol{y}_i + \beta\boldsymbol{\alpha} \odot \boldsymbol{y}_i & \leqq & \boldsymbol{Y\lambda} \\ & \boldsymbol{u}_i - \beta\boldsymbol{\delta} \odot \boldsymbol{u}_i & = & \boldsymbol{U\lambda} \\ & \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} & = & 1 \\ & \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{5.1}$$

This model can be rearranged analogous to (5.5) as

$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \mathbf{1} + \beta\boldsymbol{\alpha} & \leqq & \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\ & \mathbf{1} - \beta\boldsymbol{\delta} & = & \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\ & \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} & = & 1 \\ & \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{C.4}$$

In the optimum the restrictions for the good outputs hold with equality. To see this, consider the general case and suppose an initial solution $(\beta^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \boldsymbol{\lambda}^*)$ with all outputs exhibiting slacks.[78] Hence, initial the solution to (C.4) reads as

$$\begin{array}{rcl} \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda}^* \\ \mathbf{1} + \beta^*\boldsymbol{\alpha}^* & < & \frac{\boldsymbol{Y\lambda}^*}{\boldsymbol{y}_i} \\ \mathbf{1} - \beta^*\boldsymbol{\delta}^* & = & \frac{\boldsymbol{U\lambda}^*}{\boldsymbol{u}_i} \\ \mathbf{1}^T\boldsymbol{\alpha}^* + \mathbf{1}^T\boldsymbol{\delta}^* & = & 1 \\ \beta^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \boldsymbol{\lambda}^* & \geqq & \mathbf{0}. \end{array} \tag{C.5}$$

Now assume that $\tilde{\boldsymbol{c}} = [\tilde{c}_1, \ldots, \tilde{c}_s]^T$ serves as a vector that eliminates the slacks of the good outputs.

[78] Note that the following derivation can be analogously demonstrated assuming that the constraints for some good outputs hold with equality.

Therefore, we have

$$\begin{array}{rcl} \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\lambda}^* \\ \boldsymbol{1} + \tilde{\boldsymbol{c}} \odot \beta^* \boldsymbol{\alpha}^* & = & \frac{\boldsymbol{Y}\boldsymbol{\lambda}^*}{\boldsymbol{y}_i} \\ \boldsymbol{1} - \beta^* \boldsymbol{\delta}^* & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\ \boldsymbol{1}^T \boldsymbol{\alpha}^* + \boldsymbol{1}^T \boldsymbol{\delta}^* & = & 1 \\ \beta^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \boldsymbol{\lambda}^* & \geqq & \boldsymbol{0}. \end{array} \tag{C.6}$$

Suppose that the smallest slack occurs for the jth output. Thus, $\tilde{c}_j = \min\{\tilde{\boldsymbol{c}}\}$. In the following we show that it is possible to increase β^* to eliminate the slack in the jth output by assigning new weights to the jth output as well as to the bad outputs while holding $\boldsymbol{\lambda}^*$ and the weights for the remaining good outputs ($\boldsymbol{\alpha}^*_{-j}$) constant. Therefore, consider the following programming problem

$$\begin{array}{lrcl} \max\limits_{c_j, \alpha_j} & & c_j & \\ \text{s.t.} & 1 + c_j \beta^* \alpha_j & = & \frac{\boldsymbol{Y}_{j.}\boldsymbol{\lambda}^*}{y_{ji}} \\ & \boldsymbol{1} - c_j \beta^* \frac{\boldsymbol{\delta}^*}{c_j} & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\ & \boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j} + \frac{\boldsymbol{1}^T \boldsymbol{\delta}^*}{c_j} + \alpha_j & = & 1 \\ & c_j, \alpha_j & \geqq & \boldsymbol{0}. \end{array} \tag{C.7}$$

In this program we eliminate the slack by finding a c_j^* such that the restrictions for the bad outputs still hold while rearranging the weights between the jth good output and the bad outputs and moreover increasing the distance function from β^* to $c_j^* \beta^*$. Since the restrictions for the bad outputs are satisfied for each feasible c_j they can be excluded from the program. Rearranging the last equation and inserting it into the constraint for the jth good output leads to a single equation to find c_j^*. It is given by

$$1 + c_j \beta^* \left(1 - \boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j} - \frac{\boldsymbol{1}^T \boldsymbol{\delta}^*}{c_j}\right) = \frac{\boldsymbol{Y}_{j.}\boldsymbol{\lambda}^*}{y_{ji}}. \tag{C.8}$$

Rearranging this equation leads to the optimal c_j:

$$c_j^* = \frac{\frac{\boldsymbol{Y}_j \boldsymbol{\lambda}^*}{\boldsymbol{y}_{ji}} - 1 + \boldsymbol{1}^T \boldsymbol{\delta}^* \beta^*}{\beta^* \left(1 - \boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j}\right)}. \tag{C.9}$$

Since we assume that the jth output exhibits a slack in the initial solution, we observe that $\frac{\boldsymbol{Y}_{j.}\boldsymbol{\lambda}^*}{y_{ji}} - 1 > \beta^* \alpha_j^*$. Therefore, we find

$$c_j^* > \frac{\beta^* \alpha_j^* + \boldsymbol{1}^T \boldsymbol{\delta}^* \beta^*}{\beta^* \left(1 - \boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j}\right)} = \frac{\alpha_j^* + \boldsymbol{1}^T \boldsymbol{\delta}^*}{\left(1 - \boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j}\right)} = 1 \tag{C.10}$$

where the last equality holds because the normalization constraint of the initial solution can be written as $\boldsymbol{1}^T \boldsymbol{\alpha}^*_{-j} + \alpha_j^* + \boldsymbol{1}^T \boldsymbol{\delta}^* = 1$.

Since c_j^* is larger than one we can conclude that $\beta^{**} = c_j^* \beta^* > \beta^*$. Thus, holding the weights for all but the jth output constant and calculating new weights for the jth good and all bad outputs we can find a larger optimal value for β than in the initial solution.

Note that increasing β^* to β^{**} does not violate the restrictions on the remaining good outputs. This follows because $\tilde{c}_j \leq \tilde{\boldsymbol{c}}_{-j}$ was calculated for a given α_j^* and β^*. Since $c_j^* > 1$ we find $\frac{\boldsymbol{\delta}^*}{c_j^*} = \boldsymbol{\delta}^{**} < \boldsymbol{\delta}^*$. Moreover, from the normalization constraint in (C.7) it follows $\mathbf{1}^T\boldsymbol{\alpha}_{-j}^* + \mathbf{1}^T\boldsymbol{\delta}^{**} + \alpha_j^{**} = 1$. Since $\boldsymbol{\delta}^{**} < \boldsymbol{\delta}^*$ and $\boldsymbol{\alpha}_{-j}^*$ remains unchanged, we find that $\alpha_j^{**} > \alpha_j^*$. $\tilde{c}_j$ was calculated to remove the slack in the jth output given α_j^*. Since β^* remains unchanged and $\alpha_j^{**} > \alpha_j^*$, $c_j^* < \tilde{c}_j$ must hold to fulfill the restriction on the jth good output in program (C.7). Because $\tilde{c}_j = \min\{\tilde{\boldsymbol{c}}\}$ we can conclude that $\tilde{\boldsymbol{c}}_{-j} > c_j^*$ and the restrictions for the remaining good outputs are not violated.

Given these new optimal values we can further increase β^{**} by removing the slack for the kth output with $\tilde{c}_k = \min\{\tilde{\boldsymbol{c}}_{-j}\}$. This can be done analogously to the jth output by solving the programming problem

$$\begin{array}{lrcl} \max\limits_{c_k,\alpha_k} & & & c_k \\ \text{s.t.} & 1 + c_k\beta^{**}\alpha_k & = & \frac{\boldsymbol{Y}_{k.}\boldsymbol{\lambda}^*}{y_{ki}} \\ & 1 + c_k\beta^{**}\frac{\alpha_j^{**}}{c_k} & = & \frac{\boldsymbol{Y}_{j.}\boldsymbol{\lambda}^*}{y_{ji}} \\ & \mathbf{1} - c_k\beta^{**}\frac{\boldsymbol{\delta}^{**}}{c_k} & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\ & \mathbf{1}^T\boldsymbol{\alpha}_{-j,-k}^* + \frac{\alpha_j^{**}}{c_k} + \frac{\mathbf{1}^T\boldsymbol{\delta}^{**}}{c_k} + \alpha_k & = & 1 \\ & c_k, \alpha_k & \geqq & \mathbf{0}. \end{array} \tag{C.11}$$

Similar to the case of the jth output we find

$$c_k^* = \frac{\frac{\boldsymbol{Y}_{k.}\boldsymbol{\lambda}^*}{y_{ki}} - 1 + \alpha_j^{**}\beta^{**} + \mathbf{1}^T\boldsymbol{\delta}^{**}\beta^{**}}{\beta^{**}\left(1 - \mathbf{1}^T\boldsymbol{\alpha}_{-j,-k}^*\right)} \tag{C.12}$$

which is again larger than 1 and hence we can calculate $\beta^{***} = c_k^*\beta^{**} > \beta^{**}$.

Continuing this procedure for all outputs exhibiting slacks we find that β can be successively increased until no slacks are present anymore. Thus, the initial solution to (C.4) leading to slacks in the good outputs can not be optimal and hence in the optimum all good output constraints hold with equality.

These constraints on the good and bad outputs can therefore can be rearranged to

$$\boldsymbol{\alpha} = \frac{\boldsymbol{Y}\boldsymbol{\lambda}}{\boldsymbol{y}_i\beta} - \frac{\mathbf{1}}{\beta} \tag{C.13}$$

$$\boldsymbol{\delta} = \frac{\mathbf{1}}{\beta} - \frac{\boldsymbol{U}\boldsymbol{\lambda}}{\boldsymbol{u}_i\beta}. \tag{C.14}$$

Inserting these equalities into the normalization constraint leads to

$$\mathbf{1}^T\left(\frac{\boldsymbol{Y}\boldsymbol{\lambda}}{\boldsymbol{y}_i\beta} - \frac{\mathbf{1}}{\beta}\right) + \mathbf{1}^T\left(\frac{\mathbf{1}}{\beta} - \frac{\boldsymbol{U}\boldsymbol{\lambda}}{\boldsymbol{u}_i\beta}\right) = 1. \tag{C.15}$$

Multiplying both sides with β we obtain

$$\mathbf{1}^T\left(\frac{\boldsymbol{Y}\boldsymbol{\lambda}}{\boldsymbol{y}_i} - \mathbf{1}\right) + \mathbf{1}^T\left(\mathbf{1} - \frac{\boldsymbol{U}\boldsymbol{\lambda}}{\boldsymbol{u}_i}\right) = \beta. \tag{C.16}$$

Replacing β in the objective function of (5.1) with this expression we find

$$\begin{aligned} \max_{\boldsymbol{\lambda}} \quad & \mathbf{1}^T \left(\frac{\boldsymbol{Y}\boldsymbol{\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \right) + \mathbf{1}^T \left(\frac{\boldsymbol{U}\boldsymbol{\lambda}}{\boldsymbol{u}_i} - \mathbf{1} \right) \\ \text{s.t.} \quad & \boldsymbol{x}_i \geqq \boldsymbol{X}\boldsymbol{\lambda} \\ & \boldsymbol{\lambda} \geqq \mathbf{0}. \end{aligned} \tag{C.17}$$

Comparing (C.3) to (C.17) shows that both programs are equal and hence the optimal β from (5.1) is equal to maximal sum of slacks-based measures of program (5.5).

Table C.1: Country results of the efficiency analysis

Country	$\delta = 0$	$\delta = 1$	Country	$\delta = 0$	$\delta = 1$
AGO	1.0809	0.9189	ITA	0.0262	0.1348
ARE	0.0000	0.6299	JPN	0.2364	0.3758
ARG	0.5160	0.7584	KAZ	0.7376	0.8733
AUS	0.0660	0.5725	KOR	0.5469	0.5604
AUT	0.0000	0.0000	KWT	0.0000	0.0000
BEL	0.0000	0.0000	MAR	1.0200	0.7376
BGD	0.5831	0.7081	MEX	0.2704	0.5735
BLR	0.0077	0.5715	MYS	0.6172	0.7088
BOL	0.2092	0.7761	NGA	0.0000	0.0000
BRA	0.6056	0.7183	NLD	0.0196	0.1529
CAN	0.0375	0.5438	NZL	0.1117	0.7389
CHL	0.2123	0.5365	PAK	0.6081	0.7481
CHN	0.7799	0.8089	PER	0.5036	0.5570
COG	0.0000	0.9752	PHL	0.4488	0.6803
COL	0.3020	0.5650	POL	0.1863	0.6693
CZE	0.6065	0.7238	PRT	0.4504	0.4167
DZA	1.1244	0.7648	RUS	0.0683	0.7464
EGY	0.0506	0.3135	SAU	0.3652	0.7767
ESP	0.1625	0.3916	SDN	0.0000	0.0000
ETH	2.3036	0.9046	SGP	0.0160	0.1052
FIN	0.1706	0.4650	SWE	0.0000	0.0000
FRA	0.0215	0.1589	SYR	0.0611	0.6088
GBR	0.0000	0.0000	THA	1.0519	0.7329
GER	0.1101	0.4044	TKM	5.7316	0.9280
GRC	0.2307	0.5117	TUR	0.0395	0.1813
HUN	0.3287	0.5707	TZA	1.6338	0.8703
IDN	0.8159	0.7338	USA	0.0000	0.0000
IND	0.5762	0.7081	UZB	1.7146	0.9480
IRN	0.6050	0.7576	VEN	0.4851	0.8228
IRQ	0.8417	0.7994	VNM	0.7034	0.7643
ISR	0.2203	0.4942	ZAF	0.4269	0.8621

Appendix D

Table D.1: Median Malmquist index results for European countries (1990-2007)

Country	Productivity change	Efficiency change	Technical change
Austria	1.0148	0.9953	1.0235
Belgium	1.0032	0.9958	1.0022
Denmark	1.0334	1.0080	1.0343
Finland	1.0148	1.0112	1.0020
France	1.0038	0.9903	1.0035
Germany	1.0184	0.9926	1.0106
Greece	1.0216	1.0058	1.0205
Ireland	1.0149	1.0053	1.0046
Italy	1.0170	1.0000	1.0256
Luxembourg	1.0296	1.0000	1.0296
Netherlands	1.0103	0.9959	1.0058
Portugal	1.0028	0.9967	1.0079
Spain	1.0069	0.9888	1.0129
Sweden	1.0107	1.0055	1.0000
United Kingdom	1.0076	1.0000	1.0076

Table D.2: Median Malmquist-Luenberger index results for European countries (1990-2007)

Country	Productivity change	Efficiency change	Technical change
Austria	1.0059	1.0000	1.0059
Belgium	1.0098	0.9947	1.0108
Denmark	1.0323	0.9994	1.0287
Finland	1.0031	1.0031	1.0138
France	1.0082	1.0000	1.0043
Germany	1.0139	0.9975	1.0118
Greece	1.0185	1.0016	1.0077
Ireland	1.0147	1.0033	1.0061
Italy	1.0138	1.0000	1.0226
Luxembourg	1.1013	1.0000	1.1013
Netherlands	1.0107	0.9967	1.0109
Portugal	1.0006	1.0000	1.0021
Spain	1.0085	0.9917	1.0014
Sweden	1.0136	0.9889	1.0088
United Kingdom	1.0109	1.0000	1.0109

Table D.3: Innovating countries (Malmquist index)

Period	Innovating countries
1990-1991	Luxembourg, Portugal
1991-1992	Austria, Italy
1992-1993	Italy, Luxembourg, Turkey
1993-1994	Italy, Luxembourg, United Kingdom
1994-1995	Luxembourg, Switzerland, United Kingdom
1995-1996	United Kingdom, United States of America
1996-1997	Ireland, Italy Luxembourg,
1997-1998	Luxembourg, Switzerland, United Kingdom
1998-1999	Ireland, Luxembourg, Switzerland
1999-2000	Ireland, Luxembourg, Switzerland,
2000-2001	United Kingdom
2001-2002	Ireland, United Kingdom
2002-2003	Ireland, United Kingdom
2003-2004	Ireland, United Kingdom
2004-2005	Ireland, Luxembourg
2005-2006	Ireland, Luxembourg United Kingdom
2006-2007	Ireland, Luxembourg, United Kingdom

Table D.4: Innovating countries (Malmquist-Luenberger index)

Period	Innovating countries
1990-1991	Norway, Portugal, Switzerland
1991-1992	Austria, Norway, Turkey
1992-1993	Italy, Luxembourg, Norway, Switzerland
1993-1994	Denmark, Italy, Luxembourg, Norway, Switzerland, United Kingdom, United States
1994-1995	Luxembourg, Norway, Switzerland, United Kingdom
1995-1996	Ireland, Norway, United Kingdom, United States
1996-1997	Italy, Luxembourg, Norway, United Kingdom, United States
1997-1998	Ireland, Luxembourg, Norway, Sweden, Switzerland, United Kingdom, United States
1998-1999	Luxembourg, Sweden, Switzerland, United Kingdom, United States
1999-2000	Luxembourg, Sweden, Switzerland, United Kingdom
2000-2001	Italy, United Kingdom
2001-2002	Norway, United Kingdom
2002-2003	United Kingdom
2003-2004	Luxembourg, Norway, Sweden, Switzerland, United Kingdom
2004-2005	Luxembourg, Norway, Sweden, Switzerland
2005-2006	Luxembourg, Norway, Sweden, Switzerland, United Kingdom
2006-2007	Luxembourg, Norway, Sweden, Switzerland, United Kingdom

Table D.5: p-values of the Li test (excluding Luxembourg)

	Malmquist index		ML index	
	Before EU-BSA	After EU-BSA	Before EU-BSA	After EU-BSA
Productivity change	0.389	0.966	0.546	0.159
Efficiency change	0.056	0.046	0.081	0.026
Technical change 1	0.002	0.868	0.027	0.331
Technical change 2	0.090	0.609	0.002	0.308

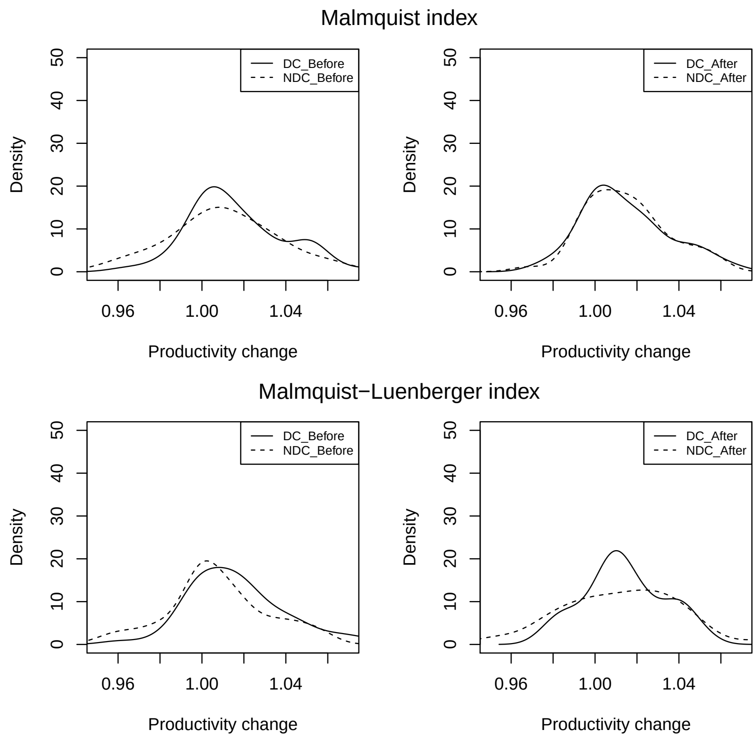

Figure D.1: Density plots of productivity change of EU15 countries (excluding Luxembourg)

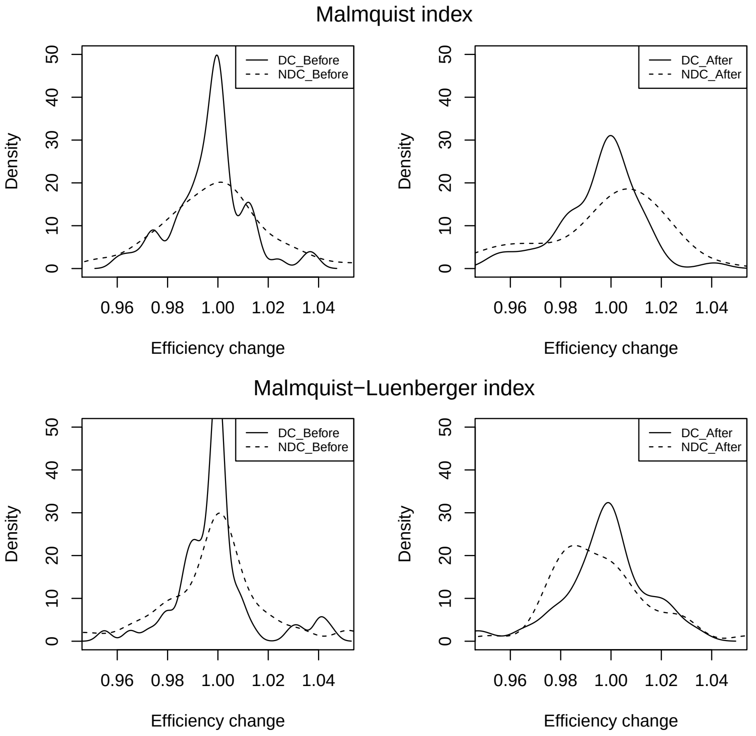

Figure D.2: Density plots of efficiency change of EU15 countries (excluding Luxembourg)

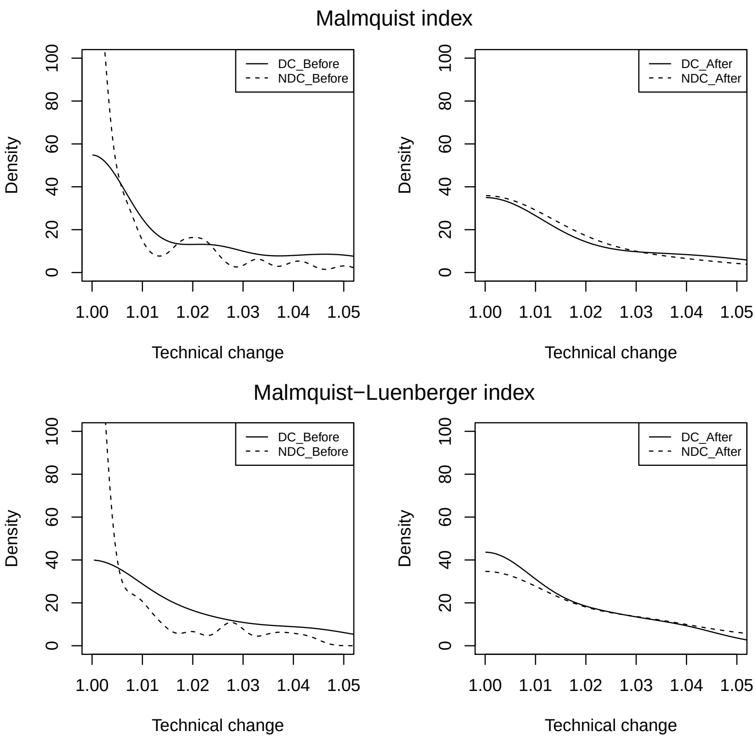

Figure D.3: Density plots of technical change of EU15 countries (excluding Luxembourg)

Affidavit

I hereby declare that the dissertation entitled

"Nonparametric Efficiency Analysis in the Presence of Undesirable Outputs"

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I have not presented this thesis or parts thereof to an university as part of an examination or degree.

Place and date

Signature