University of Neuchatel

Institute of Economic Research

IRENE, Working paper 14-01

How should we measure environmental policy stringency?

A new approach

*Caspar Sauter**

* *University of Neuchatel*

UNIVERSITÉ DE NEUCHÂTEL

Institut de recherches économiques

How should we measure environmental policy stringency ? A new approach

(IRENE Working Paper) [☆,☆☆]

Caspar Sauter[a,1,]

[a]*Institute of Economic Research, University of Neuchâtel, Abram-Louis-Breguet 2, CH-2000 Neuchâtel*

Abstract

One of the biggest obstacles in cross-country empirical research in the area of environmental economics is the absence of a sound indicator quantifying environmental policy stringency. A variety of indicators have been proposed and are currently used. Almost none of them rely on an explicitly stated methodology, violating thereby one of the most fundamental rules of index construction. To overcome this problem, this paper develops a new general methodology for the measurement of environmental policy stringency and proposes a first implementation using the example of CO_2 policy stringency. To do so it combines originally extensive databases on CO_2 emissions.

Keywords: Greenhouse gas emissions, environmental regulation, environmental policy stringency, policy stringency index, CO_2 emissions

---

1. Introduction

The absence of a broadly accepted indicator of environmental policy stringency is currently limiting applied research in several fields of economics. Words of caution mentioning limits of existing indicators can be found in a majority of papers. But surprisingly few papers have been exclusively devoted to the construction of such indexes. In one of those, Knill et al (2012) conclude that the choice of indicators of environmental policy stringency is rarely theoretically motivated but rather driven by data availability. This paper proposes and implements a new methodology attempting to overcome those problems.

[☆] JEL classification: Q50, Q53, Q58, C18

[☆☆] First Version: October 2012; This Version: August 2014

Email address: `caspar.sauter@unine.ch` (Caspar Sauter)

[1] I'm grateful for the excellent help of Alexandra Kis as well as for many helpful comments by Jean-Marie Grether, Daniel Schmitter, Nicole Mathys, Geraud Kraehenbuehl, Lionel Perini, Thiery Graf and several participants of the 2013 SSES congress. I'm also grateful for the data received by Knill et al (2012) and the World Economic Forum.

2. An overview on existing indicators

A variety of indicators have been proposed so far. Focusing on indicators available for several countries and at least one common year, one can distinguish four groups: survey indicators, monetary indicators, policy specific indicators and performance indicators. The following overview is not exhaustive. Selected examples are presented for each group of indicators.

2.1. Survey indicators

Dasgupta et al (2001) develop an index of environmental policy stringency based on reports prepared for the United Nations Conference on Environment and Development (UNCED). Those reports contain self-reported information from country officials on a variety of questions. They are complemented by responses from several NGO's, attempting to make the data less exposed to biases from self-reporting. Using the methodology of Dasgupta et al., Eliste and Fredriksson (2002) extended the database for another 31 countries but only for the agricultural sector. Recent papers (see for instance Kalamova and Johnstone (2011) or Timmins and Wagner (2009)) frequently use the indicator of environmental regulatory stringency developed by the World Economic Forum (WEF). The WEF obtains the data by asking "business leaders" the survey question: "How would you assess the stringency of your countries' environmental policy? (scale: 1=very lax − 7=among the world's most stringent)." A question asked each year in the World Executive Opinion Survey (Browne et al, 2012). This indicator is available for more than 100 countries on a yearly base since 2004. Those survey based indicators depend by construction solely on the perceptions of the survey respondents. They are not based on hard data on environmental policy.

2.2. Monetary indicators

Magnani (2000) and Pearce and Palmer (2001) use public expenditures for environmental protection as a measure of environmental policy stringency. Their data covers OECD countries during the nineties and has been collected by the OECD Environmental Program. Those indicators capture expenditure based policy instruments only, excluding the wide variety of other instruments. On top of that, due to the existence of potential efficiency differences among countries it is also unclear whether higher per capita expenditures imply stricter environmental policies. It is possible that a country with high per capita expenditures uses those expenditures in a highly inefficient way such that another country with lower per capita expenditures does a better job. Another monetary indicator which is frequently used is pollution abatement costs. Some authors, as for instance Friedman et al (1992), Crandall (1993), List and Co (2000) use total statewide pollution abatement costs as an indicator. Others like Keller and Levinson (2002) adjust them for each states' industrial composition. A third group of researchers use sectoral rather then state wide data on abatement costs (see Brunnermeier and Cohen, 2003). A fourth group uses an indicator based on capital expenditures and operating costs in environmental protection activities (see Jug and Mirza (2005)). Important disadvantages of those type of indicators have been pointed out by Copeland (2008). For firms it is difficult to correctly disentangle abatement cost from other cost. Moreover firms might have an incentive to strategically under or over report their abatement cost. Finally, reported abatement cost may be endogenous and thus induce biases in the analysis. Illustrating this argument, Copeland advances an example in which firms have heterogeneous costs of responding to environmental regulations. Assuming that the competitiveness hypothesis is correct, stringent pollution policies might drive firms with the highest abatement costs out of business. This opens the possibility that even in regions with relatively strict environmental policies, observed abatement costs may be low.

2.3. Policy specific indicators

Nakada (2006) assesses environmental policy stringency using the timing of the ratification of the Kyoto protocol. He generates a dummy variable, taking the value of zero if a country hasn't ratified the Kyoto protocol by the year 2003 and the value of 1 if a country has ratified the Kyoto protocol by the year 2003. Smarzynska and Shang-Jin (2003) use and indicator based on the ratification of four international treaties in environmental politics. They adjust their initial measure by multiplying it with the ratio of environmental NGO's per million of people in a given country, claiming that this adjustment reflects the degree of enforcement of those treaties. Those treaty based indicators remain highly specific, excluding all other instruments of environmental policy. Knill et al (2012) develop an indicator of clean air policy, capturing national statutory laws on the book. Their index is available for 24 OECD countries covering the time span from 1976 to 2003. Their indicator codifies the different clean air laws of countries either as "policy expansion" or as "policy dismantling". The data has been taken from their own database (compiled by the CONSENSUS project). Their indicator has the advantage that it relies on a well defined methodology, resulting in two indicators, one called "policy density" the other "policy intensity".

2.4. Performance indicators

Several researcher as Hilton and Levinson (1998), Deacon (1999), Damania (2001), Broner et al (2012) or Grether et al (2012) use the lead content in gasoline as an indicator. The data is taken from the Octel Worldwide Gasoline Survey. According to some of those authors, given the absence of a sound index of environmental policy stringency, their index covers at least one of the most important environmental issues of the late 20th century. A second group of performance indicators are based on Emission or Energy consumption data. Some researchers simply take emitted emission as a measure for environmental policy stringency. As an example, Xing and Kolstad (2002) use total SO_2 emissions on a country level and Smarzynska and Shang-Jin (2003) overall CO_2 emission reduction data. Again others base their index on energy intensity data (see for instance Cole and Elliott (2003)). Harris et al (2003) use energy consumption data as their measure. A third group uses the Environmental Performance Index (EPI) published by The Yale Center for Environmental Law and Policy (YCELP) (Emerson et al, 2012). Although the YCELP does not claim that their index is a measure of environmental policy stringency, some researchers use it as such. All those different approaches have one point in common: they are performance indicators. By construction, performance indicators quantify the problem environmental policies try to solve and not the stringency of the policies themselves. After all, variations in emitted emissions or in the lead content of gasoline can be due to a wide variety of factors which might be unrelated to policies.

3. Methodological framework for environmental policy indexes

3.1. What is badly defined is likely to be badly measured

Besides the specific shortcomings of the indicators discussed in section 2, a general problem is common to almost all of them. They are not constructed upon an explicitly stated methodological framework, the contribution of Knill et al (2012) constituting a notable exception. This shortcoming ignores one of the most fundamental rules found in the literature on index construction. As Nardo et al (2008) put it:

> A sound theoretical framework is the starting point in constructing (...) indicators. The framework should clearly define the phenomenon to be measured and its sub-components, selecting individual indicators and weights that reflect their relative importance and the dimensions of the overall composite. This process should ideally be based on what is desirable to measure and not on which indicators are available. Nardo et al (2008), p. 22.

A good index has to be based on a theoretical description of the phenomenon it tries to measure, making it possible to identify relevant sub-components of the main concept. Only once identified, selection criteria can be applied in order to select the underlying indicators. Good selection criteria allow you to clearly distinguish input, process and output measures of the phenomenon. According to Nardo et al (2008) this is a task which is neglected too often.

The problems due to the absence of a theoretical framework are reinforced by that what Brunel and Levinson (2013) identify as the "multidimensionality" problem. Without defining what environmental policy is, it remains ambiguous what those indicators actually intend to quantify. Environmental policy and environmental policy stringency are rather vague concepts which could cover a wide range of policies. They might include policies as diverse as the protection of a flower, the regulation of hunting or the reduction of CO_2. Hence, as Nardo et al. (2008, p. 22) put it: "What is badly defined is likely to be badly measured."

To overcome those obstacles, I develop a coherent methodological framework. Based on an explicit definition the framework allows to measure specific types of environmental policies while making a clear distinction between input, process and output measures. To mitigate the multidimensionality problem I focus on a specific type of environmental policies: pollutant policies; and on a particular case: anthropogenic CO_2 emissions.

3.2. What we should measure: input, process and output indexes

A policy can be defined as a set of government made decisions which have been implemented and which aim to solve a particular problem. The particular problem pollutant policies are dealing with is the reduction of anthropogenic emissions of the pollutant in question. Based on this general definition, one can define a CO_2 policy as a set of government made decisions which have been implemented and which aim to reduce anthropogenic CO_2 emissions.

Defined as such one can look at pollutant policies in three complementary ways by measuring the input, process or output side of the phenomenon. As economists we work every day with variables constructed on this trinity. Take our measures quantifying the phenomenon "production": to analyze the input side we use measures like the quantity of labor or capital. To analyze the process dimension we develop measures which quantify how those inputs are put together. Technology indexes and efficiency measures are examples of such process indicators. And we quantify the output dimension using indicators like GDP. Imagine for a second the mess applied research would face if we would have only one measure of production mixing the above mentioned. So why not apply the same structure to develop measures quantifying the phenomenon pollutant policy?

Given the definition of pollutant policies and the need to carefully separate input, process and output measures I propose to proceed as follows (see next section for implementation issues): To develop an *input* index we have to quantify all different decisions taken by government entities which aim to reduce the pollutant. The more decisions have been taken and the more important they were, the higher the input dimension policy stringency. To develop a *process* index of the phenomenon one has to captures how those inputs are implemented. The stronger the implementation (the less exceptions, the more inspections etc.) the higher the process dimension policy stringency. To develop an *output* index we have to quantify by how much the particular problem has been solved by the policies. The better it has been solved, the higher the output dimension stringency.

3.3. What we will measure here: input and performance indexes

Apart from space constraints, the implementation of the methodological framework is conditional on data availability. Even in the widely studies case of CO_2 emissions, reliable and comparable country-specific data is difficult to obtain. This has led to two restrictions with respect to the ideal case.

On the one hand, I do not report any process index results. This is due to the absence of the relevant information at the level of a specific pollutant. On the other hand, developing a proper output index of CO_2 policies is out of the scope of the present paper because it would mean, apart from measuring performance (e.g. CO_2 emissions per capita), estimating which part of that performance is specifically attributable to government policies (and not, say, to climate or industrial structure). Hence, I limit my objective to measuring a *performance* index, which includes the influence of other factors, and is a first step towards a real output index of pollutant policies.

With these important caveats in mind, the next two sections describe the construction of an CO_2 policy input index and the construction of a CO_2 performance index. Note that I also use this methodology to implement indexes on CH_4 and SO_2 policy stringency. The results can be found in the online appendix to this paper (Sauter, 2014).

4. Implementation of a pollutant policy input index

A CO_2 policy stringency input measure is a metric that captures the decisions taken to reduce CO_2 quantitatively. The proposed indicator captures the huge variety of CO_2

policy inputs. As such the indicator has to be seen as a so called “de jure” indicator (or in the terminology of Kaufmann and Kraay (2008) a “rule based indicator”) which captures statutory laws “on the book”. In that sense, the indicator follows partially the work of Knill et al (2012).

4.1. Approach and data sources

I use two different databases to construct two input indicators a “narrow” and a “broad” one.

Taking the definition of pollutant policies literally one can classify a policy as a CO_2 policy only if the law explicitly refers to the goal of reducing CO_2. The ECOLEX database (FAO et al, 2013) allows the extraction of such CO_2 policies. After selecting all laws in ECOLEX which contain the words CO_2 (or any derivative like carbon dioxide in any language) I had to drop some. There are laws which contain the right keywords but which cannot be classified as CO_2 policies. For example, laws on the minimum quantity of CO_2 in bottled water have been dropped. Across the 55 countries covered I identified a total of 379 narrow CO_2 measures which are (or have been) enacted, those measures are country specific. Out of those 379 measures, 35 apply on a sub-national level. Besides those sub-national measures, there are some which have been enacted by the European Union. Those supra-national measures have been attributed to the EU member countries[2]. Based on the ECOLEX data an input index is developed which is labeled “Narrow CO_2 Input Index”.

Interpreting the definition in a broader sense, one can classify a policy as a CO_2 policy if the law implicitly refers to the goal of reducing CO_2. Taxes on fuels (although not necessarily containing a paragraph specifying the goal of CO_2 reduction) can in this sense be classified as CO_2 policies. The Database for instruments of environmental policy and natural resource management published by the OECD and EEA (2012) allows to select such policies. This is a major advantage. Across the 52 covered countries, I identified a total of 1109 country-specific greenhouse gas measures which are (or have been) enacted. Out of those 1109 measures, 435 measures apply on a sub-national level. I completed the database by finding several hundreds of missing date of enactment entries in the national legislation of the concerned countries. The major advantage of this database – the possibility to include laws which can be classified in a broad sense as pollutant policies – comes at a cost. The database does not allow to make a clear distinction between pollutants. Hence the resulting input index has to be seen as a general greenhouse gas input index which is labeled “Broad GHG Input Index.”

---

[2] If a country has already been a member of the EU at the time the EU adopted the measure then the national date of enactment of the measure corresponds to the one of the EU. If a country hasn’t been a member at the time the EU adopted the measure, then the national date of enactment of the measure corresponds to the date where the country joined the EU (provided that the measure was still in place at that time).

4.2. Codification, weighting and normalization of the input indicator

Given that it is difficult to quantify juridical information, I proceed as others did in constructing “de jure” indicators[3]. In order to capture this information quantitatively I generate dummy variables. Each dummy reflects the answer to the question “Does measure j exist in country i in year t?”. A dummy variables takes the value of 1 if a measure exists in a certain country and in a given year and the value of 0 otherwise. This procedure has been applied twice using the two different databases.

I follow the “usual” equal weighting approach to construct the two input indexes. At the same time I also account for the fact that some measures are only applied on a sub-national level. The dummies are summed up by country and by year. This has been done separately for both databases leading to the following two input indicators:

$$Narrow\ Input\ Index_{i,t} = \sum_j w_{j,i,t} EcolexDummy_{j,i,t} \tag{1}$$

$$Broad\ Input\ Index_{i,t} = \sum_j w_{j,i,t} OECD/EEA\ Dummy_{j,i,t} \tag{2}$$

where j indexes the instruments, i the countries and t the time. The weight $w_{j,i,t}$ takes the value of 1 if the measure j is applied on a national scale. If the measure j is only applied in a part of the country i, then $w_{j,i,t} = \frac{n_{i,t}^{part}}{n_{i,t}}$, where $n_{i,t}$ is the population of country i at time t and $n_{i,t}^{part}$ is the population of the area of country i in which the measures is applied at time t. Proceeding in this way gives each instrument which is applied on a national scale exactly the same weight in the final index. As a further step an informed weighting approach could improve the indexes. If theoretical work will be able to rank different categories of instruments one could use those ranks to refine the weighting of the dummies. Using such an informed weighting approach, the index could be calculated as follows: $I_{i,t} = \sum_c \gamma_c \sum_j w_{j,i,t} Dummy_{j,i,t,c}$, where i indexes the countries, t the time, c the different categories, j the instruments within a given category, γ_c indicates the weights for each of the categories and $w_{j,i,t}$ the weight which accounts for federal measures as defined before. The indexes have subsequently been normalized to range between zero and one.

5. Implementation of a pollutant performance index

According to the definition of pollutant policies, CO_2 policies aim to solve the particular problem of reducing anthropogenic CO_2. A CO_2 performance indicator captures this particular problem. It is therefore also the first step in the construction of an output index of pollutant policies: if you want to know by how much the policies solved the problem you first need to quantify the problem.

---

[3] See for instance the work on the Global Integrity Index (Global Integrity, 2011) or on the Doing Business indicator (World Bank, 2012).

5.1. Approach and data sources

To implement the indexes I use the World Input Output Tables database (Stehrer et al, 2014), an extension of the National Accounting Matrix including Environmental Accounts project of Eurostat (2009)[4] . This dataset combines the conventional national accounting framework with socioeconomic as well as environmental satellite accounts. For a total of 40 major countries, and 35 sectors, input-output tables, complemented with sectoral labor and capital input data as well as sectoral emission data are available for the time span between 1995 and 2009[5].

Using this database, I develop, for each sector, a CO_2 performance indicator which is constructed on two dimensions: CO_2 intensity and CO_2 efficiency. Then I aggregate those sectoral indexes to obtain an index covering the whole economy. Conceptually, the proposed performance indicator follows and extends the work of the Yale Center for Environmental Law and Policy (YCELP) which develops an Environmental Performance Indicator (EPI)[6] (Emerson et al, 2012). There are three main differences between the climate change part of the EPI and the proposed CO_2-performance indicator: Firstly, instead of trying to quantify overall climate change performance, the CO_2-performance indicator focuses only on the performance of CO_2. Secondly, alongside CO_2 intensity CO_2 efficiency is integrated as an additional dimension of CO_2 performance. And thirdly, the proposed CO_2-performance indicator is constructed on a sectoral scale.

5.2. The construction of sectoral CO_2 performance indexes

Capturing sectoral CO_2 intensity. In accordance with the work of the YCELP I capture the sectoral CO_2 intensity dimension with two different relative measures: Sectoral CO_2 emissions per unit of sectoral GDP and sectoral CO_2 emissions per sectoral workers. Those two are common metrics used to asses the intensity in the use of carbon dioxide emissions in an economy (Emerson et al, 2012).

Capturing sectoral CO_2 efficiency. I capture the sectoral CO_2 efficiency dimension by estimating CO_2 efficiency scores. CO_2 efficiency is defined as the ratio of minimal feasible to observed use of CO_2, conditional on observed output levels and conventional inputs. CO_2-efficiency scores are therefore estimates describing how far a sectoral production process is away from the contemporary best practice. There are two different approaches in the literature on environmental efficiency. The first one treats emissions as inputs in

---

[4] Although the project which led to the elaboration of this dataset has been completed in 2012, prospects are good that the dataset will be extended in geographical and time coverage. The Statistical Division of the UN has launched the System of Environmental-Economic Accounts (SEEA) (see United Nations, 2012), which - once completed - would correspond to an extension of WIOT. Conditional on the successful implementation of SEEA, the proposed indicators could be extended, in time and in country coverage.

[5] Those 40 countries accounted for over 70% of global anthropogenic CO_2 emissions during the 00's.

[6] This indicator intends to track national environmental results on a quantitative basis. The EPI is divided into several parts, one of them measuring "climate change and energy" performance. To quantify the climate change and energy part of this index four sub-indicators are used by the YCELP: CO_2 emissions per capita, CO_2 emissions per dollar, CO_2 emissions per kWh and the percentage of renewable energy in total energy production. All those sub-indicators are constructed using aggregated data (i.e. not sectoral data). Those sub-indicators are then aggregated using weights determined by YCELP experts.

the production function while the second one considers emissions as bad outputs of the production process. Both approaches can be implemented using either Stochastic Frontier Analysis (SFA) or Data Envelopment Analysis (DEA). Treating emissions as inputs has however several theoretical shortcomings. For a review on both approaches refer to Kumar Mandal and Madheswaran (2010). I follow the second approach using DEA and treat emissions as a bad output.

The following presentation of the methodology used to estimate CO_2 efficiency scores follows closely the paper of Färe et al (2014). First some notation, assume that a decision making unit[7] produces L bad outputs $(b_1, ..., b_L) \in \mathbb{R}_+^L$, M good outputs $(y_1, ..., y_M) \in \mathbb{R}_+^M$, while using N inputs $(x_1, ..., x_N) \in \mathbb{R}_+^N$. The technology set is given by $T = \{(x, y, b) : x \textit{ can produce (y,b)}\}$. Färe et al (2014) imposes structure on the technology set by assuming that the set is closed with bounded output sets. Inputs are assumed to be strongly disposable. Good outputs (y) and bad outputs (b) are assumed null-joint: if $(x, y, b) \in T, b = 0 \Rightarrow y = 0$. Bad and good outputs are assumed being together weakly disposable: if $(x, y, b) \in T, \textit{and } 0 \leq \alpha \leq 1 \Rightarrow (x, \alpha y, \alpha b) \in T$. Finally, Färe et al (2014) assumes that good outputs are strongly disposable: if $(x, y, b) \in T, \textit{and } y' \leq y \Rightarrow (x, y', b) \in T$.

Assuming that there are I observations for a given year, (x^i, y^i, b^i) for $i = 1, .., I$, Färe et al (2014) models T in a DEA setting as follows: The pollution generating technology is given by:

$$
\begin{aligned}
T = \{(x, y, b) : \quad & \textstyle\sum_{i=1}^I \quad z_i y_{im} \geq y_m, \ m = 1, ..., M \\
& \textstyle\sum_{i=1}^I \quad z_i b_{il} = b_l, \ l = 1, ..., L \\
& \textstyle\sum_{i=1}^I \quad z_i x_{in} \leq x_n, \ n = 1, ..., N \\
& z_i \geq 0 \quad , \ i = 1, ...I\}
\end{aligned} \tag{3}
$$

The intensity variables z_i in (3) are constrained to be non-negative, imposing thereby constant returns to scale. In addition the following constraints are imposed:

$$\textstyle\sum_{i=1}^I \quad y_{im} > 0, \ m = 1, ..., M \tag{4}$$
$$\textstyle\sum_{m=1}^M \quad y_{im} > 0, \ i = 1, ..., I \tag{5}$$
$$\textstyle\sum_{i=1}^I \quad x_{in} > 0, \ n = 1, ..., N \tag{6}$$
$$\textstyle\sum_{n=1}^N \quad x_{in} > 0, \ i = 1, ..., I \tag{7}$$
$$\textstyle\sum_{i=1}^I \quad b_{il} > 0, \ l = 1, ..., L \tag{8}$$
$$\textstyle\sum_{l=1}^L \quad b_{il} > 0, \ i = 1, ..., I \tag{9}$$
$$\tag{10}$$

Constraints (4)-(7), introduced by Kemeny et al (1956) generalize the Von Neumann

---

[7] A decision making unit (DMU) may refer to an individual worker, a subsection of a firm, a firm but also - as in this paper - to a sector.

(1945) assumptions (for a discussion see (Färe et al, 2014)). Constraints (4), (5), (7) and (8) constrain good and bad outputs to be null-joint.

To obtain CO_2 efficiency scores (EEs) the following linear programming problem will be solved for each observation. Note that for a given DMU the target values (y_m^*, b_l^*, x_n^*) are set equal to the observed values:

$$EE = min\ \beta \tag{11}$$

Subject to:

$$\begin{aligned} \textstyle\sum_{i=1}^{I} \quad & z_i y_{im} \geq y_m^*,\ \ m = 1, ..., M & (12)\\ \textstyle\sum_{i=1}^{I} \quad & z_i b_{il} = \beta b_l^*,\ \ l = 1, ..., L & (13)\\ \textstyle\sum_{i=1}^{I} \quad & z_i x_{in} \leq x_n^*,\ \ n = 1, ..., N & (14)\\ z_i \geq \quad & 0,\ \ i = 1, ...I & (15) \end{aligned}$$

Note that the intensity variables z_i are jointly constrained by (12) and (14) which allows the computation of the βs. This linear programming model is separately estimated for each year and each sector. I use a sequential frontier approach assuming that all current and past observations are feasible[8]. I thereby follow Mukherjee (2008) and Kumar Mandal and Madheswaran (2010). Constructed as such, I obtain one β for each country, each sector and each year. By construction, β takes values between zero and one. A β equal to one indicates full efficiency while a β equal to zero indicates full inefficiency of the DMU.

I estimate two different models using this framework: one based on a profit function and the other based on a revenue function. In both models there is one bad output: CO_2. The profit function model uses value added as the good input, hours worked and the size of the capital stock are used as classical inputs. The model which is based on a revenue function uses gross output as good input, hours worked, the size of the capital stock and intermediate inputs are used as classical inputs. Those two models yield the two subindicators EE_t and EE_t^* used to quantify the efficiency dimension of the sectoral CO_2 performance index.

Table 1 about here

Computing the sectoral CO_2 performance indicators by weighting the four sub-indicators. The four sub-indicators listed in Table (1) - each of them standardized between zero and one - quantify the sectoral CO_2 performance. They are weighted and aggregated to obtain the sectoral CO_2 performance indicator $SPI_{i,s,t}$, where i indexes the countries, s the sectors and t the time. To weight the sub-indicators I use Principal Component Analysis (PCA). And I take the first principal component as sectoral performance index. PCA has become one of the major approaches in the construction of composite indicators.

---

[8] So, to estimate the βs of the first year only data from the first year has been used. To estimate the βs of the second year, data from the first and the second year has been used, etc.

It is now also used in social sciences in a variety of fields ranging from the computation of socio-economic well-being indexes (for an overview on weighting of such indexes see for instance Decancq and Luga (2010)), technology and science indexes (see for instance NISTEP (1995)) and sustainability indexes (see for instance Singh et al (2012)) to name only a few. PCA is generally preferred to equal weighting approaches because it doesn't impose the strong assumption that all sub-indicators are "worth" the same in the composite indicator. Even if PCA based weighting does not (necessarily) reveal the theoretical importance of the different sub-indicators, it is able to account for overlapping information between the (correlated) sub-indicators (Nardo et al, 2008), thereby avoiding the problem of "double accounting" (Decancq and Luga (2010), p. 20). The (rounded) means over all sectors of the weights per sub-indicator obtained using PCA are listed in Table (1). Note that the weights of CO_2 per capita and CO_2 per dollar are both positive. This is because those variables have been transformed as described in the footnotes in Table (1).

5.3. Computing the economy-wide CO_2 performance index by aggregating the sectoral performance indicators

The aggregation of the sectoral performance indicators is the last step in the construction of the economy-wide CO_2 performance indicator. I use a linear weighting approach to construct this final economy-wide CO_2 performance index:

$$CO_2 \ Performance \ Index_{i,t} = \sum_{s=1}^{S} \frac{(ES_{i,s,t=0} + SS_{i,s,t})}{2} SPI_{i,s,t} \tag{16}$$

Where $SS_{i,s,t}$ is country i's share of sector s in total GDP at time t, $ES_{i,s,t=0}$ is country i's share of sector s in total country emissions at time $t = 0$ and $SPI_{i,s,t}$ is country i's sectoral CO_2-performance index of sector s at time t.

I choose this weighting approach because it gives the final index several desired properties. First it respects the fact that the CO_2 performance of a country can be improved in two ways. Either by improving the within sector CO_2-performance which is captured by the SPI's. Or by reducing the share of relatively polluting sectors, a property which is respected by integrating sectoral GDP shares as a part of the weight. Second it attributes a bigger weight to a sectoral performance index the more polluting a sector is, hence the integration of the sectoral emissions share as a part of the weight. And third it allows to keep track of the history of the CO_2 performance of a country by using always the time 0 sectoral emission share.[9] The country CO_2 performance index has subsequently been normalized to range between zero and one, where one indicates the best performance and zero the worst.

---

[9] Other sectoral data based indexes, as for instance trade barrier indexes, encounter often the same problem in the aggregation phase. As an example: suppose a country improves the CO_2-performance within a given sector (and obtains a score of 1 for this sectoral CO_2 performance indicator), and in turn sectoral emissions approach zero (as an extreme case). If one now simply weights the sectoral CO_2-performance indicator (which is supposed to be one) by the corresponding contemporary sectoral emission share (which is here supposed to be zero), then the final effect on the country index would be zero. Hence the improvement in the CO_2-performance would not be reflected in the final indicator. This would be clearly undesirable.

6. Results

To obtain an overview, Figure (1) displays the evolution of the narrow CO_2 input index, the broad GHG input index and the CO_2 performance index by country. Note that due to the different data-sources, not all indexes are available for all countries. And note as well that each of the displayed indexes has been bounded to range between zero and one. In general one can observe that the CO_2 input indexes increase over time. This indicates an increase in the stringency of CO_2 policies over time in the sample. The magnitude of this increase varies however considerably among the different countries. The CO_2 performance index doesn't show such a clear pattern.

Figure 1 about here

To empirically assess whether the constructed indexes measure what they are supposed to measure one should ideally compare them to a sound benchmark. Given the absence of such a measure (i.e. given the reason why this paper has been written) I pursue two complementary evaluation strategies. First I compare the constructed input (performance) indexes to existing input (performance) indexes. Second I compare the constructed input index to the constructed performance index and verify whether the expected relationship holds, after all a higher stringency should go hand in hand with a better performance. Note that in the following I focus on the narrow CO_2 input index and the CO_2 performance index. The empirical assessment of the broad GHG input index can be found in the online appendix to this paper (Sauter, 2014). Table (2) reports the pairwise correlations of the country-means[10] of the indexes.

Table 2 about here

First look at the input indexes. The first set of benchmark indexes are the two input indexes measuring Air Policy Stringency constructed by Knill et al (2012). Both air policy indexes show a strongly positive and highly significant correlation with the narrow CO_2 input index. The higher CO_2 input policy stringency the higher air policy input stringency, a result which has been expected. As an additional benchmark the WEF survey index[11] is used (Browne et al, 2012). I expect that the opinion of the survey respondents on environmental policy stringency should be positively correlated with the CO_2 input index. This is the case, the correlation is positive and strongly significant. The benchmark comparisons seem to indicate that the narrow CO_2 input index does actually measure what he should.

The CO_2 performance index is compared to two widely used performance indexes. The Environmental Performance Index of Yale (Emerson et al, 2012) is positively and

---

[10] I use country means and not each observation available to avoid that the pairwise correlations capture trends. In the single observation case (not displayed) the correlations are stronger and more significant but the same overall tendencies hold.

[11] Even if the WEF survey index is not an input index, I use this index as a benchmark due to it's wide usage in the literature.

significantly correlated with the constructed CO_2 performance index. A better overall environmental performance parallels a better CO_2 performance. As an additional benchmark a second performance index - the index of the lead content of gasoline - is used. A negative and significant correlation is observed. Again this result is anticipated, a better CO_2 performance goes hand in hand with lower lead contents in gasoline. Overall the benchmark comparisons seem to indicate that the CO_2 performance indicator measures what he is supposed to.

Finally let's look at the relation between the narrow CO_2 input index and the CO_2 performance index. A priori, I expect that a more stringent CO_2 input policy should coincide with a better CO_2 performance. Looking simply at the correlation this seems to be the case: a positive and strongly significant correlation exists between the two indicators. Figure (2) plots the mean value of the two indexes by country, including a linear fit and the corresponding confidence interval for the mean value of the performance index given the different input index values. One can observe a rather clear tendency: the higher the mean value of the CO_2 input index, the higher the mean CO_2 performance by country. Figure (3) shows the difference between the last and the first year of the performance index on the y-axis and of the input index on the x-axis. Again a simple linear fit and the corresponding confidence interval is displayed. The results go in the expected direction, but are not strong[12]. Overall, and without making any causal statement, it seems that higher CO_2 input stringency is positively associated with a better CO_2 performance, a result which is expected.

Figure 2 about here

Figure 3 about here

Two remarks have to be made in order to illustrate both the limits and the importance of these results. First, policy unrelated factors might influence the CO_2 performance of a country. Most of the countries which show a strong increase in the CO_2 performance but only a slight increase in the policy stringency are countries from the former Soviet Union. Their position in Figure (3) might be explained by the dismantling and relocation of the heavy industry in those countries after 1991. Hence, taking performance indicators as proxies of environmental policy stringency might be quite dangerous. Second, Figure (3) underlines the importance of having clearly separated input, process and output indexes. This can be illustrated by the example of countries showing relatively big improvements of their input indexes which go hand in hand with a decline in their performance indexes. As an example take Greece. It is possible that the relatively high corruption in Greece (see for instance Transparency International (2012)) might indicate a low implementation stringency. If this is the case, this could explain Greece's position in Figure (3). Ideally, a researcher should have all three types of indicators available and use them according to his specific research question.

---

[12] In the online appendix to this paper (Autor, 2014) the same analysis has been done for other pollutants. For instance for SO_2 this result is far stronger.

7. Conclusion

In this paper I identify what I think to be the main obstacle currently limiting the development of indicators measuring environmental policy stringency: the absence of an explicitly stated methodological framework. Previous attempts violate one of the most fundamental rules found in the literature on index construction by not defining the concept they intend to quantify. What is badly defined is likely to be badly measured. The solution I propose allows to separately quantify the input, process and output dimension of various specific - hence well definable - types of environmental policies. I applied the general methodology to measure the stringency of pollutant policies, implementing a CO_2 input index and a corresponding CO_2 performance index. Additional results for SO_2 and CH_4 input and performance indicators can be found in the online appendix (Sauter, 2014) to this paper. Comparisons with available benchmark indicators suggest that the obtained indexes measure what they are supposed to.

Both of the implemented indexes can and will be extended trough time and space. In a next step it will be possible to assess to what extent policy efforts actually solved the problem they intend to solve. Using the developed input and performance indexes as well as a proxy for policy implementation stringency, it will be possible to estimate by how much policy efforts improved the CO_2 performance of a country over time. Or, in other words, it will be possible to calculate a real output index of CO_2 policy.

8. References

Broner F, Bustos P, Carvalho VM (2012) Sources of Comparative Advantage in Polluting Industries. NBER Working Papers 18337, National Bureau of Economic Research, Inc

Browne C, Geiger T, Gutknecht T (2012) The executive opinion survey: The voice of the business community. In: The Global Competitivenens Report 2012-2013, World Economic Forum, pp 69–78

Brunel C, Levinson A (2013) Measuring Environmental Regulatory Stringency. OECD Trade and Environment Working Papers 2013/5, OECD Publishing

Brunnermeier SB, Cohen MA (2003) Determinants of environmental innovation in US manufacturing industries. Journal of Environmental Economics and Management 45(2):278–293

Cole MA, Elliott RJR (2003) Do Environmental Regulations Influence Trade Patterns? Testing Old and New Trade Theories. The World Economy 26(8):1163–1186

Copeland BR (2008) The pollution haven hypothesis. In: Gallagher KP (ed) Handbook on Trade and the Environment, Edward Elgar Publishing Limited, chap 4, pp 60–70

Crandall RW (1993) Manufacturing on the Move. Brookings Institution

Damania R (2001) When the weak win: The role of investment in environmental lobbying. Journal of Environmental Economics and Management 42(1):1–22

Dasgupta S, Mody A, Roy S, Wheeler D (2001) Environmental regulation and development: A cross-country empirical analysis. Oxford Development Studies, Taylor and Francis Journals 29(2):173–187

Deacon RT (1999) The political economy of environment-development relationships: A preliminary framework. University of California at Santa Barbara, Economics Working Paper Series qt8h33b5c0, Department of Economics, UC Santa Barbara

Decancq K, Luga MA (2010) Weights in multidimensional indices of well-being: an overview. Center for Economic Studies - Discussion papers ces10.06, Katholieke Universiteit Leuven, Centrum voor Economische Studin

Eliste P, Fredriksson PG (2002) Environmental Regulations, Transfers, and Trade: Theory and Evidence. Journal of Environmental Economics and Management 43(2):234–250

Emerson J, Hsu A, Levy M, de Sherbinin A, Mara V, Esty D, Jaiteh M (2012) Environmental performance index and pilot trend environmental performance index. , Yale Center for Environmental Law and Policy

Eurostat (2009) Manual for air emissions accounts. Methodologies and working papers, European Comission

FAO, IUCN, UNEP (2013) ECOLEX - The gateway to environmental law.

Färe R, Grosskopf S, Lundgren T, Marklund PO, Zhou W (2014) Pollution-generating technologies and environmental efficiency. Journal of Chinese Economic and Business Studies 12(3):233–251

Friedman J, Gerlowski DA, Silberman J (1992) What attracts foreign multinational corporations? Evidence from branch plant location in the united states. Journal of Regional Science 32(4):403–418

Global Integrity (2011) The global integrity report 2011. Methodology white paper, Global Integrity

Grether JM, Mathys NA, de Melo J (2012) Unravelling the worldwide pollution haven effect. The Journal of International Trade & Economic Development 21(1):131–162

Harris M, Konya L, Matyas L (2003) Moddeling the impact of environmental regulations on billateral trade flows: Oecd. The World Economy 25(3):387–405

Hilton FGH, Levinson A (1998) Factoring the Environmental Kuznets Curve: Evidence from Automotive Lead Emissions. Journal of Environmental Economics and Management 35(2):126–141

Jug J, Mirza D (2005) Environmental Regulations in Gravity Equations: Evidence from Europe. The World Economy 28(11):1591–1615

Kalamova M, Johnstone N (2011) Environmental Policy Stringency and Foreign Direct Investment. OECD Environment Working Papers 33, OECD Publishing

Kaufmann D, Kraay A (2008) Governance Indicators: Where Are We, Where Should We Be Going? World Bank Research Observer 23(1):1–30

Keller W, Levinson A (2002) Pollution Abatement Costs and Foreign Direct Investment Inflows to U.S. States. The Review of Economics and Statistics 84(4):691–703

Kemeny JG, Morgenstern O, Thompson GL (1956) A generalization of the von neumann model of an expanding economy. Econometrica 24(2):115–135

Knill C, Schulze K, Tosun J (2012) Regulatory policy outputs and impacts: Exploring a complex relationship. Regulation & Governance 6(4):427–444

Kumar Mandal S, Madheswaran S (2010) Environmental efficiency of the Indian cement industry: An interstate analysis. Energy Policy 38(2):1108–1118

List JA, Co CY (2000) The Effects of Environmental Regulations on Foreign Direct Investment. Journal of Environmental Economics and Management 40(1):1–20

Magnani E (2000) The Environmental Kuznets Curve, environmental protection policy and income distribution. Ecological Economics 32(3):431–443

Mukherjee K (2008) Energy use efficiency in the Indian manufacturing sector: An interstate analysis. Energy Policy 36(2):662–672

Nakada M (2006) Distributional Conflicts and the Timing of Environmental Policy. International Environmental Agreements: Politics, Law and Economics 6(1):29–38

Nardo M, Saisana M, Saltelli A, Tarantola S, Hoffman A, Giovannini E (2008) Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD publishing

NISTEP (1995) NISTEP report no. 37: Science and technology indicators

OECD, EEA (2012) Database for instruments of environmental policy and natural resource management. Database, OECD and EEA

Pearce D, Palmer C (2001) Public and private spending for environmental protection: A cross-country policy analysis. Fiscal Studies 22(4):403–456

Sauter C (2014) Online appendix: How should we measure environmental policy stringency ? a new approach. IRENE Working Paper Series 14.2

Singh R, Murty H, Gupta S, Dikshit A (2012) An overview of sustainability assessment methodologies. Ecological Indicators 15:1:281–299

Smarzynska JB, Shang-Jin W (2003) Pollution Havens and Foreign Direct Investment: Dirty Secret or Popular Myth? The BE Journal of Economic Analysis & Policy 3(2):1–34

Stehrer R, de Vries GJ, Los B, Dietzenbacher H, Timmer M (2014) The World Input-Output Database: Content, Concepts and Applications. GGDC Research Memorandum GD-144, Groningen Growth and Development Centre, University of Groningen

Timmins C, Wagner U (2009) Agglomeration effects in foreign direct investment and the pollution haven hypothesis. Environmental and Ressource Economics 43(2):231–256

Transparency International (2012) Corruption perception index 2012

United Nations (2012) System of environmental-economic accounting framework: Central framework. White cover publication, pre-edited text subject to official editing. , European Commission, Food and Agriculture Organization, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations and World Bank

Von Neumann J (1945) A model of general economic equilibrium. The Review of Economic Studies 13(1):1–9

World Bank (2012) Methodology: Doing business indicator. http://www.doingbusiness.org/methodology

Xing Y, Kolstad C (2002) Do Lax Environmental Regulations Attract Foreign Investment? Environmental & Resource Economics 21(1):1–22

9. Tables and Figures

Table 1: Sectoral CO_2 performance sub-indicators

Indicator	**Description**	**Mean weight**	**Dimension**
$\frac{\text{sectoral } CO_2 \text{ emissions}}{\text{sectoral GDP}}$[13]	Sectoral CO_2 per sectoral GDP	0.446	CO_2 intensity
$\frac{\text{sectoral } CO_2 \text{ emissions}}{\text{sectoral work force}}$[14]	Sectoral CO_2 per sectoral workforce	0.226	
EE_t	CO_2 efficiency score (profit function)	0.592	CO_2 efficiency
EE_t^*	CO_2 efficiency score (revenue function)	0.589	

Table 2: Pairwise correlations of the means of the variables

	Narrow CO_2 II	Air Policy II 1	Air Policy II 2	WEF	CO_2 PI	EPI	Lead
Narrow CO_2 II	1						
Air Policy II 1	.633***	1					
Air Policy II 2	.687****	.905****	1				
WEF	.422***	-.139	.0466	1			
CO_2 PI	.540****	-.00243	.165	.469***	1		
EPI	.267*	.144	.273	.660****	.403**	1	
Lead	-.291*	-.0938	-.235	-.544****	-.380**	-.553****	1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Note: II stands for Input Index, PI for Performance Index. The Narrow CO_2 Input Index and the CO_2 Performance Index have been constructed by the above outlined methodology. The Air Policy Input Index 1 and 2 are taken from Knill et al (2012). The WEF survey index is taken from Browne et al (2012). The Environmental Performance Index (EPI) is taken from Emerson et al (2012) and the lead content of gasoline (Lead) index is taken from Grether et al (2012).

---

[13] Note that this variable has been re-scaled. Each observed value is subtracted from the observed maximum (max) of the variable, then the minimum (min) of the variable is subtracted: (max-observation)-min. With this transformation higher values now indicate a better performance.

[14] See: footnote 12.

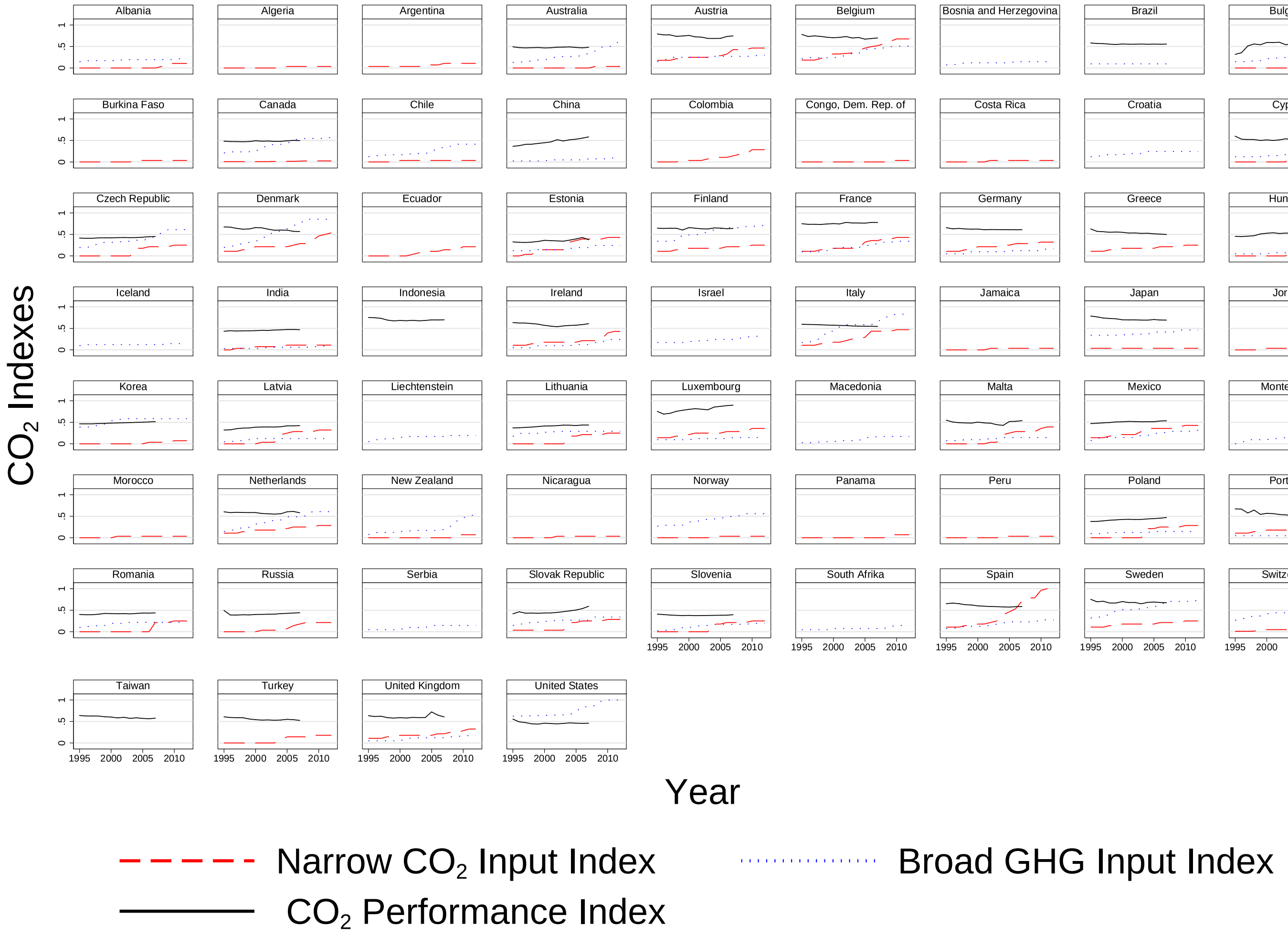

Figure 1: The CO_2 input indexes and the CO_2 performance index by country

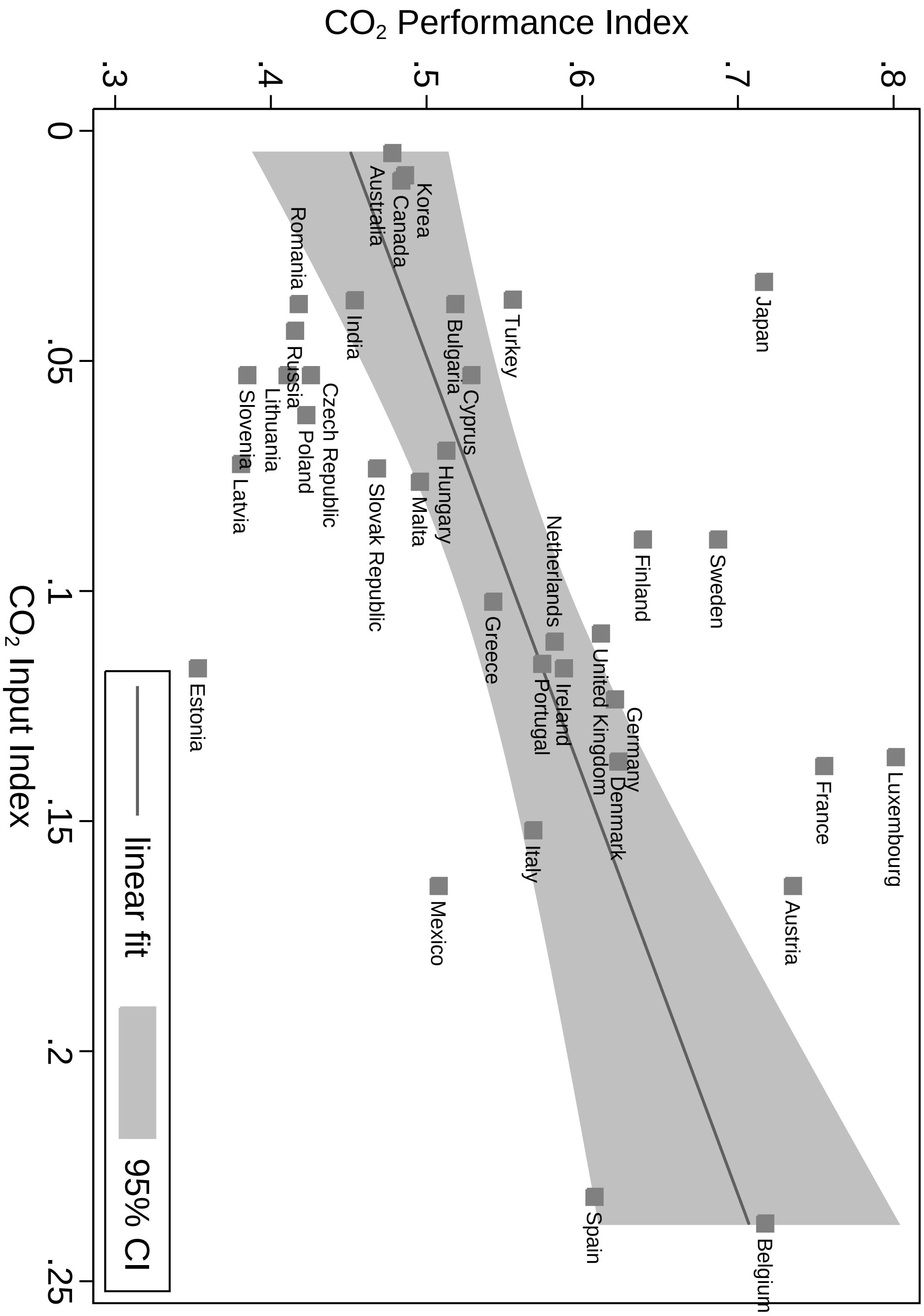

Figure 2: Means of the Narrow CO_2 input index and the CO_2 performance index by country

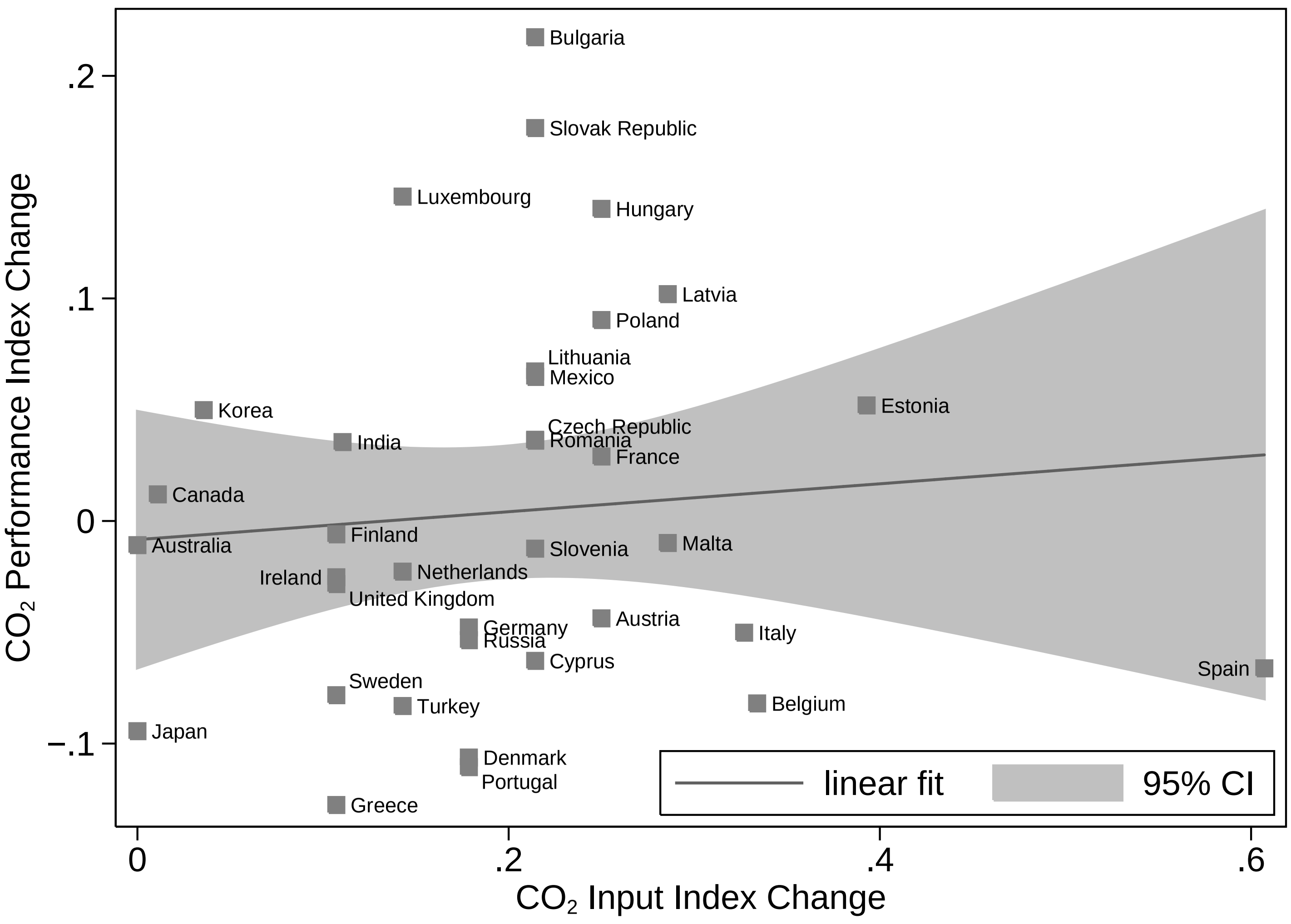

Figure 3: Change of the Narrow CO_2 input index and of the CO_2 performance index from the first to the last year in the sample