

***Nonparametric country rankings using  
health indicators and OECD health data***

*Tarmo Rätty and Kalevi Luoma*



# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Health outcomes and health care inputs</b>	<b>2</b>
<b>3. Spurious rankings</b>	<b>4</b>
<b>4. Conclusions</b>	<b>7</b>
<b>References</b>	<b>8</b>



# 1. Introduction

In their forthcoming paper in the Journal of Applied Economics António Afonso and Miguel St. Aubyn (later A&St.A.) address efficiency in education and health sectors for a sample of OECD countries applying non-parametric efficiency measures. A&St.A have chosen to use quantitative inputs and outputs only. They argue that it is “difficult to measure something as complex as the health status of a population” and choose not to be innovative but use two usual measures of health attainment infant mortality and life expectancy as health care outputs. For inputs they use the number of doctors, nurses and in-patient beds per thousand inhabitants. Using multiple inputs and outputs and producing a scalar valued efficiency requires the use of some kind of aggregator function. They choose to use a variant of algorithm known as Data Envelopment Analysis (DEA). They end up presenting rankings of 24 individual countries in health efficiency.

The data, methods and the way A&St.A interpret the results raise serious doubts on the presented rankings. Comparing the most recent OECD Health Data with the input figures used by A&St.A reveals that their numbers for all three input measures – the doctors, nurses and total hospital beds – are severely flawed. In this comment we show that results change significantly if more recent and reliable OECD Health Data are used and that efficiency results are sensitive to the choice of input variables. We also argue that the link between the health indicators and health care resource measures A&St.A use is at best weak and that their efficiency results are spurious due to omission of important socioeconomic determinants of health. A simple modification of their model show inherent problems in interpretation of DEA efficiencies. They are not comparable between countries and result in large variations in rankings.

## 2. Health outcomes and health care inputs

OECD Health Data appears as an attractive starting point for analysing the performance of health care systems since data on individual countries are easily available. In using this data one should keep in mind that according to OECD there are still important gaps with respect to international agreements on statistical methods and that the same term can refer to very different things among the OECD countries. It seems that A&St.A have not taken these warnings seriously. A quick look on their input indicators reveals anomalies. For example, within Scandinavian countries (including Finland) the number of nurses varies between 7.3 and 14.4 and the number of hospital beds from 3.7 to 14.4. Considering that basic standards of health care in these countries are more or less equal, these wide variations in health care inputs should have raised also A&St.A's concern about the comparability of the figures. They use OECD Health Data 2002. The paper is dated September 2004, thus already 2004 version should have been available to them. In our analyses we use OECD Health Data 2005 which seems to have considerably less data anomalies than the earlier health data files. For example, variation in number of nurses within Scandinavian countries is now reduced, it varies from 7.6 to 10.1.

Despite the refined and improved statistics in the new OECD Health Data we probably still are quite far from fully accurate and comparable figures. One indication for the difficulty of constructing such data is that the new OECD Health Data do not any more include the number of hospital beds for five countries, for which these numbers were available in the 2002 data. Obviously these countries have been dissatisfied with the accuracy and comparability of earlier estimates of the number of total hospital beds. It should also be noted that the proportion of nurses from the total health care work force varies widely among OECD countries. At the lowest the shares in 1999 were in Germany (20.4 %) and France (20.6 %). In Netherlands it was 47.1, in Ireland 60.6 and in Luxembourg 75.4 percent. Obviously the definition for nurses still varies greatly among OECD countries. Excluding other labour inputs than doctors and nurses leads to calculating efficiency where the proportion of inputs which are accounted for varies widely.

There is a great temptation to renew A&St.A estimates, but due to the missing data for total hospital beds the "corrected" estimates cannot be calculated. We also have to question the selection of inputs in their model. To give a fair base for the rankings the model should cover all the variables that contribute to used health status indicators. Life expectancy and infant mortality are evidently dependent on standard of living over peoples' lifetime, as well as on environmental, cultural and life style factors. Using only staffing and beds as inputs excludes major factors from the analysis.

Additionally, we can question the causality of inputs and outputs in A&St.A model. It is hard to figure out how the total number of hospital beds tells something about maternal or infant care. Nor do A&St.A provide any explanation why more hospital beds should lead to longer life expectancy. We can also argue reverse causality: healthier people give birth to healthier children, live longer and need less resources i.e. doctors, nurses and hospital beds. Or unhealthy diets, sedentary life styles and smoking can lead both to shortened life span and increased demand for medical resources.

A detailed examination of A&St.A's DEA results reveals additional problems. In Table 5 they list peers for each country, the number of peers for inefficient countries varies from two to five. The full set of peers is reached just in input oriented VRS models for France and Luxemburg. This means that in all the other solutions, some set of the inputs or outputs are not used in efficiency evaluations. For 8 countries, output oriented VRS model uses just one output and one input to calculate efficiency. The model is clearly over dimensional. As a result a large number of countries gain full efficiency and inefficient countries too high efficiencies. The most seriously, the impact of dimension on rankings of individual units is indefinite.

We do not believe that using these variables gives a fair basis for country rankings in health care, but in the next section we will show that correcting data problems and reducing the irrelevant variable hospital beds from the data results in considerable different rankings. These ranking may also change radically annually.

### 3. Spurious rankings

We demonstrate the properties of non-parametric efficiency rankings using the same set of variables as A&St.A, but excluding hospital beds from inputs. The new 2005 Health Data do not any more provide information on the number of total hospital beds for five countries as did the 2002 OECD data set used by A&St.A. Apparently this data which has been removed has been considered unreliable. To secure the maximum number of countries we pool all the country observations from 1999 to 2002. Thus, we assume that health indicators are a result of long term development and the countries should be able to reach their any past or near future productivity level. Also it is likely that rankings from pooled data set eliminate at least part of the annual variation due to reporting errors. This leaves us 84 country-years observations. As we have a lot of doubt of the specification itself, we demonstrate the results just for input oriented VRS model.

The main results are collected in Table 1. Countries are ranked according to their average efficiency score. The average rankings yield the same result. As a result of model changes the number of efficient countries reduces from 8 to 4. Japan (2000 and 2002) and Korea (1999) are classified as full efficient each year they exist in the data set. The fact that these reference countries have failed in providing data for other years raises serious doubts on the comparability of these observations. Greece (1999) and Italy (2001, 2002) are also fully efficient. Compared to A&St.A there are radical changes in efficiencies. Italy is fourth in our list compared to 12:th in A&St.A with 15 percent higher average efficiency. Both Sweden and USA loose about 35 percent of their score and their ranking drop from fully efficient to 13th and 14th respectively. Finland raises its ranking from 16th to 7th. The Spearman's rank correlation coefficient between our average and A&St.A ranking is 0.76 showing considerable differences, especially if we consider how sensitive this kind of rankings are politically.

*Table 1. Average efficiency scores for each country and pooled rankings of each country-year observation. Input oriented VRS model*

Ranking		Average score	Pooled Rankings				A&St.A ranking
			1999	2000	2001	2002	
1	Japan	1		1		1	1
2	Korea	1	1				1
3	Greece	0,988	1	8			9
4	Italy	0,981	9	7	1	1	12
5	Canada	0,815	16	15	12	11	1
6	United Kingdom	0,787	17	13	19	24	1
7	Finland	0,774	23	40	14	10	16
8	Spain	0,757	36	18	20	22	1
9	New Zealand	0,741	28	26	27		14
10	Portugal	0,727	64	39	25	21	1
11	France	0,724	32	30	31	29	11
12	Australia	0,695	35	37	34	33	13
13	United States	0,653	44	43	38	47	1
14	Sweden	0,637	41	45	56	42	1
15	Ireland	0,622	55	48	52	46	18
16	Luxembourg	0,62	51	49	50	53	19
17	Norway	0,582	54	57	63		17
18	Poland	0,562	59	60	61	62	15
19	Denmark	0,539	58	67	66	65	10
20	Austria	0,501	70	74	73	69	20
21	Czech Republic	0,491	75	71	68	82	21
22	Netherlands	0,491	72	77	79		23
23	Germany	0,483	80	81	76	78	22
24	Hungary	0,413	83			84	24

Rankings from the pooled data are controversial. The most of the countries keep their ranking rather stable, but Portugal's rankings change from 21 to 64 of 84, also Finland's ranking varies considerably. We can create alternative rankings from pooled data. For example, choosing the highest efficiency score for each country results in the 5th position in the list for Finland, right after the fully efficient countries. Creating efficiency rankings that interest policymakers or managers is easy but also hazardous. Minimal differences in efficiency may appear as a large change in the ranking, In A&St.A output oriented results Italy scores 9th just 0.5 percent short of full efficiency. Rankings are also completely conditional on the model used.

Our reduced model from A&St.A does not fix the dimensionality problem. None of the inefficient countries gets four peers, and only 19 country-year observations

have three peers. The low dimensionality and different number of peers in both ours and A&St.A results indicate problems in variables. It is related to multicollinearity in a regression model (see e.g. Olesen and Petersen 1996 or Rätty 2002), where it is not possible to separate independent impacts of two variables. In DEA, where efficiency scores for each country are solved as independent problems, this means that units are compared to each other using different variables. International rankings based on this kind of results are obviously spurious. Taking that the used health indicators and inputs are as such doubtful for efficiency measurement and serious concern on comparability of observations remains, it does not pay to search for proper dimensions of the model.

Finally, one obvious problem in our results is that the same country can exist several times as a peer for an inefficient country-year observation. In a seriously taken attempt to present country rankings this should be taken into account.

## 4. Conclusions

Specifying a model for the non-parametric efficiency measures is a demanding task that requires good understanding on both the phenomenon to be modelled and specific features of the estimator. This is especially true, when we model fuzzy production processes, like production of health, that do not have explicit inputs and outputs. The resources a country uses to organise health services can be either a reason or a consequence of the observed health status. This complex relation results in arbitrary efficiency estimates. The non-parametric DEA algorithm adds the complexity if irrelevant input or outputs are used; resulting country efficiencies are generally calculated using different sets of variables. In spite of the intensive work in efficiency measurement (see e.g. DEAZONE), researchers do not yet have a simple set of statistical indicators to describe the reliability of the results. Our illustrative model developed in spirit of A&St.A model shows how sensitive the results are, especially if rankings are concerned.

OECD Health Data is a valuable source of information, and possibly the main source of international comparisons. We see that countries work constantly to improve accuracy and comparability of data, but clearly these things depend also on the use of the figures. Therefore, using a well known data bank does not free a researcher's responsibility to check the data for country-wise comparisons and its suitability to his/her specific study.

Efficiency of health care is an interesting topic and if easy solutions to model it would exist, it certainly would have established its place in health policy managers' toolbox. Relating expected life time and infant mortality to health care resources has not yet qualified as one. One should bear in mind that life expectancy as any statistic based on mortality is at best an incomplete measure of health service performance since considerable portion of health care resources are allocated to purposes where the primary objective is relieving the pain and improving the quality of life, not extending the length of life span (see Holland and Breeze 1988). We agree with Triplett (2001) who has noted that to measure the output of the health sector we need to model health consequences of medical interventions, not to compare the aggregate level of health with the resources employed in the health care.

## References

- Afonso, António - Miguel St. Aubyn (2004): Non-parametric Approaches to Education and Health Efficiency in OECD countries, Forthcoming in the Journal of Applied Economics.
- DEAZONE, [www.deazone.com](http://www.deazone.com)
- Holland W.W. - E. Breeze (1988): "The performance of health services", in M. Keynes, D.A. Coleman and N.H. Dunsdale, eds, The political economy of health and welfare. Handsmills, Macmillan Press.
- OECD Health Data 2005 (2005): Paris, OECD.
- Olesen O. - N. C. Petersen (1996): Indicators of Ill-Conditioned Data Sets and Model Misspecification in Data Envelopment Analysis: An Extended Facet Approach", Management Science, Vol. 42(2) 205-219.
- Räty, Tarmo (2002): Efficient Facet Based Efficiency Index; A variable Returns to Scale Approach. Journal of Productivity Analysis, 17(1-2) 65-82.
- Triplett, Jack E. (2001): "What's Different About Health? Human Repair and Car Repair in National Accounts and in National Health Accounts", in D.M. Cutler and E.R. Berndt, eds., Medical Care Output and Productivity, Chicago, National Bureau of Economic Research.

## VATT-MUISTIOITA / WORKING NOTES

35. Lang Markku: Euroopan unionin jäsenyyden vaikutus maatalojen kassavirtoihin vuonna 1995. Helsinki 1998.
36. Hjerppe Reino – Kemppi Heikki: Julkisen infrastruktuurin tuottavuus. Helsinki 1998.
37. Verojärjestelmän kautta annettava tuki. Helsinki 1998.
38. Tax Expenditures in Finland. Helsinki 1998.
39. Romppanen Antti: Suomen työeläkejärjestelmä kestää vertailun. Helsinki 1999.
40. Mäki Tuomo – Romppanen Antti: Nuoret työhön – varttuneet kouluun. Näkökohtia koulutuksen kehittämistä. Helsinki 1999.
41. Niskanen Esko – Voipio Iikko B. (toim.): Seminar on Urban Economic Policy / Kaupunkipolitiikan seminaari 13.8.1998. Helsinki 1999.
42. Mäki Tuomo – Virén Matti: Käytännön kokemuksia finanssipolitiikan indikaattorien laskemisesta. Helsinki 1999.
43. Verojärjestelmän kautta annettava tuki. Helsinki 1999.
44. Tax Expenditures in Finland. Helsinki 1999.
45. Kröger Outi – Kärri Timo – Martikainen Minna – Voipio Iikko B.: Perusteellisuuden pääomahuolto 2000 – 2030. Esitutkimus. Helsinki 1999.
46. Viitamäki Heikki: Asumistuen budjetointia ja määräytymisperusteiden valmistelua koskeva selvitys. Helsinki 1999.
47. Junka Teuvo: Asuntojen hinnat vapailta markkinoilla. Helsinki 2000.
48. Kunta- ja aluetalouden seminaari 14.2.2000. Seminaarin alustukset. Helsinki 2000.
49. Viitamäki Heikki: Työmarkkinoiden verokiila vuosina 1990-2000. Helsinki 2000.
50. Valppu Pirkko: A Brief Overview of Working Conditions in the Care Sector. Helsinki 2000.
51. Hjerppe Reino: VATT-10 vuotta – tutkimustuloksia ja uusia visioita. Helsinki 2000.
52. Aronen Kauko – Järviö Maija-Liisa – Luoma Kalevi – Rätty Tarmo (toim.): Peruspalvelut 2000 seminaari 29.11.2000 Seminaarimuistio. Helsinki 2001.
53. Parkkinen Pekka: Työssä jaksaminen pitkällä aikavälillä. Helsinki 2001.
54. Lyytikäinen Teemu: Työn keskimääräisen efektiivisen veroasteen laskeminen OECD:n menetelmällä. Helsinki 2002.
55. Tukiainen Janne – Loikkanen Heikki A.: Kuudes pohjoismainen paikallisjulkistalouden konferenssi – yhteenveto. Helsinki 2002.
56. Aarnos Kari: Vihreät sertifikaatit, uusi tapa tukea sähköntuotantoa uusiutuvista energialähteistä. Helsinki 2002.
57. Kemppi Heikki: Suomen jätepoliittisten tavoitteiden mukaiset jätetutkimuksen tutkimuskokonaisuudet I. tutkimuskluusterit (Wasteprev-projekti). Liitteenä kirjallisuuskatsaus jätteiden synnyn ehkäisystä. Helsinki 2002.

58. Heikkilä Tuomo – Kilponen Juha – Santavirta Torsten (toim.): Suomen tiede- ja teknologiapolitiikan haasteet ja muuttuva toimintaympäristö. KNOGG-työpajaseminaarin yhteenveto, Innopoli 1.4.2003. Helsinki 2003.
59. Santavirta Torsten: Benchmarking the Competitiveness Strategies of Six Small European Countries: A Small Country Perspective. Helsinki 2003.
60. Antikainen Riikka – Siivonen Erkki: Julkisen sektorin verrokin periaatteet tiehankkeiden osalta. Helsinki 2003.
61. VATT:n palkkauskäsikirja. Helsinki 2003. (Sisäinen)
62. Honkatukia Juha – Kemppi Heikki – Rajala Rami: Energiaverotuksen ja päästökaupan vaikutus kilpailukykyyn. Helsinki 2003.
63. Mikrosimulointimallien välinen yhteistyö. Helsinki 2004.
64. Antikainen Riikka – Siivonen Erkki – Saltevo Anu – Salmela Vesa – Tolvanen Riku: Elinkaarimallitoteutuksen ja parhaan nykykäytännön vertailu. Helsinki 2004.
65. Mustonen Esko – Viitamäki Heikki: Työmarkkinatueltta ansiotyöhön: vaikutukset valtion ja kuntien tuloihin ja menoihin. Helsinki 2004.
66. Romppanen Antti: Maailmantalouden yhdentyminen. Helsinki 2004.
67. Mykkänen Kai: Miksi yhteisöveron tuotto moninkertaistui 1990-luvulla? Helsinki 2005.
68. Saastamoinen Hannu: Ajankäytön trendit 1999–2000. Helsinki 2005.
69. Luoma Kalevi – Moisio Antti: Kuntakoko, kuntien menot ja palvelujen tuotannon tehokkuuserot. Helsinki 2005.
70. Seppo Kari – Venetoklis Takis: Selvitys kehitysalueiden korotettujen poistojen vaikuttavuudesta. Helsinki 2005.
71. Honkatukia Juha – Perrels Adriaan: Vesidirektiivin toimeenpano Suomessa – Esitutkimus taloudellisen arvioinnin edellytyksistä. Helsinki 2005.
72. Reino Niinivaara – Heikki Viitamäki (Toimittaneet): TUJA-Käsikirja Verotuksen ja etuuksien mikrosimulointimalli. Helsinki 2005. (Sisäinen)
73. Anne Kiiskinen: Aluetekijöiden tarkastelua julkisessa työnvälityksessä. Helsinki 2005.