ISSN: 2278-3369
International Journal of Advances in Management and Economics
Available online at www.managementjournal.info

RESEARCH ARTICLE

Testing the Validity and Structure of the Data Envelopment Analysis PISA Scores

A Mete Çilingirtürk[1]*, Münevver Turanlı[2]

[1]*Department of Econometrics, Marmara University, Turkey.*

[2]*Department of Statistics, İstanbul Commerce University, İstanbul, Turkey.*

*Corresponding Author: E-mail: ***acilingi@marmara.edu.tr***

Abstract

This research aims to take efficiency as a black box between input and output, or a property of the process, and suggest an alternative way to evaluate the validity of Data Enveloping Analysis (DEA) scores. Over the last ten years, DEA has become a commonly used measure of efficiency of processes. It is offered as a non-parametric method of output oriented optimization, which is solved without any assumption under the restriction of the convexity with metric input and output measurements. Alternatively, output models like Cobb-Douglas or advanced econometric models are valid under several assumptions, and efficiency requires a significance test. Similar processes with the same input level exhibit a different level of output due to efficiency. This means that output is a function of input and efficiency. The present study models the DEA scores of an education efficiency research study by applying path analysis. Efficiency is found as moderator variable instead the expectation of mediation effect of an educational input-output process. The findings and interpretation of the structure depend on the data and its coverage, which do not lead to generalization of the efficiency concept. However, the main insight of the method is to consider real processes, not simulated model data. The purposed method enables the statistical testing of calculated DEA efficiency scores. Processes might be updated through the structured interpretation of how efficiency affects the output.

Keywords: *Data enveloping analysis, Path analysis, Education, efficiency, Score validity, Significance test.*

Introduction

Performance measurement is an analytical process for the evaluation of products or services together with the aimed goals by a responsible decision unit. It has become an accepted method of strategic management aimed at determining the level of compensation of consumed resources and produced goods and services relative to the intended goals. Effectiveness is defined as the ratio of outputs to inputs within a specific system, whereas efficiency is defined as the accessibility of the highest output with the lowest input. That means that efficiency is an intermediary effectiveness maximizing process.

The concept of performance requires a variety of measurement methods according to the activity type. Performance measurement may be divided into two types, namely, ratio analysis and efficiency frontier approaches. Data Envelopment Analysis (DEA) came into prominence as an efficiency frontier method applicable to various areas with these methodological developments.

Output is taken as a function of the input and efficiency according to this explanation, like the well-known Cobb-Douglas production function. In this case, DEA efficiency scores should explain the output level together with the input level by regression analysis. However, the model structure fits to a mediated or moderated regression.

DEA Model

A nonparametric and linear programming based method, DEA was devised to measure the relative input-output efficiency of a set of decision making units (DMU). It can apply directly to a multiple inputs and outputs system, where prices are not exactly determined. It differs from regression analysis due its frontier model development. DEA develops the frontier according to the most efficient units, whereas regression techniques model the average efficiency. The primary form of DEA, Charnes, Cooper, and Rhodes (CCR), assumes constant returns to scale (CRS) [1] while the Banker, Charnes and Cooper (BCC) model assumes variable returns to scale (VRS) [2] There are also multiplicative and additive DEA models appeared in the literature [3]. Furthermore, Andersen and Peterson [4] suggested the super efficiency model to order to maximize the efficiency of DMUs; while Sexton et. al. [5] suggested cross efficiency matrix and Li and Reeves [6] multi object DEA models.

This paper presents a returns to scale model of the k^{th} DMUs in a system with m inputs, s outputs and n units according to the CCR approach [7]. The scores h_k are solved through n linear programming problems:

$$\begin{aligned}
&\max h_k = \sum_{r=1}^{s} u_{rk} Y_{rk} \\
&\sum_{r=1}^{s} u_{rk} Y_{rj} - \sum_{i=1}^{m} v_{ik} X_{ij} \le 0 \\
&\sum_{i=1}^{m} v_{ik} X_{ik} = 1 \\
&u_{rk}, v_{ik} \ge 0 \\
&k = 1,2,\ldots,n;\ r = 1,2,\ldots,s;\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,n
\end{aligned} \tag{1}$$

The efficiency h is the ratio of the u weighted sum of the outputs (Y) to the v weighted sum of the inputs (X). The efficiency scores are solved in the same way for the input as for the output maximization. The VRS model is obtained by adding a convexity constraint to the CCR model:

$$\begin{aligned}
&\max h_k = \sum_{r=1}^{s} u_{rk} Y_{rk} - u_0 \\
&\sum_{r=1}^{s} u_{rk} Y_{rj} - u_0 - \sum_{i=1}^{m} v_{ik} X_{ij} \le 0 \\
&\sum_{i=1}^{m} v_{ik} X_{ik} = 1 \\
&u_{rk}, v_{ik} \ge 0 \\
&k = 1,2,\ldots,n;\ r = 1,2,\ldots,s;\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,n
\end{aligned} \tag{2}$$

Efficiency is defined as a linear measurement of inputs and outputs. Also the first restriction in both models could be rewritten as output being a function of efficiency and inputs. A further disadvantage of the model is that it cannot be tested for the best specification [8].

Mediation and Moderation Efficiency Models

Generally, for social research purposes the identification of causal effects is represented by P(y|x). This structure determines the sensitivity of the dependent variable Y on the changes of the explanatory variable X when all other factors are constant. Structural equation models are an appropriate technique to identify the direct effect X→Y and also the total and indirect path-specific effects of X on Y [9-12]. Despite this technique, however, mediated and moderated effects are also discussed in the social sciences [13-17]. The reason for this is that the interpretation of the causal parameters becomes confused with the linear regression coefficients [11, 18]. Fig. 1 presents the mediation and moderation models.

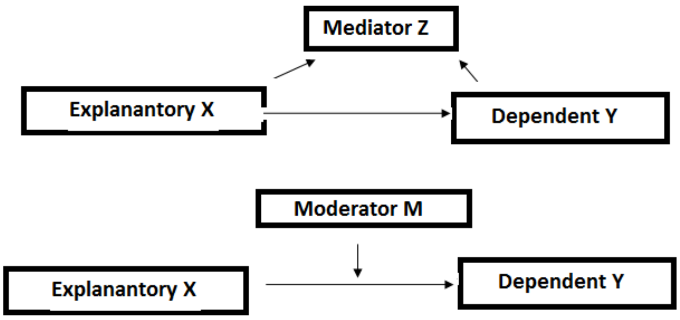

Fig. 1. Mediation and moderation models [19]

The mediation effect model is defined statistically for direct DE and indirect IE effects thus:

$$DE = \sum \left[E(Y|x+\Delta x, Z) - E(Y|x, Z)\right] P(Z|x) \tag{3}$$
$$IE = \sum E(Y|x, Z)\left[P(Z|x+\Delta x) - P(Z|x)\right] \tag{4}$$

As Equations (3) and (4) can be applied to any distribution, both type of effects can be proved for parametric and non-parametric regression models [20, 21]. Baron and Kenny [14] proposed a four-stage procedure for testing the impact of mediation:

- $Y=a_0+a_1X+u_1$ tested for statistical significance, and $R^2_{Y \bullet X}$ computed;
- $Z=b_0+b_1X+u_2$ tested for statistical significance;
- $Y=c_0+c_1X+c_2Z+u_3$ tested for statistical significance, and $R^2_{Y \bullet XZ}$ computed;
- If c_1 is not statistically significant, then Z is said to be mediator variable; otherwise it is partial mediation model.

Of course, there are exceptions in some of these stages. Multicollinearity causes an insignificant c_2 even if the model is mediated. Also, the existence of the mediator variable effects the XY relation negatively, when the sign of c_1 is reverse of the the sign of the b_1c_2 multiplication, referred to as “inconsistent mediation,” with Z a “suppressor variable” [17]. The marginal effect of determination is tested through

$$F = \frac{\left(R^2_{Y\bullet XZ} - R^2_{Y\bullet X}\right)/(k_2 - k_1)}{\left(1 - R^2_{Y\bullet XZ}\right)/(n - k_2 - 1)} \qquad (5)$$

Combined coefficients tests are suggested for mediation effects [22], where level of significance should be assumed at 0.0253 to keep type-I α error level constant.

Sobel's test : $$z = \frac{a_1 c_2}{\sqrt{c_2^2 S_{a_1}^2 + a_1^2 S_{c_2}^2}} \qquad (6)$$

Aroian's test : $$z = \frac{a_1 c_2}{\sqrt{c_2^2 S_{a_1}^2 + a_1^2 S_{c_2}^2 + S_{a_1}^2 S_{c_2}^2}} \qquad (7)$$

Goodman's test: $$z = \frac{a_1 c_2}{\sqrt{c_2^2 S_{a_1}^2 + a_1^2 S_{c_2}^2 - S_{a_1}^2 S_{c_2}^2}} \qquad (8)$$

Sobel and Aroian tests give successful results by Monte Carlo simulations with sample sizes greater than 50. The coefficient [1-($a_1c_2c_1$)] also measures the theoretical impact of the mediation. The moderator effect renews the power and also the direction of the causal relation between X and Y [23]. It weakens the causal effect of X and it is called the "fully moderated model" when this effect is expired [24]. A moderation effect regression model is thus built in addition to direct effect models $Y=a_0+a_1X+u_1$ and

$Y=d_0+d_1X+d_2M+u_4$:
$Y=g_0+g_1X+g_2M+g_3MX+u_5$ (9)

Various procedures are suggested to measure and test the moderation effect [25]:

- M will be not a moderator variable if the magnitudes and signs are compatible with $g_2\neq0$ and $g_3=0$;
- M will be a full moderator variable if the direct effect multiple regression model $R^2_{Y\bullet XZ}$ is statistically significant than model (9) according to Equation (5), as well as $g_2=0$ and $g_3\neq0$;
- M will be quasi-moderator variable if the direct effect multiple regression model $R^2_{Y\bullet XZ}$ is statistically significant than model (9) according to Equation (5), as well as $g_2\neq0$ and $g_3\neq0$.

Testing the Efficiency Scores of Secondary Data

The paper suggests a way of testing the non-parametric DEA scores, which represents the process efficiency structuring the input-output relation. A DEA research results on national education was applied to present the testing process.

The success and quality of a national education system will be affected by the government spending in social states. Efficiency in education is subject to frequent measurement by academicians and researchers with institutional and international comparisons [26-31]. Contemporary international evaluations employ PISA research results, covering the OECD countries [32-35]. PISA (Programme for International Student Assessment) began with the participation of 43 countries in 2000 and by 2009 had spread to over 68 countries. The program depends on student self-assessment in reading, math, and science skills, which are assumed to be three educational outputs. This study used the DEA scores of the research paper by Koçak and Çilingirtürk [35]. The common usage of DEA and factor analysis reduces measurement errors and input-output dimensions [36]. The output variable Y is obtained through exploratory factor analysis as standardized normal scores, which explains 90.43% (λ= 2,713) of the learning skills.

Table 1: Factor component matrix

Summary statistics		PISA scores	PISA common factor
Cronbach's α	0,969	PISA Reading	, 922
KMO	0,7670	PISA Mathematics	, 953
Bartlett χ²	124,51	PISA Science	, 977

*Prepared by the authors.

The estimated five models are presented in Table 2, where public education spending as a percentage of total GDP is independent variable X, and DEA scores are the mediator or moderator variable M.

Table 2: Regression estimates with standardized beta coefficients and summary statistics

Model	Constant	X		M		XM			
	β	β	t/p	β	t/p	β	t/p	R²	F/(p)
$Y=a_0+a_1X+u_1$	-1,21 (0,85)	0,35 (0,24)	1,45 0,16					0,058	2,10 (0,16)
$M=b_0+b_1X+u_2$	98,35 (3,53)	-1,33 (0,995)	-1,34 0,19					0,050	1,79 (0,19)
$Y=m_0+m_1M+e$	-16,85 (2,60)			0,18 (0,03)	6,48 0,00			0,553	42,03 (0,00)
$Y=c_0+c_1X+c_2M+u_3$	-21,18 (2,26)	0,62 (0,13)	4,61 0,00	0,20 (0,02)	9,01 0,00			0,711	44,13 (0,00)
$Y=g_0+g_1X+g_2M+g_3MX+u_5$	33,29 (9,47)	-15,89 (2,83)	-5,62 0,00	-0,37 (0,10)	-3,73 0,00	0,175 (0,03)	5,84 0,00	0,856	70,28 (0,00)

*Prepared by the authors.

A mediation effect is not observed as the first and second models are not statistically significant according to mediation effect tests (Table 3).

Table 3: Mediation effect significance tests

	the z-test statistic	The standard error	p-value
Sobel test	-1,3236	0,2043	0.186
Aroian test	-1,3153	0,2056	0.200
Goodman test	-1,3319	0,2030	0.183

*Prepared by the authors.

DEA scores have a significant effect in both models (3) and (4). Furthermore according to Equation (5), there is a significant change at the effect (F=74.56; p=0.00). When the marginal effect of the fifth model is tested against the direct effect model (4) according to Equation (5), the importance of the interaction XM is observed (F=33.94; p=0.00). Efficiency presents a suppressing quasi moderation effect on the educational input-output system [37-40].

Conclusion

Efficiency is usually defined as a mediating black box between input and output. However, DEA efficiency scores should modeled and tested for interpretation purposes and path specific structures. According to the results of this paper, efficiency in terms of DEA scores, in fact, have a moderating effect.

The suggested procedure could used to determine what exactly the DEA measures. Furthermore, the procedure tests the statistical significance of the DEA scores in a predefined input-output model.

References

1. Charnes, A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. European J. Operational Research, 2:429-444.
2. Boussofiane A, Dyson RG, Thanassoulis E (1991) Applied data envelopment analysis. European J. Operational Research, 52:1-15.
3. Seiford LM (1999) A cyber-bibliography for data envelopment analysis (1978–1999), in Cooper, W.W. and Seiford, L.M. (Ed.), Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Sover software, Kluwer Academic Publishers.
4. Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. Management Science, 39:1261-1264.
5. Sexton TR, Silkman RH, Hogan AJ (1986) Data Envelopment Analysis: Critique and Extensions, in Silkman, R.H. (Ed.), Measuring Efficiency: An Assessment of Data Envelopment Analysis, Jossey-Bass, San Francisco, CA.
6. Li X, Reeves GR (1999) A Multiple Criteria Approach to Data Envelopment Analysis. European J. Operational Research, 115:507-517.
7. Charnes A, Cooper WW, Rhodes E (1981) Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. Management Science, 27(6):668-697.
8. Berg S (2010) Water Utility Benchmarking: Measurement, Methodology, Performance Incentives", International Water Association.
9. Alwin D, Hauser R (1975) The decomposition of effects in path analysis. American Sociological Review, 40:37-47.
10. Graff J, Schmidt P (1982) A general model for decomposition of effects, in Joreskog K and Wold H (Ed.), Systems Under Indirect Observation: Causality, Structure, Prediction, North-Holland, Amsterdam, pp. 131-148.
11. Sobel M (2008) Identification of causal parameters in randomized studies with mediating variables", Journal of Educational and Behavioral Statistics, 33, pp. 230-231.
12. Bollen K (1989) Structural Equations with Latent Variables, John Wiley, New York.
13. Judd C, Kenny, DA (1981) Process analysis: Estimating mediation in treatment evaluations. Evaluation Review, 5:602-619.
14. Baron R, Kenny D (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J. Personality and Social Psychology, 51:1173-1182.
15. Shrout P, Bolger N (2002) Mediation in experimental and nonexperimental studies: New procedures and recommendations", Psychological Methods, 7:422-445.
16. Muller D, Judd C, Yzerbyt V (2005) When moderation is mediated and mediation is moderated. J. Personality and Social Psychology, 852-863.
17. MacKinnon DP, Fairchild AJ, Fritz MS (2007) Mediation analysis. Annual Review of Psychology, 58:593-614.
18. Holland P (1995) Some reflections on freedman's critiques. Foundations of Science, 1:50-57.
19. Pearl J (1998) Graphs, causality, and structural equation models. Sociological Methods and Research, 27:226-284.
20. Pearl J (2010) The Mediation Formula: A guide to the assessment of causal pathways in non-linear models, UCLA Computer Science Department, Technical Report R-363, January 2011.
21. Imai K, Keele L, Yamamoto T (2010) Identification, inference, and sensitivity analysis for causal mediation effects. Statistical Science, 25(1):51-71.
22. MacKinnon DP, Warsi G, Dwyer JH (1995) A simulation study of mediated effect measures. Multivariate Behavioral Research, 30:41-62.

23. Judd CM, Kenny, DA (2010) Data analysis in social psychology: Recent and recurring issues, in Fiske S, Gilbert D and Lindsay G (Ed.), The handbook of social psychology, John Wiley and Sons.

24. Kramer HC, Wilson GT, Fairburn CG, Agras WS (2002) Mediators and moderators of treatment effects in randomized clinical trials. Archives of General Psychiatry, 59:877-883.

25. Frazier PA, Tix AP, Barron KE (2004) Testing moderator and mediator effects in counseling psychology research. J. Counseling Psychology, 51:115-134.

26. Gupta, S, Verhoeven, M. (2001) The efficiency of government expenditure – experiences from Africa. J. Policy Modelling, 23:433-467..

27. Clements B (2002) How efficient is education spending in Europe? European Review of Economics and Finance, 1:3-26.

28. St. Aubyn M (2003) Evaluating efficiency in the Portuguese education sector, Economia, 26:25-51.

29. Abu-Ghaida D (2007) Education, in Gray C, Lane T, Varoudakis A (Ed.), Fiscal policy andeconomic growth. Lessons for Eastern Europe and Central Asia, The World Bank, Washington, D.C., pp. 147-178.

30. Afonso A, Schuknecht L, Tanzi V (2004) Public sector efficiency: An international comparison. Public Choice, 123:321-347.

31. Özden ÜH (2010) Veri zarflama analizi (VZA) ile Türkiye'deki vakıf üniversitelerinin etkinliğinin ölçülmesi. Istanbul University J. the School of Business Administration, 39(1):167-185.

32. Fakin B, Crombrugghe A (1997) Fiscal adjustment in Transition Economies: Social transfers and the efficiency of public spending, a comparison with OECD countries", Working paper 1803, World Bank, Washington.

33. Afonso A, St. Auby, M (2005) Non-Parametric approaches to education and health efficiency in OECD countries. J. Applied Economics, 8(2):227-246.

34. Sutherland D, Price R, Joumard I, Nicq C (2007) Performance indicators for public spending efficiency in primary and secondary education, Working paper 546, OECD Economics Department.

35. Koçak H, Çilingirtürk, AM (2011) Efficiency Analysis of OECD Public Education Spending, in Annual International Conference on Operations Research and Statistics -ORS 2011, Malaysia, 2011:83-88.

36. Yıldırım IE (2010) Veri zarflama analizinde girdi ve çıktıların belirlenmesindeki kararsızlık problemi için temel bileşenler analizine dayalı bir çözüm önerisi. Istanbul University & J. the School of Business Administration, 39(1):141-153.

37. Eugène B (2008) The efficiency frontier as a method for gauging the performance of public expenditure: A Belgian case study", Working paper 138, National Bank of Belgium, Belgium.

38. MacKinnon DP (2008) Introduction to Statistical Mediation Analysis, Lawrence Erlbaum Associates, New York.

39. Pearl J (2009) Causal inference in statistics: An overview. Statistics Surveys, 3:96-146.

40. Sobel M (1987) Direct and indirect effects in linear structural equation models. Sociological Methods & Research, 16(1):155-176.