Darmstadt Discussion Papers in ECONOMICS

TECHNISCHE UNIVERSITÄT DARMSTADT

Technical Efficiency of Automobiles – A Nonparametric Approach Incorporating Carbon Dioxide Emissions

Benjamin Hampf and Jens J. Krüger

Nr. 198

Arbeitspapiere der
Volkswirtschaftlichen Fachgebiete der TU Darmstadt

ISSN: 1438-2733

Technical Efficiency of Automobiles - A Nonparametric Approach Incorporating Carbon Dioxide Emissions

by

Benjamin Hampf*

and

Jens J. Krüger

Darmstadt University of Technology

June 2010

Abstract

We conduct an empirical analysis of the technical efficiency of cars sold in Germany in 2010. The analysis is performed using traditional data envelopment analysis (DEA) as well as directional distance functions (DDF). The approach of DDF allows incorporating the reduction of carbon dioxide emissions as an environmental goal in the efficiency analysis. A frontier separation approach is used to gain deeper insight for different car classes and regions of origin. Natural gas driven cars and sports-utility-vehicles are also treated as different groups. The results show that the efficiency measurement is significantly influenced by the incorporation of carbon dioxide emissions. Moreover, we find that there is indeed a trade-off between technological performance and environmental performance.

JEL classification: C14, L62, Q53

Keywords: nonparametric efficiency measurement, directional distance function, automobiles, air pollution

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

1 Introduction

In recent years environmental issues like global warming have gained public attention leading to a thrust in political activities to set environmental goals. Sustaining leadership in environmental standards, the European Commission started a program to limit average emissions of carbon dioxide of new cars to 130 g/km (Commission of the European Communities (2007)). The view of environmental performance is therefore based on a one-dimensional aspect that does not account for any relating performance indicators of cars (e.g. top speed, size etc.).

In this paper we conduct an analysis of the efficiency of automobiles that incorporates several inputs (e.g. fuel consumption, price etc.) and outputs (e.g. top speed, acceleration etc.) simultaneously. Our approach is in line with previous studies that evaluate car efficiency including Papahristodoulou (1997), Cantner et al. (2010) and Oh et al. (2010) but differs in two important aspects. First, while these previous studies view efficiency of cars solely from the perspectives of consumers we take a more technical perspective. Second, we estimate the efficiency including carbon dioxide emissions as undesirable outputs as proposed by Färe and Grosskopf (1983). This allows us to measure efficiency taking account of both technical as well as environmental targets and thereby obtaining a more differentiated view of the efficiency of cars.

To estimate the efficiency and hence revealing potentials for enhancements in environmental efficiency we use two approaches that by their nonparametric nature rely on very weak assumptions regarding the production process (in application with cars: transportation process). We start with using data envelopment analysis (DEA) developed by Charnes et al. (1978) to estimate the technical efficiency of cars ignoring emissions and proceed by employing a much more general approach (that includes DEA as a special case), the directional distance function (DDF) proposed by Chung et al. (1997). The DDF is applied to measure technical efficiency when emissions are included as an additional restriction and for the interesting case when they are simultaneously reduced while the desirable outputs are increased. This allows us to explore the trade-off between technical and environmental performance. We use these measures to asses how large the potential efficiency improvements can be and how they change when the direction of the efficiency measurement is amended.

To gain more insights we divide the cars in our dataset (which has been collected from an online database of the major German automobile club (ADAC)) into different broad groups (e.g. car classes, regions of origin etc.) and use the frontier separation approach first introduced by Charnes et al. (1981) to determine whether they show significant differences in their relative performance. These efficiency differences are usually assessed using nonparametric tests such as the Wilcoxon rank sum (or Mann-Whitney) test or other rank-based extensions. The application of rank-based tests in this context is criticized by Simpson in a couple of papers (Simpson (2005), Simpson (2007)). To counter this criticism we apply testing procedures for the assessment of stochastic dominance relations between the efficiency distributions of cars within different classes or regions. Stochastic dominance plays an important role in risk and decision theory but is of general applicability for one-sided comparisons of random variables such as our efficiency measures. Testing for stochastic dominance has the general advantage of being able to compare two random variables by considering their entire sampling distributions instead of just relying on specifically selected moments.

This paper is structured as follows. Section 2 presents the theoretical foundations of the nonparametric efficiency analysis ignoring and incorporating the carbon dioxide emissions. An overview of the theory of the tests for stochastic dominance is given in section 3. Section 4 introduces the data used in our analysis and in section 5 the results of the analysis are presented and discussed. Section 6 concludes the paper.

2 Theory

In this section we present the theoretical foundations and the methodology for our efficiency analysis. Our efficiency analysis incorporating carbon dioxide emissions is an extension of a standard nonparametric efficiency analysis. Thus, we use this model as a starting point. We start by introducing a formalization of the technology of the production process and an efficiency measure ignoring the emissions and then extend the model to obtain an efficiency measure that includes carbon dioxide emissions as an undesirable output.

To model the production process consider n decision making units (DMU, e.g. cars) that are using m inputs $x \in \mathbb{R}_+^m$ to produce s outputs $y \in \mathbb{R}_+^s$. The technology set of this process is the set of all attainable points (x,y) and can be stated as

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

This technology can be equivalently defined by its output sets

$$P(x) = \{y \in \mathbb{R}_+^s : x \text{ can produce } y\} = \{y \in \mathbb{R}_+^s : (x, y) \in T\}$$

which comprises all output vectors y that can be produced from a given input vector x. These output sets are assumed to satisfy the following axioms (see Färe and Primont (1995)):

1. Inactivity: $\forall x \in \mathbb{R}_+^m, 0 \in P(x)$.
   It is possible for any amount of inputs to produce no output.

2. No free lunch: $y \notin P(0)$ if $y \geq 0$.
   It is not possible to produce positive amounts of any output without using positive amounts of at least one input.[1]

3. Strong disposability of inputs: If $y \in P(x)$ and $x' \geq x$ then $y \in P(x')$.
   For a given combination (x,y) the same amount of output is attainable by using more inputs.

4. Strong disposability of outputs: If $y \in P(x)$ and $y' \leq y$ then $y' \in P(x)$.
   For a given combination of (x,y) it is possible to produce less output holding x constant.

5. Convexity: $P(x)$ is convex.
   Convex combinations of observations are possible, e.g. if y_1 and $y_2 \in P(x)$ then $\alpha y_1 + (1 - \alpha)y_2 \in P(x)$ $\forall \alpha \in [0, 1]$.

6. Boundedness: $\forall x \in \mathbb{R}_+^m, P(x)$ is a bounded set.
   It is not possible to produce infinite amounts of outputs with a given level of inputs.

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

7. Closeness: $\forall x \in \mathbb{R}_{+}^{m}, P(x)$ is a closed set.
   A technical assumption without an intuitive economic interpretation.

For evaluating the efficiency of a DMU the upper boundary of the output set is of special interest since it contains the maximum achievable output combinations for a given level of inputs. For a given input vector $x \in \mathbb{R}_{+}^{m}$ the output-oriented boundary is defined as:

$$\partial P(x) = \{y \in \mathbb{R}_{+}^{s} : y \in P(x), \delta y \notin P(x), \forall \delta > 1\}.$$

This is the so-called frontier of the output set. To evaluate the efficiency of a DMU with this model we use the Farrell-Debreu measure of output efficiency (Debreu (1951), Farrell (1957)). This measure projects the DMU under evaluation on the frontier of the output set by proportionally increasing all outputs holding input levels constant and is defined as:

$$\theta^{*} = sup\{\theta : \theta y \in P(x)\}$$

The resulting efficiency measure θ is the factor by which the DMU under evaluation has to increase the volume of all its outputs given the input vector and the technology. A DMU is part of frontier and classified as efficient if $\theta^{*} = 1$. It is part of the interior of the technology and hence inefficient if $\theta^{*} > 1$. Note that the Farrell-Debreu efficiency measure is the inverse of an analogous definition of the output distance function proposed by Shephard (1970).
To estimate the technology and obtain the efficiency measure defined above we use data envelopment analysis (DEA), a nonparametric approach developed by Charnes et al. (1978) that has been applied in many case studies of efficiency analysis (see e.g. Zhou et al. (2008a) for a survey on environmental DEA research). DEA takes the observed combinations (x_i, y_i) for all sample items $i = 1, ..., n$ and the assumptions explained above and generates a piecewise linear frontier function for the technology

$$T = \{(x, y) \in \mathbb{R}_{+}^{m+s} : x \geqq X\lambda, y \leqq Y\lambda, \lambda \geq 0\}$$

with output sets

$$P(x) = \{y \in \mathbb{R}_{+}^{s} : x \geqq X\lambda, y \leqq Y\lambda, \lambda \geq 0\}$$

where X represents the $m \times n$ matrix of inputs and Y represents the $s \times n$ matrix of outputs. λ denotes a $n \times 1$ vector of weight factors with λ positive but otherwise unrestricted implying constant returns to scale of the production process. To compute the Farrell-Debreu efficiency measure for DMU i with input-output combination (x_i, y_i) in this model one has to solve the following linear programming problem:

$$\begin{aligned} \max_{\theta,\lambda} \;& \theta \\ s.t. \quad & x_i \geqq X\lambda \\ & \theta y_i \leqq Y\lambda \\ & \lambda \geq 0 \end{aligned}$$

This linear programming problem is well-behaved and can be easily solved with the conventional simplex algorithm. In our analysis below we use this standard DEA model to evaluate the output efficiency of automobiles for the special case in which emissions are ignored.

In a second model we take into account that in the course of the production process characterized above the DMUs produce k pollutants $u \in \mathbb{R}^k_+$, e.g. carbon dioxide. Different ways to incorporate these emissions have been developed (see e.g. Scheel (2001) for a summary). These can be roughly divided into methods transforming the data and methods transforming the technology. In this paper we use an approach proposed by Färe and Grosskopf (1983) that is based on transforming the technology and treating emissions as undesirable outputs.[2] The new technology contains all attainable points (x, y, u) and can be described by

$$T = \{(x, y, u) \in \mathbb{R}^{m+s+k}_+ : x \text{ can produce } (y, u)\}$$

with output sets

$$P(x) = \{(y, u) \in \mathbb{R}^{s+k}_+ : x \text{ can produce } (y, u)\} = \{(y, u) \in \mathbb{R}^{s+k}_+ : (x, y, u) \in T\}.$$

In order to model u as undesirable outputs we require additional assumptions regarding the output sets. First, while continuing to assume the desirable outputs y to be strongly disposable we assume u to be only weakly disposable:[3]

8. Weak disposability of undesirable outputs:
   If $(y, u) \in P(x)$ and $\gamma u \leq u$ with $0 \leq \gamma \leq 1$ then $(\gamma y, \gamma u) \in P(x)$.

Given the weak disposability assumption it is only possible to produce less of the undesirable outputs if the amount of desirable outputs is decreased simultaneously. Therefore, the reduction of emissions is costly and the car owner is faced with a trade-off between speed and other performance characteristics of a car and lower emissions. Second, together with the assumption of weak disposability of outputs, we assume the output sets to be null-joint:

9. Null-jointness: If $(y, u) \in P(x)$ and $u = 0$ then $y = 0$.

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

[2] Note that simply incorporating the emissions as inputs, as it is often done in environmental economics, leads to a technology that is not bounded anymore (see Färe and Grosskopf (2003)).

[3] The concept of weak disposability was first introduced by Shephard (1970) and Shephard (1974).

The technology considering emissions as undesirable outputs can be rephrased as

$$T = \{(x, y, u) \in \mathbb{R}_+^{m+s+k} : x \geqq X\lambda, y \leqq Y\lambda, u = U\lambda, \lambda \geq 0\}$$

with output sets

$$P(x) = \{(y, u) \in \mathbb{R}_+^{s+k} : x \geqq X\lambda, y \leqq Y\lambda, u = U\lambda, \lambda \geq 0\}$$

where U is the $k \times n$ matrix of undesirable outputs with the equality constraint indicating weak disposability.[4]

To illustrate the difference between the strong disposability and the weak disposability approach, figure 1 shows a strong disposable output set with two desirable outputs (y_1, y_2) while figure 2 shows a weak disposable output set with a single desirable output (y) and a single undesirable output (u). Both output sets are generated by 3 DMUs (A,B,C). The inputs are assumed to be constant and are suppressed in the figures. Assuming strong disposability of both desirable outputs, the boundary of the output set is given by B'BB" and the output set itself is in the south-west direction of this boundary in the positive orthant. Under weak disposability of the undesirable output the boundary is given by 0ABB". Here any element (y, u) can be proportionally reduced and alway stays in the output set.

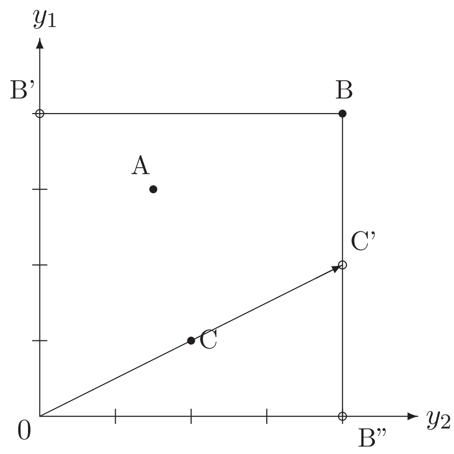

Figure 1: Strong disposable output set

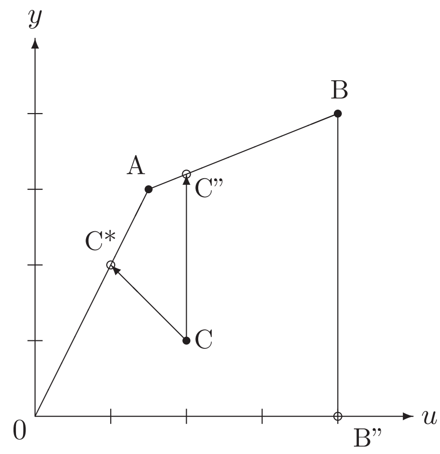

Figure 2: Weak disposable output set

For the efficiency analysis in an environmental context where emissions are modelled as undesirable outputs different measures have been developed (see e.g. Liu et al. (2010) for a survey). In our analysis we use the concept of the directional distance function (DDF) proposed by Chung et al. (1997) because its flexibility allows to measure technical together with environmental efficiency. DEA can be viewed as a special case of DDF. Therefor, DDF is also a nonparametric approach. It introduces a vector g that defines the direction

[4]Assuming constant returns to scale allows to set the scaling factor γ to 1 (see Färe and Grosskopf (2003)).

of the efficiency measurement. The resulting efficiency measure is then defined as

$$\beta^* = sup\{\beta : (y,u) + \beta g \in P(x)\}$$

where g is the directional vector and β denotes the measure of efficiency. We use the directional vectors $g_O = (y, 0)$ and $g_{UO} = (y, -u)$ to compute the efficiency of automobiles. Using g_O we measure pure output efficiency while respecting the environmental constraint, while g_{OU} measures both the technical and the environmental efficiency by the possibility of simultanously increasing the desirable and decreasing the undesirable outputs in order to reach the boundary of the output set. In this second case we use a vector specification that leads to a proportional increase of the desirable outputs and decrease of the undesirable outputs. Hence, we treat technical and environmental efficiency targets as equally important. The related efficiency measures (β_O, β_{UO}) can be obtained by solving the following two linear programming problems:

$$\begin{aligned} \max_{\beta_O,\lambda} \quad & \beta_O \\ \text{s.t.} \quad & x \geqq X\lambda \\ (1+\beta_O)y &\leqq Y\lambda \\ u &= U\lambda \\ \lambda &\geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max_{\beta_{OU},\lambda} \quad & \beta_{UO} \\ \text{s.t.} \quad & x \geqq X\lambda \\ (1+\beta_{UO})y &\leqq Y\lambda \\ (1-\beta_{UO})u &= U\lambda \\ \lambda &\geq 0 \end{aligned}$$

Again, the solution is a straightforward application of the simplex algorithm. The first model is equivalent to the output-oriented DEA with an additional environmental constraint. The second model combines the efficiency measurement with regard to both the desirable outputs and the undesirable outputs. Note that in the second case the resulting β_{UO} is implicitly bounded in the interval [0,1] since u can not get negative. To illustrate the three approaches figures 1 and 2 above show the different projections of DMU C on the boundaries. Ignoring the carbon dioxide emissions and assuming strong disposability of all desirable outputs, the reference point for DMU C is C' (see figure 1) and the efficiency of C is measured by θ. Assuming weak disposability of the undesirable output (see figure 2) and measuring the efficiency with direction g_O the reference point is C", while C* is the reference point using direction g_{UO}.
In our subsequent analysis we compare the three measures $(\theta, 1+\beta_O, 1+\beta_{UO})$ to analyse in which way the different approaches influence the results of the efficiency analysis and the ranking of the cars. In addition, we determine different groups of cars and investigate how their within and between performance changes if measure of the efficiency is changed.

3 Stochastic Dominance

Efficiency differences are usually assessed using nonparametric tests such as the Wilcoxon rank sum (or Mann-Whitney) test or other rank-based measures. The application of rank-based tests in this context is criticized by Simpson in a couple of papers (Simpson (2005), Simpson (2007)). To counter this criticism we apply testing procedures for the assessment of stochastic dominance relations between the efficiency distributions of cars within different classes or regions. Stochastic dominance plays an important role in risk and decision theory but is of general applicability for one-sided comparisons of random variables. Testing for stochastic dominance has the general advantage of being able to compare and rank two random variables by considering their entire distributions and not just relying on a few specifically selected moments.

Several notions of stochastic dominance can be distinguished (see Levy (1992)). Here, we rely on the concepts of first-order stochastic dominance (FSD) and second-order stochastic dominance (SSD). According to first-order stochastic dominance a (real-valued) random variable X stochastically dominates another random variable Y if the cumulative distribution function (CDF) of X is *completely below* that of Y over the whole support. Formally, $X \succ_{FSD} Y$ if $F_X(z) \leq F_Y(z)$ at all points z in the common support of X and Y with strict inequality for some z. Second-order stochastic dominance is less strict in relying on the area below the CDF of X (up to a certain upper bound t) being smaller than the area below the CDF of Y (up to the same t). If this requirement is satisfied for all t in the common support of X and Y then we say that X second-order stochastically dominates Y. Formally, $X \succ_{SSD} Y$ if $\int_{-\infty}^{t} F_X(z)dz \leq \int_{-\infty}^{t} F_Y(z)dz$ for all $t \in \mathbb{R}$ and strict inequality for some t. From these definitions it is clear that FSD implies SSD but not vice versa. Since FSD is a very demanding concept it is important to have a less demanding one such as SSD.[5]

For taking these concepts to the data, we rely on statistical measures proposed by Schmid and Trede (2000) as descriptive devices and use bootstrapping for computing the respective distributions and p-values. We suppose to have two samples of N observations for X (i.e. $(x_1, ..., x_N)$) and M observations for Y (i.e. $(y_1, ..., y_M)$). In our specific application these are the efficiency measures of the cars pertaining to different classes or regions or are computed by the different methods explained above. Testing FSD requires estimating the CDFs by their corresponding empirical distribution functions (EDF)

$$\hat{F}_X(z) = N^{-1} \sum_{i=1}^{N} I(x_i \leq z) \text{ and } \hat{F}_Y(z) = M^{-1} \sum_{i=1}^{M} I(y_i \leq z)$$

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

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

$$D_{FSD} = max_z\{\hat{F}_X(z) - \hat{F}_Y(z)\}$$

to test the null hypothesis $H_0 : X \succeq_{FSD} Y$ against the alternative $H_1 : Y \succ_{FSD} X$. In our implementation this statistic is evaluated over an equally spaced grid of 100 points for z spanning the whole range of all observations. The null hypothesis is rejected in favor of the alternative for large values of the test statistic. We compute the p-values from the finite sample distribution approximated by $B = 10000$ bootstrap replications.[6] The application of the bootstrap here amounts to resample B times with replacement under the null hypothesis, i.e. to resample from the joint sample $(x_1, ..., x_n, y_1, ..., y_m)$, and then to compute the test statistic repeatedly, resulting in $D^*_{FSD,1}, ..., D^*_{FSD,B}$. The p-value is subsequently computed as the fraction of the bootstrap statistics exceeding the statistic $\hat{D}_{FSD}$ computed from the original samples, i.e. by the formula

$$p = \frac{1 + \sum_{b=1}^{B} I(D^*_{FSD,b} \geq D_{FSD})}{B + 1}.$$

Applying this ordinary bootstrap approach is perfectly valid in the present cross-section setting but would be problematic in a panel or time series data context.

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

$$\int_{-\infty}^{t} F_X(z)dz = \int_{-\infty}^{t} (t - z)dF_X(z) \text{ and } \int_{-\infty}^{t} F_Y(z)dz = \int_{-\infty}^{t} (t - z)dF_Y(z)$$

with the empirical analogs

$$\hat{G}_X(z) = N^{-1} \sum_{i=1}^{N} (z - x_i)I(x_i \leq z) \text{ and } \hat{G}_Y(z) = M^{-1} \sum_{i=1}^{M} (z - y_i)I(y_i \leq z).$$

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

$$D_{SSD} = max_z\{\hat{G}_X(z) - \hat{G}_Y(z)\}$$

and reject the null hypothesis for large values of the test statistic. Again the p-values reported below are based on 10000 bootstrap replications and are computed according the procedure outlined above.

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

4 Data

In this section we give a short overview of the data used in our study. The data for technical characteristics of the cars as well as the emissions of carbon dioxide are collected from the "ADAC Autokatalog 2010" (ADAC (2010)) which is an continuosly updated database published online by the German automobile club ADAC. It contains data for 55 car producers with 403 product lines and 9686 model variants that are sold in Germany.[7] Since many of these variants only differ in aspects that are not relevant for our analysis (e.g. optional equipment like airbags) we eliminate these duplets resulting in 3961 remaining observations. These cars have been divided into 7 vehicle classes according to some of their characteristics (e.g size, price etc., see ADAC (2009)).[8] Since some of these classes only contain very few observations we aggregate them to 3 main classes: compact class, middle class and upper class. In addition to dividing the cars into classes we also take the possibility to divide them into regions of origin (Europe, Asia and United States).

To determine which inputs and outputs we use in our analysis we consider a highly simplified conception of a car. We assume that the car is bought (input variable: price) to transport a load (output variable: payload) which in combination with the engine power (output variable: engine power) and the weight (input variable: net weight) leads to the acceleration (output variable: acceleration) and the final top speed (output: top speed) of the car. This transportation process requires fuel (input variable: fuel consumption) and produces carbon dioxide emissions (undesirable output variable: CO_2). We limit our research focus on this technical view of an automobile as plainly providing transportation services and ignore variables like luxury equipment (e.g air conditioner). The following table 1 contains some descriptive statistics of the variables used in our study.

Table 1: Descriptive statistics of the data

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

[7] In this study we use the same terminology as in Cantner et al. (2010) (see table 3 in the appendix for an example).

[8] These classes are: Microwagen, Kleinstwagen, Kleinwagen, Untere Mittelklasse, Mittelklasse, Obere Mittelklasse and Oberklasse.

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

5 Results

In this section we present and discuss the results of our efficiency analysis. We start by considering the efficiency of the overall sample and later we will turn to different group-specific results. As explained above, we estimate three different efficiency measures, DEA (ignoring the CO_2 emissions), DDF_O (including the emissions as an additional restriction) and DDF_{UO} (simoultanously increasing desirable and decreasing undesirable output). For estimating the DEA efficiency scores and the density functions we use the FEAR-package provided by Wilson (2008) for the statistic software "R". The β-values of the DDF's as well as the p-values of the tests for stochastic dominance were obtained by our own programmings. Table 2 contains summary statistics of the efficiency estimates while figure 3 presents the related density functions.[9]

Table 2: Summary statistics of the efficiency measures - Overall sample

	Min.	1. Qu.	Median	Mean	3. Qu	Max.	SD	Total Eff.	% Eff.
DEA (θ)	1	1.124	1.198	1.215	1.287	1.859	0.130	82	2.07
DDF_O ($1+\beta_O$)	1	1.109	1.184	1.198	1.271	1.793	0.125	127	3.21
DDF_{UO} ($1+\beta_{UO}$)	1	1.076	1.119	1.124	1.170	1.412	0.068	112	2.83

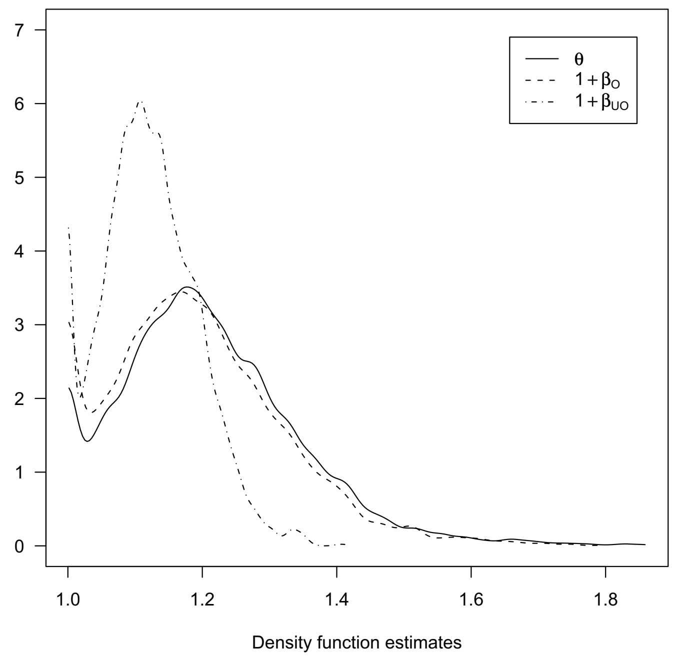

Figure 3: Density estimates of efficiency measures - Overall sample

[9] Since the efficiciency scores are bounded below at 1 we use the reflection method proposed by Boneva et al. (1971) to estimate the density functions which is implemented in the FEAR-package provided by Wilson (2008).

The results of the data envelopment analysis show that the average car of the sample has to increase all its outputs by 21.5 % holding inputs constant to become technically efficient. Rather surprising this result does not change much when emissions are introduced. The average potential for efficiency enhancement lowers only to 19.8 %.[10] This difference is also statistically significant since the test for first-order stochastic dominance of DEA to DDF_{O} does not reject the null hypothesis (see tables 9 and 10 in the appendix for the test results) and so DEA efficiency estimates stochastically dominate the DDF_{O} estimates.[11] As can be easily seen, the effect of simoultanously reducing undesirable outputs on the potential increase in technical efficiency is far greater, reducing the average potential efficiency gains to 12.4 %. We interpret the differences between the results of DDF_{O} and DDF_{UO} as the lack of performance enhancement possibilities due to simultaneously decreasing emissions (trade-off). To summarize the results for the overall dataset, the incorporation of carbon dioxide emissions lowers the technical enhancement possibilities by 1.7 % while the trade-off between improving the technical characteristics and lowering the emissions results in 7.4 % lower enhancement possibilities. This basic structure is also evident for the median as a more robust location measure and is clearly visible from the density plots.

These results are based on the whole dataset and thus apply to a quite heterogenous group of automobiles. To obtain more detailed results we analyse the efficiency of different groups in our dataset and compare the performance between the groups as well as to the overall dataset.[12] To evaluate and compare the efficiency of different groups there are two different methods. It is possible to use the efficiency results obtained above for the analysis of the overall dataset and compare them among groups respectively between a group and the remaining observations. A drawback of this approach is that the results of the analysis can be due to whithin-group or between-group differences. To illustrate this point, figure 4 shows an output set with two different groups of observations.

[10] An alternative interpretation for this result is that the average car is less inefficient since it lies nearer to the new frontier.

[11] Note that not rejecting the null hypothesis alone does not allow to make this statement since the efficiency scores could be equal (stochastically undominated). If the reverse null hypothesis is not rejected the efficiency scores are equal and we report this finding.

[12] One might argue that the resulting efficiency scores of the groups largely depend on the variable selected as inputs and outputs for our analysis. However, we want to point out that several different specifications have been tested and lead to similar results in the efficiency scores.

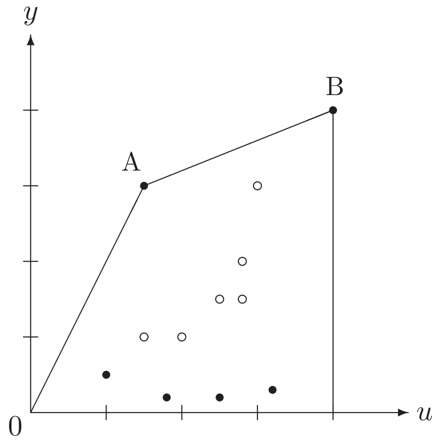

Figure 4: Weak disposable output set with two groups of DMUs

This figure shows a weak disposable output set that consists of two groups of observations (group 1 denoted by ○ and group 2 denoted by •). Since the frontier is defined only by observations from group 2 we could characterize group 2 as more efficient than group 1. But if we compare the overall effiency measures this result might be overturned since several observations from group 2 are located far below the frontier while the observations of group 1 lie more close to the boundary although no DMU of this group is classified as efficient.

To evaluate whether differences in the performance of a group depend on program or managerial aspects we use the frontier seperation approach proposed by Charnes et al. (1981) in a variant proposed by Portela and Thanassoulis (2001). In this apporach the efficiency score of a DMU estimated in the overall dataset is splitted into two components, the program efficiency and the managerial efficiency. While the managerial efficiency indicates the inefficiency of a DMU relativ to its group frontier, the program efficiency indicates the difference between the group frontier and the overall frontier hence the efficiency of the whole program. We use the terminology of managerial and program efficiency, although it does not completely suit to the analysis of cars, because it is common in the literature of the frontier seperation approach (see Thanassoulis et al. (2008)). The resulting efficiency scores are decomposed as:

- $\theta_{Ov} = \theta_{Ma} \cdot \theta_{Pr}$
- $(1 + \beta_{Ov}) = (1 + \beta_{Ma}) \cdot (1 + \beta_{Pr})$

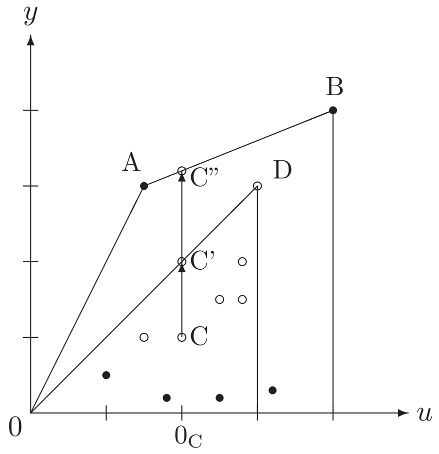

Figure 5: Frontier Seperation Approach

To illustrate this approach figure 5 shows the output set under weak disposability like figure 2. The boundary of the overall output set is still given by 0AB and the vertical extension from B. So the overall efficiency measure under weak disposability ($1+\beta_O$) of C is $\frac{0_C C''}{0_C C}$. If we compare C only to its own program than the DMUs belonging to group 1 become irrelevant and the boundary is given by 0D and the vertical extension. The efficicency score for C now is given by $\frac{0_C C'}{0_C C}$. This is called the managerial efficiency since it expresses all ineffiency that is not based on program differences. Program effiency is then given by the difference between the two boundaries and can be estimated by the residual of overall and managerial efficiency, $\frac{\frac{0C''}{0C}}{\frac{0C'}{0C}}$. We will compare the obtained efficiency scores within groups as well as between groups (for the three different measures defined above) to determine in more detail how the incorporation of emissions changes the results of the groups and to identify the sources for efficiency differences between groups.
In our paper we analyse the efficiency of the following groups:[13]

- car classes: compact class cars, middle class cars and upper class cars
- regional groups: Europe, United States, Asia
- special utility vehicles (SUVs)
- cars with natural gas engines (NGEs)

The first group refers to the general structure of the automobile market. We expect the compact cars to be more efficient as the middle and upper class vehicles because we assume that the additional luxury equipment lowers the technical and environmental performance hence resulting in another trade-off. The regional groups are based on the three largest automobile producing regions. The SUV's are of special interest since their market share in Europe has increased in the last years (see Zervas (2010)) while their environmental performance is questionable (see Plotkin (2004)). The converse holds for natural gas

[13]Note that the three groups are not disjoint, e.g. SUVs are part of the classes

engines which are supposed to be technically and environmentally superior to standard engines.

The following figures contain boxplots for the results of the efficiency analysis and the frontier seperation approach. The results for the overall dataset are also included because they provide a useful benchmark for the group results with regard to their overall efficiency.

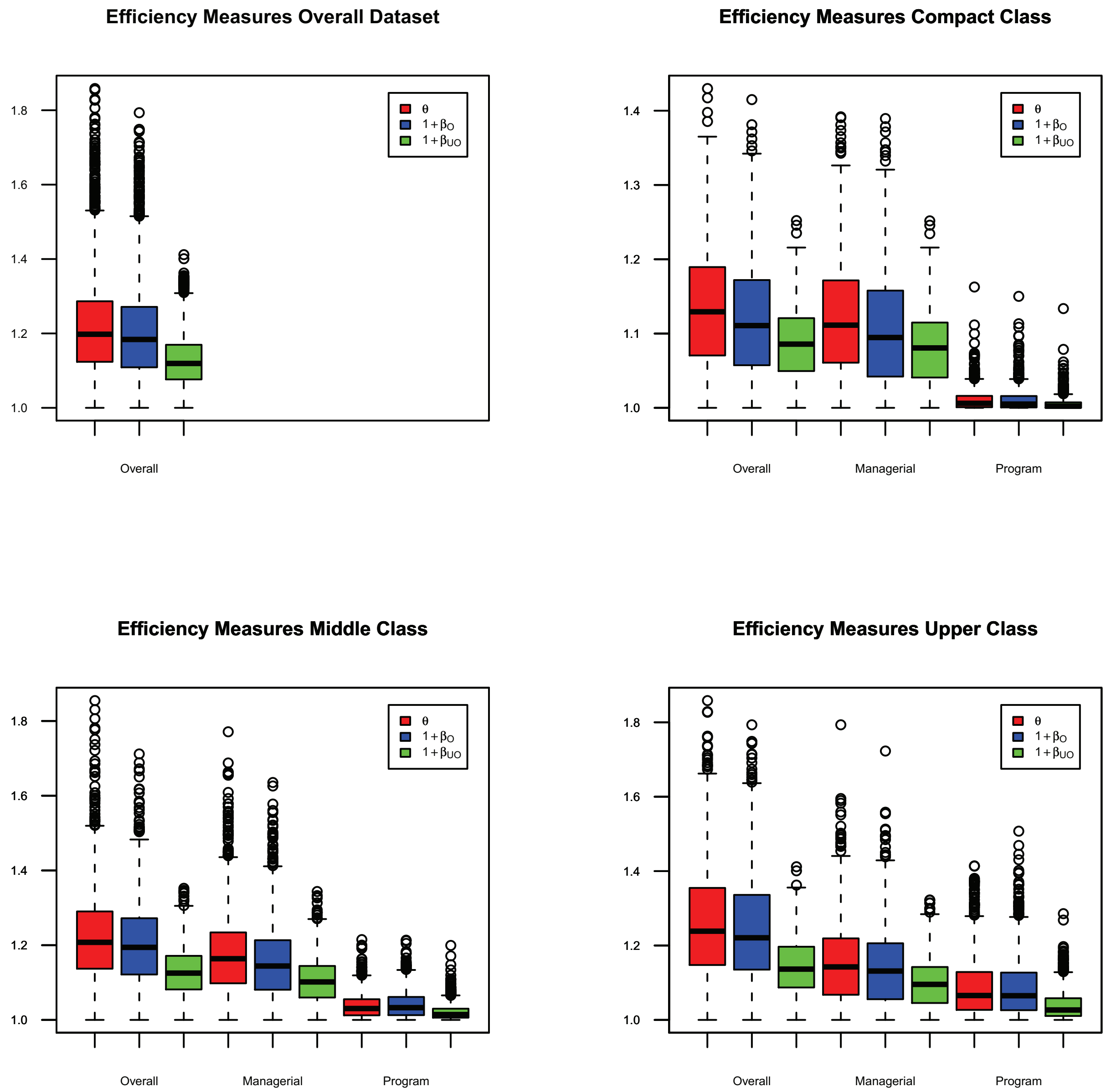

Figure 6: Boxplots for class results

Figure 6 shows the boxplots of the efficiency scores for the overall dataset as well as the results of the frontier seperation for the car classes. The summary statistics for the class results can be found in table 4 in the appendix. We start our discussion of the results with a comparison of the overall efficiency of the different groups with the results of the overall dataset. We find that compared to the overall dataset only compact cars perform better while middle and upper class cars perform worse. This result holds for all three efficiency measures and is statistically significant (see p-values in the tables 7 and 8 in the appendix). As explained above this comparison does not reveal any information about the ranking of the program efficiency of the three car classes. So from this result we can only conclude that the overall efficiency of the compact class is higher than the overall efficiency of middle and upper class cars. To obtain more detailed informations of the performance of the different classes, boxplots and summary statistics for both, managerial and program efficiency evaluations, can be found in the same figure and table as the overall efficiency results. The boxplots show, that for compact and middle class cars the overall efficiency results are mainly based on managerial aspects rather the effects due to program efficiency. For upper class cars we see that the program efficiency contributes more to the overall efficiency results. A comparison within the classes between the three efficiency measures θ, $1+\beta_O$ and $1+\beta_{UO}$ (see tables 8 and 9 in the appendix for results of the tests for stoachastic dominance) shows that while the ranking obtained for the overall dataset ($\theta \succ_{FSD} 1+\beta_O \succ_{FSD} 1+\beta_O$) can be found again for the overall and the managerial efficiency of all three classes, the results with regard to the program efficiency are different. While the results from DEA and $\mathrm{DDF_O}$ are stochastically undominated for compact and upper car class, $\mathrm{DDF_O}$ results stochastically dominate the results from DEA for the midddle class. So while the incorporation of emissions does not change the distance of the frontier functions of the compact and the upper class cars to the overall frontier, it enlarges the difference for the middle class cars indicating that this class is relatively less efficient when emissions are introduced. The between-classes analysis (for the p-values of the tests see table 13 in the appendix) shows that compact class cars are more efficient than the other classes in all efficiency types (overall, managerial and program) and for all three efficiency measures. But if we compare middle and upper class vehicles we find that while upper class performs better with regard to the managerial aspects it performs worse regarding program efficiency. Graphically speaking the observations of upper class cars lie nearer to their own (class) frontier then the middle class observations to theirs but the (class) frontier is significantly further away from the overall frontier than the middle class frontier. This latter effect dominates and so the upper class is less efficient than middle class regarding overall efficiency. If we compare the different efficiency measures between the classes, we see that the ranking of the car classes does not change, hence the ranking is neither influenced by the incorporation of emissions nor by the trade-off between environmental and technical performance.

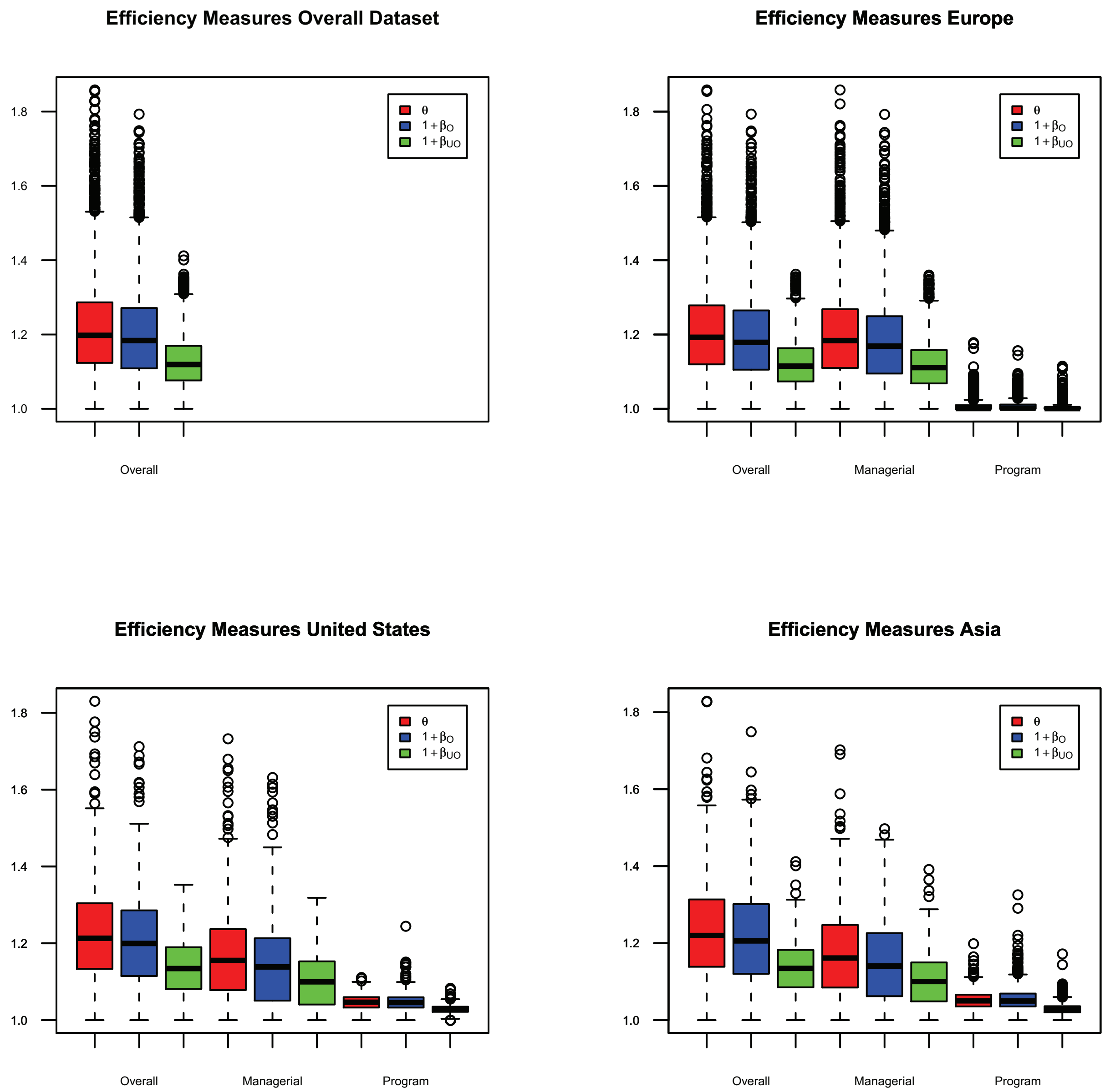

Figure 7: Boxplots for regional results

Figure 7 shows the boxplots for the results of the overall dataset as well as for the results of the regions of origin of the cars (Europe, United States, Asia) while the summary statistics for the regional results can be found in table 5 in the appendix. Comparing the overall efficiency results with the overall dataset we find that European cars perform slightly better (the efficiency results are only second-order stochastically dominated) while Asian and American cars perform worse. This result holds for all three efficiency measures. The results from the frontier seperation, presented in the same places as the previous overall results, show that for European cars the managerial aspects account for nearly all of the overall efficiency results. For the United States and Asia the effect of program aspects is larger, but the main forces that drive the overall efficiency results are also managerial aspects. The within-analysis of the regions (for test results see tables 11 and 12 in the appendix) shows for Europe the usual relation between the efficiency measures for the overall and the managerial efficiency. For program efficiency we find that the results from $\mathrm{DDF_O}$ stochastically dominate those from DEA. Therefor the European (regional) frontier lies further away from the overall frontier when emissions are incorporated. We find analogeous results for Asia but the first-order stochastic dominance of $\mathrm{DDF_O}$ to DEA in program efficiency is only significant at the 10 % level. For the United States we find weaker significance for the stochastic dominance of DEA to $\mathrm{DDF_O}$ for overall (only second-order dominance is significant) and managerial efficiency (first-order dominance is significant at the 10 % level). With regard to the program efficiency of the United States we find that both measures are stochastically undominated so the incorporation of emissions does not change the location of the program frontier relative to the overall frontier. Comparing the results between the regions (see table 13 in the appendix for the related test results) we find for all efficiency measures that Europe is stochastically dominated by Asia as well as the United States with regard to the overall as well as the program efficency. This last result shows that the regional frontier of European cars lies closer to the overall frontier than the frontiers of Asia or the United States. In fact, the DEA analysis shows that 1287 European cars (43 %) are classified efficient regarding program effiency, so we can conclude that many parts of the overall frontier are identical to the European program (regional) frontier when emissions are ignored. For the managerial efficiency we find that results for European cars dominate both the results for Asia as well as those for the United States indicating a greater distance of European cars to their regional frontier compared to the distance of Asian or United States cars to their regional frontiers. Comparing Asia and the United States we find that the overall efficiency results as well as the managerial efficiency results are undominated for all efficiency measures. Comparing the results of the program efficiency for DEA and the $\mathrm{DDF_O}$ we find that Asia stochastically dominates the United States, hence the United States frontier lies closer to the overall frontier than the Asian frontier. For $\mathrm{DDF_{UO}}$ we find no statistically significant differences. That means if we account for the trade-off between emission reduction and technical improvement then we do not find significant differences between the Asian and the United States frontier. We want to point out that although the results we obtained for the program efficiency of the regions may be often statistically significant, the program efficiency only contributes very little to the overall efficiency. Therefor, the differences of the regional frontiers to each other as well as to the overall frontier are quite small.

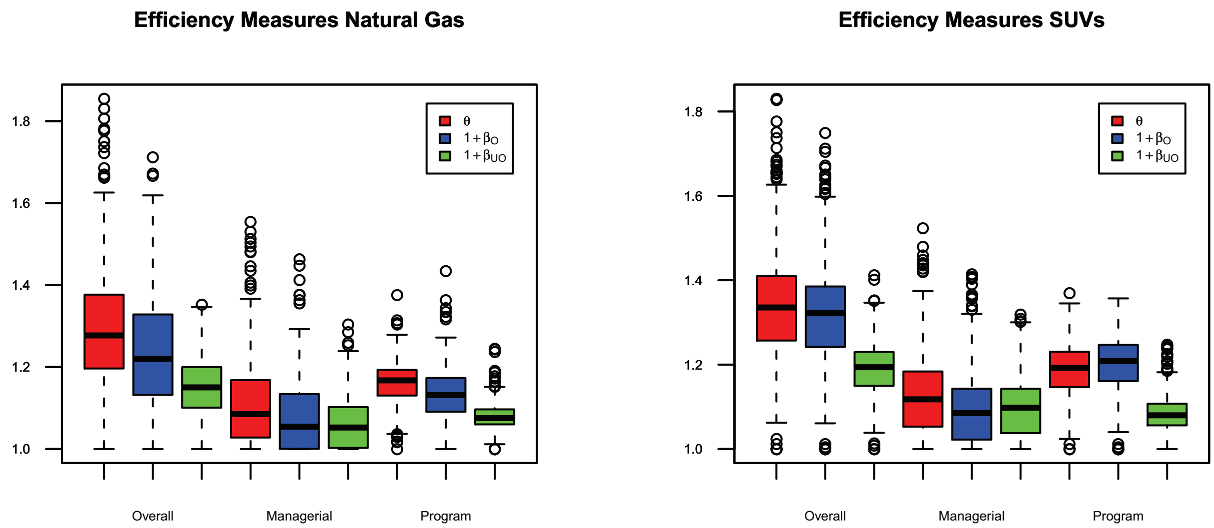

Figure 8: Boxplots for NGE and SUV results

The results from the efficiency analysis for the SUVs and especially for the NGEs which are presented in figure 8 above and table 6 in the appendix are the most surprising ones of our analysis. We present the results of both groups together, because we expected these groups to have opposite results, the NGEs being considerably more efficient than any other car group, the SUVs performing significantly worse. Interestingly the results for the NGEs do not meet our expectations. First, comparing the overall dataset to both groups we find that both groups perform significantly worse (see tables 7 and 8 in the appendix). While we expected this result for SUVs it is quite surprising for NGEs. The frontier seperation approach gives a more detailed explanation for this result. For both groups the main source for their relative ineffiency is their poor performance in program efficiency while managerial aspects only contribute little to the efficiency results (see figure 7). While we expected SUVs to have a group frontier far away from the overall frontier we are quite surprised finding nearly the same result for NGEs. The within-group analysis (see table 15 and 16 in the appendix) shows the usual dominance relation between the efficiency measures for the overall and managerial efficiency of the SUVs for the NGEs. The within-group analysis with regard to the program efficiency shows that for SUVs DEA results are dominated by the $\mathrm{DDF_O}$ results while the converse holds for the NGEs. Therfore, when emissions are incorporated the group frontier of the SUVs lies further away of the overall frontier while the group frontier of the NGEs lies closer to the overall frontier compared to the case when emissions are ignored. We interpret our finding for the NGEs in the way that although they show an enhanced performance when emissions are introduced their lack of technical effiency compared to cars with conventional engines is remarkably high (as it can be seen by comparing the θ scores of NGEs with the scores of the overall sample), overturning the environmental advantages and thus resulting in a relatively poor efficiency performance.

6 Conclusion

In this paper we analyze the technical efficiency of cars using nonparametric methods. Emphasizing the importance of incorporating carbon dioxide emissions we use a nonparametric approach and measure efficiency based on directional distance functions to evaluate the technical efficiency of the automobiles when the reduction of emissions is regarded as an equally important target. To check whether the results are significant in a statistical sense, we use tests that are based on the concept of stochastic dominance. Our results show that the incorporation of carbon dioxide emissions has a significant effect on the resulting effieciency scores of the automobiles in our dataset reducing the average potential to enhance efficiency by 1.7 %. The trade-off between reducing emissions and increasing the technical performance lowers this potential further by 7.4 %.

Dividing the dataset in several special groups and using a frontier seperation approach we get a more detailed analysis about the sources of efficiency differences between the car groups. Analysing the results for car classes, we find that the compact class cars perform significantly better then middle or upper class cars showing the trade-off between technical and environmental efficiency of a car on the one hand and energy-consuming luxury equipment of the car on the other hand. The analysis of the regions of origin shows that European cars perform better than Asian cars and cars from the United States. We emphasize that this result may be due to our dataset, that is captured from an online database of a German automobile club thereby lacking some non-European cars. The most surprising result of our analysis is the finding for cars with natural gas engine. While we expected them to be highly technically efficient, at least when incorporating carbon dioxide emissions, we found that the opposite is true. The results show that natural gas cars are significantly less efficient compared to the overall sample and that the incorporation of carbon dioxide emissions enhances their performance but ultimately has only a very limited effect on the overall effiency of this car class. We want to emphasize, that this result does not automatically imply the bold conclusion that the technology of natural gas engines is inferior to standard engines but that it also could be due to the fact that the cars that are currently saled are in an early stage of the technological development and many potential efficiency improvements appear to be unexploited yet. Furthermore we want to stress that our analysis assumes constant returns to scale and we therefor does not account for efficiency differences that are based on size effects. An extension of our analysis to the case of variable returns to scale, e.g. with the model proposed by Zhou et al. (2008b), might give further insights in the role of scale effects in the automobile technology. This is an important point on our agenda for future research.

7 References

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

— (2010). *ADAC-Autokatalog*. [Retrieved February 03, 2010]. URL: http://www1.adac.de/Auto_Motorrad/Autokatalog/default.asp?id=1356&location=2_Auto%2C+Motorrad+%26+Oldtimer.

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

Cantner, U., J.J. Krüger, and R. Söllner (2010). "Product Quality, Product Price, and Share Dynamics in the German Compact Car Market". Jena Economic Research Papers 2010 - 024.

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

— (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through". In: *Management Science* 27 (6), pp. 668–697.

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

Commission of the European Communities (2007). "Proposal for a Regulation of the European Parliament and of the Council - Setting Emission Performance Standards for New Passenger Cars as Part of the Community's Integrated Approach to Reduce CO2 Emissions from Light-Duty Vehicles". Dossier COD/2007/0297.

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

Debreu, G. (1951). "The Coefficient of Resource Utilization". In: *Econometrica* 19 (3), pp. 273–292.

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

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

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

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

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

Färe, R. and S. Grosskopf (2003). "Nonparametric Productivity Analysis with Undesirable Outputs: Comment". In: *American Journal of Agricultural Economics* 85 (4), pp. 1070–1074.

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

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

Liu, W.B., W. Meng, X.X. Li, and D.Q. Zhang (2010). "DEA Models with Undesirable Outputs". In: *Annals of Operational Research* 173 (1), pp. 177–194.

Oh, I., J.-D. Lee, S. Hwang, and A. Heshmati (2010). "Analysis of Product Efficiency in the Korean Automobile Market from a Consumers Perspective". In: *Empirical Economics* 38 (1), pp. 119–137.

Papahristodoulou, C. (1997). "A DEA Model to Evaluate Car Efficiency". In: *Applied Economics* 29 (11), pp. 1493–1508.

Plotkin, S. (2004). "Is Bigger Better?: Moving Toward a Dispassionate View of SUVs". In: *Environment: Science and Policy for Sustainable Development* 46 (9), pp. 8–21.

Portela, M.C.A.S. and E. Thanassoulis (2001). "Decomposing School and School-Type Efficiency". In: *European Journal of Operational Research* 132, pp. 357–373.

Scheel, H. (2001). "Undesirable Outputs in Efficiency Valuations". In: *European Journal of Operational Research* 13, pp. 400–410.

Schmid, F. and M. Trede (2000). "Stochastic Dominance in German Asset Return: Empirical Evidence from the 1990s". In: *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)* 220, pp. 315–326.

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

— (1974). *Indirect Production Functions*. Meisenheim am Glan: Verlag Anton Hain.

Simpson, G. (2005). "Programmatic Efficiency Comparisons between Unequally Sized Groups of DMUs in DEA". In: *Journal of the Royal Operational Research Society* 56, pp. 1431–1438.

— (2007). "A Cautionary Note on Methods of Comparing Programmatic Efficiency between Two or More Groups of DMUs in Data Envelopment Analysis". In: *Journal of Productivity Analysis* 28, pp. 141–147.

Thanassoulis, E., M.C.A.S. Portela, and O. Despić (2008). "Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis". In: *The Measurement of Productive Efficiency and Productivity Growth*. Ed. by H.O. Fried, C.A.K. Lovell, and S.S. Schmidt. Oxford University Press, pp. 251–420.

Wilson, P.W. (2008). "FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R". In: *Socio-Economic Planning Sciences* 42, pp. 247–254.

Zervas, E. (2010). "Analysis of the CO2 Emissions and of the other Characteristics of the European Market of New Passenger Cars. 2. Segment Analysis". In: *Energy Policy*. DOI: 10.1016/j.enpol.2010.02.007.

Zhou, P., B.W. Ang, and K.L. Poh (2008a). "A Survey of Data Envelopment Analysis in Energy and Environmental Studies". In: *European Journal of Operational Research* 189, pp. 1–18.

— (2008b). "Measuring Environmental Performance under Different Environmental DEA Technologies". In: *Energy Economics* 30 (1), pp. 1–14.

8 Appendix

Table 3: Used terminology (see Cantner et al. (2010))

Category	Example
Brand	VW
Produt line	Golf
Model	Golf 1.6
Model variant	Golf 1.6 Trendline

Table 4: Summary statistics of the efficiency measures - Classes

Efficiency Measure	Car Class	Type of Efficiency	Min.	1. Qu.	Median	Mean	3. Qu	Max.	SD	Total Eff.	% Eff.
DEA (θ)	Compact	Overall	1	1.071	1.129	1.134	1.189	1.43	0.083	29	4.5
		Managerial	1	1.061	1.111	1.121	1.172	1.392	0.081	42	6.52
		Program	1	1.001	1.006	1.012	1.016	1.163	0.016	110	17.08
	Middle	Overall	1	1.137	1.208	1.219	1.29	1.855	0.118	33	1.4
		Managerial	1	1.098	1.164	1.175	1.234	1.771	0.106	57	2.41
		Program	1	1.012	1.031	1.038	1.055	1.215	0.031	62	2.62
	Upper	Overall	1	1.148	1.239	1.258	1.355	1.859	0.158	20	2.1
		Managerial	1	1.068	1.142	1.156	1.219	1.793	0.117	60	6.3
		Program	1	1.027	1.066	1.088	1.129	1.415	0.082	69	7.24
DDF_O $(1+\beta_O)$	Compact	Overall	1	1.057	1.111	1.119	1.172	1.415	0.081	47	7.3
		Managerial	1	1.042	1.095	1.106	1.158	1.39	0.081	73	11.34
		Program	1	1.001	1.005	1.012	1.016	1.15	0.017	119	18.48
	Middle	Overall	1	1.122	1.194	1.202	1.272	1.712	0.111	50	2.12
		Managerial	1	1.081	1.144	1.155	1.213	1.636	0.099	90	3.81
		Program	1	1.013	1.033	1.041	1.061	1.214	0.034	63	2.66
	Upper	Overall	1	1.135	1.221	1.244	1.336	1.793	0.155	30	3.15
		Managerial	1	1.056	1.131	1.143	1.206	1.723	0.112	91	9.55
		Program	1	1.026	1.065	1.089	1.127	1.507	0.083	64	6.72
DDF_{UO} $(1+\beta_{UO})$	Compact	Overall	1	1.05	1.086	1.087	1.121	1.252	0.051	40	6.21
		Managerial	1	1.041	1.081	1.08	1.115	1.252	0.051	54	8.39
		Program	1	1	1.002	1.006	1.007	1.134	0.011	186	28.88
	Middle	Overall	1	1.082	1.125	1.127	1.171	1.352	0.063	44	1.86
		Managerial	1	1.06	1.102	1.105	1.144	1.344	0.06	86	3.64
		Program	1	1.006	1.014	1.021	1.03	1.199	0.02	58	2.45
	Upper	Overall	1	1.087	1.137	1.142	1.197	1.412	0.079	28	2.94
		Managerial	1	1.045	1.096	1.099	1.142	1.323	0.068	80	8.39
		Program	1	1.011	1.027	1.039	1.058	1.286	0.04	66	6.93

Table 5: Summary statistics of the efficiency measures - Regions

Efficiency Measure	Region	Type of Efficiency	Min.	1. Qu.	Median	Mean	3. Qu	Max.	SD	Total Eff.	% Eff.
DEA (θ)	Europe	Overall	1	1.120	1.193	1.209	1.278	1.859	0.129	67	2.24
		Managerial	1	1.110	1.184	1.199	1.268	1.859	0.128	80	2.67
		Program	1	1	1.001	1.008	1.010	1.178	0.016	1287	43.01
	United States	Overall	1	1.133	1.213	1.233	1.304	1.830	0.151	7	2.34
		Managerial	1	1.078	1.155	1.178	1.237	1.733	0.144	25	8.36
		Program	1	1.033	1.047	1.047	1.060	1.111	0.022	7	2.34
	Asia	Overall	1	1.139	1.22	1.231	1.313	1.829	0.126	8	1.19
		Managerial	1	1.085	1.161	1.169	1.247	1.702	0.113	39	5.82
		Program	1	1.036	1.05	1.053	1.066	1.199	0.026	8	1.19
DDF_O ($1+\beta_O$)	Europe	Overall	1	1.105	1.179	1.194	1.265	1.793	0.124	98	3.28
		Managerial	1	1.095	1.169	1.183	1.249	1.793	0.123	119	3.98
		Program	1	1	1.002	1.009	1.012	1.157	0.016	843	28.18
	United States	Overall	1	1.115	1.2	1.211	1.286	1.712	0.142	14	4.68
		Managerial	1	1.051	1.139	1.155	1.213	1.632	0.133	36	12.04
		Program	1	1.033	1.046	1.048	1.059	1.245	0.028	14	4.68
	Asia	Overall	1	1.121	1.206	1.213	1.301	1.749	0.12	15	2.24
		Managerial	1	1.062	1.141	1.148	1.226	1.497	0.105	62	9.25
		Program	1	1.036	1.05	1.057	1.069	1.326	0.035	15	2.24
DDF_{UO} ($1+\beta_{UO}$)	Europe	Overall	1	1.074	1.115	1.12	1.163	1.363	0.067	86	2.87
		Managerial	1	1.068	1.111	1.116	1.158	1.36	0.067	105	3.51
		Program	1	1	1.001	1.004	1.004	1.115	0.009	1031	34.46
	United States	Overall	1	1.081	1.134	1.138	1.19	1.352	0.08	13	4.35
		Managerial	1	1.041	1.1	1.107	1.153	1.319	0.077	30	10.03
		Program	1	1.022	1.029	1.029	1.035	1.083	0.014	13	4.35
	Asia	Overall	1	1.085	1.135	1.136	1.182	1.412	0.067	13	1.94
		Managerial	1	1.049	1.1	1.103	1.15	1.391	0.068	56	8.36
		Program	1	1.02	1.027	1.03	1.036	1.172	0.017	13	1.94

Table 6: Summary statistics of the efficiency measures - SUVs and NGEs

Efficiency Measure	Car Group	Type of Efficiency	Min.	1. Qu.	Median	Mean	3. Qu	Max.	SD	Total Eff.	% Eff.
DEA (θ)	SUV	Overall	1	1.257	1.335	1.343	1.409	1.83	0.135	3	0.73
		Managerial	1	1.053	1.118	1.132	1.184	1.523	0.106	37	8.98
		Program	1	1.147	1.193	1.187	1.231	1.37	0.066	3	0.73
	NGE	Overall	1	1.196	1.277	1.301	1.377	1.855	0.166	1	0.47
		Managerial	1	1.028	1.085	1.118	1.168	1.554	0.12	35	16.28
		Program	1	1.13	1.167	1.162	1.193	1.376	0.057	1	0.47
DDF_O $(1+\beta_O)$	SUV	Overall	1	1.242	1.322	1.321	1.385	1.749	0.127	5	1.21
		Managerial	1	1.022	1.085	1.099	1.143	1.414	0.089	70	16.99
		Program	1	1.161	1.208	1.201	1.247	1.357	0.067	5	1.21
	NGE	Overall	1	1.132	1.22	1.232	1.328	1.712	0.141	8	3.72
		Managerial	1	1	1.054	1.084	1.134	1.463	0.095	53	24.65
		Program	1	1.091	1.132	1.135	1.173	1.434	0.071	8	3.72
DDF_{UO} $(1+\beta_{UO})$	SUV	Overall	1	1.15	1.194	1.191	1.23	1.412	0.066	4	0.97
		Managerial	1	1.038	1.098	1.099	1.143	1.319	0.074	60	14.56
		Program	1	1.056	1.08	1.086	1.107	1.248	0.045	4	0.97
	NGE	Overall	1	1.101	1.15	1.148	1.2	1.352	0.071	4	1.86
		Managerial	1	1.003	1.052	1.063	1.102	1.304	0.062	49	22.79
		Program	1	1.06	1.075	1.081	1.097	1.245	0.038	4	1.86

Table 7: Tests of groups against overall dataset (first-order stochastic dominance)

	Compact Class	Middel Class	Upper Class	Europe	Asia	United States	SUV	Natural Gas Engine
θ	0	0.352	0.962	0.108	0.68	0.786	0.994	0.997
	0.995	0.002	0	0.993	0	0.08	0	0
$1+\beta_O$	0	0.234	0.984	0.164	0.684	0.631	0.997	0.969
	0.993	0.002	0	0.972	0	0.171	0	0.001
$1+\beta_{UO}$	0	0.52	0.937	0.036	0.852	0.797	0.997	0.993
	0.994	0.016	0	0.997	0	0.003	0	0

Table 8: Tests of groups against overall dataset (second-order stochastic dominance)

	Compact Class	Middel Class	Upper Class	Europe	Asia	United States	SUV	Natural Gas Engine
θ	0	0.798	0.736	0.049	0.8	0.695	0.794	0.795
	0.81	0.007	0	0.806	0.001	0.013	0	0
$1+\beta_O$	0	0.792	0.799	0.072	0.79	0.515	0.789	0.743
	0.794	0.007	0	0.791	0.001	0.061	0	0
$1+\beta_{UO}$	0	0.787	0.695	0.008	0.783	0.573	0.779	0.773
	0.788	0.009	0	0.789	0	0.001	0	0

Table 9: Tests for first-order stochastic dominance (overall dataset and classes - within)

H_0	Dataset	Compact Class			Middle Class			Upper Class		
		Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{FSD} (1+\beta_O)$	0.997	0.989	0.983	0.565	0.995	0.995	0.002	0.992	0.986	0.699
$(1+\beta_O) \succeq_{FSD} \theta$	0	0.003	0.001	0.43	0	0	0.911	0.152	0.022	0.833
$\theta \succeq_{FSD} (1+\beta_{UO})$	0.996	0.989	0.988	0.971	0.994	0.995	0.981	0.992	0.987	0.97
$(1+\beta_{UO}) \succeq_{FSD} \theta$	0	0	0	0	0	0	0	0	0	0
$(1+\beta_O) \succeq_{FSD} (1+\beta_{UO})$	0.886	0.797	0.37	0.974	0.945	0.97	0.976	0.982	0.789	0.986
$(1+\beta_{UO}) \succeq_{FSD} (1+\beta_O)$	0	0	0	0	0	0	0	0	0	0

Table 10: Tests for second-order stochastic dominance (overall dataset and classes - within)

H_0	Dataset	Compact Class			Middle Class			Upper Class		
		Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{SSD} (1+\beta_O)$	0.799	0.759	0.747	0.495	0.797	0.787	0.001	0.799	0.762	0.591
$(1+\beta_O) \succeq_{SSD} \theta$	0	0	0.001	0.481	0	0	0.704	0.022	0.004	0.683
$\theta \succeq_{SSD} (1+\beta_{UO})$	0.806	0.773	0.75	0.742	0.812	0.791	0.788	0.806	0.765	0.776
$(1+\beta_{UO}) \succeq_{SSD} \theta$	0	0	0	0	0	0	0	0	0	0
$(1+\beta_O) \succeq_{SSD} (1+\beta_{UO})$	0.739	0.615	0.51	0.737	0.767	0.763	0.797	0.788	0.667	0.77
$(1+\beta_{UO}) \succeq_{SSD} (1+\beta_O)$	0	0	0	0	0	0	0	0	0	0

Table 11: Tests for first-order stochastic dominance (regions - within)

H_0	Europe			Asia			United States		
	Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{FSD} (1+\beta_O)$	0.995	0.993	0	0.991	0.984	0.094	0.987	0.980	0.694
$(1+\beta_O) \succeq_{FSD} \theta$	0	0	0.834	0.069	0.005	0.708	0.111	0.094	0.639
$\theta \succeq_{FSD} (1+\beta_{UO})$	0.995	0.994	0	0.992	0.984	0.991	0.986	0.977	0.985
$(1+\beta_{UO}) \succeq_{FSD} \theta$	0	0	0	0	0	0	0	0	0
$(1+\beta_O) \succeq_{FSD} (1+\beta_{UO})$	0.903	0.834	0.983	0.943	0.861	0.973	0.944	0.776	0.935
$(1+\beta_{UO}) \succeq_{FSD} (1+\beta_O)$	0	0	0	0	0	0	0	0	0

Table 12: Tests for second-order stochastic dominance (regions - within)

H_0	Europe			Asia			United States		
	Overall	Managerial	Program	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{FSD} (1+\beta_O)$	0.788	0.786	0.004	0.807	0.742	0.005	0.789	0.754	0.33
$(1+\beta_O) \succeq_{FSD} \theta$	0	0	0.711	0.006	0	0.673	0.034	0.023	0.499
$\theta \succeq_{FSD} (1+\beta_{UO})$	0.8	0.801	0.723	0.812	0.757	0.821	0.802	0.755	0.773
$(1+\beta_{UO}) \succeq_{FSD} \theta$	0	0	0	0	0	0	0	0	0
$(1+\beta_O) \succeq_{FSD} (1+\beta_{UO})$	0.733	0.732	0.718	0.736	0.651	0.788	0.741	0.674	0.718
$(1+\beta_{UO}) \succeq_{FSD} (1+\beta_O)$	0	0	0	0	0	0	0	0	0

Table 13: Tests for stochastic dominance (classes - between)

		First-Order Stochastic Dominance			Second-Order Stochastic Dominance		
Compared Classes	Type of Efficiency	θ	$1+\beta_O$	$1+\beta_{UO}$	θ	$1+\beta_O$	$1+\beta_{UO}$
Compact $\succeq$ Middle	Overall	0	0	0	0	0	0
	Managerial	0	0	0	0	0	0
	Program	0	0	0	0	0	0
Compact $\succeq$ Upper	Overall	0	0	0	0	0	0
	Managerial	0	0	0	0	0	0
	Program	0	0	0	0	0	0
Middle $\succeq$ Compact	Overall	0.996	0.996	0.994	0.804	0.79	0.789
	Managerial	0.994	0.993	0.994	0.791	0.775	0.774
	Program	0.991	0.992	0.988	0.769	0.776	0.769
Middle $\succeq$ Upper	Overall	0	0	0	0	0	0
	Managerial	0.693	0.497	0.34	0.785	0.77	0.765
	Program	0	0	0	0	0	0
Upper $\succeq$ Compact	Overall	0.993	0.991	0.995	0.803	0.784	0.776
	Managerial	0.773	0.989	0.744	0.638	0.745	0.637
	Program	0.986	0.987	0.985	0.761	0.749	0.747
Upper $\succeq$ Middle	Overall	0.438	0.3	0.298	0.45	0.423	0.37
	Managerial	0	0	0	0	0	0.001
	Program	0.033	0.047	0.033	0.645	0.642	0.653

Table 14: Tests for stochastic dominance (regions - between)

Compared Regions	Type of Efficiency	First-Order Stochastic Dominance			Second-Order Stochastic Dominance		
		θ	$1+\beta_O$	$1+\beta_{UO}$	θ	$1+\beta_O$	$1+\beta_{UO}$
Europe $\succeq$ Asia	Overall	0	0	0	0	0	0
	Managerial	0.994	0.992	0.967	0.798	0.784	0.771
	Program	0	0	0	0	0	0
Europe $\succeq$ United States	Overall	0.01	0.04	0	0.002	0.016	0
	Managerial	0.773	0.851	0.565	0.806	0.78	0.785
	Program	0	0	0	0	0	0
Asia $\succeq$ Europe	Overall	0.686	0.654	0.932	0.802	0.788	0.777
	Managerial	0	0	0	0	0	0
	Program	0.988	0.989	0.99	0.702	0.714	0.709
Asia $\succeq$ United States	Overall	0.176	0.41	0.087	0.503	0.801	0.351
	Managerial	0.038	0.074	0.192	0.161	0.205	0.267
	Program	0.995	0.99	0.262	0.803	0.798	0.579
United States $\succeq$ Europe	Overall	0.858	0.783	0.832	0.738	0.544	0.618
	Managerial	0	0	0.001	0.001	0	0.004
	Program	0.969	0.992	0.968	0.699	0.714	0.718
United States $\succeq$ Asia	Overall	0.157	0.111	0.295	0.202	0.124	0.257
	Managerial	0.211	0.381	0.451	0.277	0.286	0.398
	Program	0.004	0.001	0.336	0	0	0.168

Table 15: Tests for first-order stochastic dominance (SUVs and NGEs - within)

H_0	SUVs			Natural Gas Engine		
	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{FSD} (1+\beta_O)$	0.99	0.978	0.001	0.986	0.965	0.848
$(1+\beta_O) \succeq_{FSD} \theta$	0.05	0	0.941	0	0.002	0
$\theta \succeq_{FSD} (1+\beta_{UO})$	0.992	0.98	0.992	0.985	0.964	0.987
$(1+\beta_{UO}) \succeq_{FSD} \theta$	0	0	0	0	0	0
$(1+\beta_O) \succeq_{FSD} (1+\beta_{UO})$	0.964	0.131	0.98	0.722	0.542	0.714
$(1+\beta_{UO}) \succeq_{FSD} (1+\beta_O)$	0	0.328	0	0	0.001	0

Table 16: Tests for second-order stochastic dominance (SUVs and NGEs - within)

H_0	SUVs			Natural Gas Engine		
	Overall	Managerial	Program	Overall	Managerial	Program
$\theta \succeq_{FSD} (1+\beta_O)$	0.81	0.742	0.001	0.807	0.733	0.778
$(1+\beta_O) \succeq_{FSD} \theta$	0.008	0	0.664	0	0	0
$\theta \succeq_{FSD} (1+\beta_{UO})$	0.826	0.733	0.829	0.83	0.728	0.804
$(1+\beta_{UO}) \succeq_{FSD} \theta$	0	0	0	0	0	0
$(1+\beta_O) \succeq_{FSD} (1+\beta_{UO})$	0.764	0.179	0.821	0.637	0.573	0.592
$(1+\beta_{UO}) \succeq_{FSD} (1+\beta_O)$	0	0.578	0	0	0.003	0