THE UNIVERSITY *of York*

CENTRE FOR HEALTH ECONOMICS

Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach

Hugh Gravelle
Rowena Jacobs
Andrew M. Jones
Andrew Street

CHE Technical Paper Series 25

CENTRE FOR HEALTH ECONOMICS TECHNICAL PAPER SERIES

The Centre for Health Economics has a well established Discussion Paper series which was originally conceived as a means of circulating ideas for discussion and debate amongst a wide readership that included health economists as well as those working within the NHS and pharmaceutical industry.

The introduction of a Technical Paper Series offers a further means by which the Centre's research can be disseminated. The Technical Paper Series publishes papers that are likely to be of specific interest to a relatively specialist audience, for example papers dealing with complex issues that assume high levels of prior knowledge, or those that make extensive use of sophisticated mathematical or statistical techniques.

The content and format of Technical Papers are entirely the responsibility of the author, and papers published in the series are not subject to peer-review or editorial control, unlike those produced in the Discussion Paper series. Offers of further papers, and requests for information should be directed to Frances Sharp in the Publications Office, Centre for Health Economics, University of York.

October 2002

Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach

Hugh Gravelle [1] - hg8@york.ac.uk

Rowena Jacobs [2] – rj3@york.ac.uk

Andrew M. Jones [3] * – amj1@york.ac.uk

Andrew Street [2] – ads6@york.ac.uk

[1] National Primary Care Research and Development Centre, Centre for Health Economics, Alcuin Block D, University of York, York YO10 5DD

[2] Centre for Health Economics, Alcuin Block D, University of York, York YO10 5DD

[3] Department of Economics and Related Studies, University of York, York YO10 5DD

* Corresponding author

Abstract

Background

The World Health Organisation (WHO) has used econometric methods to measure the efficiency of health care systems. We assess the robustness of the WHO results to definitions of efficiency and statistical procedures.

Materials and Methods

The paper uses the data originally analysed by the WHO. The data are for 1997 for 50 countries and for 1993-1997 for 141 countries. The efficiency of each country in promoting population health is estimated after taking into account health care expenditure, all other expenditure and education levels. Efficiency scores are compared under different definitions of efficiency and different estimation methods.

Results

The results show that the country rankings and efficiency scores are sensitive to the definition of efficiency and choice of model specification.

Conclusions

Econometric methods can yield insights into complex socio-economic phenomena. However, the lack of robustness of results to reasonable alternative specifications suggests that it is premature to use the methods adopted by the WHO to construct league tables of health systems.

Introduction

The World Health Organisation (WHO) has attempted to compare the performance of national health care systems along a number of dimensions including the level and distribution of population health, responsiveness and fairness in financing [1]. As part of the programme it has used econometric methods to measure the efficiency of health care systems [2]. There has been considerable criticism of the programme, including the purpose of the exercise [3], the definition of some of the performance measures [4], the quality of data [3, 5], and mixed messages [6]. We discuss a number of issues raised by the econometric methods used to derive the WHO efficiency measure. We show that the efficiency rankings and estimates of the magnitude of inefficiency in countries are not robust to other, no less reasonable, methodological choices. The WHO efficiency measures are not a reliable guide to policy for national governments concerned with the performance of their health care systems.

Materials and methods

Data

We used the data set supplied by the WHO team. It covers 141 countries for each of the five years 1993 to 1997 and 50 countries for 1997. The data set includes four variables: a measure of population health, defined as disability adjusted life years [1]; *per capita* health care expenditure; all other expenditure, defined as gross domestic product *per capita* minus health care *per capita*; and education levels, measured by the average years of schooling of adults. Health care expenditure and other expenditure are in 1997 US dollars adjusted for differences in the purchasing power of currencies.

Theory

Definitions of health system efficiency

In the WHO efficiency study it was argued that population health depends on health care expenditure *per capita* and education [2, 7]. Multiple regression analysis was used to predict expected population health given a country's health care expenditure and education.

The WHO efficiency index for a country is

$$WHO\ Efficiency\ Index = \frac{Health - minimum\ health}{Maximum\ health - minimum\ health}$$

Maximum health is the level of health achievable by the country given its health care expenditure and education if its health care system functioned as well as the most efficient country. Minimum health is defined as the population health predicted if there was no health care expenditure.

The justification for including *Minimum health* in the WHO efficiency measure is that health levels would not be zero if health care expenditure was zero [2]. But the fact that health would not be zero in the absence of a health system is irrelevant in measuring how much extra health could be achieved by improving the performance of a country's health care system. Subtracting *Minimum health* from numerator and denominator reduces the ratio and gives a misleading impression of the potential percentage health gain from improving the health care system.

The WHO study defines efficiency as "the ratio of the observed level of attainment of a goal to the maximum that could have been achieved with the observed resources" [2]. The definition focuses on what is relevant for policy: the increase in health which could be achieved if a country's system became as productive as the best system. The efficiency measure suggested by the definition is

$$\textit{Technical efficiency index} = \frac{\textit{Health}}{\textit{Maximum health}}$$

instead of the index constructed by WHO. We explored whether the results were sensitive to this definition (Model 1).

Alternative estimation methods

To calculate an efficiency index it is necessary to estimate the relationship between the "output" of population health and "inputs" or factors likely to affect health. Decisions must be made about the statistical methods, the form of the relationship, and the "inputs" to be included. We investigated the implications of alternative choices for the estimated effect of health care expenditure and for the efficiency scores of each country.

An alternative estimator.

Panel data methods are appropriate when the data are for a number of countries over a number of years since it is possible to estimate both the effects of inputs, such as health care expenditure and education, and the country-specific effects of unobserved factors which differ across countries. The two sources of variation in panel data which can provide information on the way in which the inputs affect health are known as *between variation* (across countries) and *within variation* (within countries over time).

The approach adopted by the WHO makes use of only the within country variation. In order to predict the maximum attainable level of health, each country is allowed to have a different intercept term in the panel data regression model. This is known as a country-specific *fixed effect* and it assumes that the only factor driving the country specific effects is the efficiency of the health care system. This is a poor choice of technique for the WHO data because there is very little within country variation and only five years of observation. Under these circumstances estimates will lack precision and the statistical tests that are conventionally used to discriminate between models may be misleading [8]. Moreover, unbiased estimation of the country specific effect, which is required for the efficiency score, requires a large number of time periods [9].

For 50 of the 191 countries there were data for only a single year. This has two consequences. First, the *fixed effects* regression is estimated on the data for only those 141 countries with more than one year of data. The 50 single observation countries have no impact on the estimated model. Second, efficiency scores for these 50 countries are calculated by a model that does not include them. This is inappropriate if the 50 countries are systematically different from the others - as indeed they are. They had lower health expenditure than the others.

Almost all the variation in the data (99.8% for the logarithm of the health measure and 98.4% for the logarithm of health care expenditure) arises from differences *between* countries rather than *within* them over time. Consequently it may be better to estimate the model using a *between estimator* which uses only the country means of the variables [8]. This approach is used for Model 2.

Natural units.

The WHO model fitted the logarithm of population health on the logarithms of health care expenditure and education. A logarithmic specification assumes that a country that is 80% efficient at its current level of health and health care expenditure will be 80% efficient at any other level of expenditure. If countries with lower levels of health care expenditure per capita tend to have lower levels of relative efficiency, the WHO index of performance will be biased against them. One way to test the sensitivity of results to this assumption is to estimate the model in natural units (Model 3).

Year effects.

The effect of health care expenditure and education on health may vary over time, either because of genuine changes in the underlying relationship or because of changes in other, unobserved variables that also affect health. We allowed for this possibility by including the year of observation as an explanatory variable in Model 4.

Income and health.

Although there is a large body of evidence suggesting that population health is strongly related to per capita income, the WHO model does not include per capita income as an "input" variable. It is argued that income "does not directly contribute to health" (pg. 308) and is highly correlated with health expenditure and education [2]. When income is omitted from the model, its effect on health is captured, at least partially, by health care expenditure and education. The WHO team attempted to overcome the problem by regressing income on health care expenditure and education [7]. The residuals from this regression capture the variation in income that is not correlated with health care expenditure or education. Finding that the residuals were not statistically significant when included in the regression model for health, the WHO argues that income can be omitted [7].

The WHO procedure for testing whether income should be included in the model rests on the *assumption* that health expenditure is the main determinant of population health and that only variations in income that are not correlated with health care or education influence health. We applied the WHO procedure, reversing the roles of income and health care expenditure, and found that the residuals from a regression of health care expenditure on income and education were not significant when included in a model using income and education to explain health. According to the WHO test it is arbitrary, statistically, whether either health care expenditure or income appears in the efficiency model.

We also adopted a more conventional approach to determine the relative importance of the effect of health care expenditure and income on health. Income can be spent on health care and other goods and services. We included other expenditure (income less health care expenditure) as an input in the health model, together with health care expenditure and education (Model 5).

Results

A striking feature of the data is that 99.8% of the variation in the logarithm of disability adjusted life expectancy is *between* rather than *within* countries, with similar figures for the key explanatory variables. This is evident from the model results in Table 1 which reveal high R-squared *between* values, relative to *within*, as well as high values for rho, the variation due to the time invariant country-specific effect. Effectively the panel data model collapses to a cross section comparison of countries, that can be estimated using a *between* regression of country means. Figure 1 plots curves of predicted values from a regression of health expenditure on health, treating education as constant. The predicted curve from the *between effects* estimator provides a far better fit of the data than the curve from the *fixed effects* estimator.

Table 2 reports the correlation between the efficiency scores and the country ranks estimated by the different methods. The lowest correlation is between the WHO model and the between estimator (Model 2).

Table 3 shows how the top and bottom ten countries differ under the different specifications and gives their estimated efficiency scores. Dropping the minimum from the definition of efficiency has little impact on the ranking of countries but it does have a substantial impact on the estimated efficiency scores for countries at the bottom of the table: their estimated efficiency is much higher. The ranking and scores are very sensitive to the statistical specification used to estimate the model. Note, for example, how Malta moves from the top ten under the WHO model to the bottom ten under models 2, 4 and 5. The volatility experienced by Malta is due to its actual health being close to the maximum attainable and to its having a very high level of minimum health (because of its high rate of literacy which is used to estimate minimum health). Notice also that only one country, Jamaica, appears in the top ten for both the WHO model and the between effects estimator.

Table 4 shows how the efficiency scores and rankings of the countries appearing at the top and bottom of the WHO ranking, along with the United Kingdom, fare under the alternative specifications. The impact at the top of the table is dramatic. Using between rather than within variation Oman drops from 1st to 169th. The rankings for the UK range from 23rd to 96th with efficiency scores from 0.57 to 0.94. Rankings at the bottom end of the table are much more stable.

Conclusion

The WHO team claim that their rankings are stable under 'numerous specifications' (pg. 308) [2]. In fact only six specifications were reported and these were very similar [7]. It is not good practice to rely on a single technique or model specification to test the robustness of results which may influence policy. Although econometric methods can yield insights into complex socio-economic phenomena, the issues raised in this paper suggest that it is premature to use the methods adopted by the WHO to construct league tables of health systems. The WHO should direct its efforts into developing a theory of why health system performance may differ, identifying which factors are important in explaining any differences, and ensuring that these factors are accurately measured.

Authors' Contributions: All authors initiated the study, contributed to the modelling and interpretation of results. RJ and AS analysed the data. HG, AJ and AS were responsible for the first draft of the paper. All authors approved the final version.

Competing interests: None.

Acknowledgements: We thank participants in the York Seminars in Health Econometrics and David Evans and Ajay Tandon of the WHO for providing the data. The National Primary Care Research and Development Centre and the Health Policy Team at the Centre for Health Economics are funded in part by the Department of Health. AS is funded jointly by the Medical Research Council and Northern and Yorkshire Region. The views expressed are those of the authors and not necessarily those of the funding bodies.

References

1. World Health Organization: **World Health Report 2000**. *Geneva: WHO*; 2000.
2. Evans DB, Tandon AJ, Murray CJL, Lauer JA: **Comparative efficiency of national health systems: cross national econometric analysis**. *BMJ* 2001; **323**:307-310.
3. Williams A: **Science or marketing at WHO? A commentary on 'World Health 2000'**. *Health Econ* 2001; **10**:93-100.
4. Braveman P, Starfield B, Geiger HJ: **World Health Report 2000: how it removes equity from the agenda for public health monitoring and policy**. *BMJ* 2001; **323**:678-681.
5. McKee M: **Measuring the efficiency of health systems**. *BMJ* 2001; **323**:295-296.
6. Navarro V: **Assessment of the World Health Report 2000**. *Lancet* 2000; **356**:1598-1601.
7. Evans DB, Tandon AJ, Murray CJL, Lauer JA: **The comparative efficiency of national health systems in producing health: an analysis of 191 countries.** *Geneva: WHO*; 2000.
8. Deaton A: **The Analysis of Household Surveys: a microeconometric approach to development policy**. *Baltimore, Maryland: The John Hopkins University Press*; 1997.
9. Baltagi B: **Econometric Analysis of Panel Data, second edition**. *Chichester: John Wiley*; 2001.

Figure 1: Fitted regression curves under between and fixed effects estimators

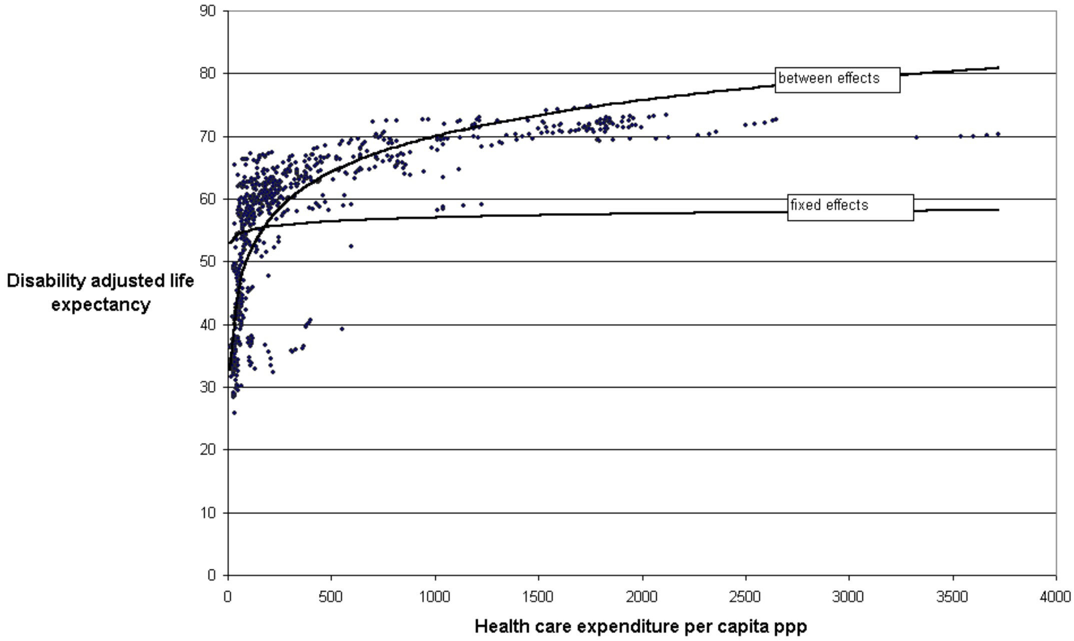

Table 1: WHO model and alternative model specifications and results

		1	2	3	4	5
		Model without minimum levels	Run as between effects model	In natural units	With year dummies	With other expenditure added
Constant	Coeff	3.799	3.233	44.514	3.911	3.633
	(P > |t|)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Health expenditure		0.009	0.080	0.001	0.006	0.005
		(0.004)	(0.000)	(0.001)	(0.044)	(0.119)
Education		0.063	0.240	2.246	0.042	0.064
		(0.084)	(0.002)	(0.000)	(0.250)	(0.073)
Education squared		0.022	-0.012	-0.036	0.002	0.015
		(0.123)	(0.647)	(0.374)	(0.870)	(0.271)
Other expenditure						0.025
						(0.000)
year 93					-0.007	
					(0.000)	
year 94					-0.006	
					(0.001)	
year 95					-0.004	
					(0.008)	
year 96					-0.002	
					(0.107)	
sigma_u		0.185		8.522	0.216	0.177
sigma_e		0.011		0.472	0.011	0.011
rho		0.997		0.997	0.998	0.996
R-squared:	within	0.138		0.260	0.159	0.161
	between	0.649		0.638	0.670	0.703
	overall	0.635	0.690	0.628	0.655	0.696
No. obs		754	754	754	754	754
No. groups		191	191	191	191	191
Estimation obs		141	141	141	141	141
F		29.80	144.62	65.42	15.02	26.75
Prob > F		0.00	0.00	0.00	0.00	0.00
RESET test	F	1.76	9.62	0.56	2.43	4.23
	Prob > F	0.17	0.00	0.64	0.09	0.02

Table 2: Score and rank correlation of WHO model with 6 alternative specifications

Scores	WHO model	1	2	3	4	5
		Model without minimum levels	Run as between effects model	In natural units	With year dummies	With other expenditure added
WHO	1.000					
1	0.974	1.000				
2	0.597	0.512	1.000			
3	0.998	0.974	0.604	1.000		
4	0.904	0.903	0.577	0.899	1.000	
5	0.637	0.629	0.675	0.629	0.832	1.000

Ranks	WHO model	1	2	3	4	5
		Model without minimum levels	Run as between effects model	In natural units	With year dummies	With other expenditure added
WHO	1.000					
1	0.987	1.000				
2	0.461	0.391	1.000			
3	0.997	0.984	0.486	1.000		
4	0.909	0.917	0.367	0.898	1.000	
5	0.955	0.947	0.537	0.953	0.949	1.000

Table 3: Top and bottom 10 countries under each model specification with efficiency scores

	WHO model		1		2		3		4		5	
			Model without minimum levels		Run as between effects model		In natural units		With year dummies		With other expenditure added	
1	Oman	0.99	Malta	1.00	Yemen	1.00	Malta	1.00	Japan	1.00	Jamaica	1.00
2	Malta	0.99	Oman	1.00	Morocco	0.92	Jamaica	0.95	France	0.97	Spain	0.97
3	Italy	0.98	Italy	0.99	Cape Verde	0.87	Spain	0.95	Spain	0.96	Andorra	0.96
4	France	0.97	France	0.99	Sol Islands	0.87	San Marino	0.94	Greece	0.96	San Marino	0.96
5	San Marino	0.97	San Marino	0.99	Bhutan	0.85	Andorra	0.94	Italy	0.96	Italy	0.96
6	Spain	0.97	Spain	0.99	Nepal	0.83	Saudi Arabia	0.94	Andorra	0.95	France	0.95
7	Andorra	0.96	Andorra	0.99	Egypt	0.82	Italy	0.93	San Marino	0.95	Greece	0.95
8	Jamaica	0.96	Jamaica	0.98	Azerbaijan	0.81	Singapore	0.93	Monaco	0.95	Japan	0.94
9	Japan	0.95	Japan	0.98	Jamaica	0.81	Portugal	0.92	Sweden	0.95	Cuba	0.93
10	Greece	0.94	Greece	0.97	Somalia	0.79	Greece	0.92	Netherlands	0.93	Portugal	0.93
182	South Africa	0.23	Swaziland	0.55	DR Congo	0.27	Rwanda	0.23	Swaziland	0.21	South Africa	0.24
183	Sierra Leone	0.23	Liberia	0.55	Malawi	0.22	Sierra Leone	0.22	Sierra Leone	0.20	DR Congo	0.24
184	Swaziland	0.23	Namibia	0.54	Swaziland	0.21	Swaziland	0.22	DR Congo	0.19	Swaziland	0.24
185	DR Congo	0.22	Botswana	0.52	Lesotho	0.20	DR Congo	0.21	Lesotho	0.19	Lesotho	0.23
186	Lesotho	0.21	Rwanda	0.51	South Africa	0.19	Lesotho	0.21	Malawi	0.17	Malawi	0.21
187	Malawi	0.20	Zimbabwe	0.51	Botswana	0.16	Malawi	0.19	Botswana	0.17	Namibia	0.20
188	Botswana	0.18	Niger	0.50	Namibia	0.16	Namibia	0.18	Namibia	0.17	Botswana	0.19
189	Namibia	0.18	Malawi	0.47	Zambia	0.11	Botswana	0.18	Zambia	0.10	Zambia	0.12
190	Zambia	0.11	Zambia	0.47	Zimbabwe	0.07	Zambia	0.11	Zimbabwe	0.07	Zimbabwe	0.08
191	Zimbabwe	0.08	Sierra Leone	0.42	Malta	-0.03	Zimbabwe	0.07	Malta	0.04	Malta	0.01

Table 4: Top and bottom 10 countries under WHO model and under different specifications, along with the United Kingdom (efficiency score and rank)

	WHO model		1		2		3		4		5	
			Model without minimum levels		Run as between effects model		In natural units		With year dummies		With other expenditure added	
Oman	0.99	1	1.00	2	0.42	169	0.90	14	0.35	169	0.81	71
Malta	0.99	2	1.00	1	0.00	191	1.00	1	0.00	191	0.00	191
Italy	0.98	3	0.99	3	0.49	158	0.93	7	0.96	5	0.96	5
France	0.97	4	0.99	4	0.47	160	0.92	11	0.97	2	0.95	6
San Marino	0.97	5	0.99	5	0.52	144	0.94	4	0.95	7	0.96	4
Spain	0.97	6	0.99	6	0.56	104	0.95	3	0.96	3	0.97	2
Andorra	0.96	7	0.99	7	0.55	112	0.94	5	0.95	6	0.96	3
Jamaica	0.96	8	0.98	8	0.81	9	0.95	2	0.87	28	1.00	1
Japan	0.95	9	0.99	9	0.56	103	0.90	13	1.00	1	0.94	8
Greece	0.94	10	0.97	10	0.62	70	0.92	10	0.96	4	0.95	7
United Kingdom	0.88	24	0.94	23	0.57	96	0.86	25	0.92	14	0.88	26
South Africa	0.23	182	0.57	176	0.18	186	0.23	182	0.22	180	0.24	182
Sierra Leone	0.23	183	0.42	191	0.27	181	0.22	184	0.20	183	0.24	181
Swaziland	0.23	184	0.55	182	0.21	184	0.23	183	0.21	182	0.24	184
Dem Rep of Congo	0.22	185	0.57	177	0.27	182	0.21	185	0.19	185	0.24	183
Lesotho	0.21	186	0.57	178	0.21	185	0.21	186	0.19	184	0.23	185
Malawi	0.20	187	0.47	189	0.22	183	0.19	187	0.17	186	0.21	186
Botswana	0.18	188	0.54	184	0.15	188	0.18	189	0.17	188	0.19	188
Namibia	0.18	189	0.52	185	0.16	187	0.18	188	0.17	187	0.19	187
Zambia	0.11	190	0.47	190	0.12	189	0.11	190	0.10	189	0.12	189
Zimbabwe	0.08	191	0.51	187	0.08	190	0.08	191	0.07	190	0.08	190