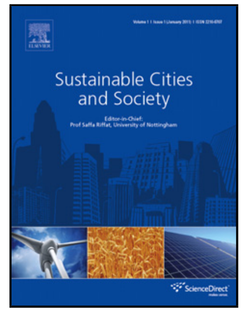
Journal Pre-proof



The effect of urban air pollutants in Germany: eco-efficiency analysis through fractional regression models applied after DEA and SFA efficiency predictions

Victor Moutinho, Mara Madaleno, Pedro Macedo

PII: S2210-6707(20)30191-8

DOI: https://doi.org/10.1016/j.scs.2020.102204

Reference: SCS 102204

To appear in: *Sustainable Cities and Society*

Received Date: 15 August 2019

Revised Date: 10 April 2020

Accepted Date: 10 April 2020

The effect of urban air pollutants in Germany: eco-efficiency analysis through fractional regression models applied after DEA and SFA efficiency predictions

Victor Moutinho,[1] Mara Madaleno[2*] and Pedro Macedo[3]

[1]Department of Management and Economics and NECE – Research Unit in Business Sciences, University of Beira Interior, Portugal, ferreira.moutinho@ubi.pt

[2] GOVCOPP – Research Unit in Governance, Competitiveness and Public Policy, Department of Economics, Management, Industrial Engineering and Tourism (DEGEIT), University of Aveiro, 3810-193 Aveiro, Portugal.

[3] CIDMA – Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, 3810-193 Aveiro, Portugal, pmacedo@ua.pt

*Corresponding author: maramadaleno@ua.pt, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

Highlights

- Analysis of the effect of urban air pollution in eco-efficiency of cities.
- Significant effect of pollutants, temperature and rainfall over technical efficiency.
- New measure of eco-efficiency ratio proposed based over pollutants.
- It is required the control of urban air pollution and temperature for sustainable cities.
- Estimations based in a two-step econometric procedure.

Abstract

Cities and living standards contribute intensively to air pollution, an environmental risk factor which causes diseases. Recently, in developed countries, the majority of cities has grown rapidly and has experienced increasing environmental problems. In this article we analyze the effect of urban air pollution considering the available data for the years 2007, 2010 and 2013 in 24 German cities. Proposing a new model, we start the analysis using data envelopment analysis (DEA) and stochastic frontier analysis (SFA) to predict eco-efficiency scores for the 24 German cities.

Afterwards, it is applied fractional regression to infer about the influencing factors of the eco-efficiency scores, at the city level. Results suggest a significant impact over eco-efficiency due to the excess of PM10, the average temperature, the average of NO_2 concentration and rainfall. The findings in this study hold important implications for policymakers and urban planners in Germany, especially those that coordinate environmental protection and economic development in cities. Therefore, interventions to reduce urban air pollution can be accomplished on different regulatory levels, leading to synergistic effects as the decrease of climate change effects and noise.

Keywords: Air pollutants; Eco-Efficiency; German Cities; Data Envelopment Analysis (DEA); Stochastic Frontier Analysis (SFA); Fractional Regression Models (FRM)

1. Introduction

This work proposes a new measure of cities' eco-efficiency by computing the ratio of GDP by a pollutant (PM10) rather than by emissions (CO_2 usually used in literature), using two methods simultaneously to evaluate differences in prediction. Then, regression modeling is implemented to identify the factors that are being able to explain cities' eco-efficiency scores.

City dwellers conditions depend over urbanization planning and management and on the use of resources (Addanki and Venkataraman, 2017). Rural-urban migrations are increasing (Silva et al., 2018; Ameen and Mourshed, 2019) and many problems arise (traffic congestion, social disorder, biodiversity reduction, air pollution, water quality deterioration, etc.) (Yin et al., 2014; Zhu et al., 2019). With urban development, air quality becomes a major problem, worth of study. In fact, in Europe, poor air quality is still classified as a huge environmental and health risk, despite recent year's improvements (Schmitz et al., 2018). According to EEA (2016), air pollution also leads to major economic costs.

Road transport was the largest source of NOx emissions in the European Union (EEA, 2017), in about 39% of total emissions, and a minor source of PM10 and PM2.5 emissions, in about 13% each. Due to the high proportion of diesel vehicles on the road in Germany, exceedances of NO_2 are a major challenge. Currently, the German federal court is ruling harder policy measures. Recently it has ruled that the implementation of bans on diesel vehicles is allowed, when these enter in highly-polluted areas in cities. This was done as last option in order for Germany to meet EU air quality regulations (Schmitz et al., 2018; Bundesverwaltungsgericht, 2018). Therefore, in order to turn cities healthier, more accessible and sustainable, and considering that air quality remains a significant challenge to cities in Europe, it is necessary to produce policy-relevant research able to guide decision making. To date there exists little research on whether air pollution concentration levels and how air pollution can deteriorate the eco-efficiency in such cities.

Eco-efficiency has been considered as an operational instrument to facilitate sustainable development and green economic transformation (Yang et al., 2020). Hence, it is quite necessary to estimate cities eco-efficiency so as to reveal a clear picture of the current situation of sustainable development, as to date there is little research exploring cities. Eco-efficiency is defined as able to generate more value through technology and process changes, whilst reducing resource use and environmental impact throughout the product or service's life (Li et al., 2018, 2018a; Yang et al., 2020).

This work contributes to the existent literature in several different ways. First, it differs from the rest of the articles which basically determine eco-efficiency scores (Li et al., 2018, 2018a). As such, we go one step further including into the analysis factors able to explain these scores. This is done with a two-step estimation procedure and by the way the city’s environmental problem was approached. Second, the greater majority of previous studies compute the eco-efficiency measure through the ratio of GDP by CO_2 emissions, while we use the ratio of GDP by PM10 as the output variable for SFA and DEA methodologies. A more recent trend of the literature pursues similar procedures and provide evidence for more rational measures using these type of methods (Piña and Martínez, 2016; Zhu et al., 2019). Third, we do not use CO_2 as usual, provided that CO_2 is not considered a pollutant (Wallace III et al., 2017; Beisner, 2019). Instead, we use PM10 to compute the ratio of regional GDP over PM10, at the city level. With this measure we are able to capture the percentage of income generated within the city in order to mitigate this pollutant. As far as we are aware, this ratio has not been used by previous literature. Furthermore, labelling CO_2 as a "pollutant" is a disservice to a gas that has played an enormous role in the development and sustainability of all life, clearly ruling out that CO_2 is not a pollutant (see European air quality reports[1]). As such, to better account for air pollution we need to take into account its constituents. Fourth, we use fractional regression models in the second step estimation to analyze how urban air pollutants and weather (number of days where PM10 exceeded 50 $\mu g/m^3$, average temperature, average concentrated NO_2, number of days that Ozone concentration exceeds 120 $\mu g/m^3$ and rainfall) influence the city's eco-efficiency.

The rest of the paper is organized as follows. Section 2 presents the data and the methodology applied, whereas Section 3 presents the main empirical results attained through fractional regression models. Section 4 presents some policy implications discussion. Section 5 concludes the work, pointing directions for future research.

2. Literature Review

2.1. Urban air pollution and cities

Climate change induced by men is a reality and has now been accepted socially and scientifically as a hard burden that the present generation and future generations need to face (greenhouse gases like carbon dioxide, nitrous oxide and methane are the main responsible for global warming).

[1] https://www.eea.europa.eu/publications/air-quality-in-europe-2018

Regionally, climate change has substantially different effects when the exposition to atmospheric pollutants of anthropogenic origin, like PM10 (particulate matter less than 10 µm), PM2.5 (particulate matter less than 2.5 µm) and NO_2 (particles concentration of nitrogen dioxide), is higher. Air pollution effects are thus clearly visible in urban areas provided the high demographic density and profound traffic circulation (Silva et al., 2018; Li et al., 2018; Qiu et al., 2019).

When EU limit values established for PM10, PM2.5 and NO_2 (particles concentrations) and ozone concentration levels are surpassed, intensive ventilation is no longer effective nor even recommended. This is also true in periods of high humidity and during heat waves.

Air pollution is any substance or matter emitted into or that otherwise enters the ambient air, caused by an agent or a combination of agents. It includes any physical, chemical, biological or radioactive substance or matter (source material, special nuclear material and byproduct material). EEA (2017) states that more than 82% of urban population in Europe is exposed to PM2.5 concentration above the guidelines of WHO (2006), as observed by Pisoni et al., (2019). This research provides a pilot study of how air pollutants and weather influence eco-efficiency in German cities.

Local authorities can play an important role, provided there are multiple efforts at the local scale which evaluate pollution concentrations (Carnevale et al., 2011) and evaluate the impact of local policies on air quality (Carnevale et al., 2014). But in order to perform good and effective policies they need to understand which factors may explain the relationship between urban air pollution and eco-efficiency of cities to develop appropriate actions. The present article tries to highlight these factors considering German cities. Results of Wang and Yuan (2018) show that in the short-term, air pollution control has a significant inhibiting effect on industrial ecological total-factor energy efficiency.

In Germany, the possible regional effects of climate change have been intensively investigated for geographic conditions and structural population conditions. Moreover, issued greenhouse gases due to production processes have a marked impact over the environment and are also the most perpetrator of climate change and global warming, especially in German cities (Moutinho et al., 2018).

2.2. Regional eco-efficiency

Eco-efficiency is the ability to produce more goods and services with less impact on the environment and less consumption of natural resources (UN, 2009). Thus, politicians, scientists and researchers have devoted increased attention to on how to reduce the environmental burden and increase eco-efficiency. Furthermore, greenhouse gases emitted as a result of production processes have a marked impact on the environment and are also the foremost culprit of global warming and climate change (Moutinho et al., 2018).

Yin et al. (2014) use eco-efficiency as an indicator to measure urban sustainable development in China, using data envelopment analysis and a super-efficiency model. Their eco-efficiency measure includes waste water, CO_2 and SO_2 emissions, industrial dust and solid waste emissions. For only one province (Guangdong), Zhou et al. (2018) test for eco-efficiency and its influencing factors based upon super-SBM and panel regression models. Results indicate that technical innovation had the greatest positive influence on eco-efficiency. Also, government regulation, openness and population density are positive influencers. Negative influencers were intensive land-use, industrial structure and per capita GDP. Bian et al. (2019) use per capita GDP and SO_2, waste water and soot/dust as variables, and super-efficiency to measure eco-efficiency in 278 Chinese cities. Huang et al. (2014) use GDP and the environment pollution index. Yu et al. (2019) provide a summary of DEA applications to eco-efficiency in China. From their Table 1 we may observe that authors use emissions instead of air pollution measures in their considerations of undesirable outputs and as desirable outputs, and mostly it is used GDP. Wand and Yuan (2018) uses a panel data of 37 subsectors of China’s industrial sector from 2003 to 2014 to examine the influence of air pollution control on ecological total-factor energy efficiency.

Despite the high number of studies testing eco-efficiency or efficiency in China, little is known considering European cities, maybe due to the lack of robust data in terms of complete periods of time. Moreover, developed countries are less studied regarding air pollution and its effects, or how urban factors may influence air quality in these cities. For these reasons and also provided it is now proposed a new measure of eco-efficiency (provided the undesirable output variable used), we ask if there are significant impacts regarding urban factors effects on air pollution, and we do that considering the EU available data, in this case for German cities.

2.3. Factors explaining eco-efficiency

For Colombia, Piña and Martínez (2016) estimate and evaluate the environmental, social and economic efficiency of cities using data envelopment analysis. The authors found differences among cities, being that the most efficient ones show adequate resource use, lower environmental impacts and better social conditions. They conclude that while city scale increases, urban sustainability declines. They reinforce that there are differences among cities that guarantee economic growth and development. As pointed by Ameen and Mourshed (2019), economic aspects of urban development are particularly important for rapidly developing countries like Iraq. However, urban development necessarily increases pollution levels and decreases air quality (Cho and Choi, 2014; Alpert et al., 2019).

In a study applied to 35 large and medium sized cities in China, Zhu et al. (2019) measure the efficiency and driving factors of urban land use. They have applied the DEA method and the PLS-SEM model to conclude that infrastructure, economic, market and land systems have significant influence on urban land use efficiency. For a complete review of urban sustainability measures for developing new cities, we refer to Addanki and Venkataraman (2017) and to Silva et al. (2018) for trends, components and open challenges in smart cities.

Lin and Zhu (2018) study the air quality in 282 Chinese cities to conclude that concentration of SO_2 and PM10 present the U-shaped inverse characteristic. They conclude that cities with higher urbanization rate tend to have lower air pollution concentration. Also for Chinese cities, Qiu et al. (2019) and Sun et al. (2018) highlight that transportation factors have the most significant effects over air quality, using panel data models. Han et al. (2014) analyze the impact of the urbanization degree on urban air quality considering fine particles like PM2.5, also for Chinese cities. Chen and Xu (2017), also for China, study the relationship between air quality and economic development in provincial capital cities. They have taken meteorological conditions and industrial structure when testing the environmental Kuznets curve (EKC) hypothesis, showing no direct relationships between PM10, SO_2 and NO_2 and gross regional product. Li et al. (2018) also analyze energy and the air quality index performance of 31 cities in China, concluding that 19 cities need to significantly improve their energy environmental efficiency and 22 cities need to significantly improve their overall efficiency.

As can be seen, many are the studies that account for emissions rather than air pollution levels in efficiency studies. Likewise, Cho and Choi (2014) show that SO_2 decreases as the proportion of green area increases in 17 cities in Korea from 1996 to 2009, while an increase in net density leads to an increase of NO_2. But, particulate matter (PM) is the main component of air pollution (Costa et al., 2017; Sun et al., 2018). Therefore, we expect that PM10, NO_2 and Ozone concentration levels do have an impact over eco-efficiency. Moreover, as urban development emerges, precipitation downwind increases (Alpert et al., 2019). Arminen and Menegaki (2019) state that climate and weather variations are important determinants of energy consumption and CO_2 emissions, leading us to include rainfall as a possible explanatory variable of eco-efficiency levels. Using SFA, Deng and Gibson (2018) analyze land use conversions and eco-efficiency in Hebei cities, China, using precipitation and average temperature as proxies for meteorological data. In conclusion, the variables used in the two-step procedure and the modeling of eco-efficiency presented are the main novelties of this article.

3. Data and Methodology

This paper analyses a set of selected German cities' performance in terms of the relative behavior of their eco-efficiencies, computed as the ratio of their gross domestic product (GDP) over particulate matter emissions (PM10 - annual average concentration in µg/m³). To perform our analysis we used data retrieved from the Eurostat municipal database[2] and from the OECD cities database[3] for the three years where there was available data (2007, 2010 and 2013). The DMUs (decision making units) are the German cities (24 DMUs).

The eco-efficiency ratio is usually measured as the ratio of the added value of what has been produced (e.g. GDP) and the added environmental impacts of the product or service produced (normally using the CO_2 emissions) (Yadong, 2013). The present article uses a new measure of cities' eco-efficiency by computing the ratio using a pollutant rather than emissions. As such, in a first step, eco-efficiency scores of the selected German cities are computed using data envelopment analysis (DEA), with variable returns to scale (VRS) and constant returns to scale (CRS), and stochastic frontier analysis (SFA) using the maximum likelihood (ML) estimator. DEA and SFA are simultaneously used here for comparison purposes and to evaluate differences in prediction. The inputs considered to predict eco-efficiency scores are the population density (persons/km^2), the labor productivity (measured in US dollars at constant values of 2010), the municipal waste (both domestic and commercial measured in thousand tons), the number of registered cars (per one thousand persons) and the number of companies in the city[4]. Table 1 presents some descriptive statistics for all the variables considered in the work (first and second step).

Table 1. Some descriptive statistics

	Minimum			Maximum			Mean			Std. Deviation		
	2007	2010	2013	2007	2010	2013	2007	2010	2013	2007	2010	2013

[2] https://ec.europa.eu/eurostat/web/cities/data/database

[3] https://stats.oecd.org/Index.aspx?DataSetCode=CITIES

[4] Since some companies pollute much more than others, information disaggregated by sector is desirable in future research.

Population density	Persons/km^2	313.0	319.0	323.0	2231.0	2196.0	2162.0	779.5	775.2	771.1	534.3	521.2	508.6
Labor productivity	US dollars at constant values of 2010	68662.6	66035.3	67234.3	132142.3	123984.9	127910.3	94071.7	89616.0	91705.9	16723.1	15157.9	15642.9
Municipal waste	1000t	69.0	68.0	60.0	1472.0	1409.0	1390.0	298.6	289.7	285.0	305.2	290.8	287.6
Number of registered cars	Per 1000 persons	290.0	289.0	302.0	467.0	431.0	443.0	363.7	356.9	372.5	37.8	34.9	36.2
Number of companies in the city	Units	5682.0	10246.0	9895.0	159617.0	164421.0	174654.0	29484.2	35170.6	36266.0	34254.4	36898.9	38989.2
GDP	Per capita	30324.0	31192.0	31570.0	64840.0	60970.0	62153.0	42939.2	42509.3	42915.3	8971.1	7879.3	7875.0
PM10	Annual average concentration in µg/m^3	18.0	18.5	16.1	28.1	27.5	24.3	22.6	22.9	20.3	2.6	2.5	2.2
PM10 > 50 µg/m^3	Number of days where PM10 exceeded 50 µg/m^3	3.0	5.5	2.0	30.0	31.8	15.0	13.3	17.0	9.0	7.5	7.2	3.9
NO_2	Annual average concentration in µg/m^3	17.3	19.4	17.3	19.6	22.7	19.7	18.4	21.3	18.5	0.7	0.7	0.8
Temperature	Average temperature	15.3	18.0	3.9	38.5	35.0	33.8	25.9	26.5	23.2	6.1	5.2	6.5
Ozone concentration (OC) > 120 µg/m^3	Number of days that OC exceeds 120 µg/m^3	7.17	13.0	3.5	33.0	1016.0	29.0	16.67	19.8	12.3	7.3	4.9	6.5
Rainfall	In l/m^2	79.7	560.4	506.7	1077.3	35.0	972.3	830.6	782.4	672.7	205.5	110.3	111.2

In a second step, it is used the fractional regression model (FRM) in its four usual forms: Logit, Probit, Log-Log and Complementary Log-Log, following Ramalho et al. (2010). Using the FRM framework, presented and developed by Papke and Wooldridge (1996), the assumption of a functional form is needed, where the dependent variable, obtained through the first step DEA and SFA scores, is limited to the interval [0, 1]. Being the eco-efficiency scores the dependent variable, the independent variables considered into the analysis are the number of days where PM10 exceeded 50 µg/m^3, the average temperature, the average concentration of NO_2 (Nitrogen Dioxide in µg/m^3), the number of days that Ozone concentration exceeds 120 µg/m^3 and rainfall (l/m^2). These five independent variables are accounted as the factors being able to explain eco-efficiency scores through cross-sectional regression models. In the second step procedure we will apply the one-part model only for efficient DMUs and a two-part models which are separated into first-part (for inefficient DMUs) and second-part (for efficient DMUs) component estimations. The R software, MATLAB and Stata 14 were used to perform all the statistical analysis in this work.

3.1 First Step: A Brief Overview of the DEA Methodology

Initially developed by Charnes, Cooper and Rhodes (1978), data envelopment analysis (DEA) is a decision-making tool based on linear programming for measuring the relative efficiencies of a set of comparable units. It consists in a nonparametric method used to estimate production frontiers and to evaluate the efficiency of the Decision Making Units (DMUs).

By applying the DEA model it is determined an envelopment surface called the empirical production function or the efficient frontier. It can be used to compute efficiency using two different specifications: in the traditional approach of Charnes et al. (1978, CCR) through constant returns to scale (CRS); or by the Banker, Charnes and Cooper (1984) extension of the original CCR approach, through variable returns to scale (VRS). The VRS specification model is sometimes known in the literature as the BCC model, named due to its authors.

The efficient DMUs cannot be compared among themselves in both DEA CCR and BCC models. In order to surpass some of these problems, DEA researchers have initiated a new era called of super-efficiency, to rank the DEA efficient DMUs. This super-efficiency, pioneered by Banker and Gifford (1988) and Banker and Datar (1989a, 1989b), infers the possible capability of a DMU in increasing its outputs and/or in reducing its inputs without becoming inefficient. Afterwards, Anderson and Peterson (1993) proposed two models (CRS and VRS) by making modifications on the original CCR and BCC models. In these, they introduced the super-efficiency as a ranking methodology to distinguish the extreme-efficient DMUs performance.

3.2 First Step: A Brief Overview of the SFA Methodology

Around the seventies of the twentieth century, Aigner et al. (1977), Battese and Corra (1977) and Meeusen and van den Broeck (1977) introduced the stochastic frontier analysis (SFA). The main feature of a SFA model is its structure of the composed error, which separates the impacts over production which are outside the producer's control (as for example, strikes, material malfunction or bad weather) from the technical efficiency. For the SFA model in this work it was assumed a log-linear Cobb-Douglas function for the production frontier. The output variable assumed, as in the DEA models, was the same (ratio of GDP per capita by the annual average concentration of PM10), as well as the input variables: labor productivity, population density, number of companies, number of registered cars and municipal waste generated.

In order to be able to predict technical efficiency scores for each producer, or DMU, distributional assumptions are required for the composed error structure. Presently, considering a random sample of producers, or DMUs, $(i = 1, 2, \dots, N)$, we use the maximum likelihood estimator and assume that v_i (noise component) is independent and identically normally distributed with zero mean and variance given by σ_v^2. It is also assumed here that u_i (inefficiency component) is independent and identically distributed, following a nonnegative half-normal distribution with zero mean and variance σ_u^2. Finally, it is also assumed that v_i is distributed independently of u_i, and both are uncorrelated with the explanatory variables. This is the well-known normal – half-normal model widely employed in empirical studies.

With respect to these model assumptions, it is interesting to note that if $u_i = 0$, then the composed error is symmetric ($\varepsilon_i = v_i$) and there is evidence of full technical efficiency, if it is assumed that the model is correctly specified. But if $u_i > 0$, we will end up with a composed error which is negatively skewed, and as a consequence, there will be evidence of technical inefficiency. As such, we need to start by testing for the presence of technical inefficiency, and here we accomplish this based on the ordinary least squares residuals, having investigating for several different models. At the end, provided the marginal density function of ε is asymmetrically distributed, we follow Kumbhakar and Lovell (2000, p. 74-80), which presented all the estimation procedures, including the log-likelihood function to be maximized using numerical optimization techniques, the prediction of technical efficiency for each producer, $TE_i := \exp(-u_i)$, and confidence intervals for efficiency scores.

3.3 Second Step: Fractional Regression Model (FRM)

By using the FRM we avoid the problems associated with the application of the linear and Tobit models considering the predictions of technical efficiency, both from the DEA and SFA methodologies, as the dependent variable. Papke and Wooldridge (1996) developed the FRM framework. It requires the assumption of a functional form, where the dependent variable, corresponding to the first stage DEA and SFA scores, is limited in the interval [0, 1]. Ramalho et al. (2010) state that this functional form enforces the desired constraints on the conditional mean of the dependent variable (y). Under this, $E(y|x) = G(x\theta)$ is, therefore, bounded to that same interval, where $G(.)$ represents a non-linear function satisfying the condition $0 \leq G(.) \leq 1$, x represents a vector with environmental variables and θ represents a vector of parameters to be estimated. Papke and Wooldridge (1996) suggest as possible specifications for the non-linear function any cumulative distribution function usually applied to model binary data. The most widely used ones are the Logit and Probit functional forms, as well as the Log-Log and the Complementary Log-Log specifications. Additionally, based on the Bernoulli log-likelihood function, the same authors propose the estimation of FRM using a quasi-maximum likelihood estimator of θ, given by $\arg\max_{\theta} \sum_i \left(y_i \log\left(G(x_i\theta)\right) + (1 - y_i)\log\left(1 - G(x_i\theta)\right)\right)$. Properties of the estimator can be found in Papke and Wooldridge (1996) and Ramalho et al. (2010).

Ramalho et al. (2010, 2011) have proposed two generalized models as an alternative to the standard models. These use an additional parameter, α, which will result in the first and second generalizations presented afterwards in the results and discussion section. So, in accordance to the Generalized Type I model we have $E(y|x) = G(x\theta)^{\alpha}$ and the partial effects of a unitary change in

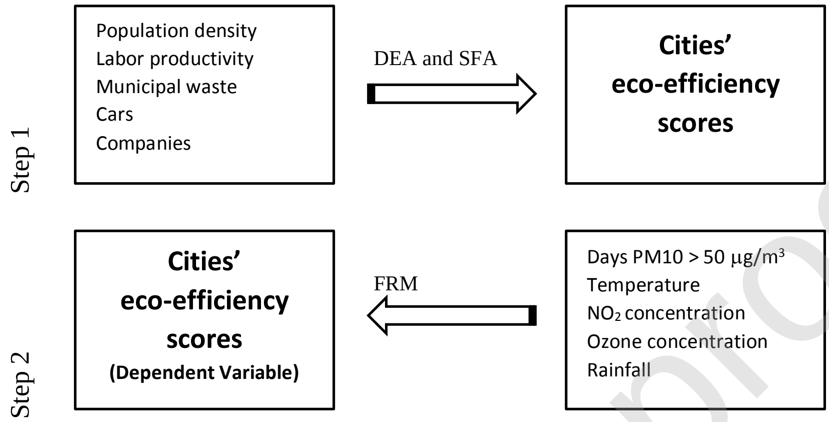
x_j is given by $\frac{\partial E(y|x)}{\partial x_j} = \theta_j g(x\theta)\alpha G(x\theta)^{\alpha-1}$, while in the Generalized Type II model we have $E(y|x) = 1-(1- G(x\theta))^{\alpha}$, where $\alpha > 0$ such that $0 < E(y|x) < 1$, and the partial effects of a unitary change in x_j is given by $\frac{\partial E(y|x)}{\partial x_j} = \theta_j g(x\theta)\alpha(1-G(x\theta))^{\alpha-1}$ (see Ramalho et al., 2010, 2011, for further details).

Additionally, there are the two part-models that should be used when the probability of observing a score of unity is relatively large. This would lead us to suspect that the sources behind efficient DMUs could differ from those associated to the inefficient DMUs (Ramalho et al., 2010). The first part of the model encompasses a standard binary choice model, which manages the probability of observing an eficient DMU, where z is a binary indicator that takes the value of 0 (for $0 < y < 1$) and 1 (for $y = 1$). The conditional probability of observing an efficient DMU, usually estimated by maximum likelihood using the whole sample, is given by $\Pr(z = 1|x) = E(z|x) = F(x\beta_{1P})$, where β_{1P} is a vector of coefficients and $F(.)$ is a cumulative distribution function. The second part of the two-part models is estimated using the sub-sample of inefficient DMUs, considering $E(y|x,\ y \in]0,1[) = M(x\beta_{2P})$, where $M(.)$ may be any of the specifications considered for $E(y|x)$ and β_{2P} is a vector of coefficients. Naturally, to correctly specify the functional form of the conditional mean $E(y|x)$ it is required a correct model specification for $G(x\theta)$ and for both $F(x\beta_{1P})$ and $M(x\beta_{2P})$ in the one- and two-part models, respectively (Ramalho et al., 2010). A possible way to accomplish this is to apply the RESET test, able to detect misspecification of the general functional form.

Davidson and MacKinnon (1981) suggested the P test which can be used to compare non-linear regression models and as such to discriminate between alternative one- and two-part FRM, as argued by Ramalho et al. (2010). Moreover, one may also apply the GOFF-I and GOFF-II tests to infer about the relevance of using either Type I or Type II generalisations, or if it is adequate just to apply the corresponding simpler standard FRM. All the mathematical details concerning FRM for second stage DEA efficiency analyses can be found in Ramalho et al. (2010).

In order to turn more evident the two-step research process adopted in this work we present a brief flowchart in Figure 1.

Figure 1. Two-step research process adopted in the work



4. Results and Discussion

Table 2 starts presenting the calculation of the geometric mean of the efficiency scores obtained to the three years of analysis 2007, 2010 and 2013, and for the methodologies applied (DEA-CRS, DEA-VRS and SFA-ML). The results of the non-parametric DEA technique, with constant and variable returns to scale, show that the German cities München, Freiburg, Karlsruhe and Saarbrücken have the score value equal to one. This means that these four cities are positioned in the frontier (maximum efficiency). Moreover, when we analyze the results of DEA-VRS, there are other German cities that are also positioned in the frontier of maximum efficiency like Berlin, Frankfurt, Bremen, Bochum, Augsburg, Münster and Aachen.

Considering the results of the parametric technique SFA-ML, there is evidence to indicate that the German cities of Stuttgart, Bochum and Hannover present the higher values of scores and closer to the frontier (maximum efficiency), revealing a second group of cities that also present high values of scores and very close to the previous ones: Freiburg, Aachen and Hamburg.

If we analyze only the last year of available data, 2013, it is observed that under the DEA-CRS technique there are two German cities, Bremen and Hannover, which are located in the frontier, while in the DEA-VRS technique, the cities Stuttgart, Leipzig and Dresden present values of maximum efficiency. In the SFA-ML technique for 2013, Bochum and Stuttgart lead the efficiency performance, with values close to the maximum efficiency, followed closely by the cities Hamburg, Hannover and München. Figure 2 presents the evolution through time of the cumulative efficiency scores attained through the SFA-ML methodology and we are able to observe that Hamburg, München, Stuttgart, Hannover, Bochum, Freiburg, Saarbrücken and Aachen are all very close to three, meaning that these are the cities which, in cumulative terms, have mostly reached maximum efficiency values through time.

Since we find several DEA scores of unity, as mentioned previously, it might be adequate to use the two-part models. In regression models we need to choose the functional form for the conditional expectation. These alternative tests are used: i) RESET test to detect badly specified functional forms; 2) the goodness-of-functional form (GOFF) test based over the generalized functional forms that contain the non-linear function G(.) ∈ [0, 1] in $E(y|x) = G(x\theta)$; iii) the P test where alternative specifications may be applied to test the complete specification of the two-part models.

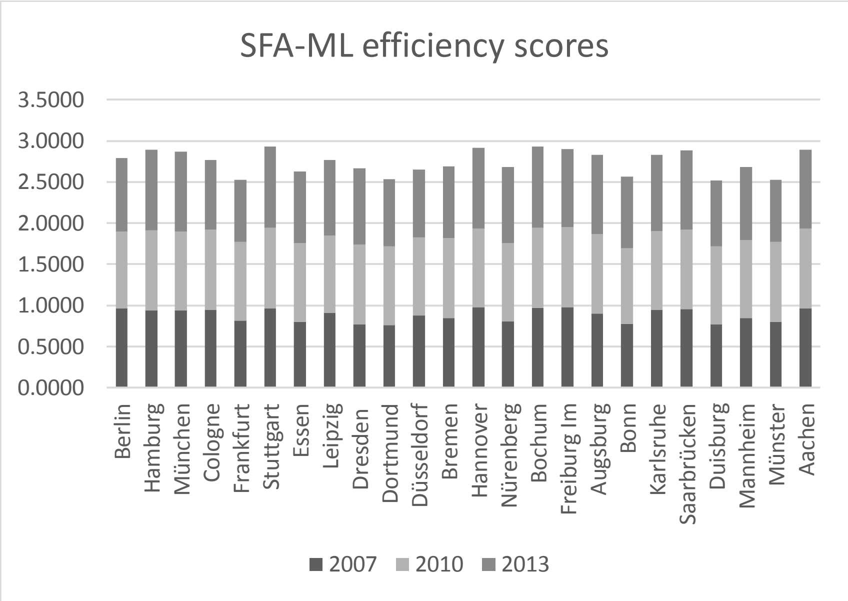
Table 2. Eco-efficiency predictions in 2007, 2010 and 2013 using DEA and SFA-ML methods

Cities/Year	DEA-CRS 2007	DEA-CRS 2010	DEA-CRS 2013	Geometric mean	DEA-VRS 2007	DEA-VRS 2010	DEA-VRS 2013	Geometric mean	SFA-ML 2007	SFA-ML 2010	SFA-ML 2013	Geometric mean
Berlin	0.7230	0.5908	0.7300	0.6781	1.0000	1.0000	1.0000	1.0000	0.9652	0.9336	0.8925	0.9300
Hamburg	0.9352	0.9588	0.8955	0.9295	0.9736	0.9717	0.9494	0.9648	0.9380	0.9732	0.9789	0.9632
München	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9365	0.9620	0.9744	0.9575
Cologne	0.8148	0.7423	0.7179	0.7572	0.8667	0.8670	0.8938	0.8757	0.9484	0.9755	0.8476	0.9222
Frankfurt	1.0000	1.0000	0.9615	0.9870	1.0000	1.0000	1.0000	1.0000	0.8131	0.9568	0.7607	0.8396
Stuttgart	0.9275	0.8999	0.9595	0.9286	0.9348	0.9044	1.0000	0.9456	0.9633	0.9793	0.9883	0.9769
Essen	0.5475	0.4636	0.6219	0.5404	0.8365	0.7723	0.8466	0.8178	0.7984	0.9597	0.8736	0.8748
Leipzig	0.8083	0.5936	0.8393	0.7385	1.0000	0.9283	1.0000	0.9755	0.9088	0.9423	0.9173	0.9277
Dresden	0.5578	0.6497	0.8205	0.6675	0.9864	0.8912	1.0000	0.9579	0.7666	0.9756	0.9216	0.8833
Dortmund	0.5441	0.4627	0.6023	0.5332	0.8309	0.7629	0.8437	0.8117	0.7582	0.9564	0.8229	0.8419
Düsseldorf	0.7971	0.7352	0.7579	0.7630	0.8706	0.8661	0.9008	0.8790	0.8749	0.9536	0.8231	0.8823

Bremen	0.9945	1.0000	1.0000	0.9982	1.0000	1.0000	1.0000	1.0000	0.8470	0.9684	0.8714	0.8941
Hannover	0.8679	0.9544	1.0000	0.9391	0.8699	0.9783	1.0000	0.9476	0.9823	0.9546	0.9783	0.9717
Nürenberg	0.7804	0.8257	0.8913	0.8312	0.8989	0.9022	0.9161	0.9057	0.8087	0.9502	0.9274	0.8932
Bochum	0.7581	0.6537	0.7865	0.7305	1.0000	1.0000	1.0000	1.0000	0.9683	0.9723	0.9896	0.9767
Freiburg Im	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9811	0.9713	0.9509	0.9677
Augsburg	1.0000	0.8367	0.9435	0.9242	1.0000	1.0000	1.0000	1.0000	0.8995	0.9632	0.9651	0.9421
Bonn	0.6633	0.6919	0.7819	0.7106	0.8301	0.8400	0.8469	0.8390	0.7787	0.9178	0.8705	0.8537
Karlsruhe	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9497	0.9578	0.9199	0.9423
Saarbrücken	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9552	0.9623	0.9684	0.9620
Duisburg	0.6086	0.5840	0.6785	0.6224	0.8940	0.8243	0.8822	0.8663	0.7671	0.9476	0.8029	0.8357
Mannheim	0.7423	0.7250	0.7912	0.7523	0.8817	0.8752	0.8887	0.8818	0.8491	0.9500	0.8795	0.8919
Münster	0.8258	0.9520	0.7924	0.8541	1.0000	1.0000	1.0000	1.0000	0.8025	0.9701	0.7570	0.8384
Aachen	0.7977	0.7637	0.8611	0.8065	1.0000	1.0000	1.0000	1.0000	0.9668	0.9662	0.9577	0.9636

The GOFF test goal, proposed by Ramalho et al. (2011), is to test the specifications of any model, allowing to determine which of the two types of proposed generalizations (type I or II) are really necessary, or among simple FRM, what is the most adequate. GOFF-I is used to introduce asymmetry in the Logit model and GOFF-II complements the asymmetry assumed by the first functional form. Both GOFF-I and II should not be applied to the Log-Log and CLog-Log models respectively, once that the proposed models possess a constant term in index (Ramalho et al., 2011). The P test contrasts by its direct applicability to the complete specification of the two-part models. Results for the one-part models and two-part models (for the first-part and second-part) are showed in Table 3 for both DEA-VRS and SFA-ML eco-efficiency scores, for the sample of German cities. It is observed from this table that the GOFF tests do not reject any of the simpler functional forms. Additionally, almost all specifications are not rejected by the P test, which uses all of the acceptable models as the alternative hypothesis (H1). Ramalho et al. (2010) suggest that when the distribution of the efficiency scores in the sample is clearly asymmetric and the number of unity scores is large, we should opt for a Complementary Log-Log (CLog-Log) specification.

Figure 2. SFA-ML efficiency scores for the three years (2007, 2010 and 2013) and by city



In fact, considering our sample of German cities, and more specifically under the DEA-VRS methodology, the number of outcomes with unity score is large as compared to the number of observations in our sample. Provided this, in the one-part models we choose a Clog-Log specification in order to be able to explain the probability of a German city to produce on the eco-efficiency frontier.

Table 3. Specification tests for one-part and two-part models for the first and second component

	One-part Models				One-part Models			
	Logit		Probit		Log-Log		CLog-Log	
	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML
RESET test	0.013	1.408	0.044	0.771	0.000	1.564	0.119	0.18
GOFF-I test	0.018	2.16	0.043	0.819	0.000	2.086	0.116	0.196
GOFF-II test	0.014	1.823	0.049	0.984	0.000	2.086	0.116	0.196
P Test								
H1: FRM II –Logit			0.052	0.473	0.000	2.544	0.211	0.034
H1: FRM II –Probit	0.004	1.561			0.003	2.204	0.181	0.117
H1: FRM II-Log-Log	0.026	2.049	0.080	0.582			0.261	0.041
H1: FRM II- CLog-Log	0.000	1.944	0.020	0.978	0.012	2.487		
	Two-part Models – First-part				Two-part Model – First-part			
	Logit		Probit		Log-Log		CLog-Log	
	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML
RESET test	0.609	1.188	0.584	0.251	0.484	0.004	0.494	1.785
GOFF-I test	0.586	1.628	0.544	0.409	0.491	0.003	0.397	2.149

GOFF-II test	0.464	1.991	0.599	0.3	0.491	0.003	0.397	2.149
P Test								
H1: FRM II-Logit			0.604	0.011	0.286	0.002	0.576	1.566
H1: FRM II-Probit	0.804	0.001			0.320	0.025	0.810	0.113
H1: FRM II-Log-Log	0.950	1.089	0.793	0.160			1.087	1.790
H1:FRM II- CLog-Log	0.227	2.623	0.326	1.959	0.168	0.003		
	Two-part Models – Second-part				Two-part Models – Second-part			
	Logit		Probit		Log-Log		CLog-Log	
	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML
RESET test	0.517	1.335	0.371	0.763	0.733	1.490	0.120	0.218
GOFF-I test	0.528	1.771	0.366	0.790	0.736	1.784	0.119	0.229
GOFF-II test	0.507	1.568	0.374	0.901	0.730	1.784	0.119	0.229
P Test								
H1: FRM II-Logit			0.310	0.537	0.790	2.078	0.080	0.082
H1: FRM II-Probit	0.515	1.456			0.816	1.931	0.098	0.164
H1: FRM II-Log-Log	0.500	1.686	0.317	0.606			0.070	0.088
H1:FRM II- CLog-Log	0.594	1.731	0.411	0.917	0.879	2.145		

Note: ***, ** and * denote coefficients or test statistics which are significant in 1%, 5% and 10%, respectively (if so, the specification should be rejected). The table presents test statistic values. FRM – fractional regression model. GOFF – generalized goodness-of-functional form test for binary and fractional regression models. P – test: Lagrange Multiplier.

Table 3 also evidences that accordingly to the results of the first-part and second-part in the two-part models, all the specification tests fail to reject any of the estimated models (if test values obtained were significant, the respective specification should be rejected). This suggests that the main issue in the regression analysis using the DEA-VRS and SFA-ML scores is not so much their bounded nature as the existence of a mass-point at unity in both their distributions. Moreover, the GOFF tests applied do not reject any of the simpler functional forms. As such, we move on in the estimations considering the four alternative two-part models: Logit, Probit, Log-Log and Clog-Log, in the following presentation of the second-part estimation.

Table 4 presents the results of the one-part (for efficient cities only) and two-part models, the latter including a first-part (inefficient cities) and a second-part (efficient cities).

The two-part models proceed with estimations in two parts (or components). The first-part considers only inefficient DMUs, i.e. those with a score value below 50%, and the second-part models will only consider the efficient DMUs derived from the first-step models implementation (DEA and SFA). The two-part fractional regression model estimation has the advantage of allowing us to analyze first why some German cities are at the efficient frontier (i.e., second-part) as compared to those that are not and to observe the distance of inefficient German cities to the border (i.e., first-part). This two-part procedure seems to be a better way to show the impact of each covariate on the eco-efficiency of German cities (i.e., DEA and SFA eco-efficiency scores).

Table 4 reports that in the two FRM there is evidence of the negative and consistent effect of the number of days where PM10 exceeded 50 $\mu g/m^3$ in both DEA-VRS and SFA-ML predictions, in the first and second-part models (except in two scenarios). In fact, the coefficients of variable X1 (number of days where PM10 exceeded 50 $\mu g/m^3$) are negative and statistically significant (at different significance levels), which means that variable X1 is related to the DEA-VRS and SFA-ML eco-efficiency scores for the efficient German cities (i.e., second-part) and to the DEA-VRS and SFA-ML scores of the inefficient German cities (i.e., first-part). We may infer from these results that, on average, at different significance levels and using this sample, the higher the PM10 concentrations are, lower will be the eco-efficiency scores within cities, ceteris paribus.

On the other hand, considering only the SFA-ML results, with respect to the impact of variable X2 (average temperature), the coefficients are not statistically significant and with an opposite sign to the theoretically expected in the first-part and in the second-part of the two-part models. However, with the DEA-VRS in the second-part, the coefficients of this variable are statistically significant at 1%.

Table 4. Estimation results for the fractional regression models

	One-part models				Two-part models											
					First-part				Second-part							
	Logit		CLog-Log		Logit		CLog-Log		Logit		Probit		Log-Log		CLog-Log	
	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML	DEA-VRS	SFA-ML
X1	-0.0207	-0.0571	-0.0093	-0.0289	0.0486	-0.0977	0.0527	-0.0804	-0.0225	-0.0343	-0.0124	-0.0206	-0.0206	-0.0285	-0.0100	-0.0197
	(0.065)*	(0.000)***	(0.024)**	(0.000)***	(0.658)	(0.058)*	(0.616)	(0.059)*	(0.034)**	(0.002)***	(0.021)**	(0.002)***	(0.041)**	(0.003)***	(0.012)**	(0.002)***
X2	-0.3852	0.0719	0.1399	0.0328	0.1549	0.2575	0.0981	0.1778	0.3728	-0.0013	0.1859	-0.0004	0.3541	-0.0037	0.1377	-0.0004
	(0.000)***	(0.344)	(0.000)***	(0.386)	(0.753)	(0.232)	(0.831)	(0.302)	(0.000)***	(0.981)	(0.000)***	(0.990)	(0.000)***	(0.935)	(0.000)***	(0.989)
X3	-0.0450	0.0117	-0.0169	0.0053	-0.1513	0.1927	-0.1562	0.0784	-0.0370	-0.0176	-0.0186	-0.0104	-0.0351	-0.0150	-0.0138	-0.0096
	(0.000)***	(0.546)	(0.000)***	(0.595)	(0.092)*	(0.091)*	(0.031)**	(0.085)*	(0.002)***	(0.260)	(0.002)***	(0.255)	(0.003)***	(0.262)	(0.001)***	(0.249)
X4	-0.0161	-0.0337	-0.0057	-0.0155	-0.0992	-0.0028	-0.0879	-0.0033	-0.0134	-0.0350	-0.0064	-0.0204	-0.0128	-0.0302	-0.0045	-0.0183
	(0.269)	(0.057)*	(0.294)	(0.062)*	(0.481)	(0.058)*	(0.510)	(0.931)	(0.346)	(0.004)***	(0.373)	(0.003)***	(0.336)	(0.003)***	(0.399)	(0.004)***
X5	0.0012	0.0006	0.0006	0.0003	0.0048	-0.0028	0.0432	0.0009	0.0011	0.0001	0.0006	0.0000	0.0010	0.0000	0.0005	0.0000
	(0.011)**	(0.232)	(0.003)***	(0.347)	(0.334)	(0.951)	(0.357)	(0.605)	(0.016)**	(0.888)	(0.011)**	(0.872)	(0.016)**	(0.913)	(0.007)***	(0.858)
Constant	-4.2014	0.7166	-1.5798	0.2168	-5.7399	-8.1604	-4.3983	-6.4845	-4.1456	2.6620	-1.9342	0.5823	-3.7218	2.6444	-1.5924	1.1364
	(0.000)***	(0.636)	(0.000)***	(0.769)	(0.574)	(0.058)*	(0.638)	(0.057)*	(0.000)***	(0.010)**	(0.000)***	(0.008)***	(0.000)***	(0.004)***	(0.000)***	(0.036)**
Observations	72	72	72	7	72	72	72	72	72	72	69	54	69	54	69	54
R^2	0.4314	0.2309	0.4417	0.2271	0.2208	0.0939	0.2617	0.0934	0.4284	0.3426	0.4328	0.3437	0.4265	0.3403	0.4376	0.3452

Notes: Dependent variable is the eco-efficiency scores based on both DEA-VRS and SFA-ML techniques; P(>|z|) values in parenthesis. *, **, *** means statistically significant coefficient at 10%, 5%, and 1%, respectively. X1 – number of days where PM10 exceeded 50 μg/m^3; X2 – average temperature; X3 – average concentrated NO_2; X4 – number of days that Ozone concentration exceeds 120 μg/m^3; X5 – rainfall.

Relatively to the estimated coefficients for the variable X3 (average concentrated NO_2), with SFA-ML predictions in the one-part models and in first-part of the two-part models, these are positive and with contrary signs to the expected ones, while in the second-part of the estimation (efficient German cities), in both the DEA-VRS and SFA-ML approaches, the coefficients present correct signs and aligned with the negative sign theoretically expected. However, the coefficients are considered statistically significant in all specifications (Logit, Probit, Log-Log and CLog-Log) only for the DEA approach.

Considering the two-part models, the variable X4 (number of days that Ozone Concentration exceeds 120 $\mu g/m^3$) presents the estimated coefficients with expected signal, which are statistical significant at the 5% significance level for the inefficient German cities (firt-part) with SFA-ML predictions and for the Logit specification, and in the second-part (efficient German cities), again with SFA-ML predictions, are statistical significant at 1% significance level in all the adopted specifications (Logit, Probit, Log-Log and CLog-Log).

Regarding the coefficients associated with the explanatory variable X5 (rainfall), there is no statistical significance (at the usual significance levels) in the first-part (inefficient cities) of the two-part models estimation for both DEA-VRS and SFA-ML approaches. However, in the second-part (efficient cities), and only with the DEA-VRS approach, there is a statistical significance of this coefficient at the 1% significance level in the Probit and CLog-Log specifications and at the 5% level in the Logit and Log-Log specifications.

The values of the R^2 coefficient are higher in the models with the DEA-VRS scores than with the SFA-ML scores, in the same model specification. However, the values are higher in the one-part models (0.4314 and 0.4417 for the Logit and CLog-Log models) than in the first-part of the two-part models (0.2208 and 0.2617 for Logit and CLog-Log, respectively). On the other hand, considering SFA-ML predictions in the two-part models, the values of the R^2 are higher in the second-part (efficient cities) than in the first-part (inefficient cities).

Table 5 reports for each model the partial effects estimated for each covariate, which were computed as the mean of the partial effects computed for each German city in our sample. A comparison of the partial effects implied by the one-part and the two-part models suggest that even the models selected by the specification tests may generate similar results in both DEA-VRS and SFA-ML models.

Table 5. Sample averages of partial effects

DEA-VRS	Two-part Model				Two-part Model			
	Logit (1st part) +		Logit (1st part) +		CLog-Log (1st part) +		CLog-Log (1st part) +	
	Logit	Probit	Log-Log	CLog-Log	Logit	Probit	Log-Log	CLog-Log
X1	-0.0141	-0.0142	-0.0141	-0.0143	-0.0139	-0.0140	-0.0139	-0.0141
X2	0.0332	0.0333	0.0331	0.0333	0.0275	0.0275	0.0273	0.0276
X3	0.0112	0.0112	0.0112	0.0112	0.0114	0.0114	0.0114	0.0113
X4	-0.0019	-0.0019	-0.0019	-0.0019	-0.0010	-0.0010	-0.0010	-0.0010
X5	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001
SFA- ML	Two-part Model				Two-part Model			
	Logit (1st part) +		Logit (1st part) +		CLog-Log (1st part) +		CLog-Log (1st part) +	
	Logit	Probit	Log-Log	CLog-Log	Logit	Probit	Log-Log	CLog-Log
X1	0.0015	0.0015	0.0015	0.0015	0.0018	0.0017	0.0018	0.0017
X2	0.0059	0.0059	0.0059	0.0059	0.0044	0.0044	0.0044	0.0044
X3	-0.0049	-0.0049	-0.0049	-0.0049	-0.0055	-0.0055	-0.0055	-0.0055
X4	-0.0032	-0.0032	-0.0032	-0.0032	-0.0031	-0.0031	-0.0031	-0.0031
X5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002

Notes: Dependent variable is the eco-efficiency scores based on both DEA-VRS and SFA-ML techniques. X1 – number of days where PM10 exceeded 50 μg/m^3; X2 – average temperature; X3 – average concentrated NO_2; X4 – number of days that Ozone concentration exceeds 120 μg/m^3; X5 – rainfall. Values based over partial effects.

5. Policy Implications

Wu et al. (2018a) evaluate the performance efficiency of photovoltaic projects in China and explore its influencing factors through a modified three-phase model. They found that the average annual temperature negatively influences efficiency, a result opposite to ours. Therefore, it seems that for

German efficient cities, areas with high temperature tend to possess high performance efficiency (as in Arminen and Menegaki, 2019), a result which depend upon the model specification adopted. If in Wu et al. (2018a), high ambient temperature leads to operating inefficiency (or to the loss of photovoltaic cells), in German cities it seems to lead to operating eco-efficiency.

Results in the first step indicate the overall top five performing German cities. SFA and DEA point that Aachen, Berlin, Bochum, Freiburg and München are the most efficient cities in each of the three time periods analyzed. In a second step, fractional regression models (FRM) are obtained using both DEA-VRS and SFA-ML predictions as the dependent variable. Results from the FRM estimation using the DEA-VRS scores as dependent variable show, in some model specifications and considering different significance levels, a significant effect of the number of the days that particulate matter PM10 concentration exceeded 50 $\mu g/m^3$, the average temperature, the average concentration of NO_2 and rainfall on eco-efficiency. It was not considered significant, at usual significance levels, the coefficient of the variable number of days the Ozone concentration exceeds 120 $\mu g/m^3$ for all FRM. Using SFA-ML efficiency scores, the FRM results point for a significant impact of the number of the days particulate matter PM10 concentration exceeded 50 $\mu g/m^3$ and the number of days the Ozone concentration exceeds 120 $\mu g/m^3$ on the eco-efficiency of the German cities, while the average temperature, the average concentration of NO_2 and rainfall do not show a significant impact over eco-efficiency. Based on the findings, policy suggestions are proposed to improve the eco-efficiency of German cities' urban characteristics and promote sustainable urban development, requiring the control of urban air pollution and temperature (fighting of emissions (GHG) and effective pollution control).

All the urban pollutants limits included as explanatory factors of German cities eco-efficiency reveal a negative influence over it (favoring the results of Costa et al., 2017; Li et al., 2018, 2018a; Sun et al., 2018; Silva et al., 2018; Qiu et al., 2019, among others). Environmental pollutants are undesirable for a city when we think from the ecological perspective (Yin et al., 2014). Even if eco-efficiency can help us measure a city's sustainability, it still remains unanswered the question on how to improve eco-efficiency. One way will be through less environmental impact (Hahn et al., 2010; Yin et al., 2014) and emission reduction movements. Sueyoshi and Yuan (2015) were the first to use DEA for assessment on China regional performance, by incorporating PM2.5 and PM10 as undesirable outputs. Wu et al. (2018a) applied a two-stage analysis model to analyze eco-efficiency of 58 Chinese coal-fired power plants including SO_2 and NOx as undesirable outputs when accounting for eco-efficiency. However, in the present article we use these pollutants as urban air pollutant factors able to explain eco-efficiency in some German cities, something that represents a novelty for the existent literature.

There is currently a major problem in many cities in Germany, and not only on these, that the limit value of nitrogen dioxide and other pollutants is often exceeded. The analysis in this study of the determinants affecting eco-efficiency in German cities provides a potential indication that particulate emissions considered as determinants of pollution and affecting economic and environmental efficiency levels can be minimized through measures to prevent pollution and emission control, including the choice of "cleaner" fuels for vehicles in and out of urban areas, with incentives for electric and/or hybrid vehicles.

The importance of the responsibility of vehicle trafficking as an endogenous determinant of pollution seems to be consensual, so environmental management measures such as improved process design, operation and maintenance, and other best practice actions can contribute to reducing PM10 emissions. Increasing combustion efficiency significantly reduces the amount of incomplete combustion products, an important component of particulate emissions. Ash reduction through fuel cleaning can also contribute to reducing particulate emissions (Andersson and Johnsson, 2006). Thus, the definition of policies and measures should be directed to anthropogenic sources, with a view to reducing particulate emissions and, consequently, to decreasing their concentration levels in ambient air, focusing in particular on traffic management and control, industry and the domestic sector, major sectors responsible for particulate emissions (Block et al, 2004).

The development of low emission zones in urban centers should be established by including targets as limit values and implementing stricter regulations for the inclusion of all relevant air pollutants and emitters, as defended by Jiang et al. (2017). In addition, for these authors the potential of other measures to prevent, change and improve traffic flows in cities should be fully exploited (e.g. banning Lorries in cities). Another option for many cities is the comprehensive optimization of traffic signal control, which needs investments in the necessary technology for this purpose, but can inhibit particle emission (Jiang et al, 2017). Thus, we consider that a broad spectrum of policy options will be available for implementation, ranging from the mandatory regulatory instruments that municipal-level economic and environmental policy makers apply to it, such as smoke-free, and particulate-free zones. Another regulatory approach might be to set targets that are adjusted over the time of the energy transition, so that new vehicle technologies and even the use of industrial or domestic equipment are required to respond to the increased urban population growth.

Rainfall intensity possesses the positive function of the component surface dust cleaning and the ambient temperature decreasing (Wu et al., 2018b). Precipitation or rainfall can diminish air pollution (Rosenfeld et al., 2007), but policy makers have no control over it, turning it into a relevant independent variable able to explain cities eco-efficiency (Saen, 2005; Arminen and Menegaki, 2019). Results in Table 4 inclusively indicate a positive coefficient of X5 meaning that higher levels of rainfall lead to higher eco-efficiency among already efficient German cities. Both the positive signs for the coefficients of X5 and X2 (although some coefficients are not statistically significant in some specifications) attained within estimations may be justified by the extreme weather conditions reported annually in Germany for the specific years under analysis, since these three years are some of the rainiest ones since 1980, except the years of 2002 and 2017, also with quite high precipitation. These correspond also to years with lowest average annual temperature.

The non-relevance of temperature (considering only the SFA-ML results) does not favor the results of both Deng and Gibson (2018) and Arminen and Menegaki (2019).

Results seem to indicate differences in terms of factors affecting eco-efficiency scores, but which should be taken into account by local policy makers. A lot is being done at this respect by limiting cars circulation in major cities and certain streets as a way to avoid air pollution. As such, more research could lead us to include other variables in the models able to explain eco-efficiency scores and which are directly related to cities. However, to develop policy directions which are more visible and sound, we should extend this work to more European cities and infer about the differences among them. This will enable us to delineate efficiency measures such as the reduction of carbon dioxide concentrations and reduce air pollution within cities. A lot in the literature has emerged recently such as to infer the impact of green measures with regard to cars circulation limitation in cities and it has been proved that these limitations increase wellbeing, reduce noise pollution, improve air quality and reduce human diseases.

Additionally, local policy makers should promote the increase of green areas within cities, provide incentives for the use of public transport and/or car-sharing initiatives, maybe ensure the fulfillment of filters use requirements and impose higher penalties for those factories which violate environmental laws, as well as to think about moving the highest polluting factories to rural areas. As to citizens, we suggest that, and not only in German cities, local policy makers should encourage the reduction of waste production and reward recycling, in order to involve citizens in these good practices and promoting eco-efficiency within cities.

6. Conclusions

This paper analyses a set of selected German cities' performance in terms of the relative behavior of their eco-efficiencies, computed as the ratio of their gross domestic product (GDP) over particulate matter emissions PM10. For this analysis, initially, eco-efficiency scores of the selected cities are computed using DEA and SFA techniques with inputs being the population density (persons/km^2), the labor productivity (measured in US dollars at constant values of 2010), the municipal waste (both domestic and commercial measured in thousand tons), the number of registered cars (per one thousand persons) and the number of companies in the city.

Results from the first stage show that the overall top five performing German cities are Aachen, Berlin, Bochum, Freiburg and München, assuming efficiency predictions from the DEA-VRS and the SFA-ML models, in the three years analyzed (2007, 2010 and 2013). In a second stage, selected fractional regression models (FRM) are discussed considering both DEA-VRS and SFA-ML efficiency scores as dependent variables. The results of the FRM estimation with the scores of DEA-VRS indicate (in most of the specifications) a significant effect of four variables on eco-efficiency: the number of the days that the particulate matter PM10 concentration exceeds 50 $\mu g/m^3$, the average temperature, the average concentration of NO_2 ($\mu g/m^3$) and rainfall (l/m^2). However, the coefficient of the variable measuring the number of the days that Ozone concentration exceeds 120 $\mu g/m^3$ is not significant at usual significance levels, for all fractional regression models. Results of the FRM estimation, according to the SFA-ML scores, in the majority of specifications, show a significant impact of the number of the days particulate matter emissions PM10 concentration exceed 50 $\mu g/m^3$ and the number of days the Ozone concentration exceeds 120 $\mu g/m^3$ on eco-efficiency of German cities, while the other independent variables (the average temperature, the average concentration of NO_2 and rainfall) do not show a relevant impact, except under the DEA-VRS models. The positive effects of rainfall and average temperature revealed in some specifications can be associated to the highly rainy years analyzed and to lower average annual temperatures verified in these years, among all the explored cities (information source: DWD - Deutscher Wetterdienst).

Results attained in this article are useful to both policy makers and decision makers within cities in Germany since they highlight the negative effect of urban air pollutants over eco-efficiency and the positive impact that rainfall (cleans dust and decreases ambient temperature) exerts over efficiency. Weather is non controllable and managers should thus control for ambient fresh air and try to improve it based upon weather (precipitation) predictions, which will be hard provided the ozone concentration pollution levels reported nowadays. For future research and for readers' convenience, a discussion about interpretation on what the environmental variables are assumed to affect should be accomplished. As mentioned by Simar and Wilson (2011), "one should carefully consider what restrictions are necessary, and whether these are reasonable. Ideally, restrictions should be tested. In addition, one should carefully consider how valid inference can be made."

Acknowledgements

This work was financially supported by the research unit on Governance, Competitiveness and Public Policy (UID/CPO/04058/2019), funded by the Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia) and by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia), references UIDB/04106/2020 and UIDP/04106/2020. Any persistent error or missing's are the authors' entire responsibility. We thank the reviewers and editor for all valuable suggestions that helped us to improve the current version of this article.

Acknowledgements

To be filled afterwards. In the title page.

References

Addanki, S. C. and Venkataraman, H. (2017). Greening the economy: A review of urban sustainability measures for developing new cities. Sustainable Cities and Society, 32, 1-8. https://doi.org/10.1016/j.scs.2017.03.009.

Aigner, D., Lovell, C. A. K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. Journal of Econometrics, 6, 1, 21-37.

Alpert, P., Radian, R., Halfon, N. and Levin, Z. (2019). Urban Rainfall Anomaly under Intensive Development, 1949–2018, Case of Tel-Aviv, Israel. Atmosphere, 10, 4, 163-175. https://doi.org/10.3390/atmos10040163

Ameen, R. F. M. and Mourshed, M. (2019). Urban sustainability assessment framework development: The ranking and weighting of sustainability indicators using analytic hierarchy process. Sustainable Cities and Society, 44, 356-366. https://doi.org/10.1016/j.scs.2018.10.020.

Andersen, P. and Petersen, N. C. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. Management Science, 39, 10, 1261-1264.

Andersson, K. and Johnsson, F. (2006). Process evaluation of an 865MWe lignite fired O2/CO2 power plant. Energy Conversion and Management, 47, 18–19, 3487-3498. https://doi.org/10.1016/j.enconman.2005.10.017.

Arminen, H., Menegaki, A. N. (2019). Corruption, climate and the energy-environment-growth nexus. Energy Economics, 80, 621-634. https://doi.org/10.1016/j.eneco.2019.02.009.

Banker, R. D. and Datar, S. M. (1989a). Analysis of Cost Variances for Management Control in Hospitals. JAI Press, 5, 268-291.

Banker, R. D. and Datar, S. M. (1989b). Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation. Journal of Accounting Research, 27, 1, 21-39. http://www.jstor.org/stable/2491205

Banker, R. D. and Gifford, J. L. (1988). A relative efficiency model for the evaluation of public health nurse productivity. Pittsburgh: Mimeo, Carnegie Mellon University.

Banker, R. D., Charnes, A. and Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science, 30, 9, 1078-1092.

Battese, G.E. and Corra, G. S. (1977). Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. Australian Journal of Agricultural and Resource Economics, 21, 3, 169-179.

Beisner, E. C. (2019). The ‘Carbon Emissions’ Deception: No, Carbon Dioxide Is Not a Pollutant. CNS News. Comment, June 26, 2019, 3:03 PM EDT. https://www.cnsnews.com/commentary/e-calvin-beisner/carbon-emissions-deception-no-carbon-dioxide-not-pollutant

Bian, J.,Zhang, Y., Shuai, C., Shen, L., Ren,H. and Wang, Y. (2019). Have cities effectively improved ecological well-being performance? Empirical analysis of 278 Chinese cities. Journal of Cleaner Production, 118913. https://doi.org/10.1016/j.jclepro.2019.118913.

Block, M.L., Wu, X., Pei, Z., Li, G., Wang, T., Qin, L., Wilson, B., Yang, J., Hong, J.S. and Veronesi, B. (2004). Nanometer size diesel exhaust particles are selectively toxic to dopaminergic neurons: the role of microglia, phagocytosis, and NADPH oxidase. The FASEB Journal, 1-22. Express article 10.1096/fj.04-1945fje.

Bundesverwaltungsgericht Luftreinhaltepläne Düsseldorf und. 2018. URL: Diesel-Verkehrsverbote ausnahmsweise möglich [WWW Document]. Press. Nr. 9/2018, Stuttgart. https://www.bverwg.de/pm/2018/9.

Carnevale, C., Finzi, G., Pederzoli, A., Turrini, E., Volta, M., Guariso, G., Gianfreda, R., Maffeis, M., Pisoni, E., Thunis, P., Markl-Hummel, L., Blond, N., Clappier, A., Dujardin, V., Weber, C. and Perron, G. (2014). Exploring trade-offs between air pollutants through an integrated assessment model. Science of the Total Environment, 481, 7-16.

Carnevale, C., Finzi, G., Pisoni, E., Singh, V. and Volta, M. (2011). An integrated air quality forecast system for a metropolitan area. Journal of Environmental Monitoring, 13, 3437-3447.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units, European Journal of Operational Research, 2, 429-444. https://www.utdallas.edu/~ryoung/phdseminar/CCR1978.pdf

Chen, N. and Xu, L. (2017). Relationship between air quality and economic development in the provincial capital cities of China. Environmental Science and Pollution Research, 24, 3, 2928–2935. https://doi.org/10.1007/s11356-016-8065-3.

Cho, H.-S. and Choi, M. J. (2014). Effects of Compact Urban Development on Air Pollution: Empirical Evidence from Korea. Sustainability, 6, 5968-5982. doi:10.3390/su6095968

Costa, L.G., Cole, T.B., Coburn, J., et al. (2017). Neurotoxicity of traffic-related air pollution. NeuroToxicology, 59, 133-139

Davidson, R. and MacKinnon, J. G. (1981). Several Tests for Model Specification in the Presence of Alternative Hypotheses. Econometrica, 49, 3, 781-793. DOI: 10.2307/1911522.

Deng, X. and Gibson, J. (2019). Improving eco-efficiency for the sustainable agricultural production: A case study in Shandong, China. Technological Forecasting and Social Change, 144, 394-400. https://doi.org/10.1016/j.techfore.2018.01.027.

EEA. (2016). Air quality in Europe — 2016 report, EEA technical report. 10.2800/413142.

EEA. (2017). Air quality in Europe — 2017 report, EEA technical report. 10.2800/358908.

Hahn, T., Figge, F., Liesen, A. and Barkemeyer, R. (2010). Opportunity cost based analysis of corporate eco-efficiency: a methodology and its application to the CO2-efficiency of German companies. Journal of Environmental Management, 91, 10, 1997-2007.

Han, L., Zhou, W., Li, W. and Li, L. (2014). Impact of urbanization level on urban air quality: A case of fine particles (PM2.5) in Chinese cities. Environmental Pollution, 194, 163-170. https://doi.org/10.1016/j.envpol.2014.07.022.

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

Jiang, W., Boltze, M., Groer, S. and Scheuvens, D. (2017). Impacts of low emission zones in Germany on air pollution levels. Transportation Research Procedia, 25C, 3374–3386. 10.1016/j.trpro.2017.05.217.

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

Li, Y., Chiu, Y.-h. and Lu, L. C. (2018). Energy and AQI performance of 31 cities in China. Energy Policy, 122, 194-202. https://doi.org/10.1016/j.enpol.2018.07.037.

Li, H., Zhang, J., Wang, C., Wang, Y. and Coffey, V. (2018a). An evaluation of the impact of environmental regulation on the efficiency of technology innovation using the combined DEA model: A case study of Xi'an, China. Sustainable Cities and Society, 42, 355-369. https://doi.org/10.1016/j.scs.2018.07.001.

Lin, B. and Zhu, J. (2018). Changes in urban air quality during urbanization in China. Journal of Cleaner Production, 188, 312-321. https://doi.org/10.1016/j.jclepro.2018.03.293.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. International Economic Review, 18, 2, 435-444.

Moutinho, V., Madaleno, M., Robaina, M. and Villar, J. (2018). Advanced scoring method of eco-efficiency in European cities. Environmental Science and Pollution Research, 25, 2, 1637-1654.

Papke, L.E. and Wooldridge, J. M. (1996). Econometric Methods for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates. Journal of Applied Econometrics, 11, 619-632.

Piña, W. H. A. and Martínez, C. I. P. (2016). Development and Urban Sustainability: An Analysis of Efficiency Using Data Envelopment Analysis. Sustainability, 8(2), 148-163. https://doi.org/10.3390/su8020148.

Pisoni, E., Christidis, P., Thunis, P. and Trombetti, M. (2019). Evaluating the impact of "Sustainable Urban Mobility Plans" on urban background air quality. Journal of Environmental Management, 231, 249-255.

Qiu, G., Song, R. and He, S. (2019). The aggravation of urban air quality deterioration due to urbanization, transportation and economic development – Panel models with marginal effect analyses across China. Science of The Total Environment, 651, Part 1, 1114-1125. https://doi.org/10.1016/j.scitotenv.2018.09.219.

Ramalho, E.A., Ramalho, J. J. S. and Henriques, P. D. (2010). Fractional regression models for second stage DEA efficiency analyses. Journal of Productivity Analysis, 34, 3, 239–255.

Ramalho, E.A., Ramalho, J.J.S. and Murteira, J.M.R. (2011). Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models. Journal of Economic Surveys, 25, 1, 19–68. https://doi.org/10.1111/j.1467-6419.2009.00602.x

Rosenfeld, D., Dai, J., Yu, X., Yao, Z., Xu, X., Yang, X. and Du, C. Inverse relations between amounts of air pollution and orographic precipitation, Science, 315, 5817, 1396-1398.

Saen, R. F. (2005). Developing a nondiscretionary model of slacks-based measure in data envelopment analysis. Applied Mathematical Computation, 169, 2, 1440-1447.

Schmitz, S., Weiand, L., Becker, S., Niehoff, N., Schwartzbach, F. and von Schneidemesser, E. (2018). An assessment of perceptions of air quality surrounding the implementation of a traffic-reduction measure in a local urban environment. Sustainable Cities and Society, 41, 525-537.

Silva, B. N., Khan, M. and Han, K. (2018). Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. Sustainable Cities and Society, 38, 697-713. https://doi.org/10.1016/j.scs.2018.01.053.

Simar, L. and Wilson, P.W. (2011). Two-stage DEA: caveat emptor. Journal of Productivity Analysis, 36, 2, 205-218. https://doi.org/10.1007/s11123-011-0230-6

Sueyoshi, T. and Yuan, Y. (2015). China's regional sustainability and diversified resource allocation: DEA environmental assessment on economic development and air pollution. Energy Economics, 49, 239-256. https://doi.org/10.1016/j.eneco.2015.01.024.

Sun, C., Luo, Y. and Li, J. (2018). Urban traffic infrastructure investment and air pollution: Evidence from the 83 cities in China. Journal of Cleaner Production, 172, 488-496. https://doi.org/10.1016/j.jclepro.2017.10.194.

UN – United Nations. (2009). Eco-efficiency Indicators: Measuring Resource-use Efficiency and the Impact of Economic Activities on the Environment. Greening of Economic Growth Series. https://sustainabledevelopment.un.org/content/documents/785eco.pdf. Accessed on 11 November 2018.

Wallace III, J. P., Christy, J. R. and d'Aleo, J. S. (2017). On the Existence of a "Tropical Hot Spot" & The Validity of EPA's CO2 Endangerment Finding. Abridged Research Report. Second Edition, April 2017. https://thsresearch.files.wordpress.com/2017/04/ef-data-research-report-second-editionfinal041717-1.pdf

Wang, Q. and Yuan, B. (2018). Air pollution control intensity and ecological total-factor energy efficiency: The moderating effect of ownership structure. Journal of Cleaner Production, 186, 373-387. https://doi.org/10.1016/j.jclepro.2018.03.106.

WHO. (2006). WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide - Global update 2005. http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/pre2009/air-quality-guidelines.-global-update-2005.-particulate-matter,-ozone,-nitrogen-dioxide-and-sulfur-dioxide.

Wu, Y., Ke, Y., Xu, C., Xiao, X. and Hu, Y. (2018a). Eco-efficiency measurement of coal-fired power plants in China using super efficiency data envelopment analysis. Sustainable Cities and Society, 36, 157-168. https://doi.org/10.1016/j.scs.2017.10.011.

Wu, Y., Ke, Y., Zhang, T., Liu, F. and Wang, J. (2018b). Performance efficiency assessment of photovoltaic poverty alleviation projects in China: A three-phase data envelopment analysis model. Energy, 159, 599-610. https://doi.org/10.1016/j.energy.2018.06.187.

Yang, Li, Ma, C., Yang, Y., Zhang, E. and Lv, H. (2020). Estimating the regional eco-efficiency in China based on bootstrapping by-production technologies. Journal of Cleaner Production, 243, 118550, https://doi.org/10.1016/j.jclepro.2019.118550.

Yadong, Y. (2013). Eco-efficiency trends in china, 1978-2010: decoupling environmental pressure from economic growth. Ecological Indicators, 24, 177–184.

Yin, K., Wang, R., An, Q., Yao, L. and Liang, J. (2014). Using eco-efficiency as an indicator for sustainable urban development: A case study of Chinese provincial capital cities. Ecological Indicators, 36, 665-671. https://doi.org/10.1016/j.ecolind.2013.09.003.

Yu, Y., Huang, J. and Zhang, N. (2019). Modeling the eco-efficiency of Chinese prefecture-level cities with regional heterogeneities: A comparative perspective. Ecological Modelling, 402, 1-17, https://doi.org/10.1016/j.ecolmodel.2019.03.012.

Zhou, C., Shi, C., Wang, S. and Zhang, G. (2018). Estimation of eco-efficiency and its influencing factors in Guangdong province based on Super-SBM and panel regression models. Ecological Indicators, 86, 67-80. https://doi.org/10.1016/j.ecolind.2017.12.011.

Zhu, X., Zhang, P., Wei, Y., Li, Y. and Zhao, H. (2019). Measuring the efficiency and driving factors of urban land use based on the DEA method and the PLS-SEM model—A case study of 35 large and medium-sized cities in China. Sustainable Cities and Society, 50, 101646. https://doi.org/10.1016/j.scs.2019.101646.