Darmstadt Discussion Papers in ECONOMICS

TECHNISCHE UNIVERSITÄT DARMSTADT

Optimal Directions for Directional Distance Functions: An Exploration of Potential Reductions of Greenhouse Gases

Benjamin Hampf and Jens J. Krüger

Nr. 216

Arbeitspapiere der
Volkswirtschaftlichen Fachgebiete der TU Darmstadt

ISSN: 1438-2733

Optimal Directions for Directional Distance Functions: An Exploration of Potential Reductions of Greenhouse Gases

by

Benjamin Hampf*

and

Jens J. Krüger†

Darmstadt University of Technology

March 2013

Abstract

This study explores the reduction potential of greenhouse gases for major pollution emitting countries of the world using nonparametric productivity measurement methods and directional distance functions. In contrast to the existing literature we apply optimization methods to endogenously determine optimal directions for the efficiency analysis. These directions represent the compromise of output enhancement and emissions reduction. The results show that for reasonable directions the adoption of best-practices would lead to sizable emission reductions in a range of about 20 percent compared to current levels.

JEL classification: C14, D24, Q54

Keywords: climate policy, nonparametric frontier functions, directional distance functions

* Department of Law and Economics, Residenzschloss, Marktplatz 15, D-64283 Darmstadt, Germany, Tel.:+49 6151 165319, Fax: +49 6151 163897, E-Mail: hampf@vwl.tu-darmstadt.de

† We am grateful to the participants of the IV Halle Workshop on Efficiency and Productivity Analysis (HAWEPA) for valuable comments. Of course, all remaining errors are ours.

1 Introduction

In the economic discussion of climate change the question by how much anthropogenic greenhouse gas emissions need to be reduced to limit global warming is a key aspect. 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).[1] 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.

Given these targets the question arises how the reductions can be achieved. One strand of the economic literature addresses this problem by discussing possibilities to spur new technologies that allow for a less emission-intensive production of outputs (see e.g. Popp et al. (2010) for an overview). In this paper we address this issue from a production-economic perspective by analyzing whether the production structure of the major greenhouse gas emitting countries is efficient given the currently available reference technology. In our analysis we identify reduction potentials for greenhouse emissions due to inefficiencies of the countries, compare these potentials with the targets presented above and thus show to which extend the exploitation of efficiency enhancement possibilities can contribute to limit global warming.

Therefore we apply nonparametric methods to estimate the technology of the countries and directional distance functions (DDF) to measure their efficiency. In contrast to parametric models like the stochastic frontier analysis (see Kumbhakar and Lovell (2000)) nonparametric approaches do not rely on a specific functional form of the production function and, moreover, allow to include emissions as undesirable outputs. This modeling captures the production process in a more realistic way than the inclusion of emissions as inputs as it is done in parametric approaches (see e.g. Reinhard et al. (2000)). The main advantage of the DDF is the possibility to define a different direction of measurement for each input or output. Thus, it is possible to analyze efficiency by increasing good outputs while simultaneously decreasing bad outputs. 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). However, the wide range 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 paper 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. 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 emission reductions.

This paper is structured as follows: section 2 presents the nonparametric approach to efficiency

[1] 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.

measurement and the directional distance functions. Section 3 describes the approaches to determine optimal directions. This is followed by the description of the data and the discussion of the results in section 4. Finally, section 5 concludes the paper.

2 Nonparametric efficiency analysis

To model the environmental production technology we assume that the decision making units (DMUs) which in our application are the major emitting countries are using m inputs $\boldsymbol{x} \in \mathbb{R}_+^m$ to produce s good outputs $\boldsymbol{y} \in \mathbb{R}_+^s$. We further assume that as a result of this production process r undesirable (or bad) outputs $\boldsymbol{u} \in \mathbb{R}_+^r$ are produced. The technology set comprises all technical feasible input-output combinations and hence reads as

$$T = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in \mathbb{R}_+^{m+s+r} : \boldsymbol{x} \text{ can produce } (\boldsymbol{y}, \boldsymbol{u})\right\}. \tag{1}$$

On this technology we impose the standard axioms by Shephard (1970), e.g. convexity, strong disposability of inputs and good outputs (see Färe and Primont (1995) for further discussions). To include emissions we assume that they are weak disposable outputs, e.g. if $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T$ and $\theta\boldsymbol{u} \leqq \boldsymbol{u}$ with $0 \leqq \theta \leqq 1$ then $(\boldsymbol{x}, \theta\boldsymbol{y}, \theta\boldsymbol{u}) \in T$ (see Färe and Grosskopf (2004) for a detailed discussion of environmental technologies). This assumption states that a reduction of emissions is costly (in terms of a loss of good outputs). Moreover, we assume that good and bad outputs are null-joint. Hence, if $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in T$ and $\boldsymbol{u} = \boldsymbol{0}$ then $\boldsymbol{y} = \boldsymbol{0}$. This indicates that good outputs cannot be produced without producing bad outputs.

The nonparametric estimation of this technology (data envelopment analysis, DEA) given a sample of $(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i)$ for $i = 1, \ldots, n$ DMUs reads as[2]

$$\widehat{T} = \left\{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \in \mathbb{R}_+^{m+s+r} : \boldsymbol{x} \geqq \boldsymbol{X\lambda}, \boldsymbol{y} \leqq \boldsymbol{Y\lambda}, \boldsymbol{u} = \boldsymbol{U\lambda}, \boldsymbol{\lambda} \geqq \boldsymbol{0}\right\}. \tag{2}$$

In this formulation $\boldsymbol{X}$ represents the $m \times n$ matrix of inputs and $\boldsymbol{Y}$ represents the $s \times n$ matrix of good outputs while $\boldsymbol{U}$ is the $r \times n$ matrix of undesirable outputs. The inequality constraints indicate strong disposability of inputs and good outputs while the equality constraint for the bad outputs indicates weak disposability. $\boldsymbol{\lambda}$ 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. Assuming constant returns to scale allows to set the scaling factor θ to 1 (see Färe and Grosskopf (2003)). This technology satisfies null-jointness of good and bad outputs if each of the DMUs uses at least one bad output and each bad output is produced by at least one DMU (see Färe (2010)).

To measure the efficiency of the DMUs given this technology we apply the directional distance function (DDF) proposed by Chambers et al. (1996). This function has been introduced into the nonparametric analysis of environmental efficiency by Chung et al. (1997) and is defined as

$$\beta(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}; \boldsymbol{g}_y, \boldsymbol{g}_u) = \sup\left\{\beta : (\boldsymbol{x}, \boldsymbol{y} + \beta\boldsymbol{g}_y, \boldsymbol{u} - \beta\boldsymbol{g}_u) \in T\right\} \tag{3}$$

[2] See e.g. Färe et al. (1985) for a general overview of nonparametric efficiency measurement.

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 (emissions) can be decreased in the direction $\boldsymbol{g}_u$, holding inputs constant, in order to reach the frontier function.

For a DMU $i \in \{1, ..., n\}$ this efficiency measure can be computed by solving the following linear program

$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\lambda} \\ & \boldsymbol{y}_i + \beta\boldsymbol{g}_y & \leqq & \boldsymbol{Y}\boldsymbol{\lambda} \\ & \boldsymbol{u}_i - \beta\boldsymbol{g}_u & = & \boldsymbol{U}\boldsymbol{\lambda} \\ & \beta, \boldsymbol{\lambda} & \geqq & \boldsymbol{0}. \end{array} \tag{4}$$

A DMU is classified as efficient if $\widehat{\beta}(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i; \boldsymbol{g}_y, \boldsymbol{g}_u) = 0$ and as inefficient if $\widehat{\beta}(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i; \boldsymbol{g}_y, \boldsymbol{g}_u) > 0$.

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 distance function the used vectors are $\boldsymbol{g}_y = \boldsymbol{y}_i$ and $\boldsymbol{g}_u = \boldsymbol{u}_i$. Thus, the directional vectors are given by the observed amounts of good and bad outputs. 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. In the following section we will present methods for determining the directional vectors endogenously. Moreover, in the empirical analysis we will compare the results given the optimized directions and a grid of directions assigning different weights to the production of good and the abatement of bad outputs.

3 Deriving Optimal Directions

In this 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.

3.1 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).[3]

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

$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \boldsymbol{y}_i + \beta\boldsymbol{\alpha} \odot \boldsymbol{y}_i & \leqq & \boldsymbol{Y\lambda} \\ & \boldsymbol{u}_i - \beta\boldsymbol{\delta} \odot \boldsymbol{u}_i & = & \boldsymbol{U\lambda} \\ & \boldsymbol{1}^T\boldsymbol{\alpha} + \boldsymbol{1}^T\boldsymbol{\delta} & = & 1 \\ & \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} & \geqq & \boldsymbol{0}. \end{array} \tag{5}$$

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 efficiency measure of this program can be denoted as $\widehat{\beta}\,(\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{u}_i)$. 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 $\boldsymbol{\mu} = \boldsymbol{\lambda}/\beta$.[4] The linear model then reads as

$$\begin{array}{lrcl} \min\limits_{\gamma,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\mu}} & \gamma & & \\ \text{s.t.} & \gamma\boldsymbol{x}_i & \geqq & \boldsymbol{X\mu} \\ & \gamma\boldsymbol{y}_i + \boldsymbol{\alpha} \odot \boldsymbol{y}_i & \leqq & \boldsymbol{Y\mu} \\ & \gamma\boldsymbol{u}_i - \boldsymbol{\delta} \odot \boldsymbol{u}_i & = & \boldsymbol{U\mu} \\ & \boldsymbol{1}^T\boldsymbol{\alpha} + \boldsymbol{1}^T\boldsymbol{\delta} & = & 1 \\ & \gamma, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\mu} & \geqq & \boldsymbol{0}. \end{array} \tag{6}$$

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 $1/(s+r)$.

[3] 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.

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

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

$$\begin{array}{lrcl} \max\limits_{\boldsymbol{\beta_y},\boldsymbol{\beta_u},\boldsymbol{\lambda}} & \mathbf{1}^T\boldsymbol{\beta_y} + \mathbf{1}^T\boldsymbol{\beta_u} & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \boldsymbol{y}_i + \boldsymbol{\beta_y} \odot \boldsymbol{e_y} & \leqq & \boldsymbol{Y\lambda} \\ & \boldsymbol{u}_i - \boldsymbol{\beta_u} \odot \boldsymbol{e_u} & = & \boldsymbol{U\lambda} \\ & \boldsymbol{\beta_y}, \boldsymbol{\beta_u}, \boldsymbol{\lambda} & \geqq & \mathbf{0} \end{array} \tag{7}$$

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

$$\begin{array}{lrcl} \max\limits_{\boldsymbol{\beta_y},\boldsymbol{\beta_u},\boldsymbol{\lambda}} & \mathbf{1}^T\boldsymbol{\beta_y} + \mathbf{1}^T\boldsymbol{\beta_u} & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \frac{\boldsymbol{y}_i}{\boldsymbol{y}_i} + \frac{\boldsymbol{\beta_y}}{\boldsymbol{y}_i} & \leqq & \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\ & \frac{\boldsymbol{u}_i}{\boldsymbol{u}_i} - \frac{\boldsymbol{\beta_u}}{\boldsymbol{u}_i} & = & \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\ & \boldsymbol{\beta_y}, \boldsymbol{\beta_u}, \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{8}$$

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 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

$$\begin{array}{lrcl} \max\limits_{\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}},\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}},\boldsymbol{\lambda}} & \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} + \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} & & \\ \text{s.t.} & \boldsymbol{x}_i & \geqq & \boldsymbol{X\lambda} \\ & \mathbf{1} + \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} & \leqq & \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\ & \mathbf{1} - \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} & = & \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\ & \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}, \boldsymbol{\lambda} & \geqq & \mathbf{0}. \end{array} \tag{9}$$

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 (7) and (9) are not equal because (7) is not independent of the units in which the inputs and outputs are measured and hence the transformation leading to (9) changes the results. In contrast, the optimal values of the objective functions of programs (5) and (9) can be shown to be equal. A proof of this equality can be found in the appendix. 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) to an analysis of greenhouse gas emissions and calculate potential reductions given the directional vectors obtained from it.

3.2 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.

Since this model is based on dynamic nonparametric productivity measures we start by briefly summarizing the main ideas behind these measures.

In dynamic nonparametric analysis ratios of distance functions can be used to estimate the productivity change of DMUs, e.g. by calculating the Malmquist index (Caves et al. (1982)). This productivity change can be decomposed into efficiency change and technical change (Färe et al. (1992)). While efficiency change measures shifts of the DMUs relatively to a frontier, technical change measures shifts of the frontier. The Malmquist-Luenberger index proposed by Chung et al. (1997) allows to measure productivity changes in the presence of undesirable outputs. With the environmental directional distance function it can be defined as

$$\mathrm{ML}^{t,t+1} = \left[\frac{1+\beta^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1+\beta^{t}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \cdot \frac{1+\beta^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1+\beta^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \right]^{1/2}. \tag{10}$$

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

$$\mathrm{MLTech}^{t,t+1} = \left[\frac{1+\beta^{t+1}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1+\beta^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)} \cdot \frac{1+\beta^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)}{1+\beta^{t}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)} \right]^{1/2} \tag{11}$$

and efficiency change

$$\mathrm{MLEff}^{t,t+1} = \frac{1+\beta^{t}\left(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{u}_t; \boldsymbol{g}_t\right)}{1+\beta^{t+1}\left(\boldsymbol{x}_{t+1}, \boldsymbol{y}_{t+1}, \boldsymbol{u}_{t+1}; \boldsymbol{g}_{t+1}\right)}. \tag{12}$$

where time indices at the distance function β indicate the time period of the technology the DMU is compared to and time indices of the arguments of the distance functions indicate the time period of the DMU. Hence, expressions with different time indices represent mixed-period distance functions, e.g. an input-output combination of a DMU in period t is evaluated to the technology in $t+1$ and vice versa. Chung et al. (1997) estimate this index by assuming that the vector of the efficiency analysis is given by $\left(\boldsymbol{y}_{t,i}, -\boldsymbol{u}_{t,i}\right)$ resp. $\left(\boldsymbol{y}_{t+1,i}, -\boldsymbol{u}_{t+1,i}\right)$ 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:

$$\widehat{\text{MLTech}}_i^{t,t+1} > 1 \tag{13}$$

$$\widehat{\beta}^t \left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i} \right) < 0 \tag{14}$$

$$\widehat{\beta}^{t+1} \left(\boldsymbol{x}_{t+1,i}, \boldsymbol{y}_{t+1,i}, \boldsymbol{u}_{t+1,i} \right) = 0. \tag{15}$$

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 1 that depicts two DMUs (A and B) and the technological frontier for two periods t and $t+1$.

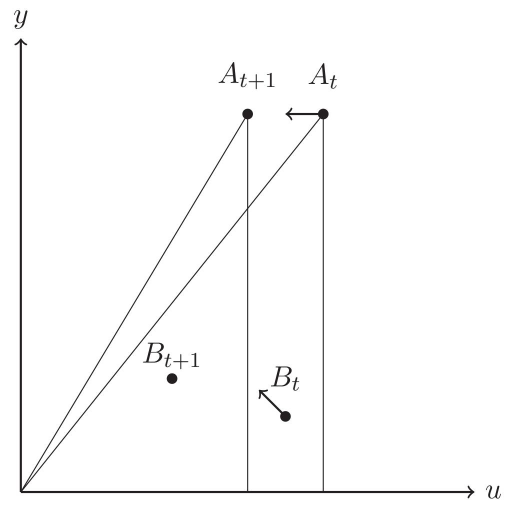

Figure 1: Example of a frontier shift

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 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:[5]

$$\boldsymbol{T}^t = \begin{bmatrix} -x_{t,11} & \dots & -x_{t,1n} \\ \vdots & \vdots & \vdots \\ -x_{t,m1} & \dots & -x_{t,mn} \\ y_{t,11} & \dots & y_{t,1n} \\ \vdots & \vdots & \vdots \\ y_{t,s1} & \dots & y_{t,sn} \\ -u_{t,11} & \dots & -u_{t,1n} \\ \vdots & \vdots & \vdots \\ -u_{t,r1} & \dots & -u_{t,rn} \end{bmatrix} \tag{16}$$

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

$$\Delta\boldsymbol{T}^{t,t+1} = \boldsymbol{T}^{t+1} - \boldsymbol{T}^t. \tag{17}$$

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.

[5] 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.

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 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:

$$\begin{array}{llcl} \max\limits_{\lambda,\boldsymbol{y},\boldsymbol{u}} & \lambda & & \\ \text{s.t.} & \boldsymbol{x}_{t,i} & \geqq & \boldsymbol{x}_{t+1}\lambda \\ & \boldsymbol{y} & \leqq & \boldsymbol{y}_{t+1}\lambda \\ & \boldsymbol{u} & = & \boldsymbol{u}_{t+1}\lambda \\ & \lambda, \boldsymbol{y}, \boldsymbol{u} & \geqq & \boldsymbol{0}. \end{array} \tag{18}$$

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

$$\widehat{\lambda} = \min_j \left\{ \frac{x_{t,ji}}{x_{t+1,ji}} \right\} \quad j = 1, \ldots, m. \tag{19}$$

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

$$\begin{array}{rrcl} \min\limits_{\beta,\boldsymbol{\alpha}_e,\boldsymbol{\delta}_e,\boldsymbol{\lambda}} & & \beta & \\ \text{s.t.} & \boldsymbol{y}_{t,ie} + \beta\boldsymbol{\alpha}_e \odot \boldsymbol{y}_{t,ie} & = & \widehat{\boldsymbol{y}}_e \\ & \boldsymbol{u}_{t,ie} - \beta\boldsymbol{\delta}_e \odot \boldsymbol{u}_{t,ie} & = & \widehat{\boldsymbol{u}}_e \\ & \mathbf{1}^T\boldsymbol{\alpha}_e + \mathbf{1}^T\boldsymbol{\delta}_e & = & 1 \\ & \beta, \boldsymbol{\alpha}_e, \boldsymbol{\delta}_e & \geqq & \mathbf{0} \end{array} \tag{20}$$

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.

4 Analysis of major GHG emitting countries

4.1 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 · pop, labor input as rgdpl · pop/rgdpwok and capital input from real investment data (ki/100) · rgdpl · pop by the perpetual inventory method.[6] Analogous to the

[6] 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

GHG emissions we average the annual values over the period 2000-2005. Descriptive statistics of the data can be found in table 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 the appendix in the order of emission volume.

Table 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 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.

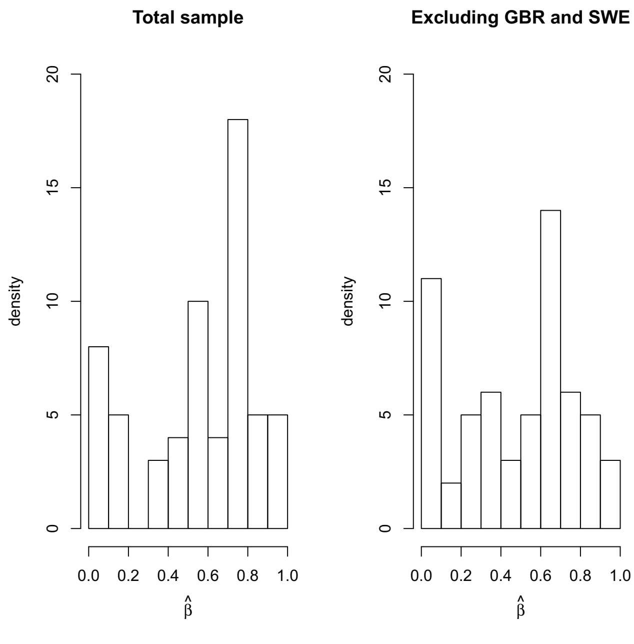

Figure 2: Effects of influential observations

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 $t = 1, 2, ...$

4.2 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 $\delta = 0, 0.1, \dots, 1$ for the reduction of the bad output (GHG).[7] 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 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 5 in the appendix 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.

[7] 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 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 2: Potentials to reduce GHG and increase GDP

Weight	Overall				Income groups				HDI groups			
	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$	$\widehat{\Delta Y}$	$\widehat{\Delta Y}/Y$	$\widehat{\Delta U}$	$\widehat{\Delta U}/U$
$\delta = 0$	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 3. For Peru, a less developed country, the results are presented in table 4.[8] 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.[9] 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 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.

4.3 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 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.

[8] Detailed results for each of the analyzed countries can be obtained from the authors upon request.
[9] Note that since the size of the peers differ, the $\widehat{\lambda}$-values can not be compared in terms of more or less important peers.

Table 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 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.[10] 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 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 2. For visualizing the differences in the results of the efficiency measurement given the different directions, figure 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 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 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 2 show that the potentials to increase

[10] 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 4.[11] 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 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 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.

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

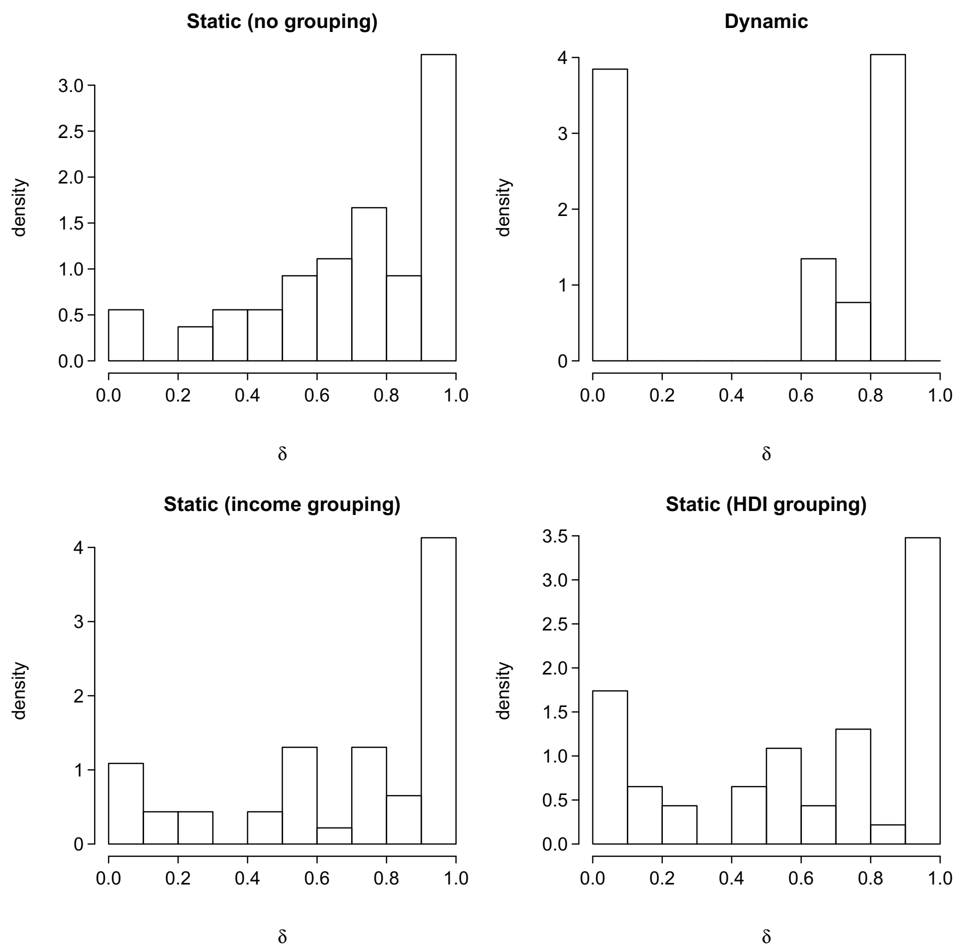

Figure 3: Histograms of optimal weights

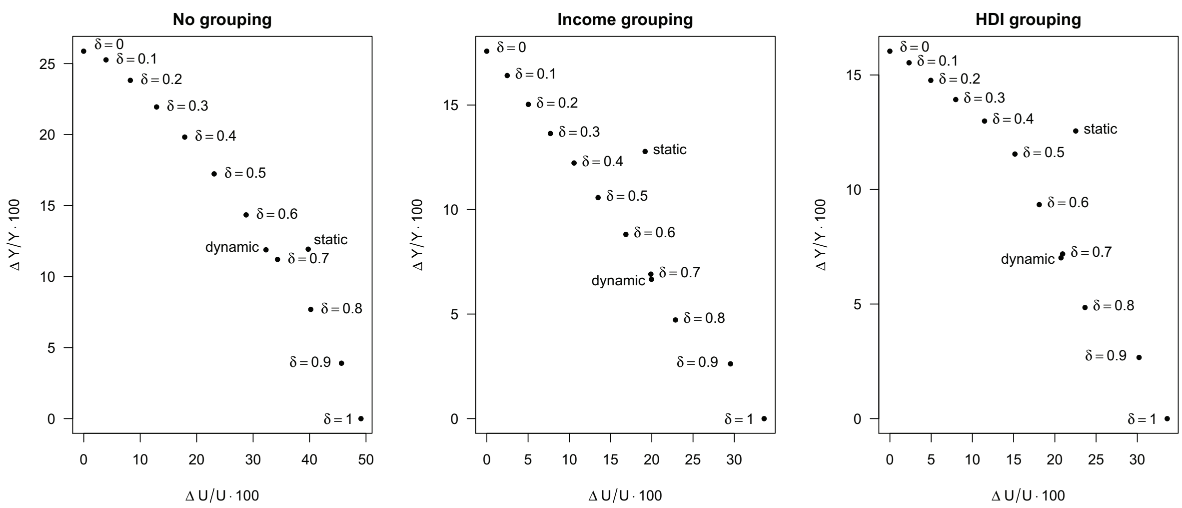

Figure 4: Comparison of results using optimized and non-optimized weights

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 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 (see Stern (2007, table 8.2)). 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 these 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. One particular problem is that the amount of emissions of a country is highly dependent on its industry structure (i.e. the relative weights of manufacturing or service industries) which is historically determined as a part of the specialization in international trade.

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 Conclusion

In this paper we have conducted a nonparametric efficiency analysis of 62 major greenhouse gas emitting countries. Accounting for the difficulties in choosing the directions of the efficiency measurement when directional distance functions are applied we have proposed two methods for determining the directions endogenously. Adapting 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.

Applying these methods to a macroeconomic analysis of environmental efficiency we have shown that different directions have indeed a significant influence on the efficiency estimates. 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. We have shown that the reduction targets mentioned in the literature are very similar to the magnitudes of reduction potentials due to inefficiencies. Therefore, eliminating these inefficiencies could contribute significantly to a reduction of greenhouse gas emissions and therefore to limit global warming.

Perspectives for future research arising from this assessment are twofold. First, we have recognized the problem that industry structure is important and needs to be accounted for in the analysis. Here we face serious data problems for less developed and also for newly industrializing countries. So a first step in this direction would be concerned with an analysis for countries of the European Union where the EU KLEMS database (see Jorgenson and Timmer (2011)) provides a valuable data source. As a second and more technical line of research we see the need to calculate measures for the precision of the estimates. This is most conveniently done in the form of confidence intervals where appropriate bootstrap based methods have been proposed recently (see Simar et al. (2012)).

References

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.

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.

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.

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

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.

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`.

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

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.

— (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, 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, 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, and C. Pasurka (2001). “Accounting for Air Pollution Measures of State Manufacturing Productivity Growth”. In: *Journal of Regional Science* 41, pp. 381–409.

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

Heston, Alan, Robert Summers, and Bettina Aten (2011). *Penn World Table Version 7.0*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, June 2011.

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.

Jorgenson, D. W. and M. P. Timmer (2011). "Structural Change in Advanced Nations: A New Set of Stylised Facts". In: *Scandinavian Journal of Economics* 113, pp. 1–29.

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

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

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.

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.

Park, W. G. (1995). "International R&D Spillovers and OECD Economic Growth". In: *Economic Inquiry* 33, pp. 571–591.

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.

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.

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.

Shephard, R. W. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.

Simar, L., A. Vanhems, and P. W. Wilson (2012). "Statistical Inference for DEA Estimators of Directional Distances". In: *European Journal of Operational Research* 220, pp. 853–864.

Simar, L. and P. W. Wilson (2002). "Non-Parametric Tests of Returns to Scale". In: *European Journal of Operational Research* 139, pp. 115–132.

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.

Stern, N. (2007). *The Economics of Climate Change - The Stern Review*. Cambridge (Mass.): Cambridge University Press.

Taskin, F. and O. Zaim (2000). “Searching for a Kuznets Curve in Environmental Efficiency Using Kernel Estimation”. In: *Economics Letters* 68, pp. 217–223.

Wilson, P. W. (1995). “Detecting Influential Observations in Data Envelopment Analysis”. In: *Journal of Productivity Analysis* 6, pp. 27–45.

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.

Zhou, P., B. W. Ang, and K. L. Poh (2008). “Measuring Environmental Performance under Different Environmental DEA Technologies”. In: *Energy Economics* 30, pp. 1–14.

Appendix

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) and (9).

To start, consider the slack-based approach of program (9) in its general formulation with m inputs, s good outputs and r bad outputs

$$
\begin{aligned}
\max_{\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}},\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}},\boldsymbol{\lambda}} \quad & \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} + \mathbf{1}^T\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} \\
\text{s.t.} \quad & \boldsymbol{x}_i \geqq \boldsymbol{X\lambda} \\
& \mathbf{1} + \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} \leqq \tfrac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\
& \mathbf{1} - \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} = \tfrac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\
& \tilde{\boldsymbol{\beta}}_{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{u}}, \boldsymbol{\lambda} \geqq \mathbf{0}.
\end{aligned} \tag{9}
$$

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 $1 + \tilde{\beta}^*_{\boldsymbol{y},j} < \frac{\boldsymbol{Y}_{j.}}{y_{ji}}\boldsymbol{\lambda}^*$ with $\boldsymbol{Y}_{j.}$ 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 (9).

Therefore, we can rearrange the constraints as

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{y}} = \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \tag{A.1}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{u}} = \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} - \mathbf{1}. \tag{A.2}$$

Substituting these conditions into the objective function of (9) leads to the transformed linear program

$$\begin{aligned} \max_{\boldsymbol{\lambda}} \quad & \mathbf{1}^T \left(\frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \right) + \mathbf{1}^T \left(\frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} - \mathbf{1} \right) \\ \text{s.t.} \quad & \boldsymbol{x}_i \geqq \boldsymbol{X\lambda} \\ & \boldsymbol{\lambda} \geqq \mathbf{0}. \end{aligned} \tag{A.3}$$

Analogously, consider program (5) 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.

$$\begin{aligned} \max_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} \quad & \beta \\ \text{s.t.} \quad \boldsymbol{x}_i &\geqq \boldsymbol{X\lambda} \\ \boldsymbol{y}_i + \beta\boldsymbol{\alpha} \odot \boldsymbol{y}_i &\leqq \boldsymbol{Y\lambda} \\ \boldsymbol{u}_i - \beta\boldsymbol{\delta} \odot \boldsymbol{u}_i &= \boldsymbol{U\lambda} \\ \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} &= 1 \\ \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} &\geqq \mathbf{0}. \end{aligned} \tag{5}$$

This model can be rearranged analogous to (9) as

$$\begin{aligned} \max_{\beta,\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\lambda}} \quad & \beta \\ \text{s.t.} \quad \boldsymbol{x}_i &\geqq \boldsymbol{X\lambda} \\ \mathbf{1} + \beta\boldsymbol{\alpha} &\leqq \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} \\ \mathbf{1} - \beta\boldsymbol{\delta} &= \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} \\ \mathbf{1}^T\boldsymbol{\alpha} + \mathbf{1}^T\boldsymbol{\delta} &= 1 \\ \beta, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\lambda} &\geqq \mathbf{0}. \end{aligned} \tag{A.4}$$

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.[12] Hence, initial the solution to (A.4) reads as

$$\begin{aligned} \boldsymbol{x}_i &\geqq \boldsymbol{X\lambda}^* \\ \mathbf{1} + \beta^*\boldsymbol{\alpha}^* &< \frac{\boldsymbol{Y\lambda}^*}{\boldsymbol{y}_i} \\ \mathbf{1} - \beta^*\boldsymbol{\delta}^* &= \frac{\boldsymbol{U\lambda}^*}{\boldsymbol{u}_i} \\ \mathbf{1}^T\boldsymbol{\alpha}^* + \mathbf{1}^T\boldsymbol{\delta}^* &= 1 \\ \beta^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \boldsymbol{\lambda}^* &\geqq \mathbf{0}. \end{aligned} \tag{A.5}$$

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.

[12] Note that the following derivation can be analogously demonstrated assuming that the constraints for some good outputs hold with equality.

Therefore, we have

$$\begin{array}{rcl} \boldsymbol{x}_i & \geqq & \boldsymbol{X}\boldsymbol{\lambda}^* \\ \mathbf{1} + \tilde{\boldsymbol{c}} \odot \beta^* \boldsymbol{\alpha}^* & = & \frac{\boldsymbol{Y}\boldsymbol{\lambda}^*}{\boldsymbol{y}_i} \\ \mathbf{1} - \beta^* \boldsymbol{\delta}^* & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\ \mathbf{1}^T \boldsymbol{\alpha}^* + \mathbf{1}^T \boldsymbol{\delta}^* & = & 1 \\ \beta^*, \boldsymbol{\alpha}^*, \boldsymbol{\delta}^*, \boldsymbol{\lambda}^* & \geqq & \mathbf{0}. \end{array} \tag{A.6}$$

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

$$\begin{array}{lrcl} \max\limits_{c_j, \alpha_j} & & c_j & \\ \text{s.t.} & 1 + c_j \beta^* \alpha_j & = & \frac{\boldsymbol{Y}_{j\cdot}\boldsymbol{\lambda}^*}{y_{ji}} \\ & \mathbf{1} - c_j \beta^* \frac{\boldsymbol{\delta}^*}{c_j} & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\ & \mathbf{1}^T \boldsymbol{\alpha}^*_{-j} + \frac{\mathbf{1}^T \boldsymbol{\delta}^*}{c_j} + \alpha_j & = & 1 \\ & c_j, \alpha_j & \geqq & \mathbf{0}. \end{array} \tag{A.7}$$

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

$$1 + c_j \beta^* \left(1 - \mathbf{1}^T \boldsymbol{\alpha}^*_{-j} - \frac{\mathbf{1}^T \boldsymbol{\delta}^*}{c_j}\right) = \frac{\boldsymbol{Y}_{j\cdot}\boldsymbol{\lambda}^*}{y_{ji}}. \tag{A.8}$$

Rearranging this equation leads to the optimal c_j:

$$c_j^* = \frac{\frac{\boldsymbol{Y}_j \boldsymbol{\lambda}^*}{\boldsymbol{y}_{ji}} - 1 + \mathbf{1}^T \boldsymbol{\delta}^* \beta^*}{\beta^* \left(1 - \mathbf{1}^T \boldsymbol{\alpha}^*_{-j}\right)}. \tag{A.9}$$

Since we assume that the jth output exhibits a slack in the initial solution, we observe that $\frac{\boldsymbol{Y}_{j\cdot}\boldsymbol{\lambda}^*}{y_{ji}} - 1 > \beta^* \alpha_j^*$. Therefore, we find

$$c_j^* > \frac{\beta^* \alpha_j^* + \mathbf{1}^T \boldsymbol{\delta}^* \beta^*}{\beta^* \left(1 - \mathbf{1}^T \boldsymbol{\alpha}^*_{-j}\right)} = \frac{\alpha_j^* + \mathbf{1}^T \boldsymbol{\delta}^*}{\left(1 - \mathbf{1}^T \boldsymbol{\alpha}^*_{-j}\right)} = 1 \tag{A.10}$$

where the last equality holds because the normalization constraint of the initial solution can be written as $\mathbf{1}^T \boldsymbol{\alpha}^*_{-j} + \alpha_j^* + \mathbf{1}^T \boldsymbol{\delta}^* = 1$.

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 (A.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 (A.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

$$\begin{array}{lrcl}
\max\limits_{c_k,\alpha_k} & & & c_k \\
\text{s.t.} & 1 + c_k\beta^{**}\alpha_k & = & \frac{\boldsymbol{Y}_{k.}\boldsymbol{\lambda}^*}{y_{ki}} \\
& 1 + c_k\beta^{**}\frac{\alpha_j^{**}}{c_k} & = & \frac{\boldsymbol{Y}_{j.}\boldsymbol{\lambda}^*}{y_{ji}} \\
& \mathbf{1} - c_k\beta^{**}\frac{\boldsymbol{\delta}^{**}}{c_k} & = & \frac{\boldsymbol{U}\boldsymbol{\lambda}^*}{\boldsymbol{u}_i} \\
& \mathbf{1}^T\boldsymbol{\alpha}_{-j,-k}^* + \frac{\alpha_j^{**}}{c_k} + \frac{\mathbf{1}^T\boldsymbol{\delta}^{**}}{c_k} + \alpha_k & = & 1 \\
& c_k, \alpha_k & \geqq & \mathbf{0}.
\end{array} \tag{A.11}$$

Similar to the case of the jth output we find

$$c_k^* = \frac{\frac{\boldsymbol{Y}_{k.}\boldsymbol{\lambda}^*}{y_{ki}} - 1 + \alpha_j^{**}\beta^{**} + \mathbf{1}^T\boldsymbol{\delta}^{**}\beta^{**}}{\beta^{**}\left(1 - \mathbf{1}^T\boldsymbol{\alpha}_{-j,-k}^*\right)} \tag{A.12}$$

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 (A.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

$$\boldsymbol{\alpha} = \frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i\beta} - \frac{\mathbf{1}}{\beta} \tag{A.13}$$

$$\boldsymbol{\delta} = \frac{\mathbf{1}}{\beta} - \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i\beta}. \tag{A.14}$$

Inserting these equalities into the normalization constraint leads to

$$\mathbf{1}^T\left(\frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i\beta} - \frac{\mathbf{1}}{\beta}\right) + \mathbf{1}^T\left(\frac{\mathbf{1}}{\beta} - \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i\beta}\right) = 1. \tag{A.15}$$

Multiplying both sides with β we obtain

$$\mathbf{1}^T\left(\frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1}\right) + \mathbf{1}^T\left(\mathbf{1} - \frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i}\right) = \beta. \tag{A.16}$$

Replacing β in the objective function of (5) with this expression we find

$$\begin{aligned} \max_{\boldsymbol{\lambda}} \quad & \mathbf{1}^T \left(\frac{\boldsymbol{Y\lambda}}{\boldsymbol{y}_i} - \mathbf{1} \right) + \mathbf{1}^T \left(\frac{\boldsymbol{U\lambda}}{\boldsymbol{u}_i} - \mathbf{1} \right) \\ \text{s.t.} \quad & \boldsymbol{x}_i \geqq \boldsymbol{X\lambda} \\ & \boldsymbol{\lambda} \geqq \mathbf{0}. \end{aligned} \tag{A.17}$$

Comparing (A.3) to (A.17) shows that both programs are equal and hence the optimal β from (5) is equal to maximal sum of slacks-based measures of program (9).

Table 5: 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