THE COMPARATIVE EFFICIENCY OF NATIONAL HEALTH SYSTEMS IN PRODUCING HEALTH:

AN ANALYSIS OF 191 COUNTRIES

David B Evans
Ajay Tandon
Christopher JL Murray
Jeremy A Lauer

GPE Discussion Paper Series: No. 29

EIP/GPE/EQC

World Health Organization

1) Introduction

For much of the last two decades, health policy makers have been concerned with the performance of their health systems and many countries have introduced reforms aimed at improving performance (1,2). Reforms have ranged across all functions of the health system – financing (e.g., social health insurance and user charges), provision (e.g., managed care, autonomous hospitals), stewardship (e.g., regulation of the private sector, health legislation) and resource development (e.g., re-training of staff). (1,3-6). An increasing number of studies are beginning to evaluate the impact of specific reforms or more general reforms in selected settings (e.g., (7-9)), but this requires an explicit framework for identifying the goals or objectives of health systems against which outcomes should be judged and a definition of health system performance which can be quantified (10).

Murray and Frenk (11) recently defined health system performance in a way which would enable performance in different countries to be compared and performance over time within a country to be monitored. To illustrate the concept, in Figure 1 the goal of the health system is measured on the vertical axis (here, labelled health) while the inputs to producing the goal are on the horizontal axis. The upper line represents the frontier, or the maximum possible level of the goal (health) that could be obtained for a given level of inputs.

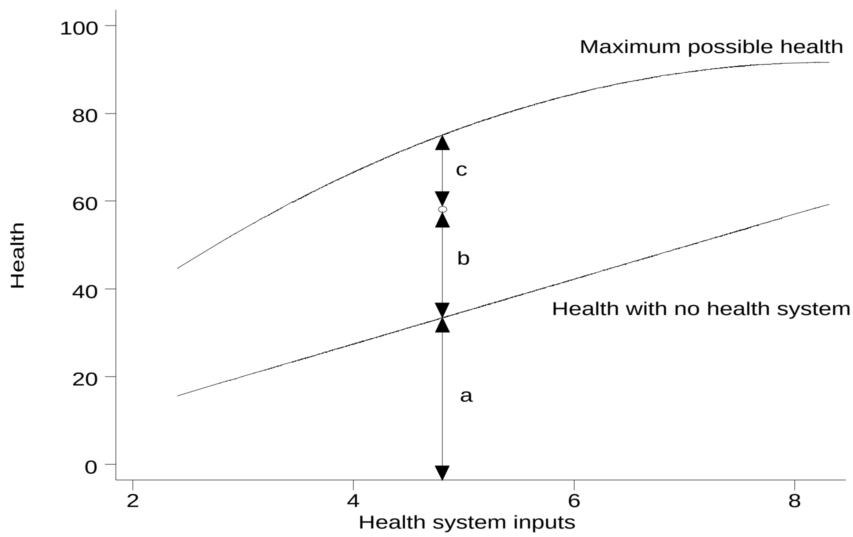

Figure 1: Health System Performance

On a farm, for example, output would be zero in the absence of inputs, but in the health sector, health levels would not be zero (i.e., the entire population would not be dead) in the absence of health expenditures and a functioning health system. The lower "frontier" in Figure 1 is defined as the health level that would occur in the absence of the system. Assume that a country is observed to have achieved (*a*+*b*) units of health.

Murray and Frenk defined system performance as $b/(b+c)$. This indicates what the system achieves compared to its potential. The challenge for health sector reform is to find a way of measuring health system performance in a systematic way, to allow comparison across countries and within countries over time. That is the purpose of this paper.

The World Health Report 2000 (12) defines three intrinsic goals of the health system – to improve health, to be responsive to the legitimate demands of the population, and to ensure that no one is at the risk of serious financial losses because of ill health. However, here we focus on performance in terms of achieving the goal of improving health -- measured in terms of healthy life expectancy using Disability Adjusted Life Expectancies (DALE) as the indicator (see (13)). Tandon et al. (14) describe how performance in terms of all health system goals is measured. Data sources have been described elsewhere (15) so the paper will focus largely on the results.

Attempts to evaluate the impact of health sector reforms, and to monitor health sector performance over time, have been hindered by the lack of agreed methods for quantifying the extent to which countries use scarce health resources to meet the goals of the health system. This paper suggests a way of doing this, and it is hoped that the resulting discussion will lead to the development of ways of routinely measuring and monitoring the performance of health systems with a resulting improvement in the health of the affected populations.

2) Estimation Methods

a) Theory

There is a long tradition of measuring efficiency in the applied economics literature, especially so in the fields of agricultural and industrial economics. The analytical framework traces its origins to Farrell (16) who introduced a methodology for measuring economic, technical and allocative efficiency. *Technical efficiency* was defined as the ability to produce the maximum possible output from a given set of inputs and is measured in terms of the relationship between observed output and the maximum attainable output for the observed inputs, for example, the ratio of actual output to maximum possible output. In Figure 1, it would be defined as $(a+b)/(a+b+c)$, assuming that a single input is measured on the horizontal axis.

Given the obvious relationship between the definition of technical efficiency and our definition of health system performance, we use the term "efficiency" to refer to system performance in the remainder of this paper. It is more easily understood, but there is one important qualification to make. Our efficiency is more than technical efficiency narrowly defined. In this paper, we use health expenditure as the indicator of inputs to the health sector, and health as the output. Efficiency then reflects not only whether health programmes and interventions are produced at the lowest possible cost (i.e., technical efficiency), but also whether the health system chooses to provide the most cost-effective set of programmes or interventions for the given level of expenditure (sometimes called allocative efficiency in the health economics literature).[1]

Given that the maximum attainable output (or the "frontier" production function) is not observable, two approaches could be used to estimate it. The first involves defining the feasible set of health interventions which could be used in that system, identifying their costs and associated outcomes if they are used in a technically efficient way, and choosing the set that maximises the objective function (here, health) for the resources available. It has not yet been widely applied – we found only two applications. For example Tengs et al. (17-20), focussing on only a small set of primary prevention interventions in the United States, showed that reallocating resources from inefficient interventions currently undertaken to those which are cost-effective but not fully undertaken, would save an additional 600,000 years of life annually for the same level of investment. Murray et al. (21) used a programming model to suggest that a typical country in sub-Saharan Africa could improve health outcomes by 40% simply by reallocating resources to the most cost-effective mix of interventions.

This type of micro-approach is very data intensive and there are not sufficient data to apply it to many countries – although it will be something that is pursued by WHO in the future. But the fact that the two existing studies show that the US and sub-Saharan Africa could achieve higher levels of health for the resources they currently have available lends important credence to our results.

The second approach is to estimate the frontier econometrically from a sample of observed inputs and outputs from different countries. In this paper, we focus on econometric estimation where the principal approaches can be broadly classified into those which use cross-sectional data and those which use panel data (a time series of cross-sectional data). Both deterministic and stochastic frontier models have been applied to cross-sectional data.

Examples of the deterministic approach include early attempts such as Farrell (16) which follow a non-parametric estimation of the frontier. A piece-wise linear "envelope" is estimated such that all observed data points lie either on this frontier or below it. In this approach, also known as free disposal hull (FDH) analysis, all points lying on the frontier are considered fully efficient. Points below the frontier are technically inefficient, with the vertical distance from the frontier measuring the degree of output inefficiency. Figure 2 illustrate this approach with the sample for 1997 of 191 countries that are members of the World Health Organization (WHO). In this case, the production of health is modelled using the log of health expenditure per capita as the only input and log of DALEs as the measure of population health. We are aware that this is not the appropriate specification of the relationship between health and the health system, and the results we report in the Annex to this paper and in the Annex tables of the World Health Report 2000 (12) are based on a more complex specification. The assumption is used here only to make it easier to illustrate the differences between the different techniques.

[1] *Allocative efficiency* refers to the optimal choice of input proportions, given their respective prices. If interventions are regarded as inputs, then this translates into choosing them according to their cost-effectiveness. Technical and allocative efficiencies put together give us the concept of *economic efficiency*, hence our use of the generic term "efficiency".

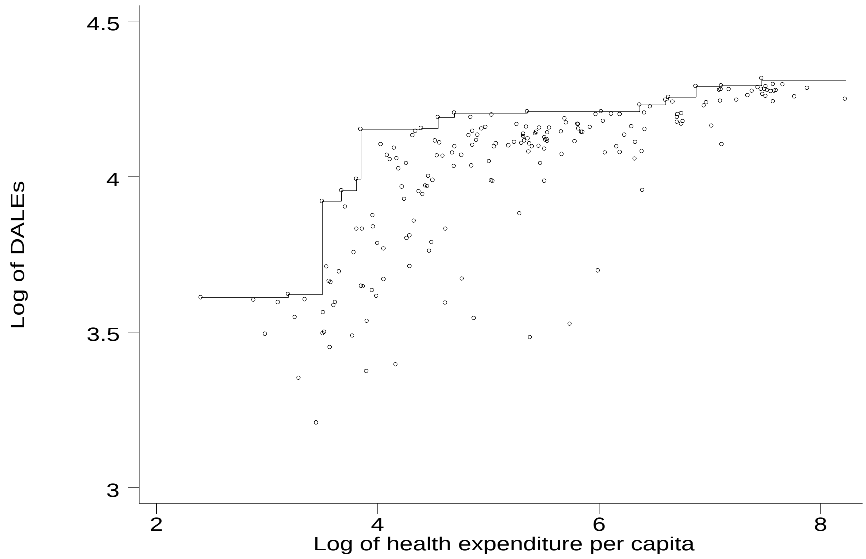

Figure 2: Frontier Production Function for 191 countries 1997: Free Disposal Hull Analysis

There are several disadvantages to FDH (Figure 2). First, multiple countries are on the frontier, each one being fully technically efficient at their level of expenditure. This makes the comparative analysis of efficiencies across the cross-section less informative. Secondly, random factors that might affect production are included in the measurement of inefficiencies (hence the categorisation as deterministic). Thirdly, this type of non-parametric analysis is not very robust to outliers or extreme data points, thereby diminishing its inferential power.

Other non-parametric deterministic approaches include data envelopment analysis (DEA). DEA uses linear programming methods to construct a linear envelope bounding the data relative to which efficiencies can be calculated. Multiple inputs and outputs can be analysed. The imputation of efficiencies can be most easily conceptualised in terms of the following linear program:[2]

$$\text{maximise}_{u,v}\left(\frac{u' y_i}{v' x_i}\right)$$

$$\text{subject to} \quad \frac{u' y_i}{v' x_i} \le 1,\ j = 1,\ldots,N,$$

$$\text{and} \qquad u, v \ge 0.$$

where x_i are inputs, y_i are outputs and u and v are scalar values chosen for each production unit such that the efficiencies of each unit are maximised and no efficiencies are greater than one. Since, however, the above problem has an infinite number of solutions,[3] an additional constraint is needed, and the program can be rewritten as:

[2] The explanation is from Coelli (70).

[3] If (u^*, v^*) is a solution, so is (αu^*, αv^*).

$$
\begin{aligned}
&\text{maximise}_{u,v}(u^{'}y_i) \\
&\text{subject to} \quad v^{'}x_i = 1, \\
&\qquad\qquad\quad u^{'}y_i - v^{'}x_i \le 0,\ j = 1,\dots,N, \\
&\text{and} \qquad\quad u, v \ge 0.
\end{aligned}
$$

Figure 3 plots the frontier implied by this approach for the same data. Fewer countries are on the frontier, but again, it is not possible to separate true inefficiency from random variation in this form of deterministic model (22,23).

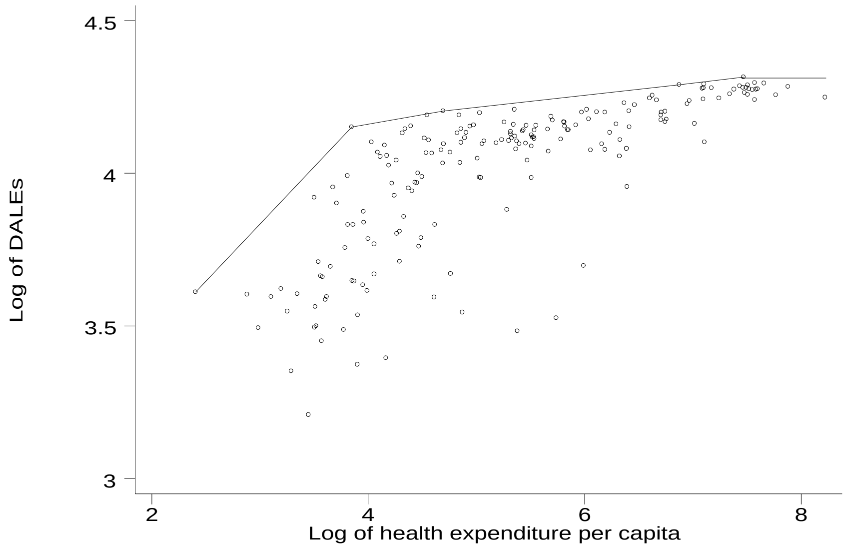

Figure 3: Frontier Production Function for 191 countries 1997: Data Envelopment Analysis

An example of a parametric deterministic approach is corrected ordinary least squares (COLS) where a production function is first estimated using ordinary least squares (OLS). The OLS intercept parameter is then shifted up by the value of the largest positive residual to give the equation for the frontier. This ensures that all data points lie below the estimated frontier. Using the same data, this approach is highlighted in Figure 4.[4] A major disadvantage of the COLS estimator, also evident in Figure 4, is that it tends not to "bound" the data as closely as possible, given that the frontier is restricted to be a parallel shift up of the estimated OLS production function (24). This can unnecessarily penalise the efficiency estimates of countries having relatively low OLS residuals, and can especially be problematic if there is heteroskedasticity in the sample, as in our case. Furthermore, COLS remains an inherently deterministic approach: technical inefficiency estimates cannot be disentangled from random errors.

[4] In the COLS example, log of DALEs is assumed to be a function of log of health expenditure per capita and its square.

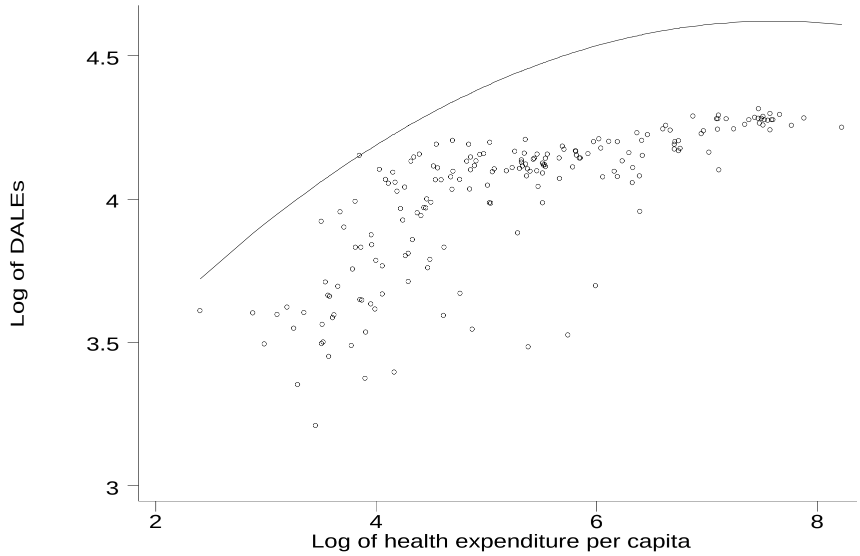

Figure 4: Frontier Production Function for 191 countries 1997: Corrected Ordinary Least Squares

Stochastic specifications of the frontier production function were pioneered by Aigner et al. (25) and by Meeusen and van den Broeck (26). In this approach it is explicitly acknowledged that some deviations from the maximum observed output are potentially due to factors unrelated to inefficiency (e.g., exogenous shocks that may be outside the control of the production system). A parametric production function is estimated and it is assumed that the error term has two components: one representing random errors, and the other representing technical inefficiency (these types of models are often referred to as "error components" models). Mathematically, this can be characterised as follows. Let Y_{it} denote output of country *i* in time period *t*. Suppose we have a production function of the following form:

$$Y_{it} = \alpha + X'_{it}\beta + v_{it} - u_i \,, \qquad \text{(i)}$$

where X_{it} is a vector of inputs and v_{it} is the error term with mean zero. The term $u_i \geq 0$ is a random variable representing country-specific technical inefficiency, and is constrained to be always non-negative.[5] Technical efficiency is then defined as the ratio of the expected value of observed output for country *i* to the expected value of the output when $u_i = 0$.[6] In other terms,

$$TE_i = \frac{E(Y_{it} \mid u_i, X_{it})}{E(Y_{it} \mid u_i = 0, X_{it})}. \qquad \text{(ii)}$$

The denominator represents the frontier since the technical inefficiency term takes the value of zero. In order to estimate such models, it is necessary to make two additional assumptions regarding: (a) the distribution of the country-specific technical inefficiency term *u*, and (b) the separation of *u*, given that the joint term (*v*-*u*) is observed. Typically, given the non-negativity constraint, the *u*'s are assumed to be distributed half-normal,

[5] In this model, u_i is assumed to be time-invariant. However, this need not be the case.
[6] This is assuming the output is measured in original units.

truncated-normal, exponential, or gamma. The choice of the distribution is somewhat arbitrary. However, despite some differences in mean efficiency scores, efficiency rankings are fairly robust across the various distributional assumptions (24). Once the distribution for *u* is specified, the model parameters can be estimated using maximum likelihood methods.

Using the same 1997 cross-sectional data, Figure 5 plots the frontier assuming the *u*'s are distributed truncated-normal. Notice that some of the data points lie above the frontier. This is different from the deterministic case precisely because of the error decomposition: the *expected* value of the output must lie on or below the frontier, however the *actual* value of output may well be above the frontier if the random error for a given country is big enough. In addition to the reliance on a distributional assumption for *u*, one additional criticism of the stochastic production function approach is that the predictions of *u*, given that (*v*-*u*) is jointly observed, can be somewhat unreliable.[7]

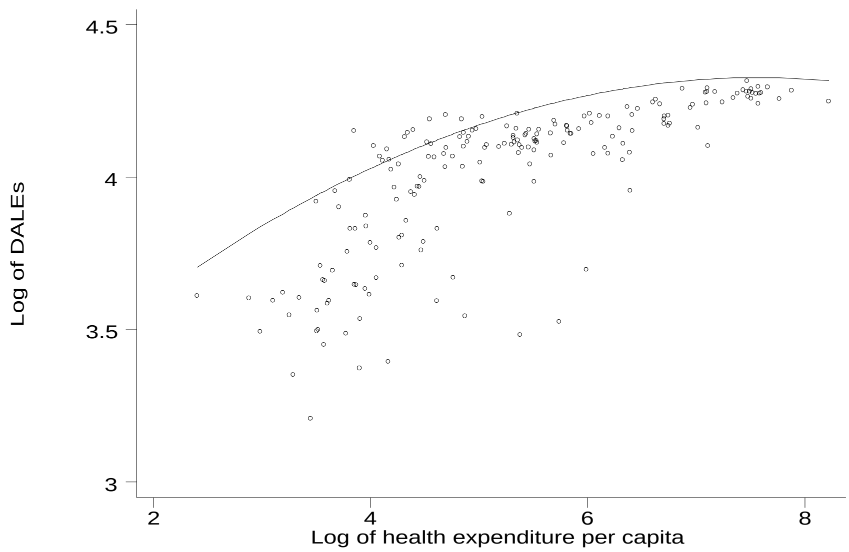

Figure 5: Frontier Production Function for 191 countries 1997: Stochastic Frontier, Truncated Normal

Panel data estimation is more efficient than models using cross-sectional data in extracting information on inefficiencies from the overall residual because of multiple observations over time and greater degrees of freedom (27). The fixed-effects approach is preferable to stochastic frontier estimation because it obviates the need for distributional assumptions for the u_i's, and it is not necessary to assume that the inefficiency terms are unrelated to the independent variables. We therefore used a fixed-effects panel data model. The fixed effect model was preferred to the random effects approach because the Hausman test rejected the hypothesis that the efficiency estimate is independent of the other inputs to the health system, as reported subsequently.

Consider the specification of the production process as before:

[7] One approach to this decomposition is to exploit the conditional distribution of *u* given (*u*-*v*) as in Jondrow et al. (71). However, this method is not perfect and some intrinsic variability in the estimation of *u* remains.

$$Y_{it} = \alpha + X'_{it}\beta + v_{it} - u_i \,. \qquad \text{(iii)}$$

This can be rewritten as:

$$Y_{it} = \alpha_i + X'_{it}\beta + v_{it} \,. \qquad \text{(iv)}$$

where the new intercept α_i = (α - u_i) is now country-specific, and estimates can be found using a fixed-effects approach. The frontier intercept is represented by α, and the u_i's are the country-specific inefficiencies. In order to ensure that all the estimated u_i's are positive, the country with the maximum α_i is assumed to be the reference and is deemed fully efficient. Mathematically,

$$\hat{\alpha} = \max(\hat{\alpha}_i) \qquad \text{(v)}$$

and

$$\hat{u}_i = \hat{\alpha} - \hat{\alpha}_i \,. \qquad \text{(vi)}$$

This normalisation ensures non-negative u_i's. Technical efficiency is defined in the same manner as before:

$$TE_i = \frac{E(Y_{it} \mid u_i, X_{it})}{E(Y_{it} \mid u_i = 0, X_{it})} \,. \qquad \text{(vii)}$$

Following our definition of efficiency in a way that is consistent with health sector performance described above, we modified this equation by subtracting out the predicted minimum level of Y_{it} (denoted by M_{it}) from the numerator and denominator. Overall efficiency, or E_i, is:

$$E_i = \frac{E(Y_{it} \mid u_i, X_{it}) - M_{it}}{E(Y_{it} \mid u_i = 0, X_{it}) - M_{it}} \,. \qquad \text{(viii)}$$

To highlight the fixed-effect panel data approach, consider now that we have data from 1993-1997 for all 191 countries. Figure 6 depicts how this approach fits the frontier, and how efficiency is determined relative to this frontier, using only health expenditure per capita as the input. Again, it should be remembered that this specification is simply for ease of exposition, and was not used to obtain the final efficiency ranks or scores.

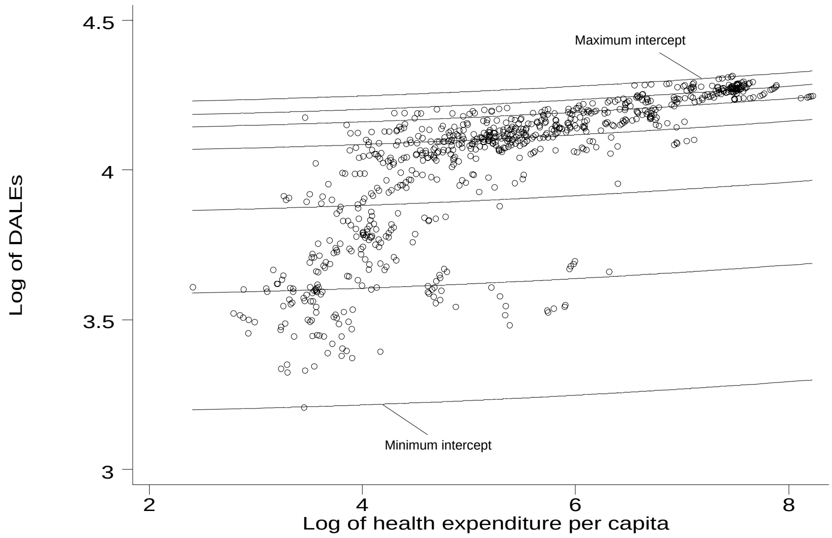

Figure 6: Frontier Production Function for 191 countries, Panel Data 1993-97: Fixed Effect Panel Data

Basically, the fixed-effect model is a variable intercept model. As can be seen in the figure, there is a different intercept for each country (the predicted values for seven countries are plotted in the figure). The country with the maximum intercept is taken as the reference country (the frontier), and distance from this maximum gives a measure of technical efficiency. There is an analogous relation with the COLS estimator: whereas in COLS the intercept was shifted up by the maximum residual, here the intercept for the predicted value of the frontier is shifted by the value of the maximum country-specific fixed effect. For cases where there are multiple years per country, the actual country data points are distributed randomly with mean zero around the country-specific intercept. For cases where there is only one data point for a given country, the intercept shift is like an observation-specific dummy variable.

b) Applications

There is a large applied literature on the measurement of technical efficiency in a variety of settings, especially in agriculture and industry, and it is now being applied to an increasing number of topics such as education and local government (28). Applications include the measurement of production inefficiencies among weaving firms in Indonesia (29), small manufacturing firms in Pakistan (30), the airline industry in the U.S. (31), dairy farms in Australia (32), and paddy farmers in India (33). A less traditional application involves a study of technical efficiency differences of senior secondary schools in Finland (34): outputs were measured in terms of number of students who passed their grade and scores on matriculation examinations. Different inputs included teaching hours per week, experience and education of teachers, and the education level of parents.

Typically, the literature has tended to follow a two-step procedure: separating the measurement of technical efficiency from its exogenous determinants. “Controllable” inputs to the production process are first related to outputs to measure efficiency. In a second stage, non-controllable exogenous determinants of differences among these estimated efficiencies are examined.

In health, the technique has been used largely for micro-level service efficiency, particularly hospitals where DEA has become relatively popular (35-37). But for comparing the efficiency of systems across countries or over time, we have found only one study in the tradition of frontier production function analysis.[8] Gupta et al. (38) applied a deterministic frontier production function (free disposable hull analysis described above) to a panel of data from 85 countries for various years between 1984 and 1995. They defined the frontier for life expectancy and infant mortality, and used average per capita government spending on health and education as inputs to the production process. They also controlled for the initial level of development, given by the initial GDP per capita in international dollars. An efficiency score for each country was estimated showing what proportion of the maximum feasible level of health each country had achieved. However, by using a deterministic frontier, the fact that some of the observed high or low efficiency could have been due to random chance was ignored. Moreover, by including the initial level of development in the production function, this departs from the tradition of including only controllable inputs in the estimation of technical efficiency described above. Uncontrollable factors such as income per capita could have better been introduced at a second stage to explain observed variations in efficiency. In addition, by only including government spending, the contribution of the private sector was ignored.

We borrow from the long tradition of applied efficiency analysis in other sectors by introducing a non-deterministic frontier approach to the analysis of health sector efficiency. This allows us to separate random error from efficiency in the estimation procedure. The model is estimated using a panel of data from all 191 countries that are members of WHO.

c) *Model Specification*

Modern production studies generally use a flexible functional form. Perhaps the most versatile is the translog model. For the two-input case (X_1, X_2), the translog model for fixed-effect panel data estimation can be written as follows (all variables in logs):

[8] An earlier version, though not in the tradition of production function analysis, was the work of Wang et al. (72) who built on the work started in the World Development Report 1993 (73). This study regressed various indicators of health (e.g. under 5 mortality, life expectancy at birth) as a function of education, income per capita, time and the respective interactions using panel data from 1960 to 1990. The difference between the observed and predicted outcome was used as the numerical estimate of country performance on health. The approach entails a number of problems. First, although their output measures are health-sector specific, they do not consider any direct inputs in the health sector. This makes it difficult to infer specific health policy implications from their analysis. Second, there is no direct indicator of inputs to the health sector in the model which implies that country performance is independent of input levels, i.e., there is no heteroskedasticity. We show subsequently that is not the case for our sample. Finally, approximately half the countries were good performers, having better than expected levels of health. This approach gives countries no incentive to improve because it does not identify the fact that most could still have done better than they did.

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 (X_{1it})^2 + \beta_4 (X_{2it})^2 + \beta_5 (X_{1it})(X_{2it}) + v_{it} .$$

In effect, the translog function is a second-order Taylor-series approximation to an unknown functional form (39,40). Both the Cobb-Douglas and the Constant Elasticity of Substitution (CES) production functions can be derived as restricted formulations of the translog functional form (41). We estimated the full translog model as well as nested versions of the model including the Cobb-Douglas log-linear formulation, and the Cobb-Douglas log-linear with each of the square terms and the interaction term separately.

d) Data

To measure efficiency using the production function approach, three general types of variable are necessary. First, it is necessary to identify an appropriate outcome indicator that represents the output of the health system. Second, it is necessary to measure the health-system inputs that contribute to producing that output, and third, it is necessary to include the effect of controllable non-health-system determinants of health.

In terms of output, it is generally agreed that one important goal of the health system is to improve population health (11). We measure health taking into account both mortality and ill health rather than using an indicator such as life expectancy at birth which relates solely to mortality. Our approach is based on an indicator of healthy life expectancy (DALE), and the rationale for it and technical details behind its derivation for the 191 countries in the sample are described elsewhere (13,42).

One option was to include the physical inputs to the health system in the production function: for example, for a farm the inputs might be the area of cultivated land and the total quantities of labour, fertiliser, water and seed employed. This type of information could not be obtained for the health sectors of the 191 countries in our sample, so we used total health expenditure per capita (public and private) in 1997 international dollars (using purchasing power parities to convert from local currency units), as a summary measure of physical inputs to the health system. This had the advantage of allowing us to interpret efficiency more broadly, as described earlier. It is also in the tradition of recent work exploring the efficiency of government expenditures in different areas, including local government authorities (28,43). The data sources and methods of calculation of health expenditure are described elsewhere (15,44). The scatter plot for DALE versus health expenditure per capita in PPPs for the 191 countries in the sample in 1997, depicted in Figure 7, shows the expected positive relationship.

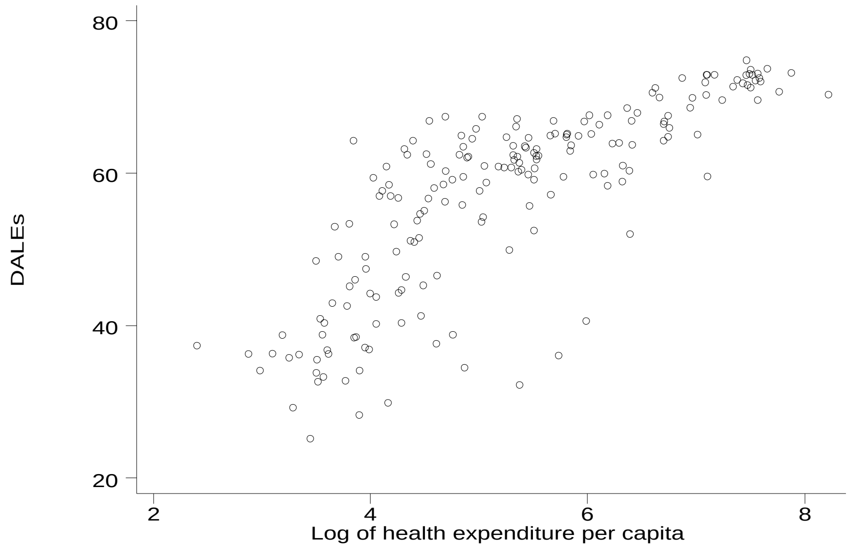

Figure 7. Scatter Plot of Health Expenditure per Capita vs. DALEs, 1997, 191 countries

It is well recognised that health is not solely a function of services provided by the health system, however broadly the system is defined (11). Indeed, some of the published work to date investigating what amounts to an aggregate health production function at the national level has focused exclusively on non-health-system determinants, such as income and educational levels, measured in various ways (45,46). While this approach can be defended in studies that aim to track mortality trends, it is not appropriate for efficiency analysis of the health system itself, where the estimation of efficiency should be based on controllable inputs to the production process. Identifying relevant variables that are available for all countries, but which are not highly correlated with health expenditure per capita, is difficult. For example, income per capita – one of the most obvious indicators of general development – is highly collinear with health expenditure per capita. While it would be possible to add income per capita directly into the estimated equations, income is not a direct determinant of the production of health. It works through other inputs such as education and housing, and it is better to capture these inputs directly. In terms of the tradition of production function analysis, income per capita can be regarded as an uncontrollable variable outside the production process which should not be included in the estimation of efficiency in the first stage.

The most widely available information on non-health-system inputs to production is for education, and there is a large literature showing that education is strongly associated with the health of both children and adults in developed and developing countries (47-61). The most sensitive indicator of the relevant kind of educational attainment is average years of schooling in the adult population. For sources and methods of calculating this educational attainment measure, see (15).

Thus, three data series comprise our model: DALE, health expenditure and average educational attainment in the adult population. Our panel covers the years from 1993 to 1997 for all 191 member countries of WHO, with some missing data for some countries and years. As mentioned above, provided there are enough countries with multiple data points (i.e., provided the data set contains enough “degrees of freedom”), an unbalanced panel does not prevent estimation of a fixed effect model, since a country with an observation for only one year will simply carry an implicit observation-specific dummy variable.[9] In our panel, every country had an observation for 1997, although about fifty countries had an observation only for that year (i.e., the remaining 141 countries were complete in all years).

e) Minimum Possible Health Output

Earlier we argued that efficiency in the health sector differs from technical efficiency as usually measured in that health levels would not be zero in the absence of health expenditures. This minimum level of health in the absence of a functioning system was estimated from a cross-section of 25 countries at around the turn of the century (average year, 1908). The assumption is that modern health systems were not functioning at that stage, so health levels could not be attributed to the system.

Following Preston (45,46), we regressed our measure of health – healthy life expectancy measured in terms of DALE – on literacy rates, as the proxy for the non-health system determinants of health, and estimated health expenditures. In order to be consistent with the estimation of the upper frontier, we did not include income per capita for reasons described elsewhere in this paper. For sources of data and methods, see (15).

We experimented with different functional forms of the model. Predicted health expenditures were not significant in most, which supports the hypothesis that the health system played little role in improving health at the turn of the century. The fact that literacy was always significant suggests that non-health system factors were predominant in determining health outcomes at the turn of the century.

A corrected ordinary least squares (COLS) procedure was used to adjust the line estimated to the minimum country, in an analogous way to the standard frontier. The relation thus obtained can be assumed to be the minimum achievable health production function today (i.e., the lower line in Figure 1), since in the absence of a health system every country today, conditional on its current literacy rate, should be able to do at least as well as would be predicted by the turn of the century relation between literacy and health.

f) Uncertainty Analysis

Results of statistical inference are routinely accompanied by a measure of their inherent uncertainty due to a finite sample size. Confidence intervals calculated on the basis of knowledge about the distribution of the error term are a standard adjunct to the presentation of econometric estimates in the scientific literature. However, a problem arises when the distribution of the quantity of interest is not known or is computationally very difficult to calculate. An example is when the quantity to be estimated is itself a function of other statistics, whose distribution may not be known. The principal statistic

[9] In instances where there is only one time period observation for a country, it is not possible to tease out the effects of random noise factors from efficiency effects.

of interest here is the efficiency of the health system in 191 countries. The critical parameters are the u_i and β_i described earlier. The estimates of E_i (equation viii), however, are a non-linear combination of these estimated coefficients.

The classic linear regression model assumes that all uncertainty (characterised as "error") has zero mean and is independently and identically distributed across observations. However, if there is uncertainty deriving from, say, a stochastic data generation process, in addition to estimation uncertainty due to the limitations of finite sample size, conventional expressions for the precision of the regression parameters need to be revised. Because practical analytical methods are generally lacking, it is almost universal to ignore this layer of uncertainty.

In many applied research situations, therefore, reported confidence intervals may be more like lower bounds of the true confidence intervals. However, the computational power afforded by computers can substitute for analytical methods by means of a simulation process repeated many times. Thus, to account explicitly for these two sources of uncertainty, and to generate realistic estimates of precision for our quantities of interest, we used the following Monte Carlo experiment.

Consider a single year in the panel data set described above:

$$Y = X\beta + \varepsilon,$$

where Y is DALE, X is the matrix of regressor variables, and β is the coefficient vector. The data sources and methods for obtaining the values in the regressor matrix are described in related publications (15,44). For estimation of the precision of the efficiency estimates, in accordance with the classical regression model we assume that the regressor matrix is fixed in repeated sampling.

To account for uncertainty surrounding DALE, for each country the relevant DALE distribution was estimated for 1997 (62). From each distribution, one observation was drawn and the frontier production function was estimated – that value of DALE, for each of the 191 countries, was regressed against the observations of health expenditure per capita and education. The process was repeated 1000 times, sampling without replacement from the DALE distribution for each country, each time running a new set of regressions of DALE on the X values. This resulted in 1000 estimates of efficiency for each country and 1000 estimates of each country's rank on efficiency.

The results on efficiency we report in the Annex are the mean values of 1000 estimates for each country, along with an 80% uncertainty interval constructed by setting the lower bound at the 10th percentile and the upper bound at the 90th percentile of the 1000 estimates obtained by simulation. Ranks are based on the mean efficiency score, while uncertainty intervals on rank are constructed by taking percentiles of the 1000 individual estimates of rank for each country, as above. Estimates of the coefficients of the frontier production function are obtained with a similar procedure, with the added feature that, because there are only four or five regression coefficients of interest, we use density estimates (smoothed histograms) to visualise the coefficient distributions (63). We report 95% uncertainty intervals for the regression coefficients, because this provides more discriminatory power against the null hypothesis that the coefficients are equal to zero.

3) Results

All possible variations of the full translog model were investigated. The chosen model is reported in Table 1, where log of DALE was estimated as a function of log of health expenditure, log of average years of schooling and the square of log of average years of schooling. We choose this parsimonious, nested version of the full translog model for reasons discussed subsequently. Table 1 also reports 95% uncertainty intervals around the estimated coefficients.

Table 1: Coefficient Estimates (Median, Mean and Uncertainty Interval) for the Frontier Health Production Function, Logged Variables, 191 Member Countries of WHO, Panel Estimates (1993 – 1997).

Coefficient estimate	Median	Mean	Uncertainty interval (95%)
Health expenditure	0.00885	0.00885	0.00868 - 0.00901
Average years of schooling	0.06296	0.06301	0.05877 - 0.06732
Square of average years of schooling	0.02166	0.02170	0.02029 - 0.02318
Constant	3.81256	3.81252	3.80862 - 3.81657
Max (u)	0.21346	0.21441	0.20730 - 0.22308

We stress again that these uncertainty intervals are not the statistical confidence intervals of the individual regressions. They were derived by omitting the lowest and the highest 2.5% of the estimates from the 1000 regressions described earlier. The distributions of the coefficients are shown in Figure 8 and are approximately normal. The constant term is the average value of the fixed effect in the sample. Max(u) is the maximum positive deviation from the average, and can be interpreted as the amount by which the best performing country does better than the average country. The sum of the constant term and the maximum positive deviation gives the value of the intercept of the health production function for the best performing country, and consequently, the intercept of the frontier. None of the uncertainty intervals of the estimated coefficients includes zero.

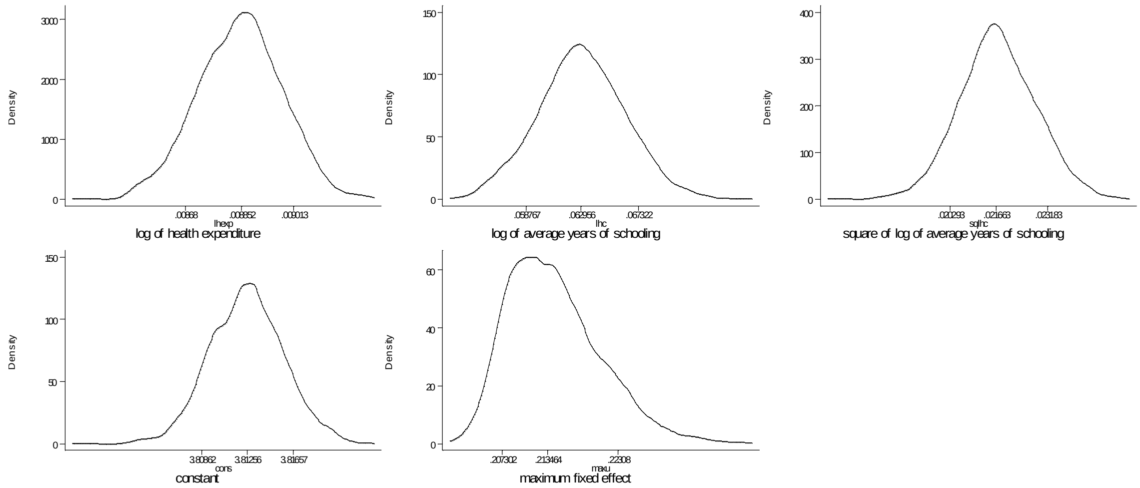

Figure 8: Distributions of the Coefficient Estimates for Log of Health Expenditure, Log of Average Years of Schooling, Square of Log of Average Years of Schooling, Constant, and Maximum Value of the Country-Specific Fixed Effect

We tested statistically whether we should use a fixed effects or random effects model. The Hausman test is a test of equality between the coefficients estimated via the fixed-effects and random-effects models. Assuming that the model is correctly specified, a significant difference in the coefficient estimates is indicative of correlation between the individual effects and the regressors. Where this correlation is present, the estimates using a random-effects model will be biased (64,65). Table 2 reports the coefficient estimates where the expected value of DALEs for each country is the dependent variable. As can be seen from the test statistic, the null of no correlation is rejected and a fixed-effects model is clearly preferable.

Table 2. Hausman Specification Test: Coefficient Estimates using Expected Value of DALEs as Dependent Variable. All Variables in Logs, 191 WHO Member Countries, Panel Estimates (1993-1997).

	Coefficients		
DALEs	Fixed-Effects	Random-Effects	Difference
Health expenditure	0.008841	0.0173468	-0.0085027
Average years of schooling	0.0629097	0.0880001	-0.0250903
Square of average years of schooling	0.0217392	0.0504634	-0.0287242
$\chi^2(3)$	106.87		
p value	0.000		

The resulting estimates of efficiency (i.e., performance) for each country are reported in the Annex, along with the uncertainty interval around the index.[10] Estimated efficiency varies from less than 0.10 to close to 1.00. In any given regression, one country has an efficiency index of 1.0 but across the 1000 regressions, no one country was consistently the most efficient in the sample. The reported efficiency index is the average of the 1000 scores from the individual regressions.

The ranks reported in column 1 of the table are based on the average efficiency index. The uncertainty interval is obtained by omitting the top and bottom 10% of ranks from the individual regressions. The distribution of ranks is asymmetric. Therefore, for instance, the uncertainty interval for Malta's rank is from 1 to 4, while for Oman it is 1 to 5, yet the overall rank of Oman is higher.

This is further illustrated in Figure 9 which shows the uncertainty intervals for the ranks graphically. Some uncertainty intervals are tight, and some are much broader. This depends largely on the uncertainty intervals around DALEs discussed earlier, which are higher for some countries. But is also depends on whether there are a number of countries clustered around a similar average efficiency index. So, for example, the uncertainty interval around the rank of Oman is very tight (1 to 5), while that of Colombia is broader (41 to 64). It is possible to claim that Oman has a significantly higher rank than Colombia, but it would not be appropriate to claim that Colombia performed significantly better than China in producing DALEs (rank: 52 to 65) even though on average Colombia had a higher rank.

[10] These are the estimates of Table 10 in the Annex to the World Health Report (12)

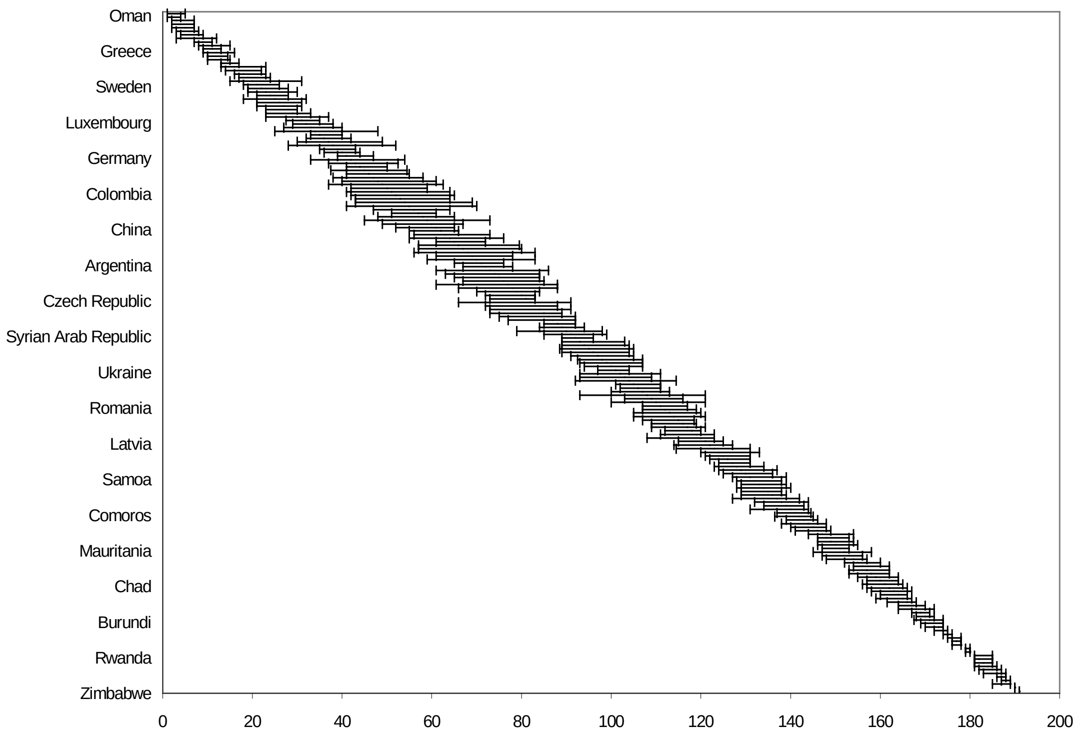

Figure 9: Uncertainty Intervals, Ranking of Performance on Health Level, 191 countries

Now it is possible to return to the choice of functional form. According to the interpretation of the translog model as a second-order approximation to an unknown functional form, the full version of the translog is, *a priori*, the reference standard. However, when parsimony is added as a criterion for model choice, the model we report -- with log of health expenditure, log of average years of schooling and the square of average years of schooling as regressors -- is both parsimonious and maps most closely to the reference standard (Figure 10). We compare rank order correlation in Figure 10, since efficiency will invariably increase with the addition of terms on the right hand side of the regression (unless the added term is completely collinear with another, it will always explain additional sample variance). Thus, the appropriate criterion for judging predictive stability across models is rank order correlation. The relevant correlation coefficients are found in Table 3.

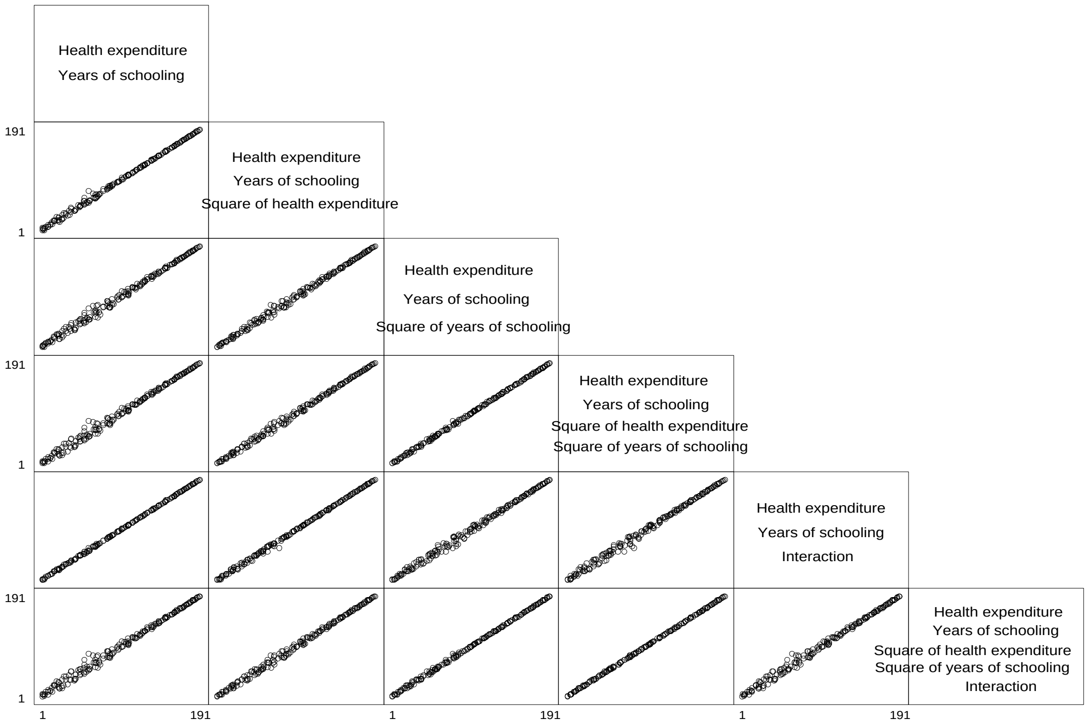

Figure 10: Correlation of Performance Ranks with Different Functional Forms

Table 3: Pearson's Correlation: Correlation of Country Ranks using Nested Variations of Translog Model Specifications versus the Most Generalised. All variables in Logs. *p* Values in Parentheses.

Independent Variables	Full translog model
Health expenditure Average years of schooling	0.9958 (0.000)
Health expenditure Average years of schooling Square of average years of schooling	0.9980 (0.000)
Health expenditure Average years of schooling Square of average years of schooling	0.9996 (0.0000)
Health expenditure Average years of schooling Square of health expenditure Square of average years of schooling	0.9999 (0.0000)
Health expenditure Average years of schooling Interaction (Health expenditure, Average years of schooling)	0.9967 (0.0000)
Health expenditure Average years of schooling Square of health expenditure Square of average years of schooling Interaction (Health expenditure, Average years of schooling)	1.0000 (0.0000)

The above results (Figures 10 and Table 3) also show clearly that the rank of the different countries is very robust to the functional form of the translog regression. The rank correlations are extremely high no matter what combination of variables is included, suggesting that poor performers perform poorly in all specifications and rank is not an artefact of the choice of model. Conversely, high performers perform well in all specifications.

To further test robustness, we explored whether the inclusion of possible other non-health system determinants of health would make a difference to the ranking – in addition to our proxy for non-health system determinants, average years of schooling. Income per capita is one option, but does not of itself improve health. It must operate through some other mechanism such as the ability to purchase better care (already in the health expenditure variable), better education (in the equation already), increased nutrition, or better housing, for example. Because possible other direct explanators were difficult to identify and measure for all countries in our sample, we defined a new variable obtained by regressing income per capita on the regressors already in our efficiency equation. The residual from the regression of income on health expenditure per capita, average years of schooling and average years of schooling squared was estimated. This residual can be interpreted as the part of income which might act through mechanisms other than health expenditures and education – or possible other pathways (called POSOTHER). POSOTHER was added to the fixed effects regression as a proxy for these possible other pathways related to income, and the efficiency analysis and ranks were recomputed.

When any new explanator is added to an equation, the residual of the dependent variable that is left unexplained is smaller, and accordingly the efficiency index is higher with POSOTHER. However, the correlation between the rankings under the two sets of estimates is very high (0.9913) showing that inclusion of POSOTHER does not have a significant impact on the relative rankings of the countries based on their efficiency in producing health. For this reason, and because it is not possible to explain which determinants of efficiency picked up by POSOTHER are controllable inputs or not, we chose to use the more parsimonious form of the equation reported above.

4) Discussion

This is the first time that the efficiency or performance of the health system in producing health has been evaluated across such a large number of countries. Previously, most of the knowledge of health sector performance has been anecdotal, or based on case studies. Received wisdom, for example, is that Sri Lanka and China have been very efficient in producing health (51,52). Our results suggest that this is no longer the case, as both Sri Lanka and China perform less well than many other developing countries. On the other hand, Oman performs extremely well – and indeed Oman has been able to reduce significantly the rate of child mortality over the last 25 years (66).

Many of the poorest performers are those in which there has been significant civil unrest over the period, which is not surprising. Many of the others are countries with a high prevalence of HIV/AIDS (Figure 11). In fact, the prevalence of AIDS can reduce DALEs by almost 15 years in some of the most highly endemic countries such as Botswana. We debated whether we should account for the presence of AIDS in the assessment of efficiency, but decided not to on the grounds that the health system should be held, at least partly, accountable for the fact that AIDS has not been controlled. We recognise that for some purposes it may be useful to evaluate efficiency in the absence of the impact of the AIDS epidemic. This type of counterfactual analysis will be undertaken in the future. Further work will also involve exploring whether factors such as geographical location or population density – uncontrollable exogenous factors – can help to explain observed differences in health system efficiency as the second phase of a standard frontier production function analysis.

While it is undeniable that other non-health-system variables may affect health (housing quality, environmental conditions, etc.), data on many of these are difficult to find or estimate. In addition, many of these will be highly correlated with educational attainment, which can be interpreted as a summary measure of non-health system inputs to health production.

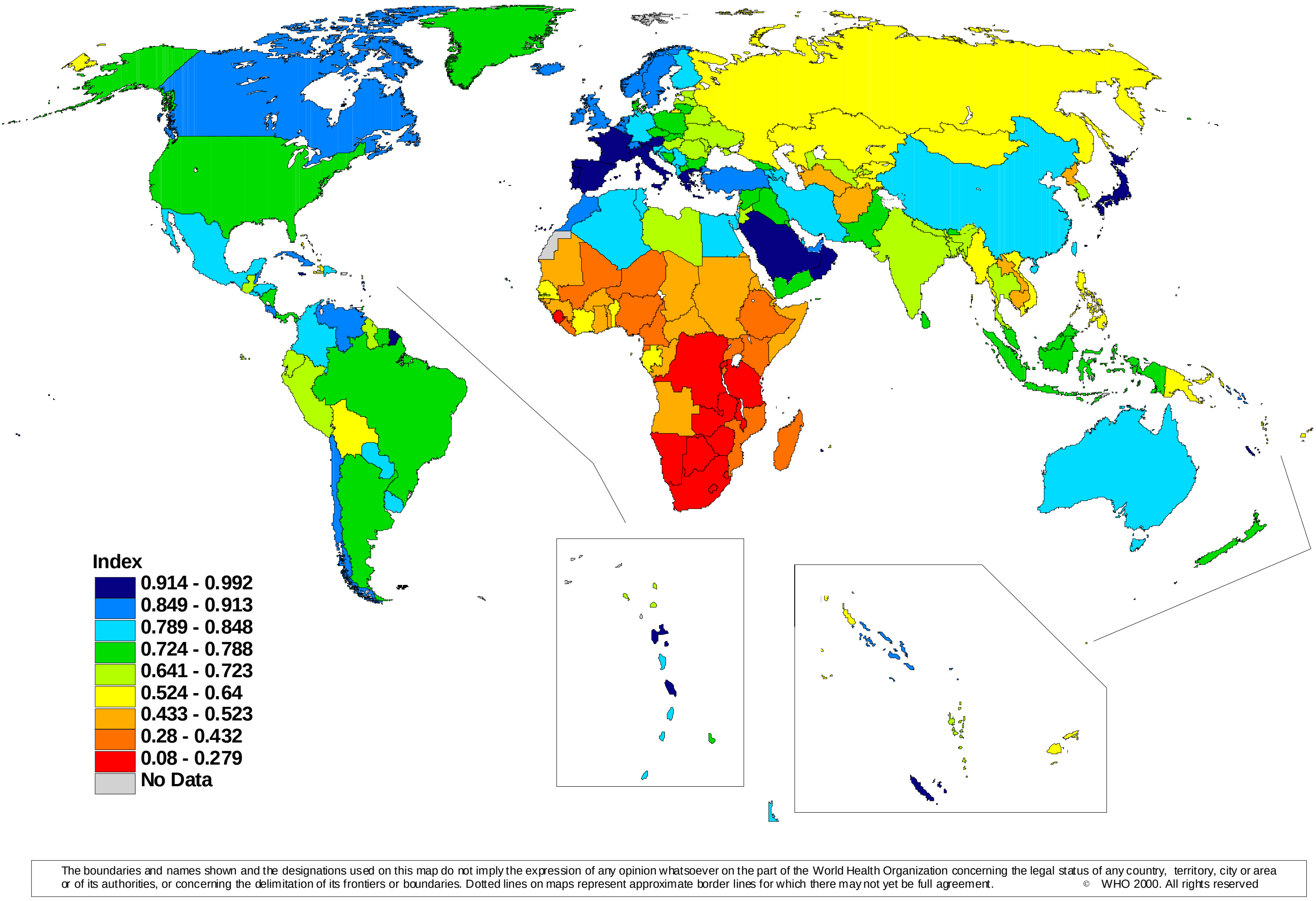

Figure 11: Global Distribution of Performance on Health

On the other hand, our estimates almost certainly overestimate efficiency. The fact that the countries ranked from 1 to 5 have efficiency indexes above 0.97 does not mean that they could only improve their systems by 3%. It means that compared to the most efficient country in the sample, they could improve by 3%. We have no way of estimating rigorously the potential of the highest performing countries to become more efficient, but micro studies suggest the potential is there.

It is also important to note that efficiency is positively related to the level of health expenditure per capita (Figure 12).[11] Indeed, the results suggests that it is very difficult for countries to be good performers below an expenditure per capita of approximately $60 in 1997 international dollars. This implies that there is an apparent minimum level of health expenditure below which the system simply cannot work well. We estimate that it would cost just over US$ 6 billion per year to bring health expenditures per capita up to this threshold in the 41 countries with lower expenditure, important information for governments, bilateral agencies and donors.

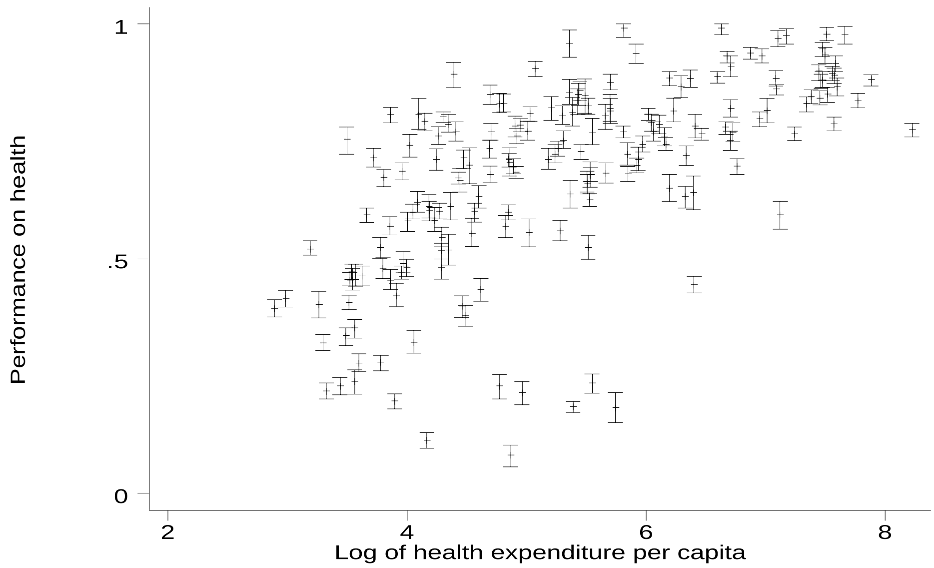

Figure 12: Efficiency (Performance) on Health Level by Health Expenditure per Capita

Despite this, there is still enough variation in efficiency at all levels of expenditure to suggest there are two critical ways of improving health outcomes. The first, which has been the focus of this paper, is to increase the efficiency of the health sector -- move toward the frontier. The second is to increase health expenditures, or move along the expansion path in the spirit of the World Health Report 1999 (67). This is illustrated in Figure 13, which plots predicted levels of health as a function of health expenditures per capita and educational attainment for our preferred frontier equation.

[11] This is the reason why the random effects model was rejected by the Hausman test.

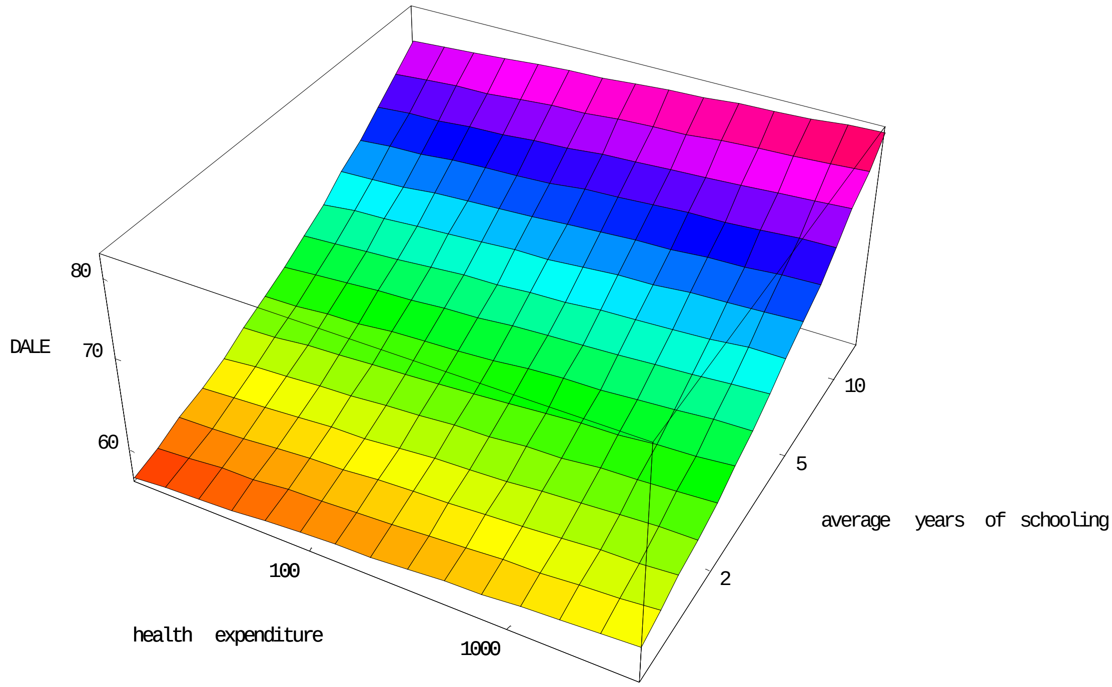

Figure 13: Logarithmic-Scale Plot of the Expected Value of the Frontier Health Production Function, 191 Member Countries of WHO, Panel Estimates (1993–1997)

The question of how to improve efficiency then becomes paramount, and that is the topic of the World Health Report 2000 itself (12). Certainly reducing waste and inefficiency is one way. But the studies of Tengs et al. (18,20) and Murray et al. (21) reported earlier show that choosing the appropriate intervention mix is also important. Indeed, their findings that the health output of the USA and sub-Saharan Africa could be significantly improved by allocating currently available resources to a different mix of interventions is consistent with our results.

These results differ from those of two recent studies. Babazono and Hillman (68) estimated the relationship between total health care spending and health outcomes for OECD countries, controlling for other factors including non-health care spending per capita and found no evidence of a relationship. However, their conclusions are based on a mechanistic stepwise regression elimination in which they included several independent variables that are highly correlated with health expenditure (e.g., number of physicians per capita, and pharmaceutical expenditure per capita). A similar conclusion was reached by Filmer and Pritchett (69) for a slightly wider sample of countries. However, their control variables included income per capita. As discussed earlier, we reject income as a possible explanatory variable in the regressions designed to estimate efficiency partly on theoretical and partly on empirical grounds. In any case, including POSOTHER in the equation did not significantly alter the efficiency rankings of the countries.

Perhaps the most important conclusion of this work is that it is possible to measure and compare the efficiency of health systems across countries and over time. This is particularly important for countries introducing health system reforms, and we hope that this study encourages all countries to routinely measure the inputs and outputs of their health systems. We intend to extend this analysis to regions within countries in the near future. An important benefit from the debate that is likely to accompany this exercise will be development of improved data sources and estimation methods. The overall aim is, of course, to stimulate action that will improve the performance of health systems and contribute to improving the welfare of people.

Reference List

1. **Maynard A, Bloor K.** Health care reform: informing difficult choices. *Int J Health Plann.Manage.* 1995;**10**(4):247-64.

2. **Collins C, Green A, Hunter D.** Health sector reform and the interpretation of policy context. *Health Policy* 1999;**47**(1):69-83.

3. **Hussein AK, Mujinja PG.** Impact of user charges on government health facilities in Tanzania. *East Afr.Med J* 1997;**74**(12):751-7.

4. **Feldman R.** The ability of managed care to control health care costs: how much is enough? *J Health Care Finance* 2000;**26**(3):15-25.

5. **Moore M.** Public sector reform; downsizing, restructuring, improving performance. Geneva, Switzerland. World Health Organization, 1996 (Forum on health sector reform No.Discussion paper 7.)

6. **Soderlund N.** An essential hospital package for South Africa--selection criteria, costs and affordability. *S.Afr.Med J* 1999;**89**(7):757-65.

7. **Almeida C, Travassos C, Porto S, Labra ME**. Health sector reform in Brazil: a case study of inequity. *Int J Health Serv* 2000;**30**(1):129-62.

8. **Durham G, Kill B**. Public health funding mechanisms in New Zealand. *Aust.Health Rev* 1999;**22**(4):100-12.

9. **Ron A.** NGOs in community health insurance schemes: examples from Guatemala and the Philippines. *Soc Sci Med* 1999;**48**(7):939-50.

10. **DeRosario JM.** Healthcare system performance indicators: a new beginning for a reformed Canadian healthcare system. *J Healthc.Qual* 1999;**21**(1):37-41.

11. **Murray CJL, Frenk J.** A WHO framework for health system performance assessment. Geneva, Switzerland. World Health Organization, 1999 (Global Programme on Evidence for Health Policy Discussion Paper No.6.)

12. **WHO.** World Health Report 2000. Geneva, Switzerland. World Health Organization, 2000

13. **Mathers C, Sadana R, Salomon J. et al.** Estimates of DALE for 191 countries: methods and results. Geneva, Switzerland. World Health Organization, 2000 (Global Programme on Evidence for Health Policy Discussion Paper No.16.)

14. **Tandon A, Murray CJL, Lauer J. et al.** Measuring overall health system performance for 191 countries. Geneva, Switzerland. World Health

Organization, 2000 (Global Programme on Evidence for Health Policy Discussion Paper No.30.)

15. **Evans DE, Bendib L, Tandon A. et al.** Estimates of income per capita, literacy, educational attainment, absolute poverty, and income Gini coefficients for The World Health Report 2000. Geneva, Switzerland. World Health Organization, 2000 (Global Programme on Evidence for Health Policy Discussion Paper No.7.)

16. **Farrell MJ.** The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series A* 1957;**120**(3):253-78.

17. **Tengs TO, Adams ME, Pliskin JS, Safran DG, Siegel JE, Weinstein MC, Graham JD.** Five-hundred life-saving interventions and their cost-effectiveness. *Risk Anal.* 1995;**15**(3):369-90.

18. **Tengs TO.** Enormous variation in the cost-effectiveness of prevention: implications for public policy. *Current Issues in Public Health* 1996;(2):13-7.

19. **Tengs TO, Meyer G, Siegel JE, Pliskin JS, Graham JD, Weinstein MC.** Oregon's Medicaid ranking and cost-effectiveness: is there any relationship. *Med.Decis.Making* 1996;**16**(2):99-107.

20. **Tengs TO.** Dying Too Soon: How Cost-Effectiveness Analysis Can Save Lives. Dallas, US. National Center for Policy Analysis, 1997 (NCPA Policy Report No.204.)

21. **Murray CJL, Kreuser J, Whang W.** Cost-effectiveness analysis and policy choices: investing in health systems. *Bulletin of the World Health Organization* 1994;**74**(4):663-74.

22. **Coelli TJ.** Recent Developments in Frontier Modelling and Efficiency Measurement. *Journal of Agricultural Economics* 1995;**39**(3):219-45.

23. **Kalirajan KP, Shand RT.** Frontier Production Functions and Technical Efficiency Measures. *Journal of Economic Surveys* 1999;**13**(2):149-72.

24. **Kumbhakar SC, Lovell CAK.** *Stochastic Frontier Analysis.* Cambridge: Cambridge University Press; 2000.

25. **Aigner D, Lovell K, Schmidt P.** Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 1977;**6**:21-37.

26. **Meeusen W, van den Broeck J.** Technical Efficiency and Dimension of the Firm: Some Results on the Use of Frontier Production Functions. *Empirical Economics* 1977;**2**(2):109-22.

27. **Schmidt P, Sickles RC.** Production frontiers and panel data. *Journal of Business and Economic Statistics* 1984;**2**(4):367-74.

28. **Grossman PJ, Mavros P, Wassmer RW.** Public Sector Technical Inefficiency in Large U.S. Cities. *Journal of Urban Economics* 1999;**46**(2):278-99.

29. **Pitt MM, Lee L-F.** The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* 1981;**9**:43-64.

30. **Burki, Abid A, Terrell D.** Measuring Production Efficiency of Small Firms in Pakistan. *World Development* 1998;**26**(1):155-69.

31. **Cornwell C, Schmidt P, Sickles R.** Production Frontiers with Cross Sectional and Time Series Variation in Efficiency Levels. *Journal of Econometrics* 1990;**46**:185-200.

32. **Battese G, Coelli T.** Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data. *Journal of Econometrics* 1988;**38**:387-99.

33. **Battese G, Coelli T.** Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* 1992;**3 1/2**:153-69.

34. **Kirjavainen T, Loikkanen HA.** Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit Analysis. *Economics of Education Review* 1998;**17**(4):377-94.

35. **Chirikos TN, Sear AM.** Measuring Hospital Efficiency: A comparison of Two Approaches. *Health Services Research* 2000;**34**(6):1389-408.

36. **Grosskopt S, Valdmanis V.** Measuring Hospital Performance: A Non-parametric Approach. *Journal of Health Economics* 1987;**6**(2):89-107.

37. **Mirmirani S, Li H-C.** Health Care Efficiency Measurement: An Application of Data Envelopment Analysis. *Rivista Internazionale di Scienze Economiche e Commerciali* 1995;**42**(3):217-29.

38. **Gupta S, Honjo K, Verhoeven M.** The Efficiency of Government Expenditure: Experiences from Africa. Washington International Monetary Fund, 1997 (IMP Working Paper No.WP/97/15.)

39. **Christensen LR, Jorgenson DW, Lau LJ.** Transcendental logarithmic production frontiers. *Review of economics and statistics* 1973;**55**(1):28-45.

40. **Christensen LR, Jorgenson DW, Lau LJ.** Transcendental logarithmic utility functions. *American economic review* 1975;**65**(3):367-83.

41. **Kmenta J.** Elements of Econometrics. New York. Macmillan Publishing, 1986

42. **Murray CJL, Salomon J, Mathers C.** A critical examination of summary measures of population health. Geneva, Switzerland. World Health

Organization, 1999 (Global Programme on Evidence for Health Policy Discussion Paper No.2.)

43. **Fakin B, de Crombrugghe A.** Fiscal Adjustments in Transition Economies - Transfers and the Efficiency of Public Spending: A Comparison with OECD Countries. Washington. World Bank, 1997 (World Bank Policy Research Working Paper No.1803.)

44. **Pouillier JP, Hernández P.** National Health accounts for 191 countries in 1997. Geneva, Switzerland. World Health Organization, 2000 (Global Programme on Evidence for Health Policy Discussion Paper No.27.)

45. **Preston SH.** The changing relation between mortality and level of economic development, Population studies. 1975 p.231 No.29.)

46. **Preston SH.** Mortality and development revisited. Population bulletin of the United Nations, 1985 p.34

47. **Bicego GT, Boerma JT.** Maternal education and child survival: a comparative study of survey data from 17 countries. *Social Science & Medicine* 1993;**36**(9):1207-27.

48. **Caldwell JC.** Education as a Factor in Mortality Decline An Examination of Nigerian Data. *Population Studies* 1979;**33**(3):395-413.

49. **Caldwell JC.** Maternal education as a factor in child mortality. *World Health Forum.* 1981;**2**(1):75-8.

50. **Caldwell JC, McDonald PF.** Influence of maternal education on infant and child mortality: levels and causes. *Health Policy and Education* 1982;**2**:251-67.

51. **Caldwell JC, Caldwell P.** Education and literacy as factors in health. *In: Good.health at.low cost., edited.by Scott.B.Halstead., Julia.A.Walsh., and Kenneth.S.Warren.New York, New York, The Rockefeller.Foundation,* 1985;181-5.

52. **Caldwell JC.** Routes to low mortality in poor countries. *Population and Development Review* 1986;**12**(2):171-220.

53. **Caldwell JC.** Cultural and social factors influencing mortality levels in developing countries. *Annals of the American Academy of Political and Social Science* 1990;**510**:44-59.

54. **Caldwell JC.** Health transition: the cultural, social and behavioural determinants of health in the Third World. *Social Science & Medicine* 1993;**36**(2):125-35.

55. **Caldwell JC.** How is greater maternal education translated into lower child mortality? *Health Transition Review* 1994;**4**(2):224-9.

56. **Farah AA, Preston SH.** Child Mortality Differentials in Sudan. *Population & Development Review* 1982;**8**(2)

57. **Hobcraft JN, McDonald JW, Rutstein SO.** Demographic Determinants of Infant and Early Child Mortality: A Comparative Analysis. *Population Studies* 1985;**. 39**(3)

58. **Pincus T, Callahan LF.** Associations of low formal education level and poor health status: behavioral, in addition to demographic and medical, explanations?. *Journal of Clinical Epidemiology* 1994;**47**(4):355-61.

59. **Ross CE, Mirowsky J.** Refining the association between education and health: the effects of quantity, credential, and selectivity. *Demography* 1999;**36**(4):445-60.

60. **Yen IH, Moss N.** Unbundling education: a critical discussion of what education confers and how it lowers risk for disease and death. *Annals of the New York Academy of Sciences* 1999;**896**:350-1.

61. **Zimmer Z, Liu X, Hermalin A, Chuang YL.** Educational attainment and transitions in functional status among older Taiwanese. *Demography* 1998;**35**(3):361-75.

62. **Salomon J, Murray CJL.** Methods for life expectancy and disability-adjusted life expectancy uncertainty analysis. Geneva, Switzerland. World Health Organization, 2000 (Global Programme on Evidence for Health Policy Discussion Paper No.10.)

63. **Härdle W.** Applied nonparametric regression. Cambridge. Cambridge University Press, 1990

64. **Kennedy P.** A Guide to Econometrics, 4th edition. Cambridge. The MIT Press, 1998

65. **Greene W.H.** Econometric Analysis, 3rd edition. Upper Saddle River, NJ. Prentice Hall, 1997

66. **Ahmad O, Lopez AD, Inoue M.** The decline of child mortality: a reappraisal. [In Press] Bulletin of the World Health Organization 2000;

67. **WHO.** The World Health Report 1999: Making a Difference. Geneva, Switzerland. World Health Organization, 1999

68. **Babazono A, Hillman AL.** A comparison of international health outcomes and health care spending. *Int J Technol.Assess.Health Care* 1994;**10**(3):376-81.

69. **Filmer D, Pritchett L.** The impact of public spending on health: does money matter? *Soc Sci Med* 1999;**49**(10):1309-23.

70. **Coelli T.** A guide to DEAP version 2.1: a data envelopment analysis (computer) program. New South Wales, Australia. Armidale, 1996 (CEPA Working Paper 96/08

71. **Jondrow J, Lovell CAK, Materov IS, Schmidt P.** On the Estimation of Technical Inefficiency in the Stochastic Production Function Model. *Journal of Econometrics* 1982;**19**(2/3):233-8.

72. **Wang J, Jamison DT, Bos E, et al.** *Measuring Country Performance on Health: Selected Indicators for 115 Countries.* Washington: The World Bank; 1999.

73. **World Bank.** World development report 1993: Investing in health. New York. Oxford University Press, 1993

Appendix: Efficiency or Performance Index and Uncertainty Intervals, 191 countries, 1993-97

Rank	Uncertainty Interval (80%)			Country	Performance Index	Uncertainty interval (80%)	
1	1	-	5	Oman	0.992	0.975	- 1.000
2	1	-	4	Malta	0.989	0.968	- 1.000
3	2	-	7	Italy	0.976	0.957	- 0.994
4	2	-	7	France	0.974	0.953	- 0.994
5	2	-	7	San Marino	0.971	0.949	- 0.988
6	3	-	8	Spain	0.968	0.948	- 0.989
7	4	-	9	Andorra	0.964	0.942	- 0.980
8	3	-	12	Jamaica	0.956	0.928	- 0.986
9	7	-	11	Japan	0.945	0.926	- 0.963
10	8	-	15	Saudi Arabia	0.936	0.915	- 0.959
11	9	-	13	Greece	0.936	0.920	- 0.951
12	9	-	16	Monaco	0.930	0.908	- 0.948
13	10	-	15	Portugal	0.929	0.911	- 0.945
14	10	-	15	Singapore	0.929	0.909	- 0.942
15	13	-	17	Austria	0.914	0.896	- 0.931
16	13	-	23	United Arab Emirates	0.907	0.883	- 0.932
17	14	-	22	Morocco	0.906	0.886	- 0.925
18	16	-	23	Norway	0.897	0.878	- 0.914
19	17	-	24	Netherlands	0.893	0.875	- 0.911
20	15	-	31	Solomon Islands	0.892	0.863	- 0.920
21	18	-	26	Sweden	0.890	0.870	- 0.907
22	19	-	28	Cyprus	0.885	0.865	- 0.898
23	19	-	30	Chile	0.884	0.864	- 0.903
24	21	-	28	United Kingdom	0.883	0.866	- 0.900
25	18	-	32	Costa Rica	0.882	0.859	- 0.898
26	21	-	31	Switzerland	0.879	0.860	- 0.891
27	21	-	31	Iceland	0.879	0.861	- 0.897
28	23	-	30	Belgium	0.878	0.860	- 0.894
29	23	-	33	Venezuela, Bolivarian Republic of	0.873	0.853	- 0.891
30	23	-	37	Bahrain	0.867	0.843	- 0.890
31	28	-	35	Luxembourg	0.864	0.847	- 0.881
32	29	-	38	Ireland	0.859	0.840	- 0.870
33	27	-	40	Turkey	0.858	0.835	- 0.878
34	25	-	48	Belize	0.853	0.821	- 0.884
35	33	-	40	Canada	0.849	0.832	- 0.864
36	32	-	42	Cuba	0.849	0.830	- 0.866
37	30	-	49	El Salvador	0.846	0.817	- 0.873
38	28	-	52	Saint Vincent and the Grenadines	0.845	0.812	- 0.876
39	35	-	43	Australia	0.844	0.826	- 0.861
40	36	-	44	Israel	0.841	0.825	- 0.858
41	39	-	47	Germany	0.836	0.819	- 0.852
42	33	-	54	Dominican Republic	0.834	0.806	- 0.863
43	37	-	53	Egypt	0.829	0.811	- 0.849
44	41	-	50	Finland	0.829	0.812	- 0.844
45	38	-	55	Algeria	0.829	0.808	- 0.850
46	41	-	55	Tunisia	0.824	0.803	- 0.844
47	38	-	58	Yugoslavia	0.824	0.798	- 0.848
48	40	-	61	Honduras	0.820	0.793	- 0.844
49	37	-	63	Grenada	0.819	0.789	- 0.850
50	42	-	59	Uruguay	0.819	0.794	- 0.842
51	41	-	64	Colombia	0.814	0.787	- 0.843
52	42	-	65	Paraguay	0.813	0.785	- 0.842
53	43	-	64	Qatar	0.813	0.786	- 0.839

54	43	- 69	Saint Lucia	0.809	0.781	- 0.837
55	41	- 70	Cape Verde	0.808	0.776	- 0.842
56	47	- 64	Armenia	0.806	0.785	- 0.823
57	51	- 61	Croatia	0.805	0.789	- 0.821
58	48	- 65	Iran, Islamic Republic of	0.805	0.783	- 0.827
59	45	- 73	Dominica	0.804	0.774	- 0.833
60	49	- 67	Azerbaijan	0.803	0.781	- 0.820
61	52	- 65	China	0.800	0.782	- 0.813
62	55	- 66	Slovenia	0.797	0.781	- 0.813
63	56	- 73	Mexico	0.789	0.771	- 0.808
64	55	- 76	Albania	0.789	0.766	- 0.808
65	61	- 72	Denmark	0.785	0.769	- 0.801
66	57	- 80	Sri Lanka	0.783	0.761	- 0.807
67	57	- 80	Panama	0.783	0.759	- 0.807
68	56	- 83	Kuwait	0.782	0.753	- 0.808
69	61	- 78	The former Yugoslav Republic of Macedonia	0.781	0.761	- 0.796
70	59	- 83	Bosnia and Herzegovina	0.780	0.754	- 0.803
71	65	- 76	Argentina	0.779	0.762	- 0.794
72	67	- 78	United States of America	0.774	0.758	- 0.789
73	61	- 86	Bhutan	0.773	0.748	- 0.797
74	63	- 84	Nicaragua	0.772	0.750	- 0.793
75	65	- 84	Iraq	0.770	0.752	- 0.791
76	67	- 85	Brunei Darussalam	0.768	0.749	- 0.787
77	61	- 88	Suriname	0.768	0.740	- 0.798
78	66	- 88	Brazil	0.767	0.745	- 0.787
79	70	- 84	Trinidad and Tobago	0.767	0.750	- 0.780
80	72	- 83	New Zealand	0.766	0.750	- 0.780
81	73	- 83	Czech Republic	0.765	0.749	- 0.779
82	66	- 91	Yemen	0.761	0.733	- 0.789
83	72	- 88	Seychelles	0.759	0.739	- 0.778
84	73	- 91	Georgia	0.758	0.736	- 0.776
85	73	- 89	Pakistan	0.757	0.738	- 0.777
86	75	- 92	Malaysia	0.751	0.731	- 0.771
87	77	- 92	Barbados	0.749	0.730	- 0.770
88	85	- 92	Slovakia	0.742	0.729	- 0.757
89	84	- 94	Poland	0.742	0.723	- 0.758
90	79	- 98	Indonesia	0.741	0.715	- 0.766
91	85	- 99	Syrian Arab Republic	0.733	0.712	- 0.755
92	89	- 96	Bulgaria	0.733	0.717	- 0.747
93	89	- 103	Lithuania	0.724	0.705	- 0.742
94	89	- 104	Libyan Arab Jamahiriya	0.723	0.699	- 0.746
95	89	- 105	Cook Islands	0.722	0.696	- 0.746
96	89	- 104	Ecuador	0.721	0.700	- 0.742
97	91	- 105	Lebanon	0.719	0.697	- 0.740
98	93	- 107	Nepal	0.714	0.691	- 0.736
99	93	- 107	Guatemala	0.714	0.691	- 0.735
100	94	- 107	Jordan	0.711	0.689	- 0.732
101	97	- 104	Ukraine	0.711	0.695	- 0.726
102	93	- 111	Thailand	0.710	0.682	- 0.736
103	93	- 109	Bangladesh	0.709	0.684	- 0.735
104	92	- 115	Guyana	0.704	0.672	- 0.738
105	101	- 111	Hungary	0.698	0.682	- 0.714
106	102	- 111	Republic of Moldova	0.696	0.680	- 0.710
107	100	- 113	Republic of Korea	0.694	0.674	- 0.711
108	93	- 121	Niue	0.693	0.650	- 0.731
109	103	- 116	Gambia	0.687	0.671	- 0.704
110	100	- 121	Micronesia, Federated States of	0.684	0.656	- 0.717
111	107	- 117	Romania	0.682	0.668	- 0.696
112	107	- 119	Uzbekistan	0.681	0.662	- 0.700
113	105	- 120	Mauritius	0.679	0.657	- 0.702
114	105	- 121	Tonga	0.677	0.651	- 0.704
115	107	- 119	Estonia	0.677	0.657	- 0.694
116	109	- 119	Belarus	0.676	0.657	- 0.692
117	109	- 121	Sao Tome and Principe	0.671	0.651	- 0.691
118	112	- 120	India	0.670	0.654	- 0.683
119	111	- 123	Peru	0.665	0.643	- 0.686
120	108	- 123	Vanuatu	0.665	0.639	- 0.689
121	115	- 125	Latvia	0.655	0.631	- 0.677
122	114	- 127	Saint Kitts and Nevis	0.650	0.621	- 0.679

123	115	- 131	Antigua and Barbuda	0.641	0.606	- 0.678
124	120	- 133	Fiji	0.632	0.600	- 0.662
125	121	- 131	Palau	0.632	0.606	- 0.656
126	122	- 131	Philippines	0.630	0.608	- 0.653
127	124	- 131	Russian Federation	0.623	0.606	- 0.638
128	123	- 134	Tuvalu	0.618	0.594	- 0.644
129	124	- 137	Myanmar	0.612	0.584	- 0.641
130	125	- 136	Viet Nam	0.611	0.587	- 0.634
131	127	- 139	Samoa	0.602	0.579	- 0.626
132	128	- 138	Senegal	0.601	0.584	- 0.620
133	129	- 139	Côte d'Ivoire	0.598	0.580	- 0.617
134	128	- 140	Kyrgyzstan	0.598	0.575	- 0.620
135	129	- 138	Kazakhstan	0.598	0.581	- 0.615
136	129	- 139	Benin	0.596	0.576	- 0.616
137	127	- 142	Bahamas	0.593	0.564	- 0.624
138	132	- 144	Mongolia	0.581	0.555	- 0.607
139	134	- 143	Haiti	0.580	0.561	- 0.599
140	131	- 144	Marshall Islands	0.579	0.549	- 0.609
141	137	- 145	Comoros	0.570	0.550	- 0.590
142	137	- 145	Bolivia	0.567	0.544	- 0.590
143	139	- 146	Gabon	0.559	0.538	- 0.579
144	138	- 148	Kiribati	0.554	0.525	- 0.581
145	140	- 148	Tajikistan	0.551	0.523	- 0.580
146	141	- 149	Papua New Guinea	0.546	0.520	- 0.572
147	144	- 154	Maldives	0.524	0.496	- 0.555
148	146	- 153	Eritrea	0.521	0.504	- 0.538
149	146	- 154	Sudan	0.519	0.496	- 0.543
150	146	- 155	Afghanistan	0.517	0.488	- 0.547
151	147	- 153	Mauritania	0.517	0.501	- 0.533
152	145	- 158	Turkmenistan	0.513	0.479	- 0.546
153	147	- 156	Democratic People's Republic of Korea	0.510	0.485	- 0.536
154	148	- 157	Somalia	0.506	0.480	- 0.530
155	152	- 160	Lao People's Democratic Republic	0.489	0.466	- 0.510
156	154	- 162	Guinea-Bissau	0.481	0.462	- 0.499
157	153	- 162	Cambodia	0.481	0.460	- 0.501
158	153	- 162	Ghana	0.479	0.457	- 0.500
159	155	- 164	Togo	0.472	0.452	- 0.492
160	157	- 164	Guinea	0.469	0.455	- 0.483
161	156	- 165	Chad	0.465	0.444	- 0.487
162	157	- 166	Burkina Faso	0.463	0.441	- 0.483
163	158	- 167	Djibouti	0.457	0.434	- 0.479
164	160	- 166	Central African Republic	0.454	0.436	- 0.470
165	159	- 167	Angola	0.453	0.433	- 0.473
166	162	- 168	Nauru	0.444	0.424	- 0.464
167	164	- 170	Congo	0.433	0.411	- 0.454
168	164	- 172	Mozambique	0.424	0.399	- 0.450
169	167	- 171	Ethiopia	0.418	0.400	- 0.435
170	168	- 172	Mali	0.410	0.393	- 0.426
171	168	- 174	Burundi	0.403	0.374	- 0.435
172	169	- 174	Cameroon	0.399	0.375	- 0.421
173	170	- 174	Madagascar	0.394	0.378	- 0.410
174	172	- 175	Equatorial Guinea	0.377	0.355	- 0.400
175	174	- 176	Nigeria	0.353	0.331	- 0.375
176	175	- 178	Liberia	0.337	0.318	- 0.355
177	176	- 178	Niger	0.323	0.306	- 0.340
178	176	- 178	Kenya	0.320	0.298	- 0.343
179	179	- 180	Uganda	0.280	0.264	- 0.295
180	179	- 180	United Republic of Tanzania	0.279	0.260	- 0.298
181	181	- 185	Rwanda	0.240	0.214	- 0.265
182	181	- 185	South Africa	0.232	0.209	- 0.251
183	181	- 185	Sierra Leone	0.230	0.213	- 0.247
184	181	- 186	Swaziland	0.229	0.205	- 0.255
185	182	- 187	Democratic Republic of the Congo	0.217	0.198	- 0.235
186	183	- 188	Lesotho	0.211	0.187	- 0.236
187	186	- 188	Malawi	0.196	0.181	- 0.211
188	187	- 189	Botswana	0.183	0.172	- 0.194
189	185	- 189	Namibia	0.183	0.152	- 0.214
190	190	- 190	Zambia	0.112	0.095	- 0.129
191	191	- 191	Zimbabwe	0.080	0.057	- 0.103