Why do some countries produce so much more output per worker than others? - A note

Gerry Boyle
National University of Ireland Maynooth*

Kieran McQuinn
Central Bank of Ireland **

Abstract

In an important paper, Hall and Jones (1999) show that international differences in output per worker across 127 countries in 1988 are fundamentally determined by variations in, what they term, a country's "social infrastructure". This paper conducts a robustness check of their findings by implementing a testing framework that is radically different to their approach. Specifically, we estimate a stochastic, rather than a deterministic, production frontier and we also model the potential role of social infrastructure in explaining productivity in a single step, rather than the statistically unsatisfactory two-step method used by Hall and Jones. We obtain two important findings that are strongly supportive of Hall and Jones' results. First, the bulk of inter-country variation in output per worker is accounted for by differences in productivity. Second, social infrastructure is found to be a highly significant variable in explaining inter-country productivity differences.

JEL classification: C33, D24, O12, O30

Key words: Productivity, Social Infrastructure, Stochastic Production Frontier.

*Department of Economics. E-mail: gboyle@may.ie.

**Economic Analysis, Research and Publications Department. E-mail: kmcquinn@centralbank.ie. The views expressed in this paper are the personal responsibility of the authors. They are not necessarily held either by the Central Bank of Ireland or the ESCB.

Why do some countries produce so much more output per worker than others? - A note

1 Introduction

In an important paper, Hall and Jones (1999) show that international differences in output per worker across 127 countries in 1988 are fundamentally determined by variations, in what they term, a country's "social infrastructure". Social infrastructure is taken to be the institutions and government policies that characterise the output generating process of an economy. Such institutions and policies, they contend, create the climate for enhanced output levels, as they provide an environment "that supports productive activities and encourages capital accumulation, skill acquisition, invention and technology transfer" (p. 2). Countries with a weak social infrastructure, on the other hand, create opportunities for diversionary activity in both the public and private sector which are not welfare enhancing.

Based on an initial simulation of a deterministic Cobb-Douglas production function, Hall and Jones identify productivity[1] as the main determinant of inter-country differences in levels of output per worker. They then go on, *inter alia*, to establish that the main driver of productivity is social infrastructure. Their key conclusion is that, while productivity is the main *proximate* cause of variations in output per worker, social infrastructure is the *fundamental* determinant. They illustrate their framework schematically as follows:

$$\text{Output per Worker} \leftarrow \text{(Inputs, Productivity)} \leftarrow \text{Social Infrastructure}$$

In words, social infrastructure determines productivity directly and output per worker indirectly.

We believe that there is substantial merit in the Hall and Jones framework. However, we contend that there is a more appropriate methodology that can be employed to test the validity of their framework than that used in their paper. Given the importance of their conclusions, it is important that they be tested for

[1] Total factor productivity.

robustness with respect to, what we believe, is a more appropriate methodology. In fact we shall show that our proposed methodology strengthens their conclusions.

An immediate issue with the Hall and Jones empirical methodology is that they employ a deterministic production frontier to estimate productivity. The productivity measure that is derived, thus, captures cyclical factors and measurement error. A preferable approach would be to estimate a stochastic production frontier. This would allow the productivity term to reflect systematic cross-country variation and, hence, would exclude purely random factors. The stochastic production frontier approach is well known (for example, Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977)). What perhaps is not as well known is the extension proposed by Battese and Coelli (1995), which puts forward an empirical framework that is almost perfectly consistent with that employed by Hall and Jones. The Battese and Coelli approach, however, is more attractive on two counts. First, it provides a method for recovering productivity estimates that is based on the estimation of a stochastic frontier and second their method allows for the estimation of the determinants of productivity in a single step unlike the two-step method employed by Hall and Jones. Their method should, thus, lead to more consistent and efficient estimates of productivity and its determinants.

The objective of our paper is, thus, to apply the Battese and Coelli methodology to the Hall and Jones data to assess the robustness of their findings. The approach put forward in this paper is consistent with that of Boyle and McCarthy (1996) and with recent research by Koop, Osiewalski and Steel (1999), Koop, Osiewalski and Steel (2000) and Kneller and Stevens (2002).

The rest of our paper is organised as follows. We first identify the proximate determinants of inter-country variations in productivity per worker by estimating a stochastic production frontier using the Hall and Jones data. We then proceed to assess the role of Hall and Jones' social infrastructure variable in explaining variations in productivity using the Battese and Coelli framework. In a final section we draw some conclusions.

2 The Proximate Cause of Variations in Output Per Worker

To enable comparisons to be drawn with the findings of Hall and Jones we first estimate productivity using their data[2] from a stochastic frontier given by the following

$$y_i = \left(\frac{K_i}{Y_i}\right)^{\alpha/(1-\alpha)} h_i e^{(\eta_i - A_i)}, \tag{1}$$

where
$y_i = Y/L$ is output per worker,
K_i is physical capital,
α is a parameter to be estimated,
$h_i = H_i/L_i$ is human capital per worker,
$\eta_i =$ independent and identically distributed errors i.e. $\eta \sim N\left(0, \sigma_\eta^2\right)$ and
$A_i =$ non-negative random variable (productivity) which is assumed to be independently distributed, such that A_i is obtained as truncations at zero of the $N\left(\mu, \sigma_A^2\right)$ distribution. It is noted that σ_η^2 and σ_A^2 are replaced with $\sigma^2 = (\sigma_\eta^2 + \sigma_A^2)$ and $\theta = \sigma_A^2/(\sigma_\eta^2 + \sigma_A^2)$, (see Battese and Corra (1977)), where θ is the fraction of the total variance of the residual that is attributable to variations in productivity.

Equation (1) was estimated by maximum likelihood and the results are given in Table 1 (insert Table 1 here).[3] The coefficient estimate for K_i/Y_i is higher than the 0.5 value assumed by Hall and Jones. The derived productivity estimates are given in Table 2 (insert Table 2 here) and we also present, for comparison purposes, the estimates produced by Hall and Jones from the deterministic frontier.

It is apparent, that while there are a few notable differences in the productivity estimates, overall, the productivity estimates are remarkably similar. While the means and standard deviations of our respective productivity estimates are very similar, it is apparent, that our estimate of productivity accounts for a lesser amount

[2] Downloaded from Charles Jones's web-site: http://elsa.berkeley.edu/ chad/datasets.html

[3] As endogeneity of the right hand side variable is a possibility, we also estimated (1) using the instrumental variables proposed by Hall and Jones for their social infrastructure variable, but the coefficient on (K/Y) was implausibly large, which could reflect significant measurement error in this variable.

of the variation between the top five and the bottom five countries (ranked in terms of output per worker).

Thus, our first important conclusion is that the employment of a stochastic, rather than a deterministic, frontier does not alter the fundamental finding of Hall and Jones, that productivity differences account for the vast bulk of the variation in output per worker across countries.

3 A Model of Productivity Differences

Here, we again follow Hall and Jones by modelling productivity in terms of cross-country differences in social infrastructure. We do so, however, by employing the single-step methodology proposed by Battese and Coelli (1995).[4] Hall and Jones generate their estimate of social infrastructure by taking the average of two variables. The first variable is an index of the anti-diversion policies of national governments created from data compiled by the group Political Risk Services. Hall and Jones follow Knack and Keefer (1995) by taking an average of 5 of the 24 different categories used to measure government's performance. These five sub-categories purport to (a) track a government's ability to protect against private diversion such as theft and fraud and (b) examine government's role as a potential diverter itself. The second variable is the Sachs and Warner (1995) index of a country's openness to trade. This index measures the fraction of years between 1950 and 1994 that a country has been open to trade. The absence of tariff quotas and associated trade barriers, it is argued, removes many oppertunities for private diversion amongst agents within the economy.

The Battese and Coelli (1995) approach involves estimating a two equation system of the form:

$$y_i = \left(\frac{K_i}{Y_i}\right)^{\alpha/(1-\alpha)} h_i e^{(\eta_i - A_i)}, \tag{2}$$

$$A_i = \rho S_i + \psi_i \tag{3}$$

[4]Estimates are obtained using the computer program FRONTIER Version 4.1, which is available on the Centre for Efficiency and Productivity Analysis (CEPA) web-site at www.uq.edu.au/economics/cepa/frontier.htm

where
S_i = social infrastructure,
ρ is a parameter to be estimated and
ψ_i = represents a random variable which is defined by the truncation of the normal distribution with a mean zero and a variance, σ^2. The point of truncation is at $-\rho S_i$ such that $\psi_i \geq -\rho S_i$.

We also follow Hall and Jones in using instruments[5] for social infrastructure, on the grounds of both measurement error and endogeneity. The instruments that they employ, broadly, seek to capture the effect of western European culture on the rest of the world, namely, the distance of a country from the equator and the extent to which the primary western European languages of English, German, French, Spanish and Portuguese are spoken as first languages today. European influence is hypothesised to have been more aimed at countries, which were rich in certain commodities, rather than those having relatively large output per worker today. Distance from the equator is taken as the absolute value of latitude in degrees divided by 90 to place it on a 0 to 1 scale. The data on languages are taken from both Hunter (1992) and Gunnemark (1991). It consists of two variables, one measuring the fraction of a country's population speaking one of the European languages as a mother tongue and another specifically measuring the faction of English spoken as a mother tongue. A fourth instrument, that is used, is a variable constructed by Frankel and Romer (1996), which is the (log) predicted trade share of an economy from a gravity model of international trade, which uses only the population and geographical features of a country as explanatory variables.

The results of estimating equation (2) using maximum likelihood are given in Table 3 (insert Table 3 here).

The most important finding is the highly significant coefficient on social infrastructure, which is strongly supportive of Hall and Jones. Our estimate may be compared with their estimate of 2.74 which is nearly 50 percent less than our coefficient. In fact, our estimate for the impact of social infrastructure on productivity is a mere 0.23 units less than their estimate for its impact on output per worker.[6] The

[5] We don't use instruments for (K/Y) because preliminary analysis using the same instruments as S yielded an implausible coefficient value.

[6] It should be noted, that, because our index of technical efficiency is defined over the [0,1] interval, it has the same interpetation as Hall and Jones, namely a 0.01 of a change in social

intercept value is relatively large, so we cannot rule out the possibility that other factors could account for variation in productivity, so the model may suffer from misspecification error. It is also worth noting, that the coefficient value on (K/Y) is now much closer to the assumed value of Hall and Jones. We also note that the estimated value of the variance parameter θ implies that the bulk of the variation in the residual of the frontier is due to systematic factors, that is, productivity.

Thus, our second important conclusion is that the employment of the Battese-Coelli methodology generates findings which are strongly supportive of Hall and Jones regarding the pivotal role of social infrastructure in driving productivity and strongly suggest that their methodology may have significantly understated its role.

4 Concluding Comments

In our opinion, Hall and Jones (1999) make a compelling case for the role played by a country's social infrastructure in determining the long-run level of output per worker, especially through its influence on cross-country productivity differences. Given the importance of their findings, this paper has conducted a stern robustness check by implementing a testing framework that is radically different to that used by Hall and Jones. Specifically, we estimate a stochastic rather than a deterministic production frontier and we also model the potential role of social infrastructure in explaining productivity in a single step, rather than the statistically unsatisfactory two-step method used by Hall and Jones.

We obtain two important findings. First, the bulk of inter-country variation in output per worker is accounted for by differences in productivity. Second, social infrastructure is found to be a highly significant variable in explaining inter-country productivity differences.

In these two key respects our results are strongly supportive of Hall and Jones. In fact, if anything, our findings provide more powerful evidence in support of the importance of social infrastructure in accounting for inter-country differences in output per worker.

infrastructure is associated with a 4.91 per cent change in productivity.

References

Aigner, D., Lovell, C. and Schmidt, P.: n.d., Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* **6**, 21–37.

Battese, G. and Coelli, T.: 1995, A model for technical inefficiency effects in a stochastic frontier production function for panel data, *Empirical Economics* **20**, 325–32.

Battese, G. and Corra, G.: 1977, Estimation of a production frontier model: With application to the pastoral zone of Eastern Australia, *Australian Journal of Agricultural Economics* **21**, 169–79.

Boyle, G. and McCarthy, T.: 1996, On and off the frontier: The impact of taxes on growth, *Economic and Social Review* **27**, 423–438.

Frankel, J. and Romer, D.: 1996, Trade and growth: An empirical investigation, *Working Paper 5476*, NBER.

Gunnemark, E.: 1991, *Countries, Peoples, and their Languages: The Geolinguistic Handbook*, Summer Institute of Linguistics Inc., Dallas, Texas.

Hall, R. and Jones, C.: 1999, Why do some countries produce so much more per worker than others?, *Quarterly Journal of Economics* **114**, 83–116.

Hunter, F.: 1992, *Ethnologue: Languages of the World*, Lanstryckeriet, Gothenburg, Sweden.

Knack, S. and Keefer, P.: 1995, Institutions and economic performance: Cross-country tests using alternative institutional measures, *Economics and Politics* **7**, 207–27.

Kneller, R. and Stevens, P.: 2002, The role of efficiency as an explanation of international income differences, *Discussion paper number 205*, National Institute of Economic and Social Research, Smith Square, London SWIP 3HE.

Koop, G., Osiewalski, J. and Steel, M.: 1999, The components of output growth: A stochastic frontier analysis, *Oxford Bulletin of Economics and Statistics* **61**, 455–487.

Koop, G., Osiewalski, J. and Steel, M.: 2000, Modelling the sources of output growth in a panel of countries, *Journal of Business and Economic Studies* **18**.

Meeusen, W. and van den Broeck, J.: 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review* **18**, 435–44.

Sachs, J. and Warner, A.: 1995, Economic reform and process of global integration, *Brookings Papers on Economic Activity* **1**, 1–95.

Table 1: Stochastic Production Frontier Parameter Estimates

Dependent Variable: $ln(Y/L) - ln(H/L)$

Regressor	
Intercept	8.851 (94.442)
$ln\,(K/Y)$	0.703 (9.777)
Log Likelihood	-127.296
Variance Parameters	
μ	0.596 (1.789)
σ^2	0.926 (3.319)
θ	0.972 (51.449)

Sample Size: = 127 observations, t-values in parentheses.

Table 2: Inter-Country Productivity Estimates from the
Estimation of a Stochastic and Deterministic Frontier (U.S.A. = 1)

Countries	Hall and Jones (A)	This Study (B)	Ratio (A/B)
U.S.A	1	1	1
Canada	1.034	1.022	1.012
Italy	1.207	1.095	1.102
West Germany	0.912	0.896	1.018
France	1.126	1.054	1.068
U.K.	1.011	1.038	0.974
Hong Kong	1.115	1.123	0.993
Singapore	1.078	1.041	1.036
Japan	0.658	0.656	1.003
Mexico	0.926	0.986	0.939
Argentina	0.648	0.689	0.940
U.S.S.R	0.468	0.454	1.031
India	0.267	0.327	0.817
China	0.106	0.122	0.869
Kenya	0.165	0.201	0.821
Zaire	0.160	0.229	0.699
Average (total sample)	0.516	0.561	0.920
Standard Deviation	0.325	0.314	1.035
Ratio top 5 to bottom 5	8.2	6.01	1.364

Table 3: Stochastic Production Frontier and Productivity Model Parameter Estimates

Dependent Variable: $ln(Y/L) - ln(H/L)$	
Regressors	
Intercept	8.978
	(86.941)
$ln\,(K/Y)$	0.513
	(5.637)
Dependent Variable: A	
Intercept	3.114
	(8.659)
Social Infrastructure	-4.914
	(-5.495)
Log-Likelihood Function	-105.964
Variance Parameters	
σ^2	0.537
	(5.010)
θ	0.963
	(45.107)

Sample Size: = 127 observations, t-values in parentheses.