Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems

William Greene[*]

Department of Economics, Stern School of Business,
New York University,

April 20, 2003

Abstract

The most commonly used approaches to parametric (stochastic frontier) analysis of efficiency in panel data, notably the fixed and random effects models, fail to distinguish between cross individual heterogeneity and inefficiency. This blending of effects is particularly problematic in the World Health Organization's (WHO) panel data set on health care delivery, which is a 191 country, five year panel. The wide variation in cultural and economic characteristics of the worldwide sample of countries produces a large amount of unmeasured heterogeneity in the data. Familiar approaches to inefficiency estimation mistakenly measure that heterogeneity as inefficiency. This study will examine a large number of recently developed alternative approaches to stochastic frontier analysis with panel data, and apply some of them to the WHO data. A more general, flexible model and several measured indicators of cross country heterogeneity are added to the analysis done by previous researchers. Results suggest that in these data, there is considerable evidence of heterogeneity that in other studies using the same data, has masqueraded as inefficiency. Our results differ substantially from those obtained by several earlier researchers.

Keywords: Panel data, fixed effects, random effects, random parameters, technical efficiency, stochastic frontier, heterogeneity, health care.

JEL classification: C1, C4

1. Introduction

The *World Health Report 2000* (*WHR*) [WHO (2000)] is a large, worldwide assessment of the effectiveness of health care delivery. Among other analyses, the study presents a rankings based comparison of the productive efficiency of the health care systems of 191 countries.[1] Predictably, the attention focused on these rankings has been considerably out of proportion to the space this section occupies in the larger report itself.

The rankings were produced using a form of the "fixed effects," stochastic frontier methodology proposed by Schmidt and Sickles (1984) and Cornwell, Sickles and Schmidt (1990) [see Evans et al. (2000a,b)] (*ETML*). The data analyzed in this econometric study were a five year (1993-1997) panel. This section of the *WHR* has been heavily criticized for numerous reasons related to the overall objectives, the quality and validity of the effectiveness measures, the input data used, and the appropriateness of the methodology. [See, e.g., Gravelle et al. (2002a,b) (*GJJS*), Williams (2001) and Hollingsworth and Wildman (2002) (*HW*).] On January 8, 2001, coincident with the 2001 meeting of the Allied Social Science Association in New Orleans, the authors of the frontier study convened a panel of researchers specifically to discuss the econometric methodology.[2] The focus of the meeting was the use of panel data, such as those in the WHR study, for measurement of efficiency in health care delivery. This paper reports subsequent research undertaken to study some of the issues raised at that meeting.

One criticism of the fixed effects methodology used (and several other related approaches) is that the model fails to distinguish between cross country heterogeneity unrelated to inefficiency and the inefficiency itself. This ambiguity is likely to be especially problematic in these data, as they are based on 191 sometimes vastly different countries; France, England and Australia appear in the sample on equal terms with Oman, Sri Lanka, Zimbabwe, the Seychelles, Colombia and Bangladesh. We undertook to examine this issue, to reanalyze the WHO data and the methods used in the study, and to propose alternative stochastic frontier based methods with greater flexibility that will allow the analyst to segregate individual, unmeasured heterogeneity and technical or cost inefficiency.

The paper is organized as follows: Section 2 reviews the WHO methodology and takes a cursory look at their results. The stochastic frontier model and strategies for modeling panel data are reviewed in Section 3.[3] The WHO data that were used in this study as well are described in Section 4. The several studies of the WHO data [*ETML* (2000a,b), *HW* (2002), *GJJS*. (2002b)] that we examined were based on two output measures, a composite measure of health care delivery (COMP) and disability adjusted life expectancy (DALE) and two inputs, health care expenditure and education levels. In this study, we consider how to use additional covariates in the data set including per

---

[1] The numerical values on which the rankings are based do not have a clear interpretation. The studies focus on the rankings, and contain only minor discussion of the quantitative efficiency measures.

[2] See WHO (2001). The participants were, in addition to Evans, et al., were William Greene of NYU, Subal Kumbhakar, University of Binghamton, Knox Lovell, University of Georgia, Kaliappa Kalirajan, ANU, Marijn Berhoeven, IMF, Paul Wilson, University of Texas, Christopher Tong, Hong Kong Baptist University and Philip Grossman, St. Cloud State University.

[3] We have focused on parametric and semiparametric stochastic frontier models. Nonparametric methods such as data envelopment analysis (DEA) are not considered here. [For commentary, see, e.g., Hollingsworth and Wildman (2002).]

capita income, income distribution, government effectiveness, and the allocation of health care expenditure between the public and private sectors to account for some of the heterogeneity noted earlier.[4] The empirical results are presented in Section 5. We begin with an examination of the production function used, and propose some results that are in broad agreement with others already in the literature regarding the impact of income and the distribution of income on health care outcomes. The second set of results will compare fixed and random effects estimates of technical inefficiency. We find that concerns of *ETML* (2000a,b) notwithstanding, for these data, a form of the random effects model appears to be a satisfactory framework for analyzing the WHO data. We then incorporate the country specific heterogeneity in the estimated distribution of technical inefficiency, then in addition, in the production function itself. Our alternative results are presented here and in the Appendices. This section also proposes some recently developed panel data techniques which, though promising for research in this area, for practical reasons, do not appear to be appropriate for the WHO data. Some conclusions are drawn in Section 6.

---

[4] Per capita income was briefly considered in each of these, but did not play a central role in any of the empirical analyses.

2. The WHO Studies of Health Care Attainment

Health policy makers are concerned with overall health system performance. Reforms are directed at all functions in the health system including financing, provision of services, management, and so on. Evaluation of the effectiveness of policies and reforms faces two large obstacles, quantifying goals and objectives so that outcomes can be measured and enumerating inputs in a way that resources can be directed toward them so as to achieve those objectives. The effectiveness study in the *WHR* is an attempt to measure health care effectiveness in a production function framework. Superficially, the setting lends itself well to that framework, with one substantive variation. In the textbook case of a production technology, zero inputs implies zero output. But, no matter how the outcome is defined, output of the "health care system" of a country would not be zero even if no resources were directed toward the health sector – {"i.e., the entire population would not be dead" [Evans et al. (2000a, page 2.)]}. This view is summarized in Figure 1, which is taken from *ETML*. (2000a, page. 2).
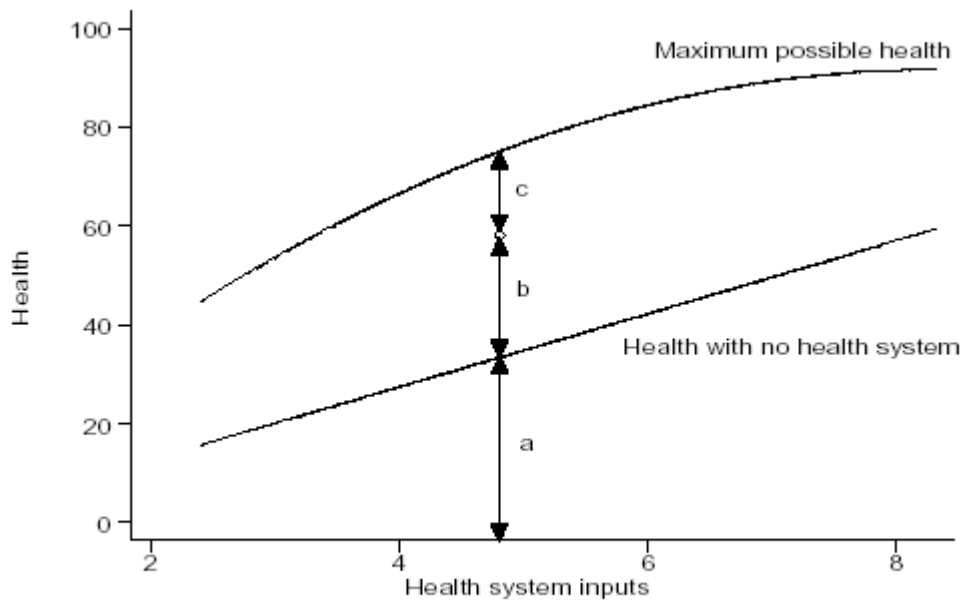
Figure 1: Health System Performance

The authors rejected several common methodologies as frameworks for modeling health system performance. Nonparametric data envelopment analysis and "free disposal hull" methods were criticized for their inability to accommodate random variation in the data. Corrected least squares [Greene (1997)] does not account for the fact that the production function is an upper bound. Finally, the stochastic frontier model [Aigner, Lovell and Schmidt (1977)] was deemed to be unduly stringent in its imposition of a specific distributional assumption on the inefficiency component of the model. ETML's (2000a,b) preferred methodology was a "panel data," production function estimator based on the framework proposed by Schmidt and Sickles (1984) and Cornwell et al. (1990).

The central feature of the estimator is a fixed effects linear regression model. It is argued that this approach brings gains in statistical efficiency while obviating assumptions about the distribution of technical inefficiency.[5]

The production function is denoted

$$(1) \qquad y_{it} \qquad = \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_i$$

where y_{it} is the (log of the) output of the system, $\mathbf{x}_{it}$ is (logs of) the set of inputs, v_{it} is the random component representing stochastic elements as well as any country (and time) specific heterogeneity, u_i is the inefficiency in the system, and i and t denote country and year, respectively. Consistent with the objectives of the study (and Figure 1), it is assumed that $u_i > 0$. The equation is rewritten

$$(2) \qquad y_{it} \qquad = (\alpha - u_i) + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it}$$

$$= \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it}.$$

Assuming that v_{it} has the familiar stochastic properties of a regression model and is uncorrelated with other components of the model, the parameters can be estimated by least squares, using the "within," or dummy variable estimator. The country specific constants embody the technical inefficiency. The inefficiencies are estimated in turn by shifting the function upward so that each constant term is measured as a deviation from the benchmark level;

$$(3) \qquad \hat{u}_i \qquad = \max_i (\hat{\alpha}_i) - \hat{\alpha}_i \geq 0.$$

(Note that by this construction, one country is measured as 100% efficient.) Technical efficiency is now measured by

$$(4) \qquad TE_i \qquad = \frac{E[y_{it} \mid \mathbf{x}_{it}, u_i]}{E[y_{it} \mid \mathbf{x}_{it}, u_i = 0]}$$

Overall efficiency is constructed by normalizing this measure to a constructed minimum level of output that would (more or less) correspond to a system with zero inputs;

$$(5) \qquad E_i \qquad = \frac{E[y_{it} \mid \mathbf{x}_{it}, u_i] - M_{it}}{E[y_{it} \mid \mathbf{x}_{it}, u_i = 0] - M_{it}}.^{[6]}$$

---

[5] The authors also rejected the random effects model based on the results of a Hausman specification test. It should be noted that the random effects *regression* model would have been counterproductive as, unlike the fixed effects model, it does not provide a readily computed estimate of the firm specific term that is the central focus of the study. We will return to this specification issue later in the text.

[6] The definition of M_{it} used differs in the two studies. In both cases, it was based on the disability adjusted life expectancy (DALE) measured in a sample of 25 countries around 1908, at a time and in places where it was felt that the influence of health service inputs on health care would have been minimal. See ETML (2000a, p. 14 and 2000b, pp. 6 and 8) for definitions.

In Figure 1, TE_i would give for each country, the ratio $(a+b)/(a+b+c)$ while E_i would measure $b/(b+c)$.

Empirical analysis in the two studies used as inputs per capita public and private health care expenditure (*EXP*) and average years of education of the population (*EDUC*). Two measures of health care attainment were analyzed, disability adjusted life expectancy (*DALE*) and a composite measure of health care delivery (*COMP*). The production function employed in both cases was

$$(6) \qquad y_{it} \quad = \alpha_i + \beta_1 logEXP_{it} + \beta_2 logEDUC_{it} + \beta_3 log^2EDUC_{it} + v_{it}$$

where y_{it} is the log of DALE in *ETML* (2000a) and the log of COMP in (2000b). (Greater detail on these is given below.) Figure 2 [Figure 6 from *ETML* (2000a)] illustrates the overall results for the first analysis.
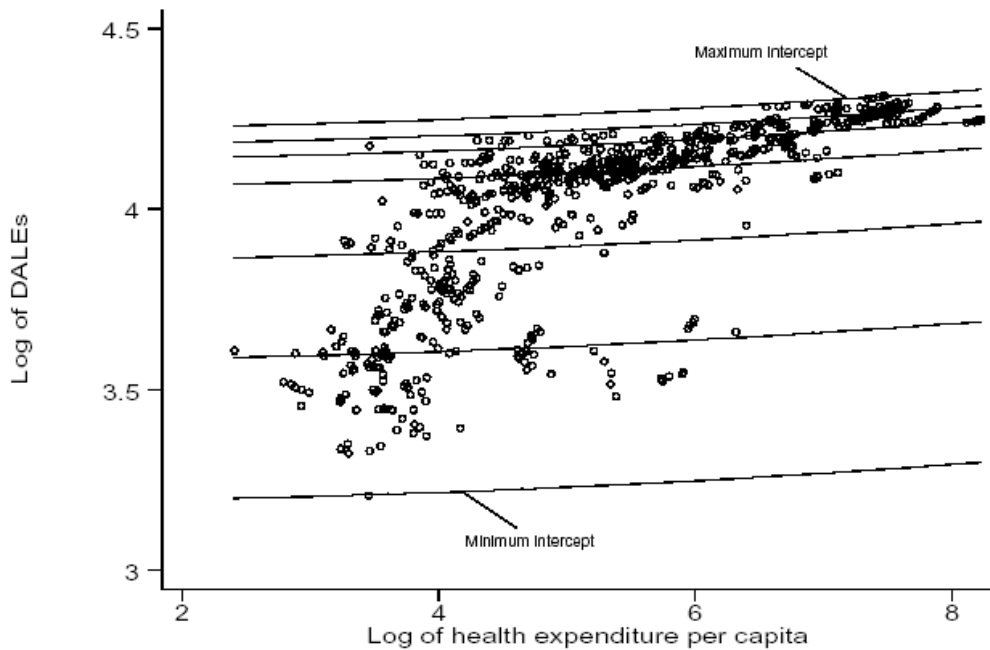
Figure 2: Estimated Efficiencies Based on DALE for 191 Countries

Table 1 shows some of the estimated results from the two studies. The specific numeric values for DALE, in years, have a ready interpretation, but those for COMP have no clear numeraire. In principle, each TE_i gives $b/(b+c)$ in Figure 1, or the percentage of maximal output above the minimum that is attained in the country. Based on the numeric values of the attainment measures, these would give the proportional potential for improvement. As noted, the focus of the studies was the rankings, not the numeric values. (Among the criticisms of the study is the lack of a clear understanding of how one might make use of these rankings. To some extent – this is suggested in the text – these studies were viewed as innovative first steps, not definitive answers to any specific question.) We note, finally, even the figures given suggest the potential for improvement in essentially

qualitative terms, especially for the COMP measure. The authors did not adopt a cost function methodology. In a cost frontier, the measured inefficiency would have implied, given observable expenditure data, a quantifiable and probably unbearable burden for large numbers of already poor countries. Of course, this raises the question noted earlier. How the results of this study or others like it can be translated into policy is an important and yet unanswered question.

Table 1: Selected WHO Estimates of Overall Efficiency

Rank	DALE[a]		COMP[b]	
	Country	E	Country	E
1	Oman	0.992	France	0.994
2	Malta	0.989	Italy	0.991
3	Italy	0.976	San Marino	0.988
4	France	0.974	Andorra	0.982
5	San Marino	0.971	Malta	0.978
25	Costa Rica	0.882	Germany	0.902
50	Uruguay	0.819	Poland	0.793
100	Jordan	0.711	St. Kitts and Nevis	0.643
150	Afghanistan	0.517	Nepal	0.457
187	Malawi	0.196	Nigeria	0.176
188	Botswana	0.183	Dem. Rep. of Congo	0.171
189	Namibia	0.183	Central African Rep.	0.156
190	Zambia	0.112	Myanmar	0.138
191	Zimbabwe	0.080	Sierra Leone	0.000
	United Kingdom (24)	0.883	United Kingdom (18)	0.925
	United States (72)	0.774	United States (37)	0.838

[a]From Evans et al. (2000a), Appendix

[b]From Evans (2000b), Annex, Table 1.

These data have been reanalyzed by several authors. A variety of strident criticisms were raised in Williams (2001), who questioned the methodology and objectives of the entire study as well as the quality of the data set and the appropriateness of the outcome measures. Several econometric studies have placed the first (DALE) study under narrower scrutiny. The second (COMP) has until now not been similarly examined.

Gravelle et al. (*GJJS*) (2000a,b) observed that in the sample of 191 countries, actually 51 of the data sets are observed for only one year (1997). As such, in computation of the fixed effects estimator, these countries fall out of the least squares sums. This implies that the estimates are actually based on these 140 countries, not 191, and the results for the 51 countries in question are computed from a model that is not based on them. This, in itself, does not necessarily taint the results if these observations are not systematically different from the others, but *GJJS* argue that they are, indeed different – they have lower health expenditures. Since the model is conditioned on health expenditure, this still does not cast doubt on the results; it does so only if the relationship between expenditure and DALE is systematically different, which is possible, but remains to be shown. Since the models estimated here involve all the data, we leave this question aside for future research. More troublesome is *GJJS*'s observation that 99.8%

of the variation in the log of the DALE variable is between, rather than within the groups (countries). The counterparts for the logs of expenditure (EXP) and education (EDUC) are 98.9% and 99.8%, respectively. Thus, there is very little actual "panel data" variation in these data – it is essentially a cross section. *GJJS* proceeded to fit several models based on the "between" estimators (group means) and computed alternative adjusted measures of efficiency. They found varying degrees of correlation between their rankings and those in *ETML* (2000a), ranging from 0.97 down to about 0.39. [See their Table 2.] This suggests, as we find below, that the specification can make a considerable amount of difference in the results. (*GJJS* also argued for inclusion of time effects and other expenditure (income minus health expense) in the model. We will return to the model specification in Section 5.)

Hollingsworth and Wildman (2002) (*HW*) also revisited the DALE results. They redid the computations using the nonparametric, data envelopment methods rejected by *ETML*. They also fit fixed effects models that extended the WHO formulation by allowing time variation in the constants of the form

(7) $\quad \alpha_{it} = \theta_{1i} + \theta_{2i}t + \theta_{3i}t^2.$

[See Cornwell et al. (1990).] The computation was done by regressing least squares residuals for each country on a constant, time, and its square. (This restricts attention to the balanced panel of the 140 countries observed in all five years.)[7] Finally, *HW* fit separate stochastic frontier models for each year, and computed technical inefficiency estimates for each country for each year. In all cases, the results obtained were fairly similar to the WHO results, though some moderately large differences did emerge. *HW* directed considerable attention to the differences between OECD and non-OECD countries in their results. Perhaps not surprisingly, their results suggest that these differences are large enough that one might want to analyze these groups separately. (Our result also suggest marked differences between the OECD and non-OECD countries.)

---

[7] Hollingsworth and Wildman (2002) argue that these estimates are consistent for increasing *T*, which would generally be the case, but in this setting, *T* is only 5, and is assumed to be fixed. The sampling variances of these country specific estimators are likely to be quite high. As they are crucial determinants of the quantity of greatest interest here, this is an important consideration.

3. The Stochastic Frontier Model

The authors of the WHO analyses questioned the distributional assumptions in the stochastic frontier formulation. However, both *GJJS* and *HW* found considerable similarity in the results across a number of different formulations. This suggests that specific assumptions about the distribution of efficiencies may be less restrictive than these views suggest. In addition, we submit that the formulations examined before were narrower than they could have been, and the stochastic frontier model allows the incorporation of cross country heterogeneity in several ways that are likely to bring large benefits in analyzing data as disparate as these. This section will lay out the stochastic frontier model in some detail. The basic formulation of the model is well documented in many sources – we suggest Kumbhakar and Lovell (2000) – so our description will be brief. We will then detail several panel data formulations, some of which have appeared in received work for some time, but several others that are quite new.

3.1. Cross Section Variants

The essential form of the stochastic production frontier model [see Aigner et al. (1977)] is

$$y_i = \alpha + \mathbf{x}_i'\boldsymbol{\beta} + v_i - u_i$$

$$v_i \sim N[0,\sigma_v^2]$$

(8)

$$u_i = |U_i|,$$

$$U_i \sim N[0,\sigma_u^2]$$

This is the canonical 'half normal' model. A central parameter in the model is the asymmetry parameter, $\lambda = \sigma_u/\sigma_v$; the larger is λ, the greater is the inefficiency component in the data. Parameters are estimated by maximum likelihood, rather than least squares. As in the earlier applications discussed here, estimation of u_i is the central focus of the analysis. With the model estimated in logarithms, u_i would correspond (to a small degree of approximation) to $1 - TE_i$ given earlier. Individual specific *efficiency* is typically estimated with $\exp(-\hat{u}_i)$. Alternatively, $\hat{u}_i$, itself, provides an estimate of proportional *inefficiency*. With parameter estimates in hand, one can only obtain a direct estimate of $\varepsilon_i = v_i - u_i$. This is translated into an estimate of u_i using Jondrow et al.'s (*JLMS*) (1982) formula,

$$(9) \quad E[u_i|\varepsilon_i] = \frac{\sigma\lambda}{1+\lambda^2}\left[z_i + \frac{\phi(z_i)}{\Phi(z_i)}\right], \ z_i = -\varepsilon_i\lambda/\sigma$$

where $\sigma = (\sigma_u^2+\sigma_v^2)^{1/2}$ and $\phi(z)$ and $\Phi(z)$ are the density and CDF of the standard normal distribution, respectively.

The narrow assumption of half normality is a viewed as significant drawback in this model.[8] *HW* and others [see Stevenson (1980)] have extended it to a *truncated* normal model by allowing the mean of U_i to be nonzero. This is a useful extension in itself (though *HW* find it makes little difference in their results). However, this extension stops short of a crucial step. The major shortcoming here is that the strict assumption suppresses individual heterogeneity in inefficiency that is allowed, for example, by the fixed effects formulation. But, as noted below, we have several indicators of this heterogeneity, such as income distribution, per capita income, OECD membership, the public share of health care expenditures, etc., and these can be incorporated into the distribution of u_i in ways that the other methods already discussed cannot accommodate. Letting $\mathbf{h}_i$ denote a set of variables that measure the group heterogeneity, we write

(10) $E[U_i] = \mu_i = \mathbf{h}_i'\delta.$

The Jondrow et al. result is now changed by replacing z_i with $z_i^* = z_i - \mu_i/(\sigma\lambda)$. With data on indicators of heterogeneity in hand, this represents a significant extension of the model.[9]

Analysts frequently subject inefficiency estimates to a second step analysis. Thus, one might explore whether per capita income or the Gini measure of income inequality provide any explanatory power in the variation of the inefficiency measures. Arguably, if such covariates do have explanatory power, then they should appear in the model at the first step (to avoid biases due to "left out variables"). [See Wang and Schmidt (2002).] Among the methods considered so far, only the truncated normal stochastic frontier in (8,10) can accommodate this sort of extension. Note, in particular, that if the covariates are time invariant, as is likely, then they cannot appear in the fixed effects model analyzed by *ETML*.

3.2. Panel Data Formulations

The Fixed Effects Model

The Schmidt and Sickles formulation,

$$y_{it} = (\alpha - u_i) + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it}$$

(11) $$= \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it}$$

$$\hat{u}_i = \max_i(\hat{\alpha}_i) - \hat{\alpha}_i \geq 0,$$

---

[8] Other distributional assumptions have been suggested, such as the normal-exponential [ALS (1977)] and the normal-gamma [Greene (1990, 2003a)]. These extensions occasionally bring noticeable changes in the results. But they are tangential for present purposes.

[9] Heteroscedasticity, of the form $\sigma_{ui}^2 = \exp(\mathbf{h}_i'\boldsymbol{\gamma})$ can also be incorporated in the model without great difficulty. Whether this is an equally beneficial extension remains to be established, and is left for further research.

has been used in a number of applications. [See Wang and Schmidt (2002).] There are two important restrictions built into this model. First, any time invariant heterogeneity will be pushed into α_i and ultimately into $\hat{u}_i$. The WHO data span a tremendous range of cultures, economies, and policy settings. This is likely to be a particularly influential aspect of the model for these data. Second, the model (and some others we will consider) assumes that inefficiency is, itself, time invariant. For short time intervals, this may be a reasonable assumption. But, five years may be long enough for this to be questionable. *HW* did find evidence to suggest that this assumption may be inconsistent with the data.

Both of these restrictions can be relaxed by placing country specific constant terms in the stochastic frontier model – we call this a 'true' fixed effects model.

(12) y_{it} $= \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_{it}$

where u_{it} has the stochastic specifications noted earlier for the stochastic frontier model. Superficially, this amounts simply to adding a full set of country dummy variables to the stochastic frontier model. The model is still fit by maximum likelihood, not least squares. Surprisingly, this has hardly been used previously[10] in spite of the fact that most of the received panel data applications involved fairly small panels.[11]

The true fixed effects model places the unmeasured heterogeneity in the production function; with a loglinear model, it produces a neutral shift of the function, specific to each country. One might, instead, have the heterogeneity reside in the inefficiency distribution. This could be accomplished with the formulation

(13) μ_i $= \delta_{0i} + \mathbf{h}_i'\boldsymbol{\delta}$

that is, by placing the country specific dummy variables in the mean of the truncated normal distribution, rather than in the production function. Once again, in a moderate sized sample, this is a minor reformulation of the familiar model. The problems noted in the next paragraph will appear, but how serious these are is an empirical issue, nor a foregone conclusion.

The true fixed effects model has the virtues mentioned above. Weighing against it are, first, the incidental parameters problem and second, the possibility that the model is now overspecified. The incidental parameters problem [Lancaster (2000)] is a persistent bias that arises in nonlinear fixed effects models when the number of periods is small. (Five is small.). It has been widely documented for binary choice models [see, e.g., Hsiao (1996) and Greene (2002b)] but not systematically examined in stochastic frontier models. In Greene (2002a), we found that the biases in coefficient estimates were surprisingly small and did not appear in the patterns predicted by received results for other models, and, moreover, that there appeared to be no biases transmitted to the estimates of technical inefficiency. The second problem is now that the model may be

---

[10] The only received application of a type of true fixed effects model in the frontiers literature is Polachek and Yoon's (1996) study of labor supply.

[11] A major practical obstacle to use of the fixed effects approach in nonlinear models such as this one is the difficulty of computing the possibly hundreds or thousands of dummy variable coefficients. See Greene (2001) for analysis of the solution to this computational problem and Greene (2002b) for an application involving direct computation of models with tens of thousands of dummy variable coefficients.

overspecified. If there is persistent inefficiency, it is now completely absorbed in the country specific constant term which is also capturing any time invariant heterogeneity. Whereas the earlier fixed effects form would tend to overestimate the inefficiency component, it is possible that this form will underestimate it. (This appears to be the case with the WHO data.) Unfortunately, this blending of the two effects is inherent in the modeling problem, and there is no simple solution that will be entirely satisfactory. Ultimately, $\alpha_i + v_{it} - u_i$ contains both country specific heterogeneity and inefficiency, and both may have invariant and time varying elements. There is no perfect way to disentangle them based on observed data.

The Random Effects Model

The random effects model is obtained by assuming that u_i is time invariant and also uncorrelated with the included variables in the model,

(14) $y_{it} \quad = \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_i$

In the linear regression case analyzed by *ETML*, the parameters are estimated by two step generalized least squares. [See Greene (2003, chapter 11).] On the basis of a Hausman specification test, *ETMA* concluded that the random effects model would not be appropriate for their data. (Their test was conducted in a model that did not include any of the observed country specific effects, so the result may have been more convincing than appropriate. We will examine a more extensive specification below.) However, even if not, the regression based random effects model has a significant drawback for present purposes; there is no implied estimator of inefficiency in this model, that is, no estimator of TE_i as in the fixed effects case. So, the model would not have been useful in any event.

Pitt and Lee (1981) showed how the time invariant composed error model could be extended to a panel data version of the stochastic frontier model. The direct extension would be of limited usefulness here, first because of the assumption of uncorrelatedness of u_i and $\mathbf{x}_i$ and, once again, because of the assumption of time invariance of the inefficiency. The first of these can be remedied in the same fashion as suggested earlier. Estimation of the random effects model with heterogeneity in E[U_i], see (10), is straightforward. [See Econometric Software, Inc. (2002).] With this extension, the *JLMS* estimator becomes

(15) $\mathrm{E}[u_i|\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_T,\mathbf{h}_i] \;=\; Z_i + \psi \left[\dfrac{\phi(Z_i/\psi)}{\Phi(Z_i/\psi)} \right]$

where $Z_i \;=\; \gamma\mu_i \;-\; (1-\gamma)\overline{\varepsilon}_i$, $\gamma = 1/(1 + T\lambda^2)$, $\psi^2 = \gamma\sigma_u^2$, and $\overline{\varepsilon}_i = (1/T)\Sigma_i\varepsilon_{it}$.[12]

The time invariance of the inefficiency component of the random effects model is a potential drawback in the random effects model. Battese and Coelli (*BC*) (1988, 1995)

---

[12] A minor adjustment must be made at this point for 'unbalanced' panels, that is data sets in which the number of observations, T, varies across individuals (countries). The adjustment is minor; in (15), ψ becomes ψ_i because it is now a function of the individual specific T_i. It is otherwise unchanged.

have proposed a modification of the model that allows some systematic variation in the model;

$$(16) \quad u_{it} \quad = \eta_t |U_i|$$

where $\eta_t = 1 + \eta_1(t\text{-}T) + \eta_2(t - T)^2$ and $U_i \sim N[\mu,\sigma_u^2]$. Various other forms of the function η_t have been proposed, such as $\eta_t = \exp[-\eta(t\text{-}T)]$. [See *BC* (1995, p. 1995) for discussion.] Kumbhakar and Orea (2003) suggest a more general form, $\eta_{it} = \exp(\mathbf{g}_{it}'\boldsymbol{\pi})$ which subsumes *BC*'s formulation as well as many others. Greene (2003c) added the heterogeneous truncation form, $E[U_i] = \mu_i$ to this as well. Let $\boldsymbol{\eta}_i = (\eta_{i1},\eta_{i2},\ldots,\eta_{iT})$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},\varepsilon_{i2},\ldots,\varepsilon_{iT})$. Estimates of technical inefficiency based on this model follow the same form as those in the Pitt and Lee model, where Z_i is now

$$(17) \quad Z_i* \quad = \frac{\sigma_v^2\mu_i + \sigma_u^2\boldsymbol{\eta}_i'\boldsymbol{\varepsilon}_i}{\sigma_v^2 + \sigma_u^2\boldsymbol{\eta}_i'\boldsymbol{\eta}_i}$$

and ψ becomes

$$(18) \quad \psi_i* \quad = \frac{\sigma_u^2\sigma_v^2}{\sigma_v^2 + \sigma_u^2\boldsymbol{\eta}_i'\boldsymbol{\eta}_i} \, .$$

The degree of complexity of the model obviously increases with the added generality. The payoff is that the initial assumptions of homogeneity and time invariance are bypassed in the process.

A random effects counterpart to the true fixed effects model would be a 'true random effects' stochastic frontier model,

$$(19) \quad y_{tt} \quad = (\alpha + w_i) + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_{it}$$

The time invariant, random constant term embodies the cross country heterogeneity. The one sided inefficiency component now varies freely across time and country. This form of the model overcomes both of the drawbacks noted earlier. Estimation of this model by simulated maximum likelihood methods is discussed in Greene (2001, 2002a). Measured heterogeneity (income distribution, public contribution to health care financing, etc.) can enter this model through two avenues. Simple cross country heterogeneity may affect the location of the frontier, which would be modeled in the form

$$(20) \quad w_i \quad = \mathbf{f}_i'\boldsymbol{\theta} + \omega_i$$

This produces a 'hierarchical' or 'multilevel' model. The heterogeneity may also enter the distribution of u_{it} which can, as before, have mean μ_i or, in principle, even μ_{it} with time variation in the covariates. Country specific estimates of inefficiency are computed using the *JLMS* formulation, though simulation methods are needed to integrate out the unmeasured random effects.

A Random Parameters Model

The preceding has suggested various ways to accommodate both cross country heterogeneity and time variation in inefficiency in the stochastic frontier model. The most general of those considered thus far are constructed to accommodate the cross country heterogeneity. Time variation in inefficiency is achieved by removing restrictions on u_{it} and allowing it to vary unsystematically through time. We now consider two more complete specifications of the stochastic frontier which continue along these lines.

The 'true' random effects model suggested above is a special case of a fully specified random parameters (or hierarchical or multilevel) model,

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\boldsymbol{\beta}_i + v_{it} - u_{it}$$

$$\begin{pmatrix} \alpha_i \\ \boldsymbol{\beta}_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \boldsymbol{\beta} \end{pmatrix} + \begin{pmatrix} \Delta_\alpha \\ \Delta_{\boldsymbol{\beta}} \end{pmatrix} \mathbf{h}_i + \begin{pmatrix} \omega_{i\alpha} \\ \omega_{i\boldsymbol{\beta}} \end{pmatrix}$$

$$
\begin{aligned}
(21) \quad v_{it} &\sim N[0, \sigma_v^2] \\
u_{it} &= |U_{it}| \\
U_{it} &\sim N[\mu_i, \sigma_u^2] \\
\mu_i &= \mu + \mathbf{h}'_i \boldsymbol{\delta}_i \\
\boldsymbol{\delta}_i &= \boldsymbol{\delta} + \Delta_{\boldsymbol{\delta}} \mathbf{h}_i + \omega_{i\boldsymbol{\delta}} \\
\omega_i &= (\omega_{i\alpha}, \omega'_{i\boldsymbol{\beta}}, \omega'_{i\boldsymbol{\delta}}) \sim N[\mathbf{0}, \Omega].
\end{aligned}
$$

Estimation of the structural parameters is done by maximum simulated likelihood. As before, the end result is estimates of u_i which are computed using the *JLMS* formulation, from the reduced form, after integrating out the random effects. (The model as stated can be restricted in various ways, for example to place the heterogeneity only in the production function or only in the distribution of the inefficiency.)

Latent Class Models

The random parameters model allows the heterogeneity to take the form of continuous parameter variation across countries. An alternative, formulation that imposes somewhat more structure is the finite mixture, or latent class model [see Kumbhakar and Orea (2003) and Greene (2003c) for discussion]. The latent class form of the model posits that there are a finite number of structures underlying the data. Each country belongs to one class, though class membership is unknown to the analyst. On first glance, for example, the WHO data might seem naturally to be segregated between OECD and non-OECD countries, or, alternatively, between countries whose main focus in health care priorities is AIDS and other serious epidemics and those more focused

cancer and other quality of life issues. (Our investigation of these data did not suggest that either of these hypotheses seemed to be valid, however.)

The latent class model would appear as a finite mixture of stochastic frontier models,

(22) $(y_{it}|\text{class} = j)\quad = \alpha_j + \mathbf{x}_{it}'\boldsymbol{\beta}_j + v_{it}|j - u_i|j$

and a model for the mixing probabilities

(23) $\text{Prob}[\text{country i is a member of class j}]\quad = F_j(\mathbf{h}_i, \boldsymbol{\theta}), \; 0 \le F_{ij} \le 1.$

Heterogeneity enters this model through the prior mixing probabilities. As before, it can also enter through the distribution of u_{it}. The latent class model is an alternative to the random parameters model described in the preceding section. With a sufficient number of classes, the finite mixture can provide a good approximation to continuous parameter variation. In practical terms, this model is somewhat less flexible than the random parameters model discussed above. Kumbhakar and Orea (2003) and Greene (2003c) have extended it to the most general variant of the Battese and Coelli formulation of the random effects model. Since this approach is new to the literature, its usefulness as an empirical tool remains to be established.

3.3. Estimating the Stochastic Frontier Models

The initial criticisms of the fixed and random effects stochastic models, and to the stochastic frontier model in general, were founded on very narrow forms of the model. With panel data, sufficiently flexible and general forms of the model can be designed that should overcome these limitations and, in the process, produce effective estimates of technical inefficiency. All of these models can be estimated with existing, widely available software, some in Coelli's FRONTIER package (2000) and all of them in LIMDEP [Econometric Software, Inc. (2003).]

As noted by *GJJS*, the WHO data set, though it is a five year panel, actually contains very little within group variation. As a consequence, many of the model forms noted here are, in fact, not estimable with these data. The results below will present estimates based on a few of the formulations.

4. The World Health Organization WHR Data Set

The data set used in this analysis were used in Evans et al. (2000a,b).[13] The full data set is a panel of data observed for 191 member countries of the WHO. Observations were also provided for the 8 states of Australia and 32 of Mexico, 12 provinces of Canada, 26 internal jurisdictions of Sweden and 8 internal regions of the United Kingdom for a total of 277 observational units and 840 observations in total.[14] The panel data are observed for 5 years, 1993 to 1997, though 51 of the 191 countries and all of the aforementioned internal political units are observed in only one year.[15] The data are more fully described in the World Health Report and in numerous publications that can be obtained from the WHO website.

Two outcome variables were observed,

DALE = Disability Adjusted Life Expectancy.[16]

COMP = composite measure of success in 5 health goals, by year health, health distribution, responsiveness, responsiveness in distribution, fairness in financing. The composite is an equally weighted composite of the five attainment variables. The components of this variable were constructed from survey data gathered by WHO in the respective countries.

Natural logs of both outcome variables were used in the analysis to follow. (All references to these in regression results are based on logs.) *GJJS* expressed some skepticism about using logs for the DALE variable. In the interest of comparability, we have maintained the forms used by the researchers at WHO in this study.

The first of the outcome variables is the familiar output measure that was analyzed by *HW*, *GJJS* and Williams in addition to *ETML* (2000a). The second variable was analyzed in *ETML* (2000b), but were not analyzed by the other authors mentioned.

Two variables are modeled as the inputs variables to the production process of health care attainment:

HEXP = Health expenditure per capita in 1997 ppp$,

EDUC = Average years of schooling.

Both input variables entered the production function in log form. The translog model, with squares and the cross product was also considered. A restricted form of this model that has appeared in the earlier studies is discussed below.

---

[13] The assistance of researchers at WHO, especially David Evans and Ajay Tandon is gratefully acknowledged. The data set used here is an expanded version of the data used in ETML (2000a,b). The earlier papers did not use the invariant measures, per capita GDP, income distribution, etc.

[14] One of the issues discussed at the meeting mentioned in the introduction was whether it would be useful and appropriate to extend the analysis to the subcountry level units. For purposes of this paper, these data will be left unused. WHO has committed to ongoing study of methods and results of these studies, so these data may yet prove useful.

[15] One country, Algeria, was only observed four times.

[16] A variable DIFF which gives an estimate of the adjustment of DALE for AIDS was given for some countries. This was not used in the analysis.

Several variables that provide indicators of cross country and timewise heterogeneity were also analyzed:

GINI   = Gini coefficient, income inequality,
VOICE   = World Bank measure of democratization and freedom of political unit,
GEFF   = Measure of government effectiveness, World Bank measure,
TROPICS   = Dummy variable for tropical location,
POPDEN   = Population density, people per Km^2,
PUBFIN   = Percentage of health care paid by the government,
GDPC   = Per capita GDP in 1997 ppp$,
OECD   = Dummy variable for OECD membership[17].

The population density and per capita GDP variables appear in logs in all model results to follow.

Some of the data series were incomplete and had to be interpolated or otherwise filled in. Details on the data construction are provided in Appendix A. Descriptive statistics for the data used here are given in Table 2. (These are the 1997 values of these variables.)

Table 2: Descriptive Statistics for Variables, 1997 Observations

	Non-OECD		OECD		All	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev
DALE	54.32	11.73	70.27	3.01	56.83	12.29
COMP	70.30	10.96	89.42	3.97	73.30	12.34
HEXP	249.17	315.11	1498.27	762.01	445.37	616.36
EDUC	5.44	2.38	9.04	1.53	6.00	2.62
GINI	0.399	0.0777	0.299	0.0636	0.383	0.0836
VOICE	−0.195	0.794	1.259	0.534	0.0331	0.926
GEFF	−0.312	0.643	1.166	0.625	−0.0799	0.835
TROPICS	0.596	0.492	0.0333	0.183	0.508	0.501
POPDEN	757.9	2816.3	454.56	1006.7	710.2	2616.5
PUBFIN	56.89	21.14	72.89	14.10	59.40	20.99
GDPC	4449.8	4717.7	18199.07	6978.0	6609.4	7614.8
Sample	161		30		191	

---

[17] Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, The Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

5. Technical Inefficiency Estimates from the WHO Data

5.1. The Production Function

A number of researchers have studied health outcomes in a production function framework. Early contributions to this literature are Auster et al. (1969) and Grossman (1972), who analyzed mortality and suggested the approach used by a number of subsequent researchers. Berger and Messer (2002) provide a recent survey of the numerous cross country comparisons. Mortality and life expectancy have provided a focus of this literature. As noted by ETML (2000b), theirs is an innovation in its analysis of a broader measure of health care.

Health expenditure is the most visible input to the process, and public health expenditure is a major component of health care policy. There is tremendous variation in the public share of financing of health expenditure across countries and, as noted by Berger and Messer, across time as well. As can be seen in Table 2, the mean and standard deviation of 59.4 and 21.0 for PUBFIN, respectively suggest a range of variation of at least 20% to 100%. The interquartile range in our data is 45.4% to 76.2%; in the raw data, the actual range is from 9% to 100%. Some specific values for the larger economies include 72% for Canada, 25% for China, 92% for the Czech Republic, 77% for France, 78% for Germany, 13% for India, 57% for Italy, 82% for Norway, 78% for Sweden, 97% for the United Kingdom and 44% for the United States. Berger and Messer report comparable figures for some of these countries and note striking changes through time, as well, such as 86% for the UK in 1993 and an increase from 24% in 1960 to 41% in 1985 for the US. It is unclear, however, how or whether at all changes in the public financing of health care will translate to changes in health outcomes. Berger and Messer suggest, for example, that the degree of public financing could lead to either improvement or worsening of the *efficiency* of health care delivery. As they note, this aspect has not received much attention in the empirical literature. The data and models framed for this study will allow us to examine this issue. They also suggest a range of other factors that might affect health outcomes, such as income, demographic factors, and behavior. Again, the received results on these factors are fairly thin.

A number of researchers have examined the responsiveness of health expenditure to increases in income. Newhouse (1977) reported estimates of the income elasticity of health expenditures in the range of 1.15 to 1.31. Subsequent researchers have examined cross section and pooled time series-cross section data sets with similar results. [See BM (2002) for a survey.] Our own results based on the WHO data for 1997, shown in Table 3, are consistent, with elasticity estimates of 1.08 for the full sample and 1.23 for the OECD countries. The values in excess of 1.0 suggest that populations value health care as a normal good. How or whether those expenditures translate into health outcomes remains to be established, however.

Table 3: Income Responsiveness of Health Care Expenditure
(Estimated Standard Errors in Parentheses)

	Constant	Income	Education	R^2
Non-OECD	−3.67 (0.251)*	1.02 (0.037)*	0.275 (0.616)*	0.902
OECD	−5.57 (0.747)*	1.23 (0.078)*	0.347 (0.196)*	0.917
All	−4.14 (0.201)*	1.09 (0.031)*	0.268 (0.059)*	0.936

* indicates t statistics larger than 2.0.

BM recount a series of cross section studies that have found small and insignificant relationships between income levels and health outcomes. Subsequent analyses of income distribution as an alternative explanation have likewise concluded that the distribution of income adds little explanatory power. Our results of regressions of our two health outcomes, DALE and COMP (in logs) on health care expenditure, education, the Gini measure of income inequality, and the log of per capita income and its square, shown in Table 4, do not agree with these findings. The results suggest that for the poorer, non-OECD countries, there are significant relationships in the expected directions both for per capita income and for the income distribution. The results also suggest that the relationships are stronger for non-OECD countries than for OECD countries. The pattern of the quadratic relationships suggests that the improvement in health outcomes provides greater benefits at lower incomes than at higher ones. For example, for the non-OECD countries, the quadratic relationship suggests that the marginal improvement in logCOMP becomes zero at income of approximately exp(.474/(2(.050))) = \$ppp 13,000.[18] This is roughly the 95th percentile of the income distribution for the 161 nonOECD countries in the sample.

The production relationship suggests, not surprisingly, that health care expenditure should be an input to health outcomes.   Authors differ on what else should appear in the function, though education is a common component. Other variables suggested by *BM* would include health behavior (smoking, eating habits, etc.) demographic composition of the population, other market goods (for which we might use per capita income as a proxy), and country specific terms that should capture other unmeasured heterogeneity. *GJJS* suggested the difference between per capita GDP and health care expenditure as a measure of this effect. Since our data include only the 1997 per capita GDP, we have used it, rather than the difference, as this variable. There does appear to be something of a consensus in the literature on the direct inputs, health expenditure and education. Evans et al. omitted these from their production function for the practical reason that their fixed effects model could not accommodate time invariant covariates. Our results thus far suggest that a significant amount of explanatory power may have been foregone in the process. Precisely how these covariates should enter the model is another question. Only one of the several panel data models we propose above, the true fixed effects model, is constrained to omit time invariant effects. However, whether such shift factors should enter the production function or the inefficiency distribution is an important question without an obvious answer.

---

[18] This approximation ignores the scaling by $\exp(\sigma^2/2)$ necessary in the translation from the lognormally distributed variable COMP to the normal variate, logCOMP.

Table 4: Health Care Outcome Regressions[a]
(Estimated Standard Errors in Parentheses.)[b]

		Constant	Exp	Educ	Gini	Income	Income^2	R^2
NonOECD	**DALE**	3.04 (0.176)*	0.187 (0.028)*	0.017 (0.033)	-0.578 (0.148)*	0.098 (0.037)*		0.676
		-0.712 (0.832)	0.130 (0.029)*	0.033 (0.313)	-0.703 (0.142)*	1.046 (0.209)*	-0.058 (0.013)*	0.715
	COMP	3.56 (0.101)*	0.100 (0.016)*	0.036 (0.191)	-0.426 (0.086)*	0.064 (0.0214)*		0.769
		1.94 (0.495)*	0.075 (0.017)*	0.043 (0.019)*	-0.481 (0.085)*	0.474 (0.124)*	-0.025 (0.008)	0.784
OECD	**DALE**	3.82 (1.72)*	-0.042 (0.271)	0.046 (0.244)	-0.102 (0.078)	0.0157 (0.0316)		0.725
		-0.261 (2.25)	-0.041 (0.026)	0.036 (0.024)	-0.103 (0.074)	-0.896 (0.477)	-0.046 (0.024)	0.760
	COMP	3.91 (0.137)*	0.035 (0.216)	0.341 (0.195)	-0.119 (0.621)	0.031 (0.025)		0.846
		-0.507 (1.69)	0.036 (0.195)	0.023 (0.180)	-0.120 (0.056)*	0.967 (0.358)*	-0.049 (0.019)*	0.880
All	**DALE**	3.01 (0.162)*	0.197 (0.026)*	-0.0078 (0.029)	-0.397 (0.125)*	0.105 (0.034)*		0.714
		-0.481 (0.627)	0.128 (0.026)*	0.031 (0.028)	-0.625 (0.122)*	0.980 (0.156)*	-0.054 (0.009)*	0.757
	COMP	3.54 (0.093)*	0.107 (0.015)*	0.024 (0.017)	-0.338 (0.072)*	0.068 (0.196)		0.823
		1.97 (0.371)*	0.076 (0.016)*	0.041 (0.0165)*	-0.440 (0.072)*	0.462 (0.092)*	-0.024 (0.006)*	0.840

[a] DALE, COMP, Expenditure, Education, and Per Capita Income are all in logs. Results are all based on the 1997 data.

[b] * Indicates *t* ratios larger than 2.0

The basic production function analyzed here (and in *HW*) is of the simple form

(24)   *Health$_{it}$= f(Education$_{it}$, Expenditure$_{it}$) + v$_{it}$ - u$_{it}$.*

Additional influences on health outcomes that appear in the distribution of the inefficiency term are

(25)   **h**$_i$ = [*GEff$_i$, Voice$_i$, Gini$_i$, Tropics$_i$, PopDen$_i$, GDP$_i$, PubFin$_i$, OECD$_i$*]

These will also appear in the production function itself. Finally, in our preferred model, we will allow for time variation with the year specific dummy variables,

(26)   **t**   = year$_{1994}$, year$_{1995,}$ year$_{1996}$, year$_{1997}$.

The functional form of the production model remains to be determined. *ETML* specified a translog model,

(27)   *LogHealth* = α + $\beta_1 logExp$ + $\beta_2 logEduc$ + $\beta_{11}[(log^2 Educ)/2]$ + $\beta_{22}[(log^2 Exp)/2]$
                  + $\beta_{12}[(logEduc)(logExp)]$ + v_{it} - u_{it},

then, in the interest of parsimony, dropped the last two terms. Hollingsworth and Wildman, in the interest of comparability, adopted the same functional form. To maintain continuity of this strand of analysis, we will do likewise.[19]

5.2. Random and Fixed Effects Estimates of Inefficiency

Table 5 presents the estimated production functions based on the simplest panel data specifications, *ETML*'s fixed effects model and the Pitt and Lee random effects model.[20] Neither model has any built in accommodation for cross country heterogeneity. The two estimated models differ substantially, though as shown below, the differences in the estimates of the structural parameters is misleading – the estimated inefficiency terms (estimates of u_i) are nearly identical. The estimated random effects models are consistent with Gravelle et al.'s observation, that there is little within group variation. The variance decomposition is dominated by u_i; the estimates of $\lambda = \sigma_u/\sigma_v$ are 25.42 for DALE and 32.16 for COMP. Both of these are quite large by common standards.

Table 5: Estimated Frontier Production Models with Time Invariant Effects
(All variables in logarithms. Estimated standard errors in parentheses.)

		Constant	Expenditure	Education	Education^2	
DALE	Fixed Effects		0.00884 (0.00305)*	0.0629 (0.0363)	0.0435 (0.0285)	s=0.242 R^2=0.999
	Random Effects	3.94 (0.0159)*	0.0178 (0.00214)*	0.106 (0.0175)*	0.00672 (0.0145)	s_u=.2777 s_v=.0109
COMP	Fixed Effects		0.00654 (0.00185)*	0.0495 (0.0221)*	0.0455 (0.0171)*	s=0.172 R^2=0.999
	Random Effects	4.23 (0.0093)*	0.0116 (0.00108)*	0.0638 (0.0100)*	0.0398 (0.0878)*	s_u=.2132 s_v=.00663

* indicates the *t* ratio is larger than 2.0 in absolute value. Significant at the 95% level.

Estimated inefficiencies are computed using the methods discussed earlier. In all cases, to simplify comparisons, we have used the direct estimate of inefficiency, $\hat{u}_i$. For the fixed effects estimator, this is simply $\max(a_i) - a_i$. The random effects estimator is computed using the Jondrow et al. estimator in (9). Table 6 presents an analysis of these effects. The correspondence between the two sets of estimates, for both health outcomes is striking. The simple correlations between the pairs of estimates is almost 1. The random effects estimators also reproduce the rankings of the fixed effects estimator which, in turn, gives the same results as ETML obtained with their normalized version. (France remains fourth in the DALE results and first in the COMP results, for example.) This degree of correspondence between these two estimators has been observed

---

[19] Based on only the 191 observations for the year 1997, we did obtain *F* statistics for testing the hypothesis that the two second order terms for expenditure could be omitted from the equation of 9.54 for logDALE and 4.82 for logCOMP. The 99% critical value from F[2,185] is 4.72, so both hypotheses would be rejected.

[20] Likelihood functions are not presented with the maximum likelihood estimates in Table 5 or those to follow. Since we are not engaging in any type of specification search, these values would not be useful. Our interest here is in the estimates themselves and in the estimates of technical inefficiency. Fit measures for particular models are tangential to that analysis.

elsewhere. This finding suggests that the impact of the specific distributional assumption of the stochastic frontier model is not so severe as suggested earlier.

Table 6: Analysis of Estimated Technical Inefficiencies.
(Estimated standard errors in parentheses)

	DALE		COMP	
	Fixed Effects	**Random Effects**	**Fixed Effects**	**Random Effects**
Mean	0.2287	0.2121	0.1893	0.1772
Standard Dev.	0.1853	0.1793	0.1262	0.1187
Correlation	0.9984		0.9985	
Second Step Regression Results				
Constant	1.070　(0.139)*	0.9749 (0.1377)*	0.7584　(0.080)*	0.6826 (0.0801)*
Gini	0.440　(0.1359)*	0.4599 (0.1349)*	0.3223　(0.079)*	0.3167 (0.0784)*
Voice	-0.0230 (0.0163)	-0.0237 (0.0161)	-0.0261 (0.009)*	-0.025 (0.0094)*
GEFF	0.0280 (0.0211)	0.0288 (0.0209)	0.0065 (0.0122)	0.00654 (0.012)
LogPopDensity	-0.0095 (-.0054)	-0.0095 (0.0054)	-0.0061 (0.0031)	-0.0059 (0.0031)
LogIncome	-0.121　(0.0146)*	-0.111　(0.0145)*	-0.0806 (0.0085)	-0.073　(0.0085)*
Tropics	0.0506 (0.0230)*	0.0515　(0.0228)	0.0136　(0.0133)	0.0121 (0.0132)
Public Finance	-0.00024 (0.001)	-0.0003 (0.0005)	-0.0002 (0.0003)	-0.0002 (0.0003)
OECD	0.1014 (0.0365)*	0.0989 (0.0363)	0.0538 (0.0212)	0.0522 (0.0211)*
R^2	0.5533	0.5299	0.6774	0.6371

* indicates t ratios greater than 2.0 in absolute value. Statistical significance at 95% level.
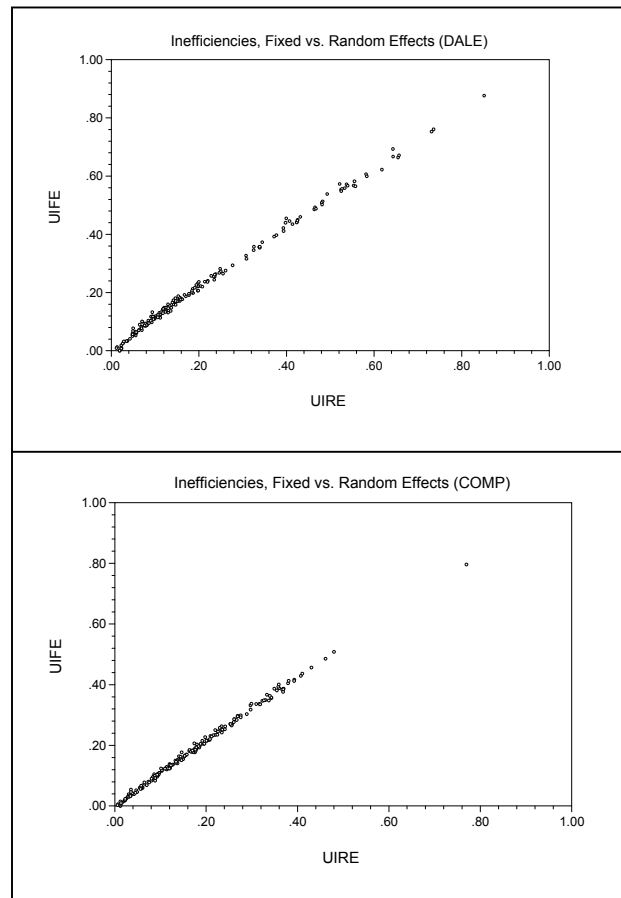
Figure 3: Estimated Fixed and Random Effects Inefficiencies

The lower panel of Table 6 presents a second step analysis of the estimated inefficiencies from the two models. They suggest that income and the distribution of income are both significant in explaining variation in efficiency. Since u_i is in proportional terms, the absolute magnitudes of the coefficients give the proportional impacts. It appears that the most important determinant is the distribution of income, with larger values of the gini coefficient (less equal income distribution) having a major negative impact on health outcomes however measured. (Increases in u_i imply lower efficiency.) The second largest determinant is per capita income, which works in the expected direction – higher income is associated with more efficient delivery of health care and achievement of higher life expectancy. (These results are not interpretable as direct impacts on the health outcomes.) Surprisingly, OECD status is associated with less efficient production. Note, though that per capita income is already in the equation, so whatever effect is at work in this persistent result is net of the impact of per capita income.

5.3. Incorporating Measured Heterogeneity in the Estimates of Inefficiency

The second step analysis of estimated inefficiencies such as that shown in Table 6 is a common exercise. In principle, if there are significant determinants of u_i in the observed information set, then they should appear in the model at the outset, and it may well be that estimates obtained without them suffer from the usual sorts of left out variable biases that are widely acknowledged in regression analysis. [Schmidt and Wang (2002) discuss this issue at length with respect to stochastic frontier analysis.] ETML's fixed effects estimator provides no method of accommodating time invariant indicators of heterogeneity in the model, but most of the others discussed to. Table 7 presents estimates of a random effects truncated normal stochastic frontier in which the time invariant covariates appear in the underlying mean of U_i. This is the heterogeneous form, with $E[U_i] = \mu_i = \mathbf{z}_i'\boldsymbol{\delta}$. The pattern in these results is similar to the preceding outcome. Once again, per capita income and the distribution of income appear to be significant determinants of the mean level of inefficiency, again in the expected direction. The inefficiency in the life expectancy equation is also significantly affected by tropical location.

The incorporation of cross country heterogeneity in the model has also produced the expected result with respect to the variation in inefficiency. In Table 2, in the simple random effects model, the standard deviation of the underlying distribution of U_i, σ_u, is estimated as 0.2777 and 0.2132 for DALE and COMP, respectively. In the expanded model in Table 7, the counterparts are 0.1567 and 0.07669. A large proportion of the variation in 'inefficiency' appears to be explainable as heterogeneity in the mean, instead. The estimates of the residual variation, σ_v, are identical in the two formulations.

We have included in Table 7 a comparison to the fixed effects estimates produced by *ETML*. As before, the results are strikingly similar, though noticeable differences can be seen as well. Figures 4 and 5 show the correspondence graphically. In fact, the statistical results in Table 7 do not reveal some rather large adjustments. For many countries, this expansion of the model appears to be fine tuning, but for a large number of others, quite substantial differences emerge. Table 8 shows the effect of the respecified model on the ranks of the top ten countries in the original fixed effects form.

Table 7: Estimated Heterogeneous Truncated Normal Stochastic Frontiers (Estimated Standard Errors in Parentheses)

		DALE		COMP	
		Truncation	**Batt./Coelli**	**Truncation**	**Batt./Coelli**
Production Function	**Constant**	3.96 (0.0196)*	3.87 (0.0156)*	4.27 (0.0117)*	4.29 (0.0188)*
	Exp	0.0162 (0.0023)*	0.0169 (0.0021)*	0.00950 (0.00129)*	0.0091 (0.0013)*
	Educ	0.1238 (0.0166)*	0.2037 (0.0179)*	0.0867 (0.00987)*	0.0790 (0.0134)*
	Educ^2	-0.0204 (0.0120)	-0.0667 (0.0118)*	-0.000827 (0.00856)	0.0019 (0.0089)
Mean Inefficiency	**Constant**	8.65 (1.88)*	1.20 (0.310)*	10.568 (1,50)*	0.829 (0.115)*
	Gini	6.06 (1.46)*	0.985 (0.239)*	6.58 (1.13)*	0.508 (0.0804)*
	Voice	-0.2312 (0.207)	-0.366 (0.0350)	-0.448 (0.146)*	-0.0349 (0.0117)*
	GEFF	0.00534 (0.307)	-0.0016 (0.0514)	-0.0722 (0.183)	-0.00513 (0.0145)
	Pop. Density	-0.0767 (0.0577)	-0.0118 (0.0095)	-0.0829 (0.0474)	-0.00643 (0.00371)
	GDP Per capita	-1.25 (0.207)*	-0.182 (0.0378)*	-1.27 (0.153)*	-0.0994 (0.0121)*
	Tropics	0.6402 (0.280)*	0.0997 (0.049)*	0.287 (0.202)	0.0229 (0.0158)
	Public Finance	-0.00253 (0.00656)	-0.0004 (0.0011)	-0.00497 (0.00519)	-0.0004 (0.0004)
	OECD	-0.0965 (1.032)	-0.0480 (0.177)	-0.272 (0.696)	-0.0207 (0.0537)
Distributions of u and v	σ_u	0.1567	0.1564	0.07669	0.07681
	σ_v	0.01092	0.01065	0.00664	0.00664
	λ	14.349	14.685	11.542	11.566
	Year 1994		0.0081 (0.026)*		-.000051 (0.018)
	Year 1995		0.0148 (0.0208)		-.00089 (0.0159)
	Year 1996		0.0227 (0.0174)		-0.0019 (0.0165)
	Year 1997		0.0301 (0.0129)*		-0.00323 (0.0109)
Estimated Inefficiency (In Comparison to Fixed Effects Estimates)	**Correlation**	0.884	0.993	0.891	0.993
	Rnk Corr.	0.891	X	0.895	X
	Mean	0.196(0.229)	0.195(0.229)	0.160(0.189)	0.178(0.189)
	Std. Deviation	0.189 (0.185)	0.181(0.185)	0.149(0.126)	0.128(0.126)

* indicates t ratios greater than 2.0 in absolute value. Statistical significance at 95% level.

Table 8: Selected Country Ranks Based on the Random Effects Model

	DALE		COMP	
Rank in FE	Country	New Rank	Country	New Rank
1	Malta	3	France	1
2	Oman	7	Italy	2
3	Italy	4	San Marino	72
4	France	5	Andorra	74
5	San Marino	40	Malta	8
6	Spain	2	Singapore	7
7	Andorra	42	Spain	4
8	Jamaica	8	Oman	25
9	Japan	1	Austria	6
10	Greece	6	Japan	3

The second set of estimates in each grouping in Table 7 are for Battese and Coelli's extension of the random effects model. This form of the model incorporates some time variation in the inefficiency; the generic specification is

(28) U_{it} $= g(\mathbf{z}_i, \mathbf{t}, \boldsymbol{\theta}) \, |U_i|$

This is a hybrid. The time variation is systematic. The underlying random component of the inefficiency remains time invariant. The effect of this on the overall nature of the inefficiency is a bit ambiguous. We have adopted an extension of Kumbhakar and Orea's (2003) form, with separate time effects (dummy variables) and a linear function of the country specific effects. Thus, this model includes measured country heterogeneity both directly in the production function and in the inefficiency distribution. The results in Table 7 suggest that both production and the efficiency of production are significantly affected by both the distribution of income and the level of per capita income. This persistent result has been hinted at in previous analysis. We do find considerable evidence of the result in our estimates. Figure 6 plots the estimates of u_i for this model. As before, it produces quite similar results to the fixed effects estimator.

Country Rank, Heterogeneous RE Model vs. Fixed Effects
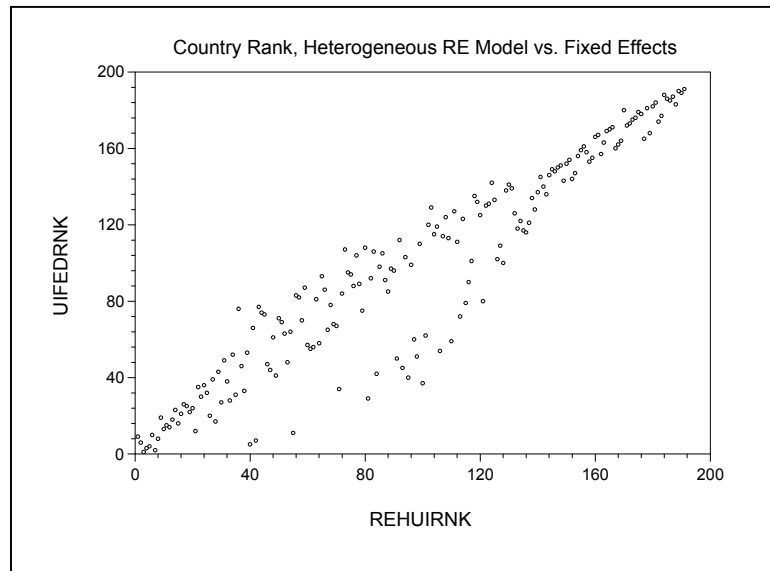
Figure 4a: Ranks Based on Fixed Effects and Random Effects Models

Estimated Inefficiencies: FE vs. Heterogeneous RE
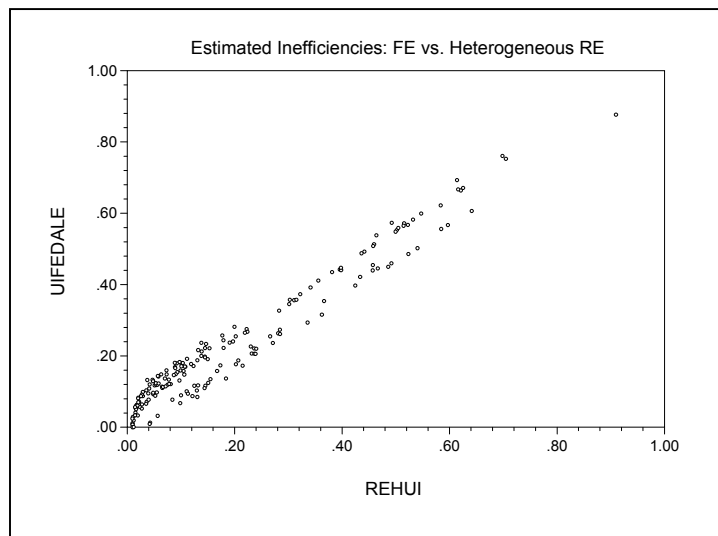
Figure 4b: Efficiencies Based on Fixed Effects and Random Effects Models

Figure 4: Comparison of Heterogeneous Random Effects to Fixed Effects Estimates, DALE
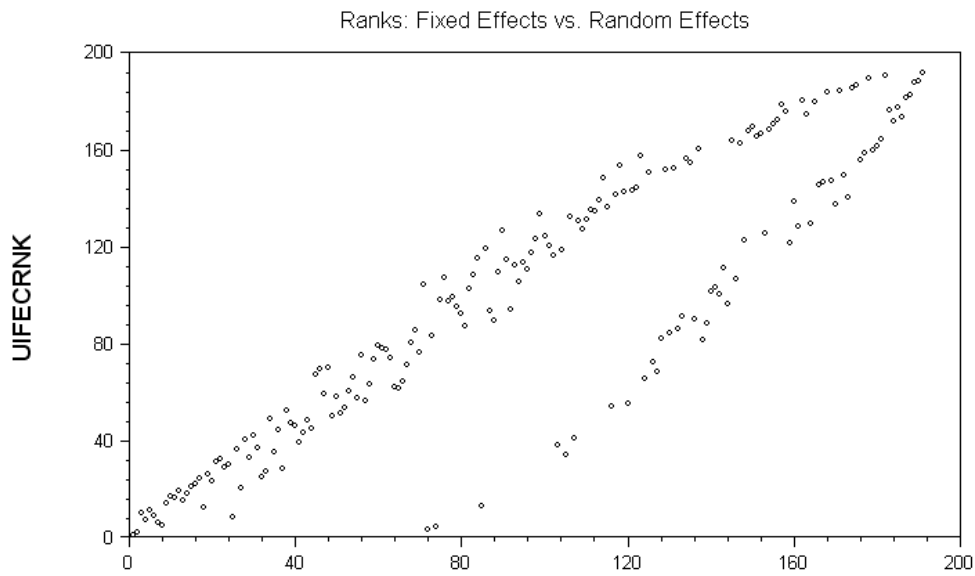
Figure 5a: Ranks Based on Fixed Effects and Random Effects Models
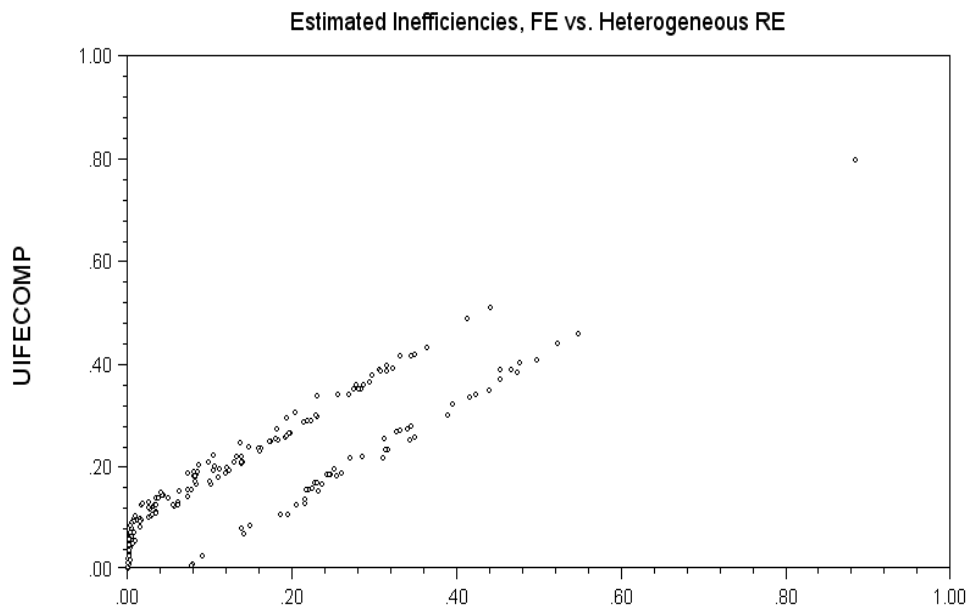
Figure 5b: Efficiencies Based on Fixed Effects and Random Effects Models

Figure 5: Comparison of Heterogeneous Random Effects to Fixed Effects Estimates, COMP
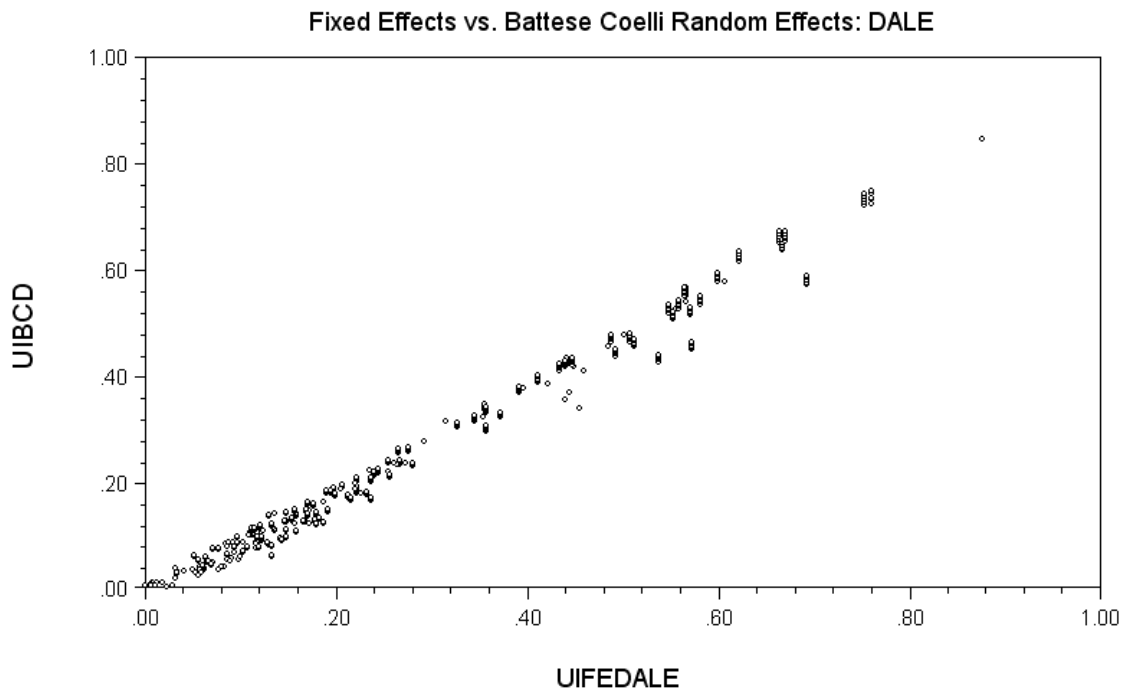
Figure 6a: Fixed Effects and Battese-Coelli Estimates, DALE
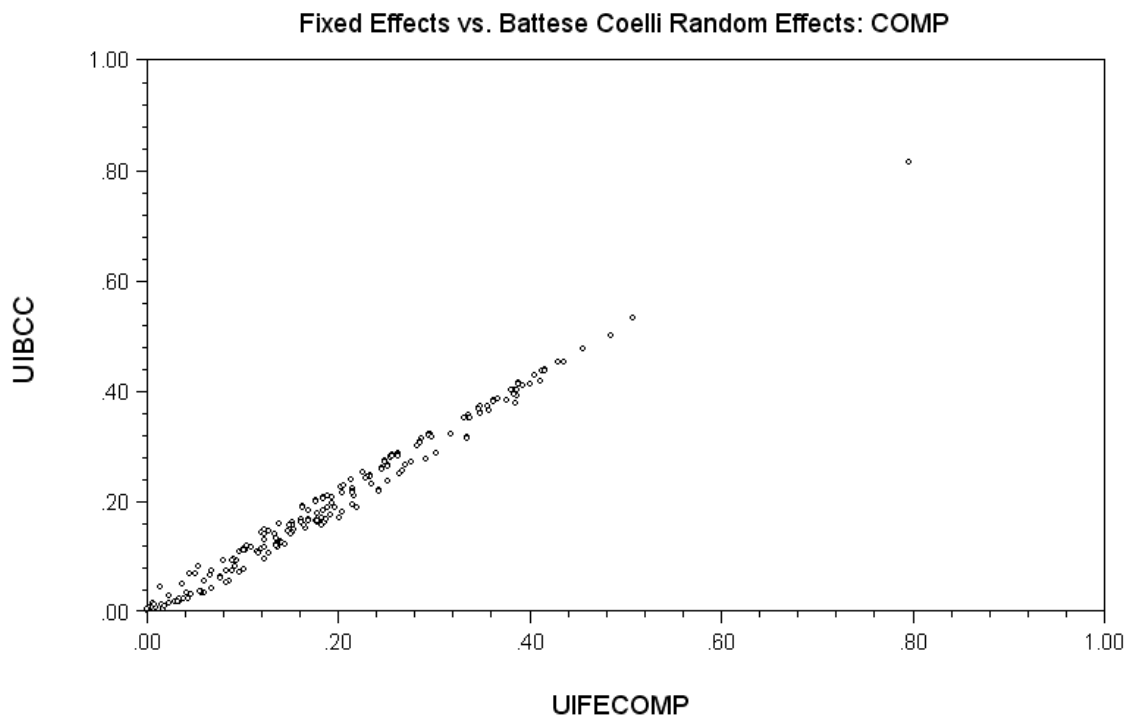
Figure 6b: Fixed Effects and Battese-Coelli Estimates, COMP

5.4. Incorporating Measured Heterogeneity in the Production Function

A primary motivation for the model extensions presented here is the possibility that the models with time invariant inefficiencies (random and fixed effects) might be forcing other unmeasured heterogeneity unrelated to technical efficiency into the estimates of u_i. The extensions in Sections 4.2 and 4.3 move some indicators of heterogeneity into the distribution of u_i, but retain the feature that other country characteristics are omitted from the model. We can accommodate this to some degree by retaining the random effects formulation, but building the cross country variation into both the production function and the distribution of inefficiency. The resulting structure is quite general;

$$y_{it} = \alpha + \mathbf{z}_i'\gamma + \mathbf{t}'\theta + \mathbf{x}_{it}'\beta + v_{it} - u_i$$

(29) $\quad u_i = |U_i|$

$$U_i \sim N[\mu_i, \sigma_u^2], \; \mu_i = \mathbf{z}_i'\delta$$

This extends the original *ETML* model in two ways, neither of which can be accommodated in the fixed effects formulation. We have also added time variation to the shift factors in the production. (The model is inestimable with these included in the mean of U_i.)

Table 9 presents parameter estimates for this model. (In the interest of brevity, at this point, we have omitted the estimated standard errors. The patterns of statistical significance are indicated in the results as in the preceding tables.) In terms of its impacts in the model, we now see that per capita income remains an important determinant both of the level of and efficiency of production. But, the distribution of income (Gini) appears to be an important determinant only of the efficiency of production. In view of our previously estimated models, this appears to uncover an important specification problem. In models less general than this one, it does appear that significant influences on efficiency, such as Gini, could be misplaced, and appear to be significant at points where they in fact might not be. Figures 7a and 7b display the estimates of u_i from this model, once again in comparison to the initial estimates in *ETML*. We have segregated the estimates by OECD and non-OECD countries in the figures and in the descriptive statistics in Table 9. This partition of the sample reveals a major aspect to all of the results. In the figures, the OECD countries are packed tightly in the lower left corner of the plots. From the descriptive statistics, as borne out by the figures, we now see that nearly all of the variation in the estimated inefficiencies in both *ETML*'s fixed effects estimates, and these, our final proposed alternatives, is attributable to the non-OECD countries in the sample.

It can be seen in the figures that this model brings some rather large changes in the results. To explore this further, we have plotted the country ranks for the DALE estimates in Figure 8. (The figure for COMP is similar.) The results have changed considerably. Table 10 lists the to 25 countries from the original fixed effects model, and their ranks produced by this more general model.

Table 9: Estimated Heterogeneous Truncated Normal Stochastic Frontiers

		DALE		COMP	
		Production	**Inefficiency**	**Production**	**Inefficiency**
Production Function	**Constant**	3.931*		3.953*	
	Exp	0.00716*		0.00547*	
	Educ	0.101*		0.0836*	
	Educ^2	-0.0462*		-0.0269*	
Technological Change	**Year 1994**	0.00173		0.00146	
	Year 1995	0.00363		0.00306	
	Year 1996	0.00558		0.00472	
	Year 1997	0.00799		0.00663	
Heterogeneity	**Constant**		9.58*		6.38*
	Gini	-0.112	5.86*	0.0338	8.94*
	Voice	0.00668	-0.250	-0.0137	-0.878*
	GEFF	-0.00716	-0.123	-0.00324	-0.101
	Pop. Density	-0.00272	-0.129	0.000249	-0.0804
	GDP Per capita	0.0226*	-1.41*	0.0426*	-0.970*
	Tropics	-0.00960	0.764*	-0.0336*	-0.0264
	Public Finance	0.000109	-0.00297	0.00007	-0.00700
	OECD	0.0143	-0.208	0.0126	-1.073
Distributions of u and v	σ_u		0.17398		0.08422
	σ_v		0.01073		0.00647
	λ		16.221		13.012
Estimated Inefficiency (In Comparison to Fixed Effects Estimates)	**Correlation**	0.971		0.928	
	Rnk Corr.	0.915		0.910	
	Mean OECD **Non** **All**	0.0266 (0.0793) 0.213 (0.256) 0.183 (0.229)		0.0126 (0.0558) 0.1543 (0.214) 0.132 (0.189)	
	Std.Dev.OECD **Non** **ALL**	0.183 (0.0473) 0.198 (0.188) 0.193 (0.185)		0.0230 (0.0400) 0.124 (0.121) 0.125 (0.126)	

* indicates t ratios greater than 2.0 in absolute value. Statistical significance at 95% level.
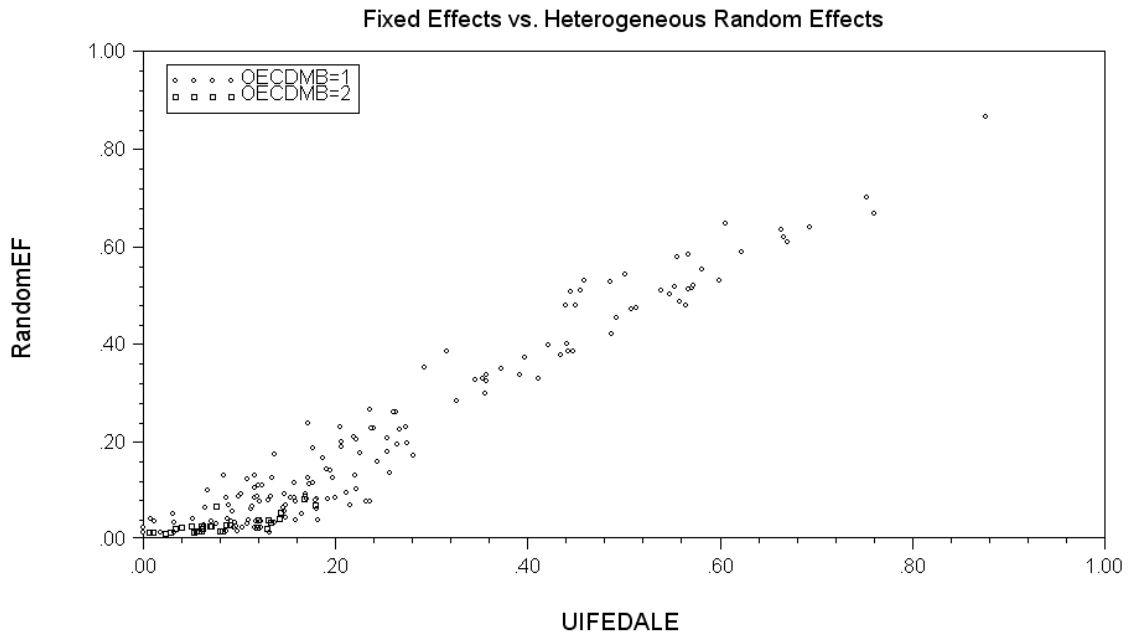
Figure 7a: Fixed Effects and Heterogeneous Random Effects Estimates, DALE
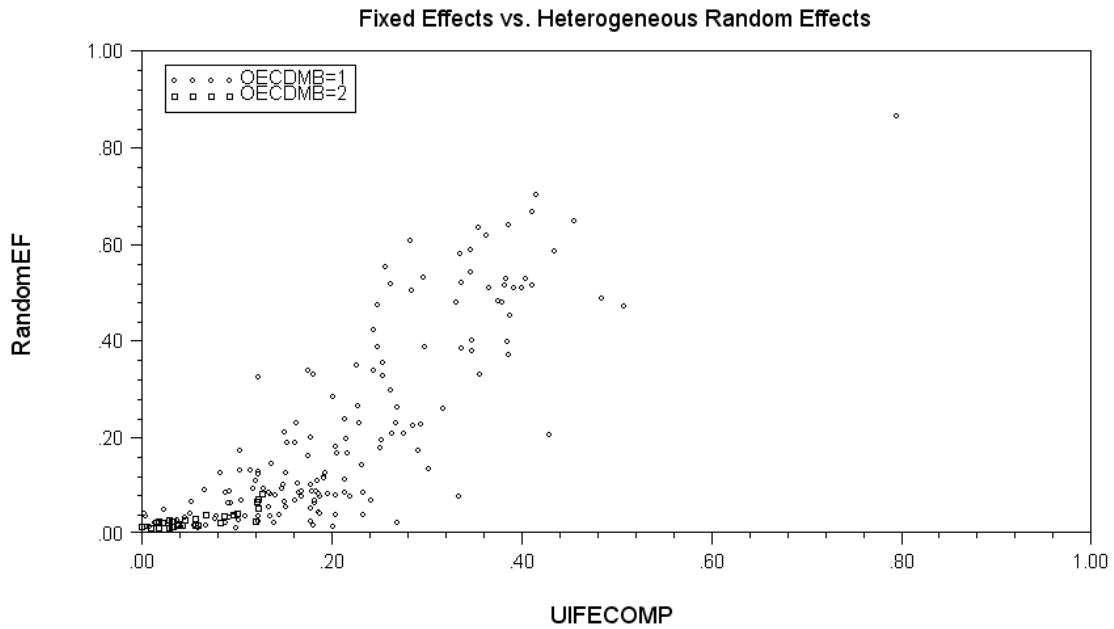
Figure 7b: Fixed Effects and Heterogeneous Random Effects Estimates, COMP
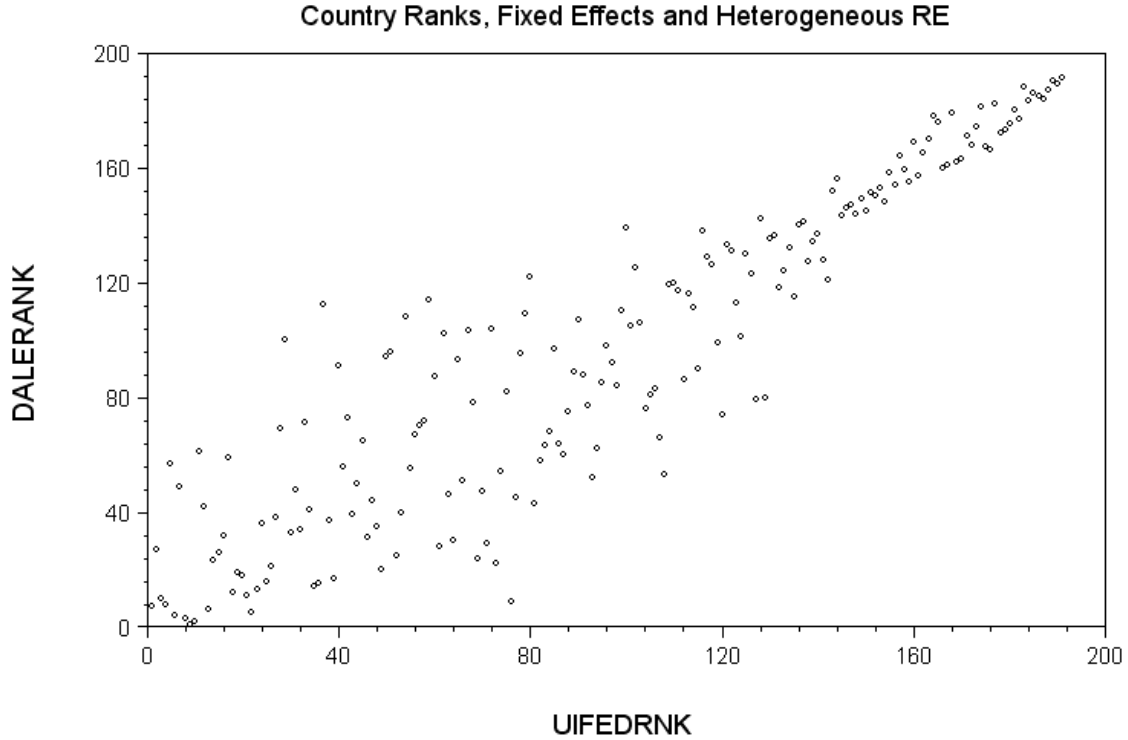
Figure 8: Country Ranks: Fixed Effects and Heterogeneous Random Effects, DALE

5.5 Alternative Estimates of Technical Efficiency

More detailed and elaborate models are discussed in Section 5.6. But, for reasons discussed below, these methods appear to be unsuitable for this data set. The model in Section 5.4 is our preferred specification. As a benchmark against alternative specifications, the results given there compare measures from these models to the fixed effects model based on otherwise the same specification. Save for the normalizations discussed earlier, however, these are comparable to Evans et al.'s results. The results shown in the preceding section are our alternative estimates. The direct comparison to Evans et al's estimates is presented in Table 11. The values in the table are based on their efficiency measure. The comparison to Evans' total effectiveness measure would be

$$(30) \quad \widehat{TE_i} \quad = \exp\left(-\hat{u}_i\right)$$

As before, most of the variation in the measured values comes from the non-OECD countries. The correlation of the ranks is compute using

$$(31) \quad Rank\ Correlation = 1 - \frac{6\sum_{i=1}^{191}(Rank_{ETML,i} - Rank_i)^2}{191(191^2 - 1)}.$$

Table 10: Country Ranks for the top 25 Countries in *ETML* Sample

Rank	Country (DALE)	New Rank	Country (COMP)	New Rank
	DALE		COMP	
1	Malta	7	France	15
2	Oman	27	Italy	11
3	Italy	10	San Marino	59
4	France	8	Andorra	58
5	San Marino	57	Malta	6
6	Spain	4	Singapore	8
7	Andorra	49	Spain	2
8	Jamaica	3	Oman	41
9	Japan	1	Austria	17
10	Greece	2	Japan	3
11	Monaco	61	Norway	10
12	Saudi Arabia	42	Portugal	32
13	Singapore	6	Monaco	69
14	Portugal	23	Greece	1
15	Austria	26	Iceland	21
16	Norway	32	Netherlands	7
17	United Arab Emir.	59	Luxembourg	30
18	Netherlands	12	Ireland	25
19	Sweden	19	United Kingdom	12
20	Costa Rica	18	Colombia	14
21	Cyprus	11	Switzerland	19
22	Chile	5	Belgium	22
23	United Kingdom	13	Cyprus	9
24	Iceland	36	Sweden	5
25	Switzerland	16	Saudi Arabia	79

Table 11. WHO and Stochastic Frontier Estimated Technical Efficiency

		DALE			COMP		
		OECD	Non	All	OECD	Non	All
Evans	Mean	0.851	0.643	0.676	0.883	0.597	0.637
	S.D.	0.079	0.199	0.200	0.078	0.200	0.214
	Min	0.694	0.080	0.800	0.734	0.000	0.000
	Max	0.976	0.992	0.992	0.994	0.998	0.998
Greene	Mean	0.973	0.823	0.847	0.987	0.859	0.879
	S.D.	0.018	0.148	0.146	0.022	0.117	0.118
	Min	0.923	0.422	0.422	0.879	0.080	0.080
	Max	0.990	0.990	0.990	0.996	0.995	0.996
Sample Size		30	161	191	30	161	191
Correlation		0.757	0.906	0.912	0.566	0.799	0.829
Rank Correlation		0.886			0.914		

The individual results for DALE are shown in Figure 9. The layer of points at the upper right is the OECD observations. The figure bears out the impression in the descriptive statistics, that the variation in technical efficiency is associated with the non-OECD countries. Appendices B and C give the full listing of our results in opposition to those of Evans et al. As is evident from the various descriptors, the change in the results it is fairly substantial.
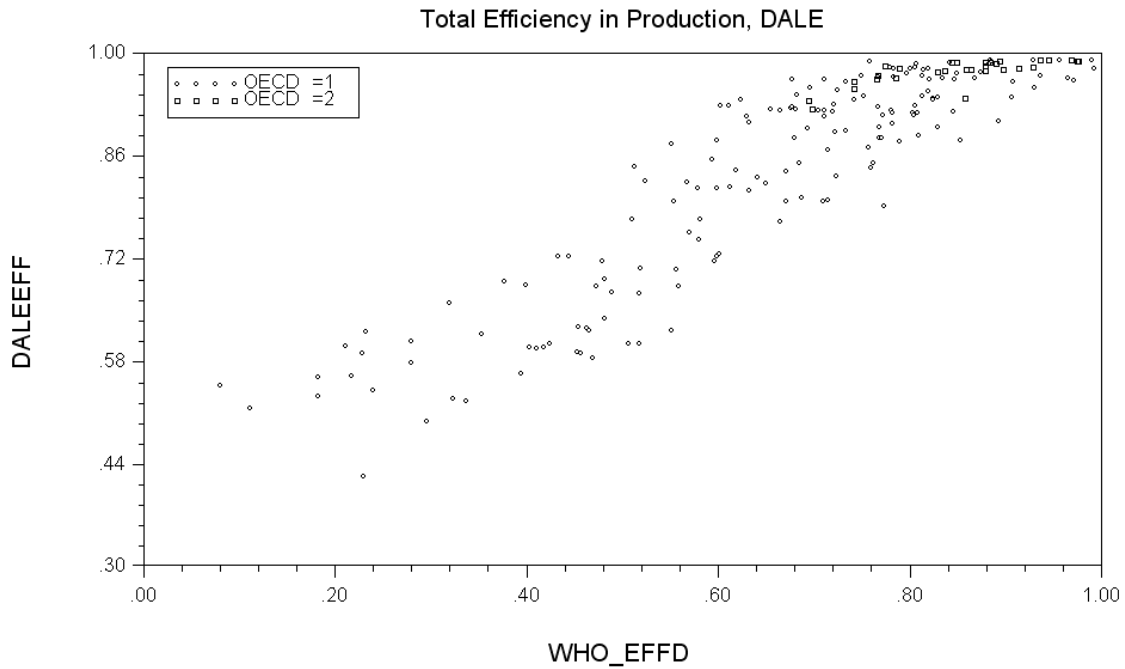
Figure 9: Total Efficiency, DALE

5.6. Models with Unmeasured Heterogeneity and Time Varying Inefficiency

We now examine some formulations that move all time invariant effects out of the distribution of u_{it}. These could be viewed as the most general specifications to be considered. However, in this data set, with almost no within group variation, it seems likely that, in fact, this approach will overspecify the country specific heterogeneity. We present these results to illustrate what should be useful extensions of the stochastic frontier model in data sets with greater within group variation. Results are presented for DALE only. Some of the models are inestimable with COMP (the latent class model, for example). Where both can be fit, the results are similar. In the interest of brevity, only one outcome variable is analyzed.)

A 'true' fixed effects model would appear as

$$y_{it} = \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_{it}$$

$$(32) \quad u_{it} = |U_{it}|$$

$$U_{it} \sim N[0, \sigma_u^2].$$

This differs from the model initially specified in that the inefficiency term u_{it} is unique in every period for every country. In principle, this is more general than the Schmidt and Sickles formulation – individual heterogeneity and time variation in u are both accommodated. Accommodating measured indicators of heterogeneity, such as per capita income, remains a problem. As we have seen, the random effects formulation can be made general enough that its narrow distributional assumption is probably not restrictive. A 'true ' random effects model at this level of generality would be

$$(33) \quad y_{it} = (\alpha + w_i) + \mathbf{z}_i'\boldsymbol{\gamma} + \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_{it}$$

This model is fit by maximum simulated likelihood. [See Greene (2001, 2002a).] Figure 10 compares the inefficiency estimates obtained with the two fixed effects models. They are nearly uncorrelated. The random effects estimates bear even less resemblance. Based on all the preceding evidence, it appears that further development of these models is called for. As noted, it seems that these model forms are not suited for this data set. Note the difference in the range of variation in the two sets of values. An appropriate decomposition between heterogeneity and inefficiency likely falls between the two.

True Fixed Effects vs. Fixed Effects, DALE
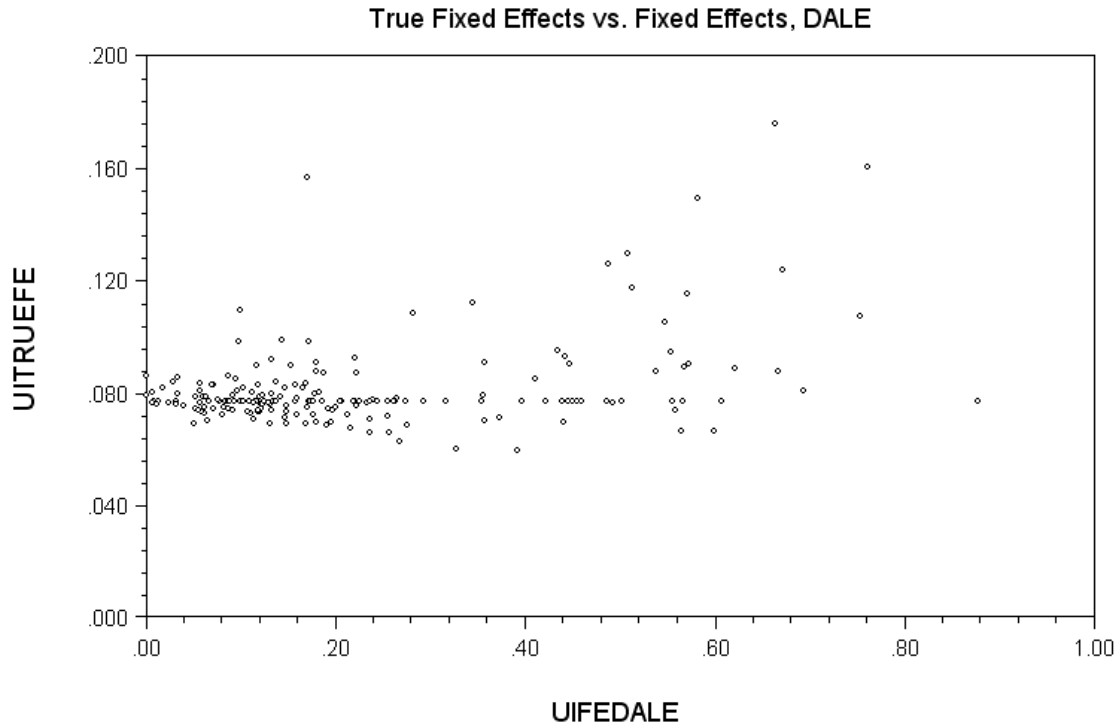
Figure 10: Fixed Effects Estimates

A full random parameters specification of the frontier model would, in principle, capture yet more country specific variation, and isolate the inefficiency in the time varying u_{it}. The estimated structure of a random parameters model for DALE is shown in Table 12. (Annotations are included to clarify the underlying structure.)

Table 12. Estimated Coefficients of a Random Parameters Model for DALE

	Random Coefficient Multiplies Variable			
	Constant	Expenditure	Education	Education[2]
Constant term in the mean of the parameter distribution				
Constant	1.084	0.181	0.164	0.059
Coefficients on variables that enter the mean of the parameter distribution				
Gini	−0.083	−0.141	−0.880	0.529
GEFF	−0.0617	0.00690	0.162	−0.206
Voice	0.127	−0.00719	−0.123	0.0849
Pop. Density	0.0112	0.00586	0.121	−0.131
Per Capita GDP	0.345	−0.0291	−0.0686	0.104
Tropics	−0.000036	0.0236	−0.566	0.0454
Public Finance	0.00603	0.00118	−0.00172	−0.00301
OECD	−0.509	0.00666	0.528	−0.390
Standard deviation of the random distribution of the parameter				
Std. Deviation	0.01362	0.05605	0.09876	0.14298
Parameters of underlying random components				
σ_u	0.01892			
σ_v	0.00910			
λ	2.0776			

The estimated correlation matrix of the random parameters is

$$\mathbf{R} = \begin{bmatrix} 1 & 0.987 & 0.417 & -0.715 \\ 0.987 & 1 & 0.504 & -0.691 \\ 0.417 & 0.504 & 1 & -0.500 \\ -0.715 & -0.691 & -0.500 & 1 \end{bmatrix}$$

A major element of the development in this paper is the accommodation of heterogeneity, both measured and unmeasured, in the production parameters and in the inefficiency distribution. Williams (2001) and others have argued that this a particularly important consideration in cross country comparisons of the production of health outcomes. The random parameters model provides an individual specific, posterior estimate of the parameter vector. [See Train (2002) and Greene (2002a) for discussion of the computations. Figure 11 presents kernel density estimators of the posterior distributions of the parameters of the production function across countries. These estimates embody all the sample information on the measured and unmeasured effects in the data. The figures suggest, as have other sources, that there is quite a large range of variation across countries in the parameters of the production function.
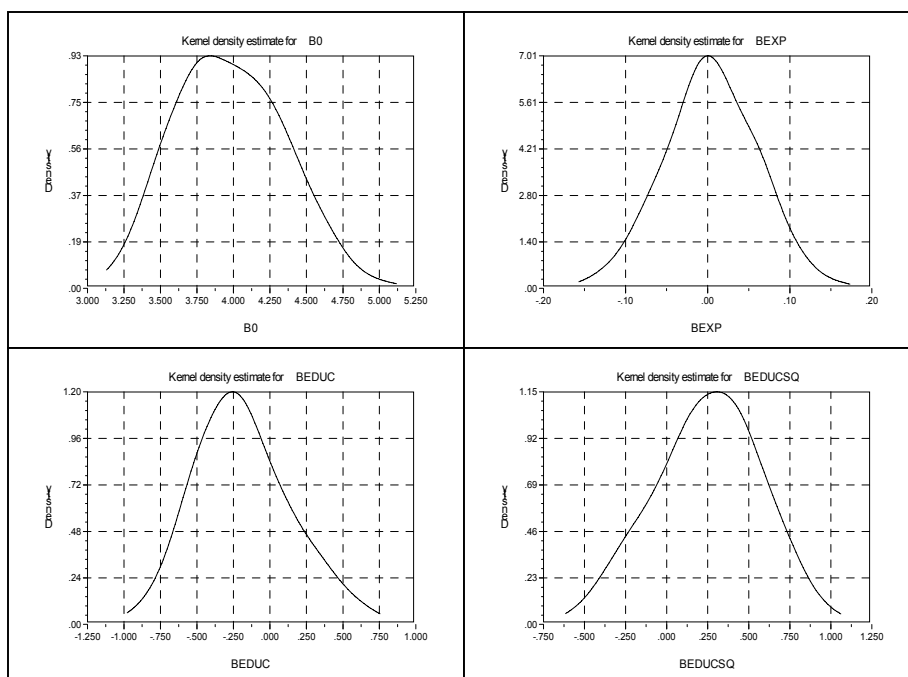
Figure 11: Kernel Density Estimates for Random Parameter Distributions

Estimates of u_i from the random parameters model are computed using the posterior, country specific parameter estimates. The differences between the fixed effects estimates and these in the descriptive statistics in Table 13 suggests the degree to which the random parameters model has absorbed the cross country variation. The correlation between these two sets of estimates is -0.1; they are essentially unrelated.

Table 13: Inefficiency Estimates

	Mean	Std. Deviation	Minimum	Maximum
Fixed Effects	0.222	0.186	0.000	0.876
Random Parameters	0.0246	0.0226	0.00130	0.0492

The finite mixture (latent class) model is yet another approach to modeling heterogeneity. In this case, the latent heterogeneity is modeled with discrete support, rather than continuous as above. It was not possible to fit any form of finite mixture model with the WHO data; evidently, this model is inconsistent with the data.

6. Conclusions

The WHO data analyzed here include most of the world's countries and embody nearly its entire population. The original studies by Evans et al. were an innovative, large scale effort to compare these countries on the effectiveness of delivery of health care using two measures, disability adjusted life expectancy (DALE) and a recently developed composite measure of health care delivery (COMP). The authors used a panel data technique based on the fixed effects regression model to obtain both quantitative measures of effectiveness and simple rankings for each country. Subsequent researchers have analyzed these same data, but generally provided descriptive statistics for the same measures that more or less agreed with the original results. This study has continued that analysis, with several extensions. First, the model used is considerably more general than those used previously. Second, we have incorporated measured indicators of cross country heterogeneity in the estimates. Third, we have produced an alternative set of individual country specific efficiency measures and country ranks. Our results differ substantially from those obtained by the authors of the original studies.

This reanalysis of the WHO data was motivated by several considerations.

- Both the fixed effects approach used earlier, and several others that have appeared in received studies make no distinction between technical inefficiency and cross country heterogeneity. Thus, some of what was reported as measured inefficiency should instead be treated as heterogeneity. Decomposing a two part unmeasured effect is a difficult exercise. The several models proposed here do so to varying degrees, ranging from none in the received models to, probably, excessively in the new techniques presented in the final section of this paper. We find, in our preferred specification, that making this distinction brings a substantial change in the estimated results.

- The fixed effects model used in Evans et. al. (2000a,b) does not allow the analyst to make use of measured effects that capture some of the cross country heterogeneity in the data. The models proposed here can accommodate these variables. We find including these measured indicators, such as per capita income and a measure of income distribution, noticeable changes in the estimated results.

- Being a rich panel data set, the WHO data present an opportunity to examine the behavior of new techniques which are better suited to studying heterogeneous data than were used earlier. The methodological sections of this paper have described a wide variety of panel data estimators for the stochastic frontier model.

We have obtained a number of empirical results. In formulating the production function, we confirm what earlier researchers have found with respect to the income elasticity of health expenditure. Our estimates range from 1.08 to 1.23, with much larger values for the wealthier OECD countries than for the remaining countries in the sample. A number of earlier studies have examined the relationship between income (measured

by per capita GDP) and health outcomes, and found a weak to nonexistent support. We find a persistent, significant impact of income, both on outcomes as well as the efficiency of distribution. Likewise, the distribution of income has been suggested as an important influence, but prior results were somewhat inconclusive. As in the case of income, itself, we find significant explanatory power in the distribution of income. It should be noted that in both these cases, and throughout this study, the distinction between OECD and non-OECD countries, which explains much of the variation in these measures, also explains much of the variation in health outcomes and in the efficiency of delivery of them.

Our final results are presented in Section 5.5 and in Appendices B and C. These are our alternatives to the those in the original studies. It should be emphasized, the measures and rankings in the tables pertain to the efficiency of delivery of health outcomes, not to absolute levels of those outcomes. Our results do not comment on the levels of health attainment in the countries contained in this data set.

Appendix A. Missing and Imputed Data in The World Health Organization WHR Data Set

The country level sample contained missing values for three variables, population density (PopDen), government effectiveness (GEFF) and democratization (Voice).

Data on population density were missing for the 18 countries listed in Table A.1 below. The website http://www.geography.about.com/library/cia/blcindex.htm contains year 2000 data on land area and population for the countries for which our data were missing. The table lists the country names and extracted data. The population density was calculated by drawing the year 2000 estimated population back to 1997 using the estimated population growth rate for three years, then computing the density in persons per square kilometer.

Table A.1. Extracted Data for Constructing Missing Population Density			
Country	Land Area	Population (2000)	Pop. Growth Rate
Bosnia	51,129	3,835,777	+3.1%
Belarus	207,600	10,366,719	-0.2%
Cook Islands	240	20,407	+1.6%
Czechoslovakia	72,276	10,272,179	-0.8%
Croatia	56,410	4,282.216	+0.9%
Kyrgisztan	191,300	4,685,230	+1.4%
Macedonia	24,586	2,041,467	+0.1%
Marshall Islands	181.3	18,126	+3.9%
Nauru	21	11,845	+0.2%
Niue	260	26,937	+1.5%
Palau	458	18,766	+1.8%
Slovakia	48,800	5,407,956	+0.1%
Slovenia	20253	1,927,593	+0.1%
Tuvalu	26	10,838	+1.4%
Turkmenistan	488,100	4,518,268	+1.9%
Yugoslavia	102,136	10,662,087	+0/8%
San Marino	60.5	2,113	+0.5%
Dem. Rep. Congo	2,267,600	51,964,999	+3.2%

The government effectiveness and popular voice variables are measures of the responsiveness of the government to the underlying populace. They are discussed, for example, at www.worldbank.org/wbi/governance. The government effectiveness variable (GEFF) is missing from the WHO data set for the 39 countries listed in Table A.2 below. Of these 39, 21 are also missing the VOICE variable. No countries are missing VOICE but not GEFF. A total of 83 observations are missing GEFF and 29 are missing VOICE. The website www.worldbank.org/wbi/governance/zips/aggind-data.xls provides data from several extensive studies of governance of 174 countries. Some of these were useable for filling in the gaps in the preceding data. Data for GEFF for the following countries were obtained from this source: Afghanistan (a), Barbados (b), Belize (a), Djibouti (b), Dominica, Eritrea (a,b), Equatorial Guinea (b), Cambodia (a), Laos (a),

Mauritania, Nepal, Rwanda, Burundi, Cape Verde, where (a) indicates that the value was from a year 2000 study and (b) indicates that a proxy, 'Rule of Law' was used.

Table A.2. Missing Data on Government Effectiveness (GEFF)		
Cook Islands (2)	Dominica (2), (3)	Mauritania (1), (3)
Andorra (2)	Eritrea (3)	Nepal (1), (3)
Afghanistan (3)	Micronesia (2)	Nauru (2)
Antigua (2)	Equatorial Guinea (1), (3)	Palau (2)
San Marino (2)	Grenada (2)	Rwanda (1), (3)
Burundi (1), (3)	Cambodia (3)	Solomon Islands
Belize (3)	Kiribati (2)	Sao Tome
Barbados (1), (3)	St. Kitts (2)	Seychelles (2)
Bhutan	Laotian PDR (3)	Tuvalu (2)
Niue (2)	St. Lucia (2)	Tonga (1), (2)
Comorros (1)	Monaco (2)	St. Vincent (2)
Cape Verde (1), (3)	Maldives (1)	Vanuatu (2)
Djibouti (3)	Marshall Islands (2)	Samoa (1), (2)
(1) Country observed for 5 years 1993-1997. (Others only observed for one year)		
(2) Also missing VOICE		
(3) Data for this country completed from additional sources noted in the text.		

No source for filling in the remaining observations for these variables could be located. The remaining 41 observations on GEFF and 28 on Voice were filled in by using the predictions produced by a linear regression of these variables on the 1997 values all of the other exogenous variables in the data set, including the inputs in the production function. (The R^2s in these regressions were 0.61 for Voice and 0.61 for GEFF.) This will produce a small amount of measurement error in two of the eight covariates. Since these observations constitute only about 5.4% of the sample of 754 observations and represent the very smallest countries in the data set, this should not materially affect the empirical results.

The sample contains data on 191 countries. Of these 191 countries, 140 were observed in all 5 years (11 of them incompletely as shown in Table A.2), one (Algeria) was observed in 4 years (1996 is missing) and 50 were observed only in 1997. We elected not to fill in the 1996 data for Algeria, primarily because none of the estimators discussed in the text or estimated in the empirical results actually requires a balanced panel and, secondarily, because the crucial health expenditure variable does not vary smoothly over the sample period for this country. Interpolation methods based on constant growth rates for this small number of years seemed likely to be quite inaccurate. The final sample contains a total of 754 observations.

Appendix B. Country Ranks and Performance Indices Based on the Stochastic Frontier Model, with WHO Results, DALE

Country	RANK	Efficiency	WHO Rank	WHO Efficiency
Japan	1	.98995	9	.94500
Greece	2	.98986	11	.93600
Jamaica	3	.98972	8	.95600
Spain	4	.98954	6	.96800
Chile	5	.98944	23	.88400
Singapore	6	.98896	14	.92900
Malta	7	.98868	2	.98900
France	8	.98857	4	.97400
Georgia	9	.98816	84	.75800
Italy	10	.98814	3	.97600
Cyprus	11	.98777	22	.88500
Netherlands	12	.98754	19	.89300
United Kingdom	13	.98657	24	.88300
Canada	14	.98647	35	.84900
Australia	15	.98605	39	.84400
Switzerland	16	.98572	26	.87900
Israel	17	.98568	40	.84100
Costa Rica	18	.98529	25	.88200
Sweden	19	.98510	21	.89000
Armenia	20	.98492	56	.80600
Belgium	21	.98117	28	.87800
United States of America	22	.98083	72	.77400
Portugal	23	.98033	13	.92900
Argentina	24	.97996	71	.77900
Croatia	25	.97963	57	.80500
Austria	26	.97871	15	.91400
Oman	27	.97865	1	.99200
China	28	.97863	61	.80000
Panama	29	.97827	67	.78300
Mexico	30	.97822	63	.78900
Uruguay	31	.97777	50	.81900
Norway	32	.97714	18	.89700
Luxembourg	33	.97662	31	.86400
Ireland	34	.97632	32	.85900
Colombia	35	.97560	51	.81400
Iceland	36	.97490	27	.87900
Germany	37	.97479	41	.83600
Venezuela	38	.97387	29	.87300
Finland	39	.97253	44	.82900
Slovenia	40	.97156	62	.79700
Cuba	41	.97127	36	.84900
Saudi Arabia	42	.96880	10	.93600
Barbados	43	.96857	87	.74900
Paraguay	44	.96791	52	.81300
New Zealand	45	.96754	80	.76600
Sri Lanka	46	.96678	66	.78300
Trinidad and Tobago	47	.96651	79	.76700
Bahrain	48	.96572	30	.86700
Andorra	49	.96565	7	.96400
Dominican Republic	50	.96547	42	.83400
Denmark	51	.96498	65	.78500
Ukraine	52	.96364	101	.71100
Tonga	53	.96346	114	.67700
Czech Republic	54	.96326	81	.76500
Honduras	55	.96319	48	.82000
El Salvador	56	.96248	37	.84600
San Marino	57	.96180	5	.97100
Poland	58	.96066	89	.74200
United Arab Emirates	59	.96066	16	.90700
Bulgaria	60	.95935	92	.73300
Monaco	61	.95229	12	.93000
Republic of Moldavia	62	.95160	106	.69600
Slovakia	63	.94969	88	.74200
Lithuania	64	.94784	93	.72400
Grenada	65	.94761	49	.81900
Romania	66	.94188	111	.68200
Qatar	67	.94110	53	.81300
Malaysia	68	.94037	86	.75100

Morocco	69	.93864	17	.90600
Algeria	70	.93821	45	.82900
Turkey	71	.93697	33	.85800
Tunisia	72	.93644	46	.82400
Yugoslavia	73	.93556	47	.82400
Russian Federation	74	.93531	127	.62300
Indonesia	75	.93527	90	.74100
Republic of Korea	76	.93361	107	.69400
Ecuador	77	.92845	96	.72100
Iran (Islamic Republic)	78	.92804	58	.80500
Viet Nam	79	.92799	130	.61100
Samoa	80	.92768	131	.60200
Estonia	81	.92615	115	.67700
Brazil	82	.92591	78	.76700
Belarus	83	.92337	116	.67600
Uzbekistan	84	.92269	112	.68100
Hungary	85	.92256	105	.69800
Latvia	86	.92241	121	.65500
Bosnia and Herzegovina	87	.92102	70	.78000
Thailand	88	.92081	102	.71000
Guyana	89	.92081	104	.70400
Peru	90	.92009	119	.66500
Saint Vincent	91	.91975	38	.84500
Lebanon	92	.91875	97	.71900
Cape Verde	93	.91813	55	.80800
Azerbaijan	94	.91794	60	.80300
Kuwait	95	.91773	68	.78200
Dominica	96	.91470	59	.80400
Nicaragua	97	.91410	74	.77200
Jordan	98	.91287	100	.71100
Philippines	99	.91221	126	.63000
Solomon Islands	100	.90578	20	.89200
Fiji	101	.90471	124	.63200
The Former Yugoslavian Rep	102	.90286	69	.78100
Egypt	103	.89824	43	.82900
Suriname	104	.89732	77	.76800
Niue	105	.89580	108	.69300
Syrian Arab Republic	106	.89256	91	.73300
Cook Islands	107	.89213	95	.72200
Saint Lucia	108	.88615	54	.80900
Brunei Darussalam	109	.88375	76	.76800
Iraq	110	.88364	75	.77000
Mauritius	111	.88315	113	.67900
Belize	112	.88015	34	.85300
Kazakhstan	113	.88006	135	.59800
Albania	114	.87901	64	.78900
Tajikistan	115	.87539	145	.55100
Pakistan	116	.86955	85	.75700
Guatemala	117	.86736	99	.71400
Bahamas	118	.85409	137	.59300
Micronesia	119	.84947	110	.68400
Yemen	120	.84835	82	.76100
Turkmenistan	121	.84404	152	.51300
Seychelles	122	.84309	83	.75900
Tuvalu	123	.83999	128	.61800
India	124	.83700	118	.67000
Libya	125	.83116	94	.72300
Antigua and Barbuda	126	.83023	123	.64100
Maldives	127	.82478	147	.52400
Bolivia	128	.82294	142	.56700
Saint Kitts and Nevis	129	.82132	122	.65000
Myanmar	130	.81612	129	.61200
Kyrgyzstan	131	.81475	134	.59800
Marshall Islands	132	.81450	140	.57900
Palau	133	.81203	125	.63200
Gambia	134	.80110	109	.68700
Nepal	135	.79902	98	.71400
Bangladesh	136	.79749	103	.70900
Kiribati	137	.79723	144	.55400
Sao Tome	138	.79710	117	.67100
Bhutan	139	.79090	73	.77300
Democratic P.R. Korea	140	.77266	153	.51000

Mongolia	141	.77230	138	.58100
Vanuatu	142	.76857	120	.66500
Comoros	143	.75474	141	.57000
Haiti	144	.74412	139	.58000
Senegal	145	.72461	132	.60100
Cote d'Ivoire	146	.72247	133	.59800
Nauru	147	.72161	166	.44400
Congo	148	.72142	167	.43300
Benin	149	.71497	136	.59600
Ghana	150	.71478	158	.47900
Sudan	151	.70629	149	.51900
Papua New Guinea	152	.70407	146	.55600
Cambodia	153	.69115	157	.48100
Equatorial Guinea	154	.68726	174	.37700
Cameroon	155	.68217	172	.39900
Gabon	156	.68150	143	.55900
Togo	157	.68140	159	.47200
Lao People's Dem. Rep.	158	.67291	155	.48900
Mauritania	159	.67145	151	.51700
Kenya	160	.65768	178	.32000
Guinea-Bissau	161	.63717	156	.48100
Central African	162	.62496	164	.45400
Burkina Faso	163	.62334	162	.46300
Eritrea	164	.62065	148	.55100
Chad	165	.62047	161	.46500
South Africa	166	.61981	182	.23200
Nigeria	167	.61557	175	.35300
United Rep. of Tanzania	168	.60578	180	.27900
Afghanistan	169	.60279	150	.51700
Somalia	170	.60202	154	.50600
Mozambique	171	.60188	168	.42400
Lesotho	172	.59987	186	.21100
Ethiopia	173	.59846	169	.41800
Burundi	174	.59793	171	.40300
Mali	175	.59596	170	.41000
Angola	176	.59144	165	.45300
Swaziland	177	.59040	184	.22900
Djibouti	178	.58910	163	.45700
Guinea	179	.58299	160	.46900
Uganda	180	.57682	179	.28000
Madagascar	181	.56203	173	.39400
Democratic Rep. of Congo	182	.55903	185	.21700
Namibia	183	.55690	189	.18300
Zimbabwe	184	.54566	191	.08000
Rwanda	185	.53973	181	.24000
Botswana	186	.53128	188	.18300
Niger	187	.52835	177	.32300
Liberia	188	.52480	176	.33700
Zambia	189	.51470	190	.11200
Malawi	190	.49676	187	.29600
Sierra Leone	191	.42205	183	.23000

Appendix C. Country Ranks and Performance Indices Based on the Stochastic Frontier Model with WHO Results, COMP

Country	RANK	Efficiency	WHO Rank	WHO Efficiency
Greece	1	.99563	14	.93300
Spain	2	.99558	7	.97200
Japan	3	.99547	10	.95700
Costa Rica	4	.99539	36	.84900
Sweden	5	.99528	23	.90800
Malta	6	.99524	5	.97800
Netherlands	7	.99515	17	.92800
Singapore	8	.99514	6	.97300
Cyprus	9	.99495	24	.90600
Norway	10	.99486	11	.95500
Italy	11	.99474	2	.99100
United Kingdom	12	.99471	18	.92500
Philippines	13	.99467	60	.75500
Colombia	14	.99463	22	.91000
France	15	.99460	1	.99400
Israel	16	.99460	28	.88400
Austria	17	.99449	9	.95900
Australia	18	.99438	32	.87600
Switzerland	19	.99430	20	.91600
Canada	20	.99418	30	.88100
Iceland	21	.99407	15	.93200
Belgium	22	.99407	21	.91500
Finland	23	.99403	31	.88100
Barbados	24	.99399	46	.80800
Ireland	25	.99376	19	.92400
Denmark	26	.99331	34	.86200
Germany	27	.99325	25	.90200
New Zealand	28	.99323	41	.82700
Slovenia	29	.99318	38	.83800
Luxembourg	30	.99310	16	.92800
Czech Republic	31	.99295	48	.80500
Portugal	32	.99268	12	.94500
Poland	33	.99228	50	.79300
Croatia	34	.99145	43	.81200
United States of America	35	.99018	37	.83800
Chile	36	.98872	33	.87000
Thailand	37	.98871	47	.80700
Jamaica	38	.98762	53	.78200
Ukraine	39	.98752	79	.70800
Slovakia	40	.98483	62	.75400
Oman	41	.98411	8	.96100
Trinidad and Tobago	42	.98403	67	.74200
Sri Lanka	43	.98361	76	.71600
Malaysia	44	.97855	49	.80200
Lithuania	45	.97675	73	.72200
Hungary	46	.97655	66	.74300
Dominican Republic	47	.97643	51	.78900
Republic of Korea	48	.97475	58	.75900
Belarus	49	.97321	72	.72300
Estonia	50	.97189	77	.71400
Tonga	51	.96930	116	.60700
Kazakhstan	52	.96788	64	.75200
Samoa	53	.96550	119	.58900
Nicaragua	54	.96534	71	.73300
Mexico	55	.96528	61	.75500
Republic of Moldavia	56	.96366	101	.63900
Georgia	57	.96341	114	.61500
Andorra	58	.96270	4	.98200
San Marino	59	.96003	3	.99800
Morocco	60	.95909	29	.88200
Cuba	61	.95889	39	.83400
Romania	62	.95757	99	.64500
Uruguay	63	.95673	65	.74500
Venezuela	64	.95665	54	.77500
Fiji	65	.95648	96	.65300
Armenia	66	.95552	104	.63000
Panama	67	.95319	95	.65600
Dominica	68	.95300	35	.85400

Monaco	69	.95125	13	.94300
Senegal	70	.95024	59	.75600
Bulgaria	71	.94738	102	.63900
Latvia	72	.94234	105	.63000
Jordan	73	.94147	83	.69800
Argentina	74	.94104	75	.72200
Paraguay	75	.94030	57	.76100
Mauritius	76	.93542	84	.69100
Tunisia	77	.93226	52	.78500
United Arab Emirates	78	.92971	27	.88600
Saudi Arabia	79	.92795	26	.89400
Kuwait	80	.92692	45	.81000
Ecuador	81	.92499	111	.61900
Guyana	82	.92319	128	.55400
Indonesia	83	.92173	92	.66000
Qatar	84	.91881	44	.81200
Bahrain	85	.91658	42	.82400
Brunei Darussalam	86	.91510	40	.82900
Egypt	87	.91062	63	.75200
Benin	88	.91018	97	.64700
Guatemala	89	.90991	78	.71300
Cape Verde	90	.90790	113	.61700
Bolivia	91	.90318	126	.57100
Lebanon	92	.90057	91	.66400
Albania	93	.89972	55	.77400
Comoros	94	.89914	118	.59200
Grenada	95	.89711	85	.68900
India	96	.89579	112	.61700
Uzbekistan	97	.89430	117	.59900
Solomon Islands	98	.89329	80	.70500
Niue	99	.89247	121	.58400
Bangladesh	100	.88931	88	.67500
Saint Vincent	101	.88918	74	.74400
Bahamas	102	.88712	94	.66700
Algeria	103	.88461	81	.70100
Palau	104	.88457	82	.70000
El Salvador	105	.88376	115	.60800
Cook Islands	106	.88292	107	.62800
Saint Lucia	107	.88284	68	.74000
Belize	108	.88176	69	.73600
Yemen	109	.88090	120	.58700
Bosnia and Herzegovina	110	.87976	90	.66400
Turkey	111	.87929	70	.73400
The former Yugoslav Rep.	112	.87888	89	.66400
Peru	113	.87702	129	.54700
Seychelles	114	.87663	56	.77300
Russian Federation	115	.87444	130	.54400
Honduras	116	.87434	131	.54400
Antigua and Barbuda	117	.87247	86	.68800
Ghana	118	.87084	135	.52200
Iran (Islamic Republic)	119	.86903	93	.65900
Kenya	120	.86894	140	.50500
Suriname	121	.86669	110	.62300
Azerbaijan	122	.85956	109	.62600
Syrian Arab Republic	123	.85939	108	.62800
Nauru	124	.85802	98	.64700
Brazil	125	.85474	125	.57300
Iraq	126	.85245	103	.63700
Pakistan	127	.85225	122	.58300
Micronesia	128	.85142	123	.57900
Burkina Faso	129	.85032	132	.54300
Viet Nam	130	.84961	160	.39300
Yugoslavia	131	.84912	106	.62900
Saint Kitts and Nevis	132	.84414	100	.64300
Maldives	133	.84381	147	.47700
Sudan	134	.84211	134	.52400
Haiti	135	.84187	138	.51700
Kiribati	136	.84174	142	.50500
Burundi	137	.84171	143	.49400
United Republic Tanzania	138	.84096	156	.42200
Tajikistan	139	.84034	154	.42800
Uganda	140	.83767	149	.46400

Marshall Islands	141	.83152	141	.50500
Tuvalu	142	.82982	136	.51800
Turkmenistan	143	.82702	153	.44300
Cote d'Ivoire	144	.82596	137	.51700
China	145	.82267	144	.48500
Zimbabwe	146	.81330	155	.42700
Libya	147	.81291	87	.68300
Togo	148	.81132	152	.44900
Gambia	149	.80615	146	.48200
Sao Tome	150	.80366	133	.53500
Vanuatu	151	.80149	127	.55900
Mongolia	152	.80027	145	.48300
Mali	153	.78950	163	.36100
Nepal	154	.78636	150	.45700
Kyrgyzstan	155	.78295	151	.45500
Congo	156	.77985	166	.35400
Cameroon	157	.77430	164	.35700
Bhutan	158	.77246	124	.57500
Rwanda	159	.77204	172	.32700
Zambia	160	.76247	182	.26900
Mauritania	161	.76243	162	.38400
Namibia	162	.76179	168	.34000
Papua New Guinea	163	.75984	148	.46700
Equatorial Guinea	164	.75644	171	.33700
Botswana	165	.75222	169	.33800
Madagascar	166	.74572	159	.39700
Gabon	167	.74050	139	.51100
Mozambique	168	.73993	184	.26000
Ethiopia	169	.73841	180	.27600
Niger	170	.73772	170	.33700
Malawi	171	.73614	185	.25100
Guinea-Bissau	172	.73200	176	.31400
South Africa	173	.72956	175	.31900
Djibouti	174	.72606	157	.41400
Democratic Rep. Korea	175	.72590	167	.35300
Lao People's Dem	176	.72555	165	.35600
Eritrea	177	.72455	158	.39900
Cambodia	178	.72114	174	.32200
Lesotho	179	.72040	183	.26600
Swaziland	180	.70747	177	.30500
Guinea	181	.70066	161	.38500
Chad	182	.68813	178	.30300
Nigeria	183	.67582	187	.27600
Myanmar	184	.67293	190	.13800
Somalia	185	.66585	179	.28600
Afghanistan	186	.66296	173	.32500
Angola	187	.65609	181	.27500
Central African	188	.64982	189	.21500
Democratic Repub. Congo	189	.64534	188	.17100
Liberia	190	.63272	186	.20000
Sierra Leone	191	.45812	191	0

References

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

Auster, R., I. Levinson and D. Sarachek, "The Production of Health: An Exploratory Study," *Journal of Human Resources*, 4, 1969, pp. 411-436.

Battese, G. and T. Coelli, "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India," *Journal of Productivity Analysis*, 3, 1, 1992, pp. 153-169.

Battese, G. and Coelli, T., "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical Economics*, 20, 1995, pp. 325-332.

Berger, M. and J. Messer, "Public Financing of Health Expenditures, Insurance, and Health Outcomes," *Applied Economics*, 34, 2002, pp. 2105-2113.

Coelli, T., "A Guide to FRONTIER, Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation," CEPA Working Paper 96/7, Department of Econometrics, University of New England, Armidale, New South Wales, Australia.

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

Econometric Software, Inc., "LIMDEP, Version 8.0," ESI, New York, 2002.

Gravelle, H., R. Jacobs, A. Jones, and A. Street, "Comparing the Efficiency of National Health Systems: Econometric Analysis Should be Handled with Care," University of York, Health Economics, UK, Manuscript, 2002

Gravelle, H., R. Jacobs, A. Jones, and A. Street, "Comparing the Efficiency of National Health Systems: A Sensitivity Approach, University of York, Health Economics, Manuscript, UK, 2002.

Greene, W. "Maximum Likelihood Estimation of Econometric Frontier Functions," *Journal of Econometrics*, 1980.

Greene, W., "A Gamma Distributed Stochastic Frontier Model," *Journal of Econometrics*, 46, 1, 1990, pp. 141-164.

Greene, W., "Frontier Production Functions," in M. Pesaran and P. Schmidt, eds., *Handbook of Applied Econometrics, Volume II: Microeconometrics*, Oxford, Blackwell Publishers, 1997.

Greene, W., Fixed and Random Effects in Nonlinear Models, Stern School of Business, Department of Economics, Working Paper 01-01, 2001 on simulated MLE

Greene, W., "Fixed and Random Effects in Stochastic Frontier Models, Stern School of Business, Department Of Economics, Working Paper 02-16, 2002a.

Greene, W., "Fixed Effects and the Incidental Parameters Problem in the Tobit Model," Manuscript, Department of Economics, Stern School of Business, New York University, 2002b.

Greene "Simulated Maximum Likelihood Estimation of the Normal-Gamma Stochastic Frontier Model," *Journal of Productivity Analysis*, 2003a, forthcoming.

Greene, W., *Econometric Analysis, 5th ed.*, Prentice Hall, Upper Saddle River, 2003b.

Greene, W., "Estimation of a Latent Class Stochastic Frontier Model," manuscript, Department of Economics, Stern School of Business, New York University, 2003c.

Grossman, M., "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy,* 80, 1972, pp. 223-255.

Hollingsworth, J. and B. Wildman, "The Efficiency of Health Production: Re-Estimating the WHO Panel Data Using Parametric and Nonparametric Approaches to Provide Additional Information," *Health Economics*, 11, 2002, pp. 1-11.

Hsiao, C, " Logit and Probit Models. In *The Econometrics of Panel Data: Handbook of Theory and Applications, Second Revised Edition,* Matyas, L. and Sevestre, P. (eds.). Kluwer Academic Publishers: Dordrecht, 1996.

Jondrow, J., I. Materov, K. Lovell and P. Schmidt, "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 19, 2/3, 1982, pp. 233-238.

Kim, Y. and P. Schmidt, "A Review and Empirical Comparison of Bayesian and Classical Approaches to Inference on Efficiency Levels in Stochastic Frontier Models with Panel Data," *Journal of Productivity Analysis*, 14, 2, 2000, pp. 91-118.

Kumbhakar, S. and K. Lovell, *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge, 2000.

Kumbhakar, S. and L. Orea, "Efficiency Measurement Using a Stochastic Frontier Latent Class Model," *Efficiency Series Papers*, University of Oviedo, 2002.

Lancaster, T. "The Incidental Parameters Problem Since 1948," *Journal of Econometrics*, 95, 2000, pp. 391-414.

Newhouse, J. "Medical Care expenditure: A Cross National Survey," *Journal of Human Resources*, 12, 1977, pp. 115-124.

Pitt, M. and L. Lee, "The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry," *Journal of Development Economics*, 9, 1981, pp. 43-64.

Polachek, S. and B.Yoon, "Panel Estimates of a Two-Tiered Earnings Frontier," *Journal of Applied Econometrics*, 11, 1996, pp. 169-178.

Schmidt, P. and R. Sickles, "Production Frontiers with Panel Data," *Journal of Business and Economic Statistics*, 2, 4, 1984, pp. 367-374.

Segal, D. and W. Greene "Profitability, Growth, and Efficiency in the U.S. Life Insurance Industry," *Journal of Productivity Analysis*, 2004, forthcoming.

Stevenson, R., "Likelihood Functions for Generalized Stochastic Frontier Functions," *Journal of Econometrics*, 13, 1980, pp. 57-66.

Train, K., *Discrete Choice: Methods with Simulation*, Cambridge, Cambridge University Press, 2002.

Tandon, A., Murray, C., Lauer, J. and Evans, D., "The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries," World Health Organization, GPE Discussion Paper, No. 29, EIP/GPE/EQC

Tandon, A., Murray, C., Lauer, J. and Evans, D., "Measuring Overall Health System Performance for 191 Countries," World Health Organization, GPE Discussion Paper, No. 30, EIP/GPE/EQC

Wang, H. and P. Schmidt, "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels," *Journal of Productivity Analysis*, 18, 2002, pp. 129-144.

Williams, A., "Science of Marketing at WHO? A Commentary on World Health 2000," Health Economics, 10, 2001, pp. 93-100.

World Health Organization, The World Health Report, 2000, Health Systems: Improving Performance, 2000, Geneva.