A Continuous State Space Approach to “Convergence by Parts”

Paul A. Johnson*
Department of Economics
Vassar College
Poughkeepsie NY 12604

March 2004

Using a continuous state space approach, this note extends Feyrer's [2003] study of the proximate determinants of the shape of the long-run distribution of income per capita. Contrary to Feyrer's finding of the primacy of TFP, the results here imply that traps in both TFP growth and capital accumulation may matter.

JEL Classification: O40, O57

Keywords: twin peaks, convergence club, discretisation, development accounting

*Department of Economics, Vassar College, Poughkeepsie NY 12604-0708. Email: pajohnson@vassar.edu. Telephone: 845-437-7395. Fax: 845-437-7576. This note was written while I was an Honorary Felllow in the Department of Economics, University of Wisconsin, Madison. Their hospitality is gratefully acknowledged. I thank Steven Durlauf, Christopher Kilby, Jens Krueger, Joy Lei, Grazia Pittau and an anonymous referee for comments on an earlier draft. I am obliged to James Feyrer for graciously allowing me to use his data. All errors are mine.

Vassar College Economics Working Paper # 54

1. Introduction

The "development accounting" literature attempts to discover, and in some cases explain, the contributions of differences in inputs per capita and technology to cross-country differences in output per capita.[1] For example, Klenow and Rodríguez-Clare [1997] challenge the "neoclassical revival" begun by Mankiw, Romer, and Weil [1992] with the finding that cross-country variations in productivity explain a good deal more than than the 22% of the cross-country variation in output per capita found by the latter authors. Prescott [1998] finds a similarly important role for productivity differences which, he argues, cannot be explained by cross-country differences in technical knowledge alone. Hall and Jones [1999] also demonstrate the importance of productivity disparities and argue that differences in social infrastructure drive cross-country variation in both factor accumulation and productivity In addition, Henderson and Russell [2003] document the emergence of a second mode in the cross-country distribution of output per worker between 1965 and 1990 and, using data envelopment analysis, find changes in efficiency (the distance from the world tecnological frontier) and physical capital accumulation to be primarily responsible. Adding to this literature, Feyrer [2003] finds that the bimodalility in the long-run (ergodic) distribution of per capita output is due to bimodality in the ergodic distribution of productivity rather than in those of the quantities of per capita inputs. As he notes, this result has potentially important implications for theoretical modeling of development traps as it suggests that they are more due to traps in productivity growth rather than to the traps in physical capital accumulation often stressed in the development literature.[2]

The purpose of this note is to extend Feyrer's analysis using a continuous state-space approach. The contribution is that arbitrary discretisation of the state space and its

[1]The term "development accounting" is due to King and Levine [1994] who introduced it to differentiate this literature from the older growth accounting literature which focuses on the decomposition of output growth rates into contributions from technological progress and growth in inputs.
[2]In the spirit of Romer [1993], these could be referred to as "idea traps" and "object traps" resprectively.

possible effects on the results are avoided. Contrary to Feyrer's finding of the primacy of TFP, the results here imply that development traps may be due to traps in both TFP growth and capital accumulation.

2. Analysis

Feyrer [2003] uses the discrete Markov chain methods introduced to the empirical growth literature by Quah [1993] to compute estimates of the ergodic distributions of output per capita, the capital-output ratio, human capital per worker, and a measure of total factor productivity (TFP). He finds that the implied ergodic distributions of both output per capita and TFP are bimodal while those of both the capital-output ratio and human capital per worker are unimodal and so concludes "… that the origin of the twin peaks result for income is a result of productivity differences and not the accumulation of the factors of production" (p. 22).[3] This note extends Feyrer's analysis by using a continuous state-space method to analyze the transition dynamics and estimate the implied long-run distributions. This extension is important because, as Quah [1997] and Bulli [2001] discuss, the process of discretising the state space of a continuous variable is necessarily arbitrary and can alter the probabilistic properties of the data. In particular, as Reichlin [1999] demonstrates, the inferred dynamic behavior of the distribution in question and the apparent long-run implications of that behavior are sensitive to the discretisation. Especially relevant in the current context is the fact that the shape of the ergodic distribution – whether it is single or twin-peaked, for example – can be altered by changing the discretisation scheme.[4]

The data used here are exactly those used in Feyrer [2003], where a full discussion of sources, construction methods, and caveats can be found. Briefly, output

[3]This is consistent with Quah's [1996] finding that conditioning on measures of physical and human capital accumulation and a dummy variable for the African continent has little effect on the dynamics of the cross-country income distribution.

[4]See Quah [2001] for a discussion of all of these points and an advocacy of the continuous state space approach employed in this note.

per capita, y, is measured by RGDPC from the Penn World Tables, the capital-output ratio, k/y, is computed using capital stock data from Easterley and Levine [2001], and human capital per worker, h, is constructed following the approach in Hall and Jones [1999]. Following Klenow and Rodríguez-Clare [1997] and Hall and Jones [1999], for each country, Feyrer uses the assumed common world-wide production function $y = k^{\alpha}(Ah)^{1-a}$, with $\alpha = \frac{1}{3}$, written in the form $y = (k/y)^{\frac{\alpha}{1-a}}Ah$ so that A, the measure of TFP used here, is calculated $A = y/[(k/y)^{\frac{\alpha}{1-a}}h]$. As in Feyrer, each variable is expressed as a ratio to the corresponding within-period world mean prior to further analysis.

To estimate the long-run distributions of y, k/y, h, and A, I suppose that the time-t cross-country distribution of a variable x can be described by the density function $f_t(x)$, where x is variously y, k/y, h, or A. In general, this distribution will evolve over time so that the density prevailing at time $t+\tau$ for $\tau > 0$ is $f_{t+\tau}(x)$. Assuming that the process describing the evolution of the distribution is time-invariant and first-order, the relationship between the two densities can be written as $f_{t+\tau}(z) = \int_0^{\infty} g_{\tau}(z|x) f_t(x) dx$ where $g_{\tau}(z|x)$ is the τ-period-ahead density of z conditional on x.[5] After dividing the state space into 5 intervals based on the quintiles of the initial distribution of each variable, Feyrer computes 1-year Markov transition matrices and uses them to compute the implied ergodic distributions of y, k/y, h, and A. Accordingly, I estimate a $g_1(z|x)$ for these variables using the data described above and the adaptive kernel method described in Silverman [1986, Section 5.3].[6] So long as they exist, the ergodic (long-

[5]While the basic idea here is the same as that in Quah [1996, 1997], I simplify the presentation by assuming that the marginal and conditional income distributions have density functions. Quah's development of the approach avoids these assumptions and is far more general. Also, I have also abused notation slightly in the interests of simplifying the exposition.

[6]The adaptive kernel estimator is a kernel estimator with a window width that decreases as the local density of the data increases. In the first step of this 2-step estimator, a “pilot” estimate of the density is found. In the second step this density is used to vary the window width in an otherwise standard kernel estimator. I use an Epanechnikov kernel estimator with a (fixed) window width as given on pages 86-7 of Silverman [1986] to find the pilot estimate of the joint density. The adaptive kernel estimator of the joint density of z and x also employs the Epanechnikov kernel. Throughout, Silverman's suggested value of the “sensitivity parameter”, 0.5, is used. The estimated joint density of z and x is integrated over z to give the marginal

run) densities implied by each of the estimated $g_1(z|x)$ functions, $f_\infty(z)$, can be then found as the solution to $f_\infty(z) = \int_0^\infty g_1(z|x) f_\infty(x) dx$.[7] Figure 1 plots those densities.

Consistent with the results of Feyrer's discrete state space approach, and with the work of Quah and others, the estimated ergodic distribution of output per capita is bimodal with a mode at about half of mean income and another at about $2\frac{1}{4}$ times mean income. Similar to Feyrer, the estimated ergodic distribution of TFP is *almost* bimodal[8] and, I suggest, consistent with the hypothesis that the actual distribution is bimodal.[9] However, contrary to Feyrer's results, the estimated ergodic density of capital-output ratio is also bimodal, admitting the possibility that cross-country differences in the long-run behavior of income per capita can be explained by a model with multiple steady states in factor accumulation.

The estimated density of human capital per worker is strongly single peaked although the peak occurs close to the mean rather than well above the mean as found by Feyrer. Neither this nor the other differences between the results here and those of Feyrer are resolved by integrating the estimated ergodic density functions over the intervals used by Feyrer to construct his discretised data.[10] The point, as discussed by Quah [2001], is that arbitrary discretisation of the data alters its probabilistic properties. Bulli [2001] shows how to discretise the state space in a way that preserves these properties and finds that when this method is applied to cross-country data on income per

---

density of x. The ratio of the former to the latter provides the estimate of $g_1(z|x)$ used to calculate $f_\infty(z)$. All computations in this paper were performed using GAUSS.

[7]The solution method is outlined in the appendix. Johnson [2000] uses the approach employed in this paper to investigate the transition dynamics and implied long-run behavior of income per capita in the US states.

[8]By this I mean that only a little extra mass would have to be added to the $f_\infty(x)$ for A in a neighborhood of $x = 1.4$ for the density to become bimodal.

[9]As Quah [2001] notes, there is "as yet" no theory of inference for this issue but it seems clear that any confidence bands around the $f_\infty(x)$ for A would not need to be very wide in order for a bimodal null density to be drawn within them.

[10]For example, Feyrer divides the data on the capital-output ratio (relative to the within-period mean) into the intervals 0 to 55%, 55% to 83%, 83% to 111%, 111% to 147%, and 147% to ∞, and finds the corresponding values of the ergodic distribution to be 0.12, 0.18, 0.25, 0.26, and 0.19 respectively. Integrating the ergodic density for k/y found here over these intervals gives $\int_0^{0.55} f_\infty(x)dx = 0.22$, $\int_{0.55}^{0.83} f_\infty(x)dx = 0.22$, $\int_{0.83}^{1.11} f_\infty(x)dx = 0.17$, $\int_{1.11}^{1.47} f_\infty(x)dx = 0.15$, and $\int_{1.47}^{\infty} f_\infty(x)dx = 0.23$.

capita the estimated ergodic distribution is quite different from that found by arbitrary discretisation as well as being an accurate approximation to the distribution computed using a continuous state space method.

3. Conclusions

The results in this note do not support the conclusion that the long-run twin peaks in output are due solely to twin peaks in TFP.[11] Rather, these results are consistent with the view that the apparent bimodality in long-run distribution of output per capita is the product of bimodality in the long-run distributions of both the capital-output ratio and TFP. Instead of TFP playing an exclusive role, the effects of TFP and the capital-output ratio seem to reinforce each other with regard to the shape of the long-run distribution of output per capita. An important caveat on these results arises because, as is often the case in the development accounting literature, TFP is measured here as a residual under the assumption of a common world-wide production function. Durlauf and Johnson [1995] present evidence contrary to that assumption and in support of the implied multiple steady states in the growth process. As Graham and Temple [2003] show, the existence of multiple steady states can increase the variance and accentuate bimodality in the observed cross-country distribution of TFP in such circumstances. The extent to which the shape of the ergodic distribution of TFP presented here reflects this influence remains a matter for future inquiry. Finally, nothing in this note should be taken to imply anything about the relative contribution of factors of production or productivity to the cross-country variation in output per capita.

[11]The shapes of the estimated ergodic densities are, of course, sensitive to the window widths used in computing the underlying estimated joint density functions. As Silverman [1986, Section 2.4] explains, wider windows will tend to obscure detail in the shapes while narrower windows tend to increase it but possibly spuriously so. This sensitivity is of little concern for the conclusions reached here as equiproportionate increases in the window widths will remove any tendency to bimodality in the ergodic density of A before doing so in that of k/y. Similarly, equiproportionate decreases in window widths will make the bimodality in A more pronounced without removing that in k/y.

References

Bulli, Sandra, [2001], "Distribution Dynamics and Cross-Country Convergence: A New Approach," *Scottish Journal of Political Economy*, 48:226-43.

Durlauf, Steven N., and Paul A. Johnson, [1995], "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, 10:365-84.

Easterly, William and Ross Levine, [2001], "It's Not Factor Accumulation: Stylized Facts and Growth Models," *World Bank Economic Review*, 15:177-219.

Feyrer, James, [2003], "Convergence by Parts," manuscript, Dartmouth College. (available at http://www.dartmouth.edu/~jfeyrer/parts.pdf)

Graham, Bryan S., and Jonathan R. W. Temple, [2003], "Rich Nations, Poor Nations: How Much Can Multiple Equilibria Explain?," manuscript, University of Bristol. (available at http://www.ecn.bris.ac.uk/www/ecjrwt/abstracts/richpoor.htm)

Hall, Robert E., and Charles I. Jones, [1999], "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114:83-116.

Henderson, Daniel J., and R. Robert Russell, [2003], "Human Capital and Convergence: A Production-Frontier Approach," manuscript, SUNY Binghamton. (available at http://bingweb.binghamton.edu/~djhender/pdffiles/hc.pdf)

Johnson, Paul A. [2000], "A Nonparametric Analysis of Income Convergence across the US States," *Economics Letters*, 69:219-23.

King, Robert G., and Ross Levine, [1994], "Capital Fundamentalism, Economic Development, and Economic Growth," *Carnegie-Rochester Conference Series on Public Policy*, 40:259-92.

Klenow, Peter J., and Andres, Rodríguez-Clare, [1997], "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," NBER Macroeconomics Annual 1997, 73-103, Ben S. Bernanke, and Julio J. Rotemberg, eds., MIT Press, Cambridge.

Mankiw, N Gregory, David Romer, and David N. Weil, [1992], "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, CVII:407-37.

Prescott, Edward C., [1998], "Needed: A Theory of Total Factor Productivity," *International Economic Review*, 39:525-51.

Quah, Danny, [1993], "Empirical Cross-section Dynamics in Economic Growth", *European Economic Review*, 37:426-34.

Quah, Danny, [1996], “Convergence Empirics Across Economies with (Some) capital Mobility”, *Journal of Economic Growth*, 1:95-124.

Quah, Danny, [1997], “Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs”, *Journal of Economic Growth*, 2:27-59.

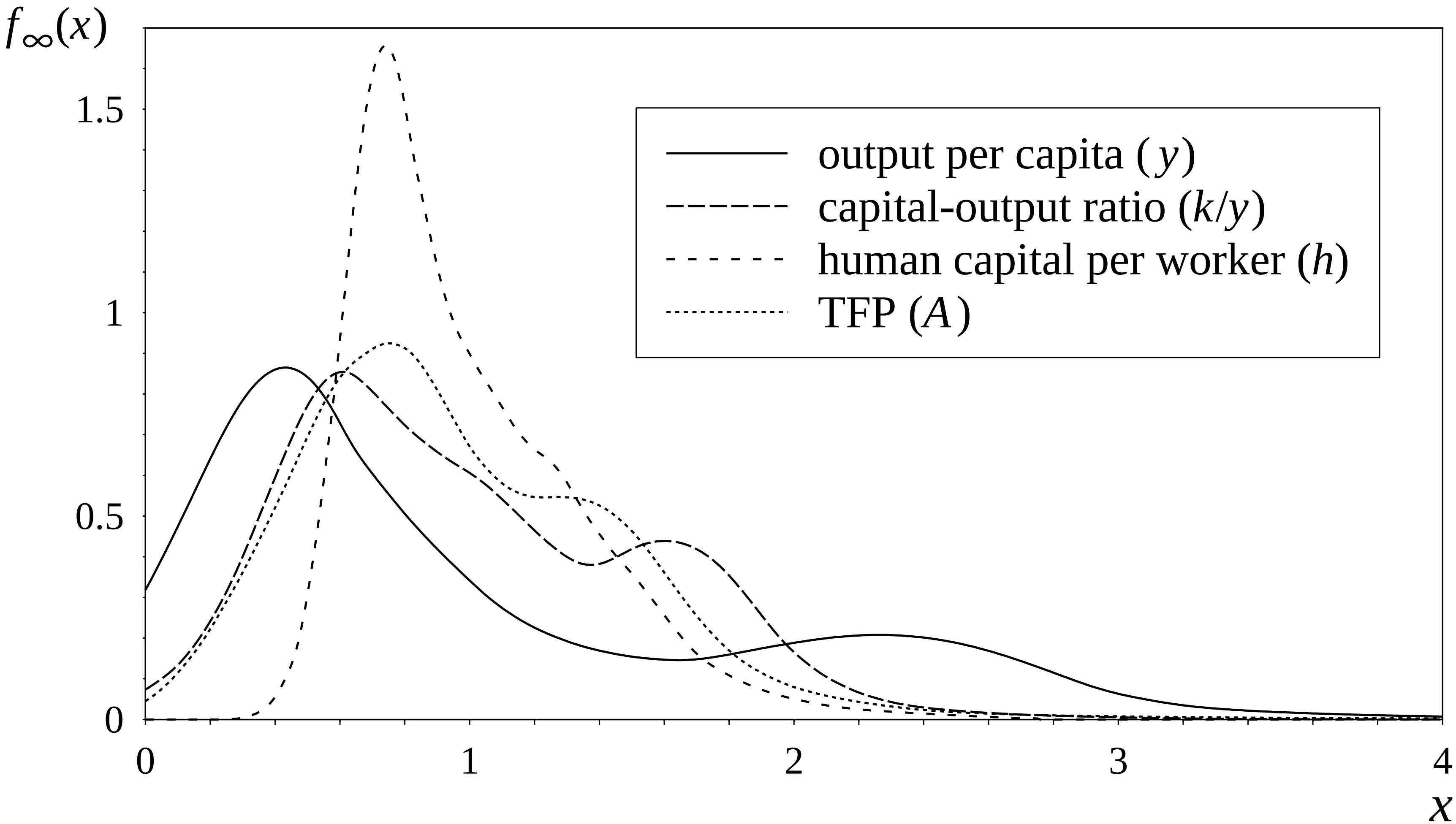
Quah, Danny, [2001], “Searching for Prosperity: A Comment,” *Carnegie-Rochester Conference Series on Public Policy*, 55:305-19.

Reichlin, Lucrezia., [1999], “Discussion of ‘Convergence as Distribution Dynamics’,” (by Danny Quah), *In Market Integration*, *Regionalism, and the Global Economy*, 328-35, Richard Baldwin, Daniel Cohen, Andre Sapir, and Anthony Venables, eds., Cambridge University Press, Cambridge.

Romer, Paul [1993], “Idea Gaps and Object Gaps in Economic Development,” *Journal of Monetary Economics*, 32:543-73.

Silverman, B.W., [1986], *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

Figure 1: Estimated Ergodic Densities



Appendix: Solving $f_\infty(z) = \int_a^b g_\tau(z|x) f_\infty(x)dx$

Assume that the solution exists and partition $[a, b]$ into n non-overlapping intervals $[s_{i-1}, s_i]$, $i = 1, 2, \ldots n$, such that $s_i = s_{i-1} + \frac{b-a}{n}$ with $s_0 = a$. Define z_j to be the midpoint of $[s_{j-1}, s_j]$. For any x, $g_\tau(z|x)$ is a probability density function implying $\int_a^b g_\tau(z|x)dz = 1$ so that we can write

$$\sum_{j=1}^{n} g_\tau(z_j|x)\frac{b-a}{n} \approx 1 \tag{1}$$

for any $x \in [a, b]$ where the approximation can be made arbitrarily accurate by taking n sufficiently large. Take $x = x_i$, the midpoint of $[s_{i-1}, s_i]$ and define $p_{ij} = g_\tau(z_j|x_i)\frac{b-a}{n} \geq 0$ for $j = 1, 2, \ldots n$. By virtue of (1) and the nonnegativity of the p_{ij}, we can, for any i, treat $\{p_{ij}\}_{j=1}^n$ as a (conditional) probability mass function. Define the matrix P by

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

and note that P has the same structure as the transition matrix of a Markov chain. We can use an argument similar to that used to motivate (1) to write $f_\infty(z) = \int_a^b g_\tau(z|x) f_\infty(x)dx$ as

$$f_\infty(z_j) \approx \sum_{i=1}^{n} g_\tau(z_j|x_i) f_\infty(x_i)\frac{b-a}{n} \tag{2}$$

and also to write

$$\sum_{j=1}^{n} f_\infty(z_j)\frac{b-a}{n} \approx 1.$$

Define $\phi_i = \frac{b-a}{n} f(x_i) = \frac{b-a}{n} f(z_i)$ for $i = 1, 2, \ldots n$ and write (2) as

$$\phi_j = \sum_{i=1}^{n} p_{ij}\phi_i. \tag{3}$$

By defining $\phi = (\phi_1, \phi_2, \ldots, \phi_n)'$, (3) is recognized as the expression for the product of ϕ' and the i^{th} column of P so that we have $\phi' = \phi' P$. As P has the same structure as the transition matrix of a Markov chain, we recognize ϕ to be the ergodic mass function associated with that chain. Given P, it is straightforward to find ϕ (if it exists) and then use $f(x_i) = \phi_i / \frac{b-a}{n}$ $i = 1, 2, \ldots n$ to get a vector of values of the ergodic density, $f_\infty(x)$, evaluated at a set of points $\{x_i\}_{i=1}^n$.