Club Convergence: Some Empirical Issues*

Carl-Johan Dalgaard
Institute of Economics
University of Copenhagen

Abstract

This note discusses issues related to testing for club-convergence. Specifically some empirical results from Durlauf and Johnson (1995) are presented and discussed. Finally, the note is also concerned with the problem of having a country specific rate of convergence, due to variation in e.g. growth rates of the labor force.

1 Theoretical preliminaries

Inspired by Azariadiz and Drazen (1990) we might assume that the level of productivity depends on the stage of development, as proxied by the capital stock in efficiency units, $\tilde{k}$:

$$B\left(\tilde{k}\right) = \begin{cases} \bar{B} \text{ if } \tilde{k} \geq \phi \\ \underline{B} \text{ if } \tilde{k} < \phi \end{cases},$$

where $\bar{B} > \underline{B}$ while ϕ denotes a critical level of development beyond which the economy becomes more productive. Assuming that the aggregate production function is on the form $Y = B\left(\tilde{k}\right) K^{\alpha} (AL)^{1-\alpha}$, where $A(t) = A(0)\, e^{xt}$ reflect exogenous technical progress, we can write output in efficiency units of labor in the following way:

$$\tilde{y} = B\left(\tilde{k}\right) \tilde{k}^{\alpha}$$

*Lecture Notes: Economic Growth, Spring 2005.

So essentially we are assuming that, for some reason or other, countries with more capital are more efficient in producing output.[1] Finally, we assume the world works as described by a standard Solow model (for brevity). This structure may lead to multiple steady states. Accordingly we assume that the B's and ϕ is chosen such that this is the case.[2]

2 Evidence

Approaching this model from an empirical angle, we start by noting that the steady state level of income per efficiency units of labor is given by

$$\tilde{y}^* = \begin{cases} \bar{B}^{\frac{1}{1-\alpha}} \left(\frac{s}{n+\delta+x}\right)^{\frac{\alpha}{1-\alpha}} & \text{if } \tilde{k}(0) \geq \phi \\ \underline{B}^{\frac{1}{1-\alpha}} \left(\frac{s}{n+\delta+x}\right)^{\frac{\alpha}{1-\alpha}} & \text{if } \tilde{k}(0) < \phi \end{cases}$$

In the vicinity of either steady state, we can perform a log-linearization of the dynamic system, in the usual way (cf. Lecture note #1), yielding:

$$\ln \tilde{y}(t) - \ln \tilde{y}(0) = \left(e^{-\lambda t} - 1\right) \ln \tilde{y}(0) + \left(1 - e^{-\lambda t}\right) \ln \tilde{y}^*,$$

or, using what we know about y^* and converting things into income per worker ($y(t) = A(t)\,\tilde{y}(t)$):

$$\begin{aligned} \ln y(t) - \ln y(0) &= xt + \left(1 - e^{-\lambda t}\right) \ln A(0) + \frac{1 - e^{-\lambda t}}{1-\alpha} \ln B\left(\tilde{k}\right) \\ &\quad - \left(1 - e^{-\lambda t}\right) \ln y(0) + \left(1 - e^{-\lambda t}\right) \frac{\alpha}{1-\alpha} \ln\left(\frac{s}{n+\delta+x}\right) \end{aligned}$$

Essentially then we have two regimes, where income per worker follows:

$$\ln y(t) - \ln y(0) = \begin{cases} \beta_0^H + \beta_1 \ln y(0) + \beta_2 \frac{\alpha}{1-\alpha} \ln\left(\frac{s}{n+\delta+x}\right) & \text{for } \tilde{k}(0) \geq \phi \\ \beta_0^L + \beta_1 \ln y(0) + \beta_2 \ln\left(\frac{s}{n+\delta+x}\right) & \text{for } \tilde{k}(0) < \phi \end{cases}.$$

[1]There are many ways to rationalize this sort of an assumption. See Azariadis (1996) for an overview. Moreover, we could add more heterogeneity, by assuming that the production function, i.e. α in the Cobb-Douglas case, differs across countries. Durlauf and Johnson considers this possibility. But the above is the simplest way to illustrate the theoretical framework, so we stick with that.

[2]Its a good exercise to convince yourself that it need not be the case that a Solow model – modified by the above production technology – lead to multiple steady states. One alternative is that of a unique steady state, but where the dynamics of the system are characterized by "stages of development" (see Durlauf and Johnson, 1995, Section 6.1. for a geometric characterization of "stages").

where

$$\beta_0^i = \begin{cases} xt + \left(1 - e^{-\lambda t}\right) \ln A(0) + \frac{1-e^{-\lambda t}}{1-\alpha} \ln \bar{B} \text{ for } i = H \\ xt + \left(1 - e^{-\lambda t}\right) \ln A(0) + \frac{1-e^{-\lambda t}}{1-\alpha} \ln \underline{B} \text{ for } i = L \end{cases}.$$

Theoretically, the top equation should hold for countries with sufficiently favorable initial conditions ($\tilde{k}(0) \geq \phi$), whereas the lower one relates to initially "poor" economies ($\tilde{k}(0) < \phi$). If the above represents the "true model of the world", the MRW estimation equation is misspecified: the same equation does not hold for all the countries in the sample. Or, to put it somewhat differently, we can argue that the MRW analysis suffers from an omitted variable problem: $B\left(\tilde{k}\right)$ is not taken care of in the econometric analysis. If $B\left(\tilde{k}\right)$ is ignored in the analysis the term $\left(\frac{1-e^{-\lambda t}}{1-\alpha}\right) \ln B\left(\tilde{k}\right)$ end up in the disturbances. And since B depends essentially on $\tilde{k}(0)$, we would expect the covariance between the disturbances and $\ln y(0)$ to be positive. Consequently, our estimate for β_1 will receive an upward bias, and the estimate for the rate of convergence becomes biased downward. How do we correct for this?

One possibility would be to break our data set down into parts. To be specific, believing in the theoretical structure above: Break it into two parts according to initial conditions. Assuming one is able to choose the groupings correctly (i.e. figuring out what ϕ approximately is), such that, in fact, within sub-groups β_0 is constant, then the estimate for β_1 should become more negative.

This is essentially the approach taken by Durlauf and Johnson (1995). The table below show the results from reestimating the MRW convergence equation on two sub-samples: One consisting of initially rich countries, with high levels of human capital, while the other consists of 42 countries that all were initially poor and featured low levels of human capital, as measured by the literacy rate.[3]

[3] They split their data both "exogenously" (the results from which we discuss here), and by using a mechanical methodology. See Durlauf and Johnson Section 5 for details on the latter.

Table II. Cross-section regressions: initial output and literacy-based sample breaks: dependent variable: $\ln(Y/L)_{i,1985} - \ln(Y/L)_{i,1960}$

	M–R–W	$(Y/L)_{i,1960} < 1950$ and $LR_{i,1960} < 54\%$	$1950 \leq (Y/L)_{i,1960}$ and $54\% \leq LR_{i,1960}$
Observations	98	42	42
		Unconstrained regressions	
Constant	3·04[a]	1·40	0·450
	(0·831)	(1·85)	(0·723)
$\ln(Y/L)_{i,1960}$	−0·289[a]	−0·444[a]	−0·434[a]
	(0·062)	(0·157)	(0·085)
$\ln(I/Y)_i$	0·524[a]	0·310[a]	0·689[a]
	(0·087)	(0·114)	(0·170)
$\ln(n+g+\delta)_i$	−0·505	−0·379	−0·545
	(0·288)	(0·468)	(0·283)
$\ln(SCHOOL)_i$	0·233[a]	0·209[a]	0·114
	(0·060)	(0·094)	(0·164)
$\bar{R}^2$	0·46	0·27	0·48
σ_ε	0·33	0·34	0·30
		Constrained regressions	
Θ	−2·56[a,b]	2·29	−0·395
	(1·14)	(1·17)	(1·24)
α	0·431[a]	0·275[a]	0·509[a]
	(0·061)	(0·097)	(0·098)
γ	0·241[a]	0·217[a]	0·108
	(0·046)	(0·061)	(0·094)
$\bar{R}^2$	0·42	0·28	0·50
σ_ε	0·34	0·34	0·29

[a] Significance at asymptotic 5% level.

Figure 1: Sample splitting and the "convergence equation". Source: Durlauf and Johnson, 1995.

Column 1 show the results from reestimating the convergence equation on the full sample of 98 countries. As can be seen, the coefficient for initial income rises when we move to the sub-sample results (in absolute value), implying faster convergence. This is consistent with our priors based on theoretical considerations of the issue at hand. In addition, its interesting to note that the coefficients on investments and human capital are different across country groupings; both being numerically smaller in the "bad initial conditions" sample. Durlauf and Johnson argue (p. 371) that "*these estimates suggest that the aggregate production function differs substantially across subgroups*". What they have in mind is that the estimate for I/Y reflects the rate of convergence, and $\alpha/(1-\alpha)$ (in the pure Solow case; in the MRW model it also depends on the human capital-output elasticity).

If the rate of convergence is constant within subgroups, this would imply that α must vary across subgroups to make sense of the result above.[4] But if α varies, this is the same as saying that the Cobb-Douglas functions are different in the two subgroups.

In sum these findings are consistent with the "club convergence" view of the world. That is, the estimation results suggest that the dynamic system governing the growth process differ across groups of countries, even when relevant structural characteristics are controlled for. If true, then common structural characteristics is not sufficient to ensure convergence.

This is not, however, the only feasible interpretation of their findings since what we basically are considering is the possibility that the intercept in the MRW regression varies across countries. This variation could be due to the presence of "$B\left(\tilde{k}\right)$". But another reason why the intercept could vary is that $\ln A(0)$ varies across countries. Panel studies like Islam (1995) and Caselli et al (1996) also find support for variation in "intercepts" (cf. Lecture note #5). But variation in $\ln A(0)$ is essentially harmless, in the sense that such variation in no way leads to multiple equilibria. And the problem is that we cannot tell these two interpretations apart: we cannot separately identify "$A(0)$" and "$B\left(\tilde{k}\right)$".[5] This does not mean that Durlauf and Johnson are wrong. Just that its not obvious that they are right.[6]

What should we make of the implied variation in the parameters of the production function? The sceptically inclined would argue that something

[4]The rate of convergence depends on α. So if α varies across groups, so should the rate of convergence. But within groups α is thought to be the same, and, as a result, so is the rate of convergence. Durlauf and Johnson's result do not really lead us to believe that the latter is significantly different across sub-samples however; the point estimates are essentially identical and certainly not significantly different from one another. We return to the issue of variable rates of convergence below in Section 3.

[5]Actually, a third interpretation is that x varies across countries in a systematic way, since this growth rate also enters into β_0. That doesn't exactly make life easier since x too is unobservable. Consequently we cannot even identify the sum $A+B$.

[6]There are, of course, other ways to obtain multiple equilibria rather than appealing to technology like above. For example, the rate of labor force (or population) growth may depend on income. If this relationship is sufficiently nonlinear, multiple equilibria arises (Buttrick, 1958). But in this case n is not exogenous, and the estimation strategy of both MRW and Durlauf and Johnson is flawed.

else could be driving these results: measurement error. Suppose investment data, and data on schooling are accurate in rich countries, but that the data is of very poor quality in poor countries. As we saw in Lecture note #3, this implies that the coefficients on investment, in the poor-country sample, could be biased towards zero.[7] Consequently, the results displayed above need not imply that the aggregate production function differs across countries.[8]

Finally, its worth noting that Durlauf and Johnson, like MRW, are not dealing with the problem of endogenous regressors. This too will bias their results. The direction and size of the bias is, however, impossible to assess without further data analysis.

Over-all, it seems fair to conclude that the Durlauf and Johnson study does not provide decisive evidence in favor of club convergence.

At the same time, it is important to realize that the structural equation derived under the hypothesis of club convergence above is consistent with the structural equation estimated in Islam (1995) and Caselli et al (1996) (which allow for country specific intercepts). Consequently, these studies cannot be taken as evidence *uniquely* in favor of "conditional convergence"! Indeed, the results are also consistent with the club convergence hypothesis. The door swings both ways.

There is a seperate issue, raised by Durlauf and Johnson, which is worth reflecting upon: Why would the rate of convergence be the same across

[7] Although not necessarily, as shown in the same set of lecture notes. While the case of "random" measurement error (typically referred to as "classical" measurement error) will always lead to the conclusion that the OLS estimator is biased toward zero, its also possible for the reverse to arise if measurement error is related to some underlying variable - say income. Hence, measurement error may in fact lead the OLS estimator to be biased upward. The fact that the coefficient of "school" is (numerically) larger in the "poor initial conditions" sample, can therefore also — potentially – be ascribed to measurement error.

[8] In all fairness it should noted that their regression results stemming from estimating the convergence equation on four sub-samples of varying sizes (determined mechanically) are more difficult to explain in this way (see Table V in the paper). In this exercise there are other findings that seem hard to understand. For example, the authors find no evidence of conditional-β-convergence among 21 rich countries like Denmark, Austria and the US ($\hat{\beta}_1$ actually comes out positive albeit insignificant). Similarly, it is difficult to understand why investments should be harmfull to growth in poor places (column 1 in Table V).

countries? As an example, in the Solow model

$$\lambda = (1-\alpha)(n+\delta+x).$$

Even if α, δ and x are the same across countries, as MRW assume, this still leaves us with the growth rate of the labor force, which is observably different from country to country. Consequently, the parameter for $\ln y(0)$ is not a constant. This may seem disturbing, as MRW are assuming parameter constancy in their analysis.

As it turns out, this is not nessesarily an invalidating problem for cross-section regressions, but requires us to think about the estimate for β_1 in a specific way – as the average value in the sample.[9] The next section conveys the basic intuition behind this result by appealing to a simple cross-section model. As noted at the end of Section 3, the result also hold for "dynamic equations", like the convergence equation.

3 Heterogenous parameters in a cross-section regression

Imagine we have the following cross-country model, i denoting the individual observations:

$$y = a + b^i x^i + \varepsilon^i$$

where ε is mean zero noise. Then it seems that we are violating a fundamental assumption (parameter constancy) if we estimate the equation by OLS. However, all that happens here is that we need to change our interpretation of our estimate for x^i. Note that our model can be rewritten to yield

$$y = a + \beta x^i + \eta^i$$

where

$$\eta^i = \left(b^i - \beta\right) x^i + \varepsilon^i,$$

[9] I was blissfully unaware of the econometric literature cited below until Henrik Hansen told me about it.

and β is the average value of b^i across countries, i.e. $E(b^i)$. Now, our OLS estimate for β is

$$\hat{\beta} = \beta + \frac{cov(x^i, \eta^i)}{var(x^i)}.$$

The covariance between x^i, η^i can also be stated as

$$cov\left(x^i, \eta^i\right) = E\left(x^i \eta^i\right) - E\left(x^i\right) E\left(\eta^i\right).$$

Assuming x^i is non-stochastic it follows that

$$E\left(x^i \eta^i\right) = x^i E\left(\eta^i\right) = 0,$$

where the last equality follows from:

$$\begin{aligned} E\left(\eta^i\right) &= E\left(\left(b^i - \beta\right) x^i + \varepsilon^i\right) = E\left(\left(b^i - \beta\right) x^i\right) + E\left(\varepsilon^i\right) \\ &= x^i E\left(b^i\right) - x^i \beta + 0 = 0. \end{aligned}$$

The bottom line is that since $E(x^i \eta^i) = E(x^i) E(\eta^i) = 0$, $cov(x^i, \eta^i) = 0$, and so the OLS estimate of β (again: the *average* value of b^i in the sample) is unbiased.

For this illustration I assumed that x^i is non-stochastic. This is not necessary in order to prove essentially the above result (but it sure makes it simpler to illustrate). This is demonstrated in Zellner (1969). Things get a bit more complicated if we have a dynamic equation (i.e. the lagged value of the left hand side variable entering as a regressor - like the convergence equation). But Pesaran and Smith (1995) have showed that the above result basically can be extended to this case as well, in the cross-section scenario. This means for example, that the estimate for the rate of convergence in, say, the MRW study, should be interpreted as the *average rate* of convergence in the sample.

References

[1] Azariadis,C., 1996. The Economics of Poverty Traps Part One: Complete Markets. Journal of Economic Growth, 1, p. 449-96

[2] Azariadis, C. and A. Drazen, 1990. Threshold Externalities in Economic Development. Quarterly Journal of Economics. 105, 501-26

[3] Buttrick, J., 1958. A Note on Professor Solow's Growth Model. Quarterly Journal of Economics, 72, pp. 633-636.

[4] Caselli, F, G. Esquivel and F. Lefort, 1996. Reopening the convergence debate: A new look at Cross-Country Growth Empirics. Journal of Economic Growth, 1, 363-89.

[5] Durlauf,-Steven-N; Johnson,-Paul-A, 1995. Multiple Regimes and Cross-Country Growth Behaviour. Journal of Applied Econometrics. 10, p. 365-84

[6] Islam, N., 1995. Growth Empirics: A Panel Data Approach. Quarterly Journal of Economics, 110, 1127-1170.

[7] Pesaran M.H and R. Smith, 1995. Estimating Long-run relationsships from dynamic heterogenous panels. Journal of Econometrics, 68, 79 - 113.

[8] Zellner, A., 1969. On the aggregation problem: A new approach to a troublesome problem. In M. Beckmann and H.P. Künzi (eds.) :" Economic Models, Estimation and Risk Programming – Essays in honor of Gerhard Tintner". Springer Verlag, Berlin.