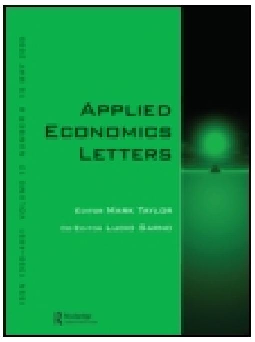

Applied Economics Letters

ISSN: 1350-4851 (Print) 1466-4291 (Online) Journal homepage: https://www.tandfonline.com/loi/rael20

Assessing health efficiency across countries with a two-step and bootstrap analysis

António Afonso & Miguel St. Aubyn

Applied Economics Letters, 2011, **18**, 1427–1430

Routledge
Taylor & Francis Group

Assessing health efficiency across countries with a two-step and bootstrap analysis

António Afonso[a,b] and Miguel St. Aubyn[a,*]

[a] *Research Unit of Complexity and Economics, Department of Economics, School of Economics and Management, Technical University of Lisbon, Lisbon, Portugal*
[b] *Fiscal Policies Division, Directorate General Economics, European Central Bank, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany*

Regressing Data Envelopment Analysis (DEA) output efficiency scores on nondiscretionary variables, with a two-stage DEA/Tobit and bootstrap procedures, we show that health inefficiency in Organization for Economic Co-operation and Development (OECD) countries is related to Gross Domestic Product (GDP) per head, education level, obesity and smoking habits.

I. Introduction

Health is one of the most important services provided by governments. Generally, health provision is efficient if its producers make the best possible use of available inputs, and the sole fact that health inputs weigh heavily on the public purse would call for a careful efficiency analysis. Existing research suggested that important inefficiencies are at work for public expenditure in the Organization for Economic Co-operation and Development (OECD), notably for health (Evans *et al*., 2000; Afonso and St. Aubyn, 2005; Afonso *et al*., 2005; Spinks and Hollingsworth, 2009).

Using a two-stage Data Envelopment Analysis (DEA)/Tobit and a bootstrap procedure, we show that most OECD countries perform below the frontier, and that inefficiency in health is strongly related to factors that are in the short to medium run, beyond the control of governments. These variables are Gross Domestic Product (GDP) per capita, education level, obesity and smoking habits.

II. Methodology

We employ the more usual DEA/Tobit approach and the bootstrap procedures suggested by Simar and Wilson (2007).[1] The fact that DEA output scores are biased, and that the environmental variables are correlated to output and input variables, recommend the use of bootstrapping techniques. Our article is one of the first applications of this recent technique.

DEA assumes the existence of a convex production frontier. An output-oriented study estimates by how much outputs can be increased without changing inputs, the perspective taken here.[2]

The linear programming problem assuming variable returns to scale is sketched in Equation 1 for the i-th Decision Making Unit (DMU):

$$\begin{aligned} &\text{Max}_{\lambda,\delta_i}\delta_i \\ \text{subject to } &\delta_i y_i \leq Y\lambda \\ &x_i \geq X\lambda \\ &n1'\lambda = 1 \\ &\lambda \geq 0 \end{aligned} \qquad (1)$$

*Corresponding author. E-mail: mstaubyn@iseg.utl.pt
[1] See Simar and Wilson (2007).
[2] See Farrell (1957) and Coelli *et al*. (2005).

Applied Economics Letters ISSN 1350–4851 print/ISSN 1466–4291 online © 2011 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/13504851.2010.541149

There are p inputs and q outputs for n DMUs, y_i being the vector of outputs and x_i the vector of inputs. X is the $(p \times n)$ input matrix and Y is the $(q \times n)$ output matrix. δ_i, a scalar, satisfies $\delta_i \geq 1$. This efficiency score measures technical efficiency of the i-th unit as the distance to the efficiency frontier (a linear combination of best practice observations). With $\delta_i > 1$, the DMU is below the estimated frontier (inefficient), whereas $\delta_i = 1$ implies that the DMU is on the estimated frontier (efficient). λ is a $(n \times 1)$ vector of constants that measures the weights used to compute the location of an inefficient DMU if it were to become efficient, and 1 is a vector of ones.

As nondiscretionary and discretionary inputs jointly contribute to each DMU's outputs, the use of two-stage models is suitable.[3] Let z_i be a $(1 \times r)$ vector of nondiscretionary variables:

$$\hat{\delta}_i = z_i\beta + \varepsilon_i \tag{2}$$

where $\hat{\delta}_i$ is the efficiency score from solving Equation 1. β is a $(r \times 1)$ vector of parameters to be estimated in step two associated with each nondiscretionary input. As $\hat{\delta}_i \geq 1$, researchers estimate Equation 2 using censored regression techniques (Tobit).

The two-stage DEA/Tobit method is likely to be biased in small samples. Firstly, because output scores are jointly estimated by DEA, the error term ε_i in Equation 2 is serially correlated. Secondly, nondiscretionary variables z_i are correlated with the error term ε_i, because these inputs are correlated to the outputs, and consequently to estimated efficiency scores. Therefore, Simar and Wilson (2007) propose bootstrap methods, which we use in this article.[4]

III. Empirical Analysis

OECD (2005) is our health database for OECD countries.[5] Output is measured by life expectancy and infant mortality to assess potential years of added life.

Because outputs are measured in such a way that 'more is better', and given that Infant Mortality Rate (IMR) is equal to: (Number of children who died before 12 months)/(Number of born children) $\times$ 1000, we have calculated an 'Infant Survival Rate', ISR,

$$ISR = \frac{1000 - IMR}{IMR} \tag{3}$$

which is the ratio of children that survived the first year to the number of children that died; therefore, it increases with a better health status.

We have also considered the Potential Years of Life *Not* Lost, $PYLNL$, computed on the basis of the indicator Potential Years of Life Lost, $PYLL$ (OECD, 2005). $PYLL$ equals the number of life years lost due to all causes before the age of 70 and that could be, *a priori*, prevented. Therefore, a transformation provides an increasing monotonic relation between $PYLNL$ and health status:

$$PYLNL = \lambda - PYLL \tag{4}$$

where λ = 3 618 010 is an estimate of the number of potential years of life per 100 000 population under 70 years.[6]

Inputs are the number of practising physicians, nurses, acute care beds per thousand habitants and Magnetic Resonance Imagers (MRI). Data are averaged for the period 2000 to 2003.

To deal with eventual difficulties posed when there are too many inputs and/or outputs, we used Principal Component Analysis (PCA). Applying PCA to the input variables, we used the first three principal components, which explains 88% of the variation of the four variables. For the output variables, we selected the first principal component, accounting for 84% of the variation of the three variables.

Table 1 reports the results for the DEA variable-returns-to-scale technical efficiency output scores. Seven countries would be located on the production possibility frontier: Canada, Finland, Japan, Korea, Spain, Sweden and the United States.[7] Canada, Finland, Japan, Spain and Sweden perform quite well in the output indicator, getting above-average results. On the other hand, Korea and the United States are generally below average regarding the use of resources in all the first three components selected. Hungary, the Slovak Republic and Poland are located on the opposite end. Their output could be substantially increased if they were to become located on the efficiency frontier. On average and as a conservative estimate, countries could have increased their results by 40% using the same resources.

[3] See Ruggiero (2004).
[4] Afonso and St. Aubyn (2006) explain the algorithms (available on request), implemented in Matlab.
[5] Data sources are detailed in Afonso and St. Aubyn (2006).
[6] See Afonso and St. Aubyn (2006).
[7] Afonso and St. Aubyn (2005) addressed health efficiency for 2000 without PCA. Countries labelled as efficient were in line with our results.

Table 1. DEA output oriented: inputs (PCA on doctors, nurses, beds and MRI), output (PCA on life expectancy, infant survival rate and potential years of life not lost)

Country	VRS TE	Rank	Peers	Rank
Australia	1.101	10	Canada, Sweden, Korea, Finland	10
Austria	1.304	15	Sweden, Japan	15
Canada	1.000	1	Canada	6
Czech Republic	1.592	18	Japan, Sweden	18
Denmark	1.368	16	Korea, Japan, Sweden, Finland	16
Finland	1.000	1	Finland	4
France	1.106	11	Sweden, Spain	11
Germany	1.282	14	Sweden, Japan	14
Hungary	4.386	21	Sweden, Japan, Korea	21
Italy	1.143	12	Sweden, Japan	12
Japan	1.000	1	Japan	2
Korea	1.000	1	Korea	3
Luxembourg	1.372	17	Korea, Japan, Sweden	17
Poland	1.876	19	Spain, Korea	19
Portugal	1.083	9	Korea, Spain	9
Slovak Republic	2.667	20	Korea, Sweden, Japan	20
Spain	1.000	1	Spain	4
Sweden	1.000	1	Sweden	1
Switzerland	1.166	13	Sweden, Japan	13
United Kingdom	1.070	8	Canada, Sweden, Korea, Finland	8
United States	1.000	1	United States	7
Average	1.406			

Notes: VRS TE, variable returns to scale technical efficiency. Rank, taking into account the number of times the efficient countries are peers of inefficient countries.

We proceed to evaluate the importance of nondiscretionary inputs. Even if Tobit results are possibly biased, it is not clear that bootstrap estimates are necessarily more reliable, being based on a set of assumptions concerning the data generation process and the perturbation term distribution that may be disputed. Therefore, we apply both methods.

We regress the efficency scores on GDP per capita, Y, educational level, E, obesity, O and tobacco consumpion, T,[8]

$$\hat{\delta}_i = \beta_0 + \beta_1 Y_i + \beta_2 E_i + \beta_3 O_i + \beta_4 T_i + \varepsilon_i \tag{5}$$

Table 2 reports results from the censored normal Tobit regressions. Inefficiency is strongly related to variables that are, in the short to medium run, beyond the control of governments: economic background, proxied by the country GDP per capita, level of education, smoking habits and obesity. The estimated coefficients of the first two nondiscretionary inputs are statistically significant and negatively related to the efficiency measure. An increase in education achievement reduces inefficiency, implying that the relevant DMU moves closer to the production possibility frontier. The same applies to GDP, with higher GDP per capita resulting in more efficiency. On the other hand, efficiency is lower the stronger smoking habits are and the higher the percentage of obese population is.

Table 2. Censored normal Tobit results (19 countries)

	Model 1	Model 2	Model 3	Model 4
Constant	−3.2574	9.0162	−1.1185	9.9146
	(0.000)	(0.029)	(0.092)	(0.009)
Y	−4.38E-05		−4.44E-05	
	(0.000)		(0.000)	
Log(Y)		−1.2476		−1.1546
		(0.000)		(0.000)
E			−0.1060	−0.0891
			(0.010)	(0.034)
O	0.0895	0.0783	0.0946	0.0841
	(0.000)	(0.001)	(0.000)	(0.000)
T	0.1708	0.1453	0.1463	0.122
	(0.000)	(0.000)	(0.000)	(0.001)
$\hat{\sigma}_\varepsilon$	0.5677	0.5600	0.4759	0.5088
	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Y, GDP per capita; E, educational level; O, obesity; T, tobacco consumption; $\hat{\sigma}_\varepsilon$, estimated SD of ε.
p-Values are denoted in parentheses.

[8] Educational level, percentage of population that achieved tertiary education in 2000–2003; GDP per capita, PPP USD in 2003; obesity, percentage of obese population in 2002; smoking, percentage of population that consumed tobacco in 2003.

Table 3. Bootstrap results (19 countries)

	Model 1	Model 2	Model 3	Model 4
Constant	−6.9657	6.6360	−1.8317	10.1002
	(0.007)	(0.005)	(0.009)	(0.000)
Y	−1.0697E-04		−0.6383E-04	
	(0.028)		(0.000)	
Log(*Y*)		−1.4625		−1.4967
		(0.001)		(0.000)
E			−0.1800	−0.0962
			(0.000)	(0.007)
O	0.1555	0.1376	0.1080	0.1229
	(0.011)	(0.008)	(0.000)	(0.000)
T	0.29480	0.2596	0.2050	0.2076
	(0.011)	(0.008)	(0.000)	(0.002)
$\hat{\sigma}_\varepsilon$	0.5085	0.4155	0.4279	0.3759
	(0.000)	(0.000)	(0.000)	(0.000)

Notes: *Y*, GDP per capita; *E*, educational level; *O*, obesity; *T*, tobacco consumption; $\hat{\sigma}_\varepsilon$, estimated standard deviation of ε. *p*-Values are denoted in parentheses.
Using algorithm 1 from Simar and Wilson (2007).

Table 3 reports the estimation results from the bootstrap procedures employing algorithm 1 from Simar and Wilson (2007). Estimated coefficients are close to the estimates derived from the more usual Tobit procedure and are highly significant. Significance across different models and estimation methods confers robustness to our empirical evidence.

IV. Conclusion

We have evaluated efficiency in health services in OECD by assessing outputs (life expectancy, *ISR*, *PYLNL*) against inputs directly used in the health system (doctors, nurses, beds, MRI units) and environment variables (wealth and country education level, smoking habits and obesity).

Results from the first stage imply that inefficiencies are quite high. On average, countries could have increased their results by 40% using the same resources. Our second-stage procedure shows that GDP per head, educational attainment, tobacco consumption and obesity are highly and significantly correlated to output efficiency scores – a wealthier and a more cultivated environment are important conditions for a better health performance, whereas a more obese population and prevalence of smoking habits worsen health performance.

We have applied both the usual DEA/Tobit procedure and the recently proposed bootstrap algorithms. Results were strikingly similar with these different estimation processes, which brings increased confidence to obtained conclusions.

Acknowledgements

The authors are grateful to I. Joumard, K. Wendorff and the participants at workshops at the Zentrum für Europäische Wirtschaftsforschung (ZEW) and at the German Ministry of Finance for helpful comments. The opinions expressed herein are those of the authors and do not necessarily reflect those of the authors' employers.

References

Afonso, A. and St. Aubyn, M. (2005) Non-parametric approaches to education and health efficiency in OECD countries, *Journal of Applied Economics*, **8**, 227–46.

Afonso, A. and St. Aubyn, M. (2006) Relative efficiency of health provision: a DEA approach with non-discretionary inputs, Department of Economics, ISEG-UTL, Working Papers No. 33/2006/DE/UECF, Technical University of Lisbon, Lisbon.

Afonso, A., Schuknecht, L. and Tanzi, V. (2005) Public sector efficiency: an international comparison, *Public Choice*, **123**, 321–47.

Coelli, T., Rao, P., O'Donnell, C. and Battese, G. (2005) *An Introduction to Efficiency and Productivity Analysis*, Kluwer, Boston.

Evans, D., Tandon, A., Murray, C. and Lauer, J. (2000) The comparative efficiency of national health systems in producing health: an analysis of 191 countries, GPE Discussion Paper No. 29, World Health Organisation, Geneva.

Farrell, M. (1957) The measurement of productive efficiency, *Journal of the Royal Statistical Society, Series A*, **120**, 253–90.

OECD (2005) *OECD Health Data 2005*, OECD, Paris.

Ruggiero, J. (2004) Performance evaluation when non-discretionary factors correlate with technical efficiency, *European Journal of Operational Reserach*, **159**, 250–7.

Simar, L. and Wilson, P. (2007) Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, **136**, 31–64.

Spinks, J. and Hollingsworth, B. (2009) Cross-country comparisons of technical efficiency of health production: a demonstration of pitfalls, *Applied Economics*, **41**, 417–27.