Mendel University in Brno
Faculty of Business and Economics

Svatopluk Kapounek
Hana Vránová (eds.)

38th INTERNATIONAL CONFERENCE ON MATHEMATICAL METHODS IN ECONOMICS (MME 2020)

Conference Proceedings

September 9–11, 2020
Mendel University in Brno
Czech Republic

International Conference on Mathematical Methods in Economics 2020 (MME 2020) was organized by the Faculty of Business and Economics, Mendel University in Brno.

Responsible Editors:

Svatopluk Kapounek
Hana Vránová

Programme Committee:

prof. Dr. Ing. Miroslav Plevný
President of the Czech Society for Operations Research
University of West Bohemia, Faculty of Economics

prof. Ing. Josef Jablonský, CSc.
Vice-president of the Czech Society for Operations Research
University of Economics in Prague, Faculty of Informatics and Statistics

prof. RNDr. Helena Brožová, CSc.
University of Life Science in Prague, Faculty of Economics and Management

prof. Ing. Mgr. Martin Dlouhý, Dr., MSc.
University of Economics in Prague, Faculty of Informatics and Statistics

doc. Ing. Jan Fábry, Ph.D.
University of Economics in Prague, Faculty of Informatics and Statistics

prof. RNDr. Ing. Petr Fiala, CSc., MBA
University of Economics in Prague, Faculty of Informatics and Statistics

Ing. Ludvík Friebel, Ph.D.
University of South Bohemia, Faculty of Economics

prof. Ing. Jana Hančlová, CSc.
Technical University of Ostrava, Faculty of Economics

doc. Ing. Svatopluk Kapounek, Ph.D.
Mendel University in Brno, Faculty of Business and Economics

doc. RNDr. Jana Klicnarová, Ph.D.
University of South Bohemia, Faculty of Economics

Ing. František Koblasa, Ph.D.
Technical University of Liberec, Faculty of Mechanical Engineering

doc. RNDr. Ing. Miloš Kopa, Ph.D.
Charles University, Faculty of Mathematics and Physics

doc. RNDr. Ing. Ladislav Lukáš, CSc.
University of West Bohemia, Faculty of Economics

prof. RNDr. Jan Pelikán, CSc.
University of Economics in Prague, Faculty of Informatics and Statistics

prof. RNDr. Jaroslav Ramík, CSc.
Silesian University in Opava, School of Business Administration in Karviná

Ing. Karel Sladký, CSc.
Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation

Ing. Eva Šírová, Ph.D.
Technical University of Liberec, Faculty of Economics

doc. Ing. Tomáš Šubrt, Ph.D.
University of Life Science in Prague, Faculty of Economics and Management

doc. RNDr. Jana Talašová, CSc.
Palacký University in Olomouc, Faculty of Science

Organisation Committee:

Technical Editors:

ISBN 978-80-7509-734-7

Use of the BCC and Range Directional DEA Models within an Efficiency Evaluation

Michal Houda[1]

Abstract. The contribution deals with two data envelopment analysis (DEA) models, in particular the BCC model (radial DEA model with variable returns to scale), and the range directional model. The mathematical description of the models are provided and several properties reported. A numerical comparison of the two models on real industrial data is provided with discussion about possible drawbacks of simplifying modeling procedures.

Keywords: Data Envelopment Analysis, BCC Model, Range Directional Model

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Question of measuring productive efficiency is classically based on the use of the Pareto efficiency notion to define the production function. Traditionally, the inefficiencies in input/output usage are neglected and total production growth is represented by shift in technologies. The pioneering works of Shephard [10] and Farrell [9] in measuring the productive efficiency using all inputs and outputs in order to prevent the index number problem (inadequacy of separate indices for labor and capital productivity), and in introducing the conceptual use of various types of efficiencies, were followed by a successful attempt to compute the productivity efficiency using a linear optimization model by Charnes, Cooper, and Rhodes [5]. Their non-parametric approach to estimate the production function as the efficiency frontier made up as the boundary of the convex hull of data points adopted the name of *Data Envelopment Analysis* (DEA) and spread around the scientific world quickly.

The classical CCR DEA model implicitly assumes constant returns-to-scale and continuous linear production possibility set. Many extensions to this original model were adopted, e. g. the widely used variable returns-to-scale model [1], discrete and continuous additive models [2], [4], slack-based measure models [11], or stochastic extensions [6]. We refer the reader to the monographs [8] and [7] for further information.

In our contribution, we want to point out that the classical DEA models are not able to work with negative data. We want emphasize, in particular, that units with negative data cannot be simply dropped from investigation as the results will become distorted. In Section 2, we provide definitions of basic notions and several DEA models used for comparison. Section 3.2 present some insights to numerical results based on real dataset, discussed and concluded in Section 4.

2 DEA Models

2.1 Production Possibility Set

Consider K decision-making units denoted $DMU_k,\ k = 1,\dots,K$. Each unit k is characterised by a collection of m inputs $x_{ik},\ i = 1,\dots,m$ and outputs $y_{jk},\ j = 1,\dots,n$; the input and output matrix are then denoted $X = (x_{ik})$ and $Y = (y_{jk})$, respectively. For the sake of convenience, let also denote $x_k = (x_{1k},\dots,x_{mk})^T$ and $y_k = (y_{1k},\dots,y_{nk})^T$ the input and output vectors of DMU_k. A unit under actual investigation (for which the efficiency is evaluated) is denoted DMU_0 thorough the paper.

The set of all combinations of allowed input and output vectors are known as *production possibility set* (*PPS*) and its correct specification plays an elementary role in DEA analysis. In general, *PPS* is defined through

$$PPS = \{(x,y) \mid y \text{ may be produced from } x\}. \tag{1}$$

A unit may be then characterized as efficient in Pareto–Koopmans dominance sense with respect to such defined production possibility set:

[1] University of South Bohemia in České Budějovice, Faculty of Economics, Department of Applied Mathematics and Informatics, Studentská 13, České Budějovice, and Czech Academy of Sciences, Institute of Information Theory, Pod Vodárenskou věží 4, Praha 8, Czech Republic, houda@ef.jcu.cz

Definition 1. *DMU_1 dominates DMU_2* if $x_1 \leq x_2, y_1 \geq y_2$, and at least one (one-dimensional) inequality is strict.

Definition 2. *DMU_0* is *efficient* with respect to *PPS* if there is no (real or virtual) unit with $(x,y) \in PPS$ dominating DMU_0.

2.2 BCC Model

Giving an example, Banker, Charnes, and Cooper's (BCC) model [1] assuming variable returns to scale is related to the following *continuous convex PPS*

$$PPS_C = \left\{ (x,y) \mid x \leq X\lambda,\ y \geq Y\lambda,\ \sum_k \lambda_k = 1,\ \lambda \geq 0 \right\} \tag{2}$$

It is not hard to see that PPS_C represents the convex hull of all available input-output data points. Let s^- and s^+ be the slack (surplus) for the inequalities $X\lambda \leq \theta x_0$ and $Y\lambda \geq y_0$ with some $\theta \in \mathbb{R}$. The efficiency of DMU_0 with respect to PPS_C may verified solving the *BCC input oriented* model (in envelopment form)

$$\begin{aligned} \min \theta + \epsilon \left(\sum_i s_i^+ + \sum_j s_j^+ \right) & \text{ subject to} \\ X\lambda + s^- &= \theta x_0 \\ Y\lambda - s^+ &= y_0 \\ \sum_k \lambda_k = 1, \lambda \geq 0, s^-, s^+ &\geq 0,\ \theta \text{ unconstrained}\ , \end{aligned} \tag{3}$$

where ϵ is so-called non-Archimedean infinitesimal (a positive number smaller than any other positive number). DMU_0 is efficient with respect to PPS_C if the optimal solution to (3) has $\theta^* = 1$ and $s^{-*} = s^{+*} = 0$, implying that DMU_0 is lying on the boundary of PPS_C and is (in fact) an extreme point of it.

Remark 1. Although the model (3) is not linear in principle due to the infinitesimal part, it may be solved by two-stage procedure with two linear optimization problems, see e. g. [8]. The infinitesimal part is present to ensure that some boundary points with $\theta^* = 1$ but non-zero slacks (also known as weakly efficient) are excluded from the final optimal solution.

2.3 Directional Distance Model

A single factor θ in (3) ensures that all the inputs are improved proportionally when projecting onto the efficient frontier. This is the property that characterizes so-called *radial* DEA models: the actual efficiency of the unit is furthermore a proportion of input values of DMU_0 and its peer unit. This implies a main drawback of the presented model: it may work well only if the data—the matrices X and Y—contain only positive elements. With negative data, the efficient measure is not well defined and the model cannot give appropriate results.

To generalize the radial feature and overcome the inability to work with negative data, Chambers, Chung, and Färe [3] proposed a so-called *directional distance model*:

$$\begin{aligned} \max \beta & \text{ subject to} \\ X\lambda &\leq x_0 - \beta g^x \\ Y\lambda &\geq y_0 + \beta g^y \\ \sum_k \lambda_k = 1, \lambda &\geq 0,\ \beta \geq 0 \end{aligned} \tag{4}$$

where g^x and g^y are pre-specified vectors of improvement directions and β is called *directional distance measure*. DMU_0 is efficient with respect to PPS_C if $\beta^* = 0$ in (4). The input oriented BCC model is then seen as a special case of (4) with $g^x = x_0, g^y = 0$ and $\theta = 1 - \beta$.

2.4 Range Directional Model

The free choice of possible improvement directions g^x and g^y gives the decision maker a great flexibility to represent many particular kinds of preference strengths inside input and output vectors. On the other hand, this may represent a shortcoming too, especially in the case of missing or limited information about the nature of the data (the input/output improvement directions should be given in advance). This potentially unwanted freedom may be solved by choosing a particular reference point I (sometimes called "ideal point") with respect to which the improvements are considered. A frequent choice for such point is $I = (\min_k x_{\cdot k}; \max_k y_{\cdot k})$ (the minimums of inputs and maximums of outputs are taken component-wise in the above notation). The improvement directions are then defined by

$$g^x = x_0 - \min_k x_{\cdot k},$$
$$g^y = \max_k y_{\cdot k} - y_0,$$

and the corresponding directional distance model, called the *range directional model*, takes the form

$$\begin{aligned} &\max \beta \text{ subject to} \\ &\quad X\lambda \leq (1-\beta)x_0 + \beta \min_k x_{\cdot k} \\ &\quad Y\lambda \geq (1-\beta)y_0 + \beta \max_k y_{\cdot k} \\ &\sum_k \lambda_k = 1, \lambda \geq 0, \beta \geq 0 \end{aligned} \tag{5}$$

Again, DMU_0 is efficient with respect to PPS if the optimal solution of (5) is $\beta^* = 0$. (Adaptation of the two-stage technique to deal with non-efficient boundary points is straightforward and we will not provide additional details here.) The range directional model is still radial but the reference point is now I and not the origin as in the BCC model. Furthermore, the particular choice of I ensures that the improvement directions are meaningfully defined for all kind of the data and there is no need to restrict the DEA investigation to nonnegative inputs and outputs only.

3 Numerical Illustration

In this section we will compare two particular DEA models in order to point out the danger of bad model specification when working with real economic data and problems.

3.1 Example Setting

We have considered annual accounts of 380 Czech companies from the food industry (NACE C.10) from the year 2014. For the purpose of this paper, wee have chosen to evaluate the companies using

1. the input oriented model with variable returns to scale (BCC model), and
2. the range directional model (RD model).

The analysis was based on four inputs: *SPMAAEN* (material and energy consumption), *ON* (personnel costs), *STALAA* (fixed assets), and *POSN* (percentage of personnel costs); and two outputs *VYKONY* (business performance), and *ROA* (return on assets). Among 380 companies, 89 reported no material and energy costs and were removed from investigation. Further, 70 companies have negative *ROA* and cannot be analysed using BCC model. This resulted into 244 feasible observations for the BCC model and 291 feasible observation for the RD model.

3.2 Numerical Results

Among the 244 companies, 22 of them (9%) were evaluated as BCC efficient; additional three companies have the efficiency score higher than 95%. The alternative RD model evaluated only 10 companies (3.4%) to be efficient. The histograms of efficiency scores for both models are given in Figure 1 (BCC model is on the left hand side, RD model on the right hand side of the figure).

Furthermore, only four from nine NACE subgroups comprise an RD efficient company, while the distribution of BCC efficient companies is more widespread. Figure 2 provides an additional insight into the relationship between the subgroup size and the relative number of efficient companies in the overall model.

Another interesting comparison of the two models was made for selected input and output of both models. Figure 3 provides the empirical distribution plot for the personnel cost percentage (variable *POSN*). Note, in particular, the noticeable difference in distribution of this input for efficient (cyan colored) units – units with small percentages were marked more often as efficient in BCC model. Similar plot is provided as Figure 4 for

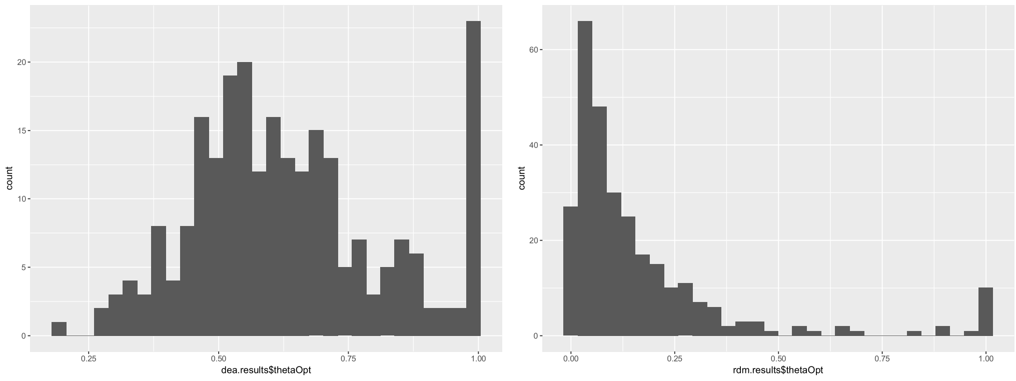

Figure 1 Efficiency Scores for BCC and RD Models

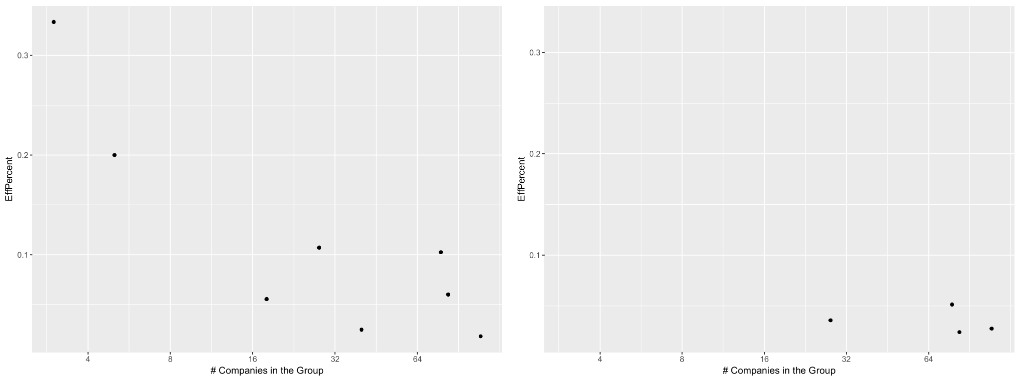

Figure 2 Relationship between the Size of the Group and the Relative Number of Efficient Units

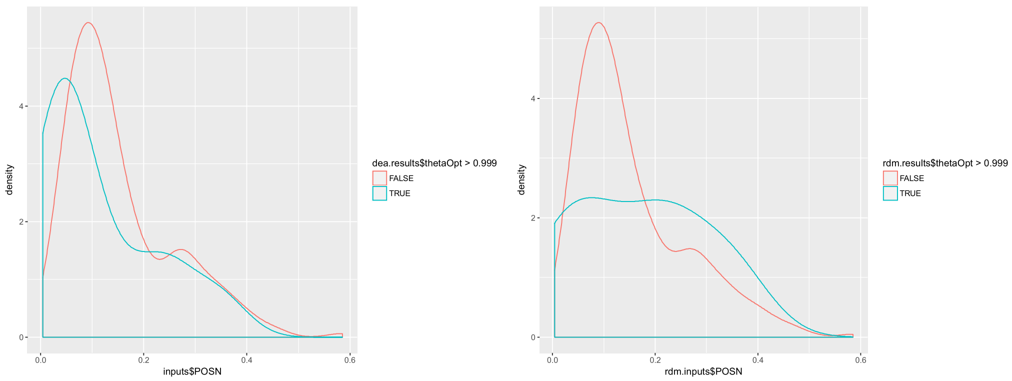

Figure 3 Empirical Distribution of Personnel Cost Percentage

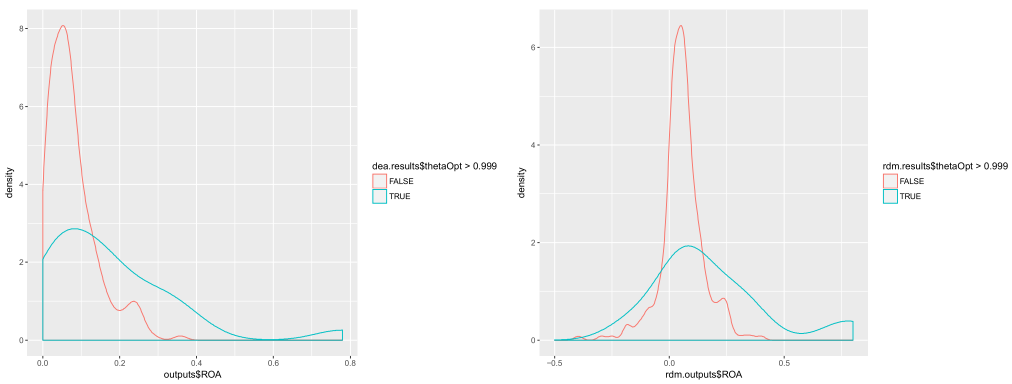

Figure 4 Empirical Distribution of Return on Assets

the return on assets (variable *ROA*). For the BCC model, all units with negative *ROA* must be excluded from the computations.

A scatter plot of a (partial) relationship of the input and output mentioned above is provided in Figure 5. Interestingly, one unit with negative *ROA* and average *POSN* was made efficient by RD model. Again, only the upper part of the right hand side plot (units with *ROA* above zero) is comparable with the left hand side plot.

Figure 5 Relationship between the Personnel Cost Percentage and Return on Assets

4 Discussion and Conclusion

The comparison made in Section 3.2 clearly demonstrates that choice of a good model in production analysis is crucial. For example, the distribution of the optimal efficiency θ^* among the investigated units shows that BCC model (probably inadequately) overvalued efficiencies of the remaining units (with positive *ROA*). Notice that only a small part of the units exceeded the efficiency of 0.50 for the RD model. Furthermore, analysing Figure 5 one may notice a great number of BCC efficient units with relatively small *ROA* which were not marked as efficient by the RD model.

To conclude: simple deleting the observations which do not conform to the assumption of the chosen model (as done, for example, by the BCC case) ineluctably leads to false conclusions concerning efficiency of the remaining units and any subsequent results.

Acknowledgements

This paper was partially supported by the Czech Science Foundation This research was partially supported by the Czech Science Foundation under Grant No. GA18-02739S (Stochastic Optimization in Economic Processes).

References

[1] Banker, R. D., Charnes, A., and Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30** (1984), 1078–1092.

[2] Bowlin, W. F., Brennan, J., Cooper, W. W., and Sueyoshi, T. *DEA models for evaluating efficiency dominance*. Research report, The University of Texas, Center for Cybernetic Studies, Austin, Texas, 1984.

[3] Chambers, R. G., Chung, Y. H., and Färe, R. Benefit and distance functions. *Journal of Economic Theory* **70** (1996), 407–419.

[4] Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M., and Stutz, J. D. Foundations of data envelopment analysis for pareto-koopmans efficient empirical production functions. *Journal of Econometrics* **30** (1985), 91–107.

[5] Charnes, A., Cooper, W. W., and Rhodes, E. Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–444.

[6] Cooper, W. W., Deng, H., Huang, Z., and Li, S. X. Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. *Journal of the Operational Research Society* **53** (2002), 1476–9360.

[7] Cooper, W. W., Seiford, L. M., and Joe, Z., eds. *Handbook on Data Envelopment Analysis*, volume 164 of *International Series in Operations Research & Management Science*. 2 edition. Springer, New York, 2011.

[8] Cooper, W. W., Seiford, L. M., and Tone, K. *Data Envelopment Analysis*. 2 edition. Springer, New York, 2007.

[9] Farrell, M. J. The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)* **120** (1957), 253–290.

[10] Shephard, R. W. *Cost and Production Functions*. Princeton University Press, Princeton, 1953.

[11] Tone, K. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* **130** (2001), 498–509.