CONSISTENT ESTIMATION OF GROWTH REGRESSIONS

Paul Evans*

Original Version: July 1994

Current Version: September 1997

Abstract

Ordinary least squares is unlikely to estimate growth regressions consistently. A simple alternative method is formulated. Applying ordinary least squares and this method to data on 85 countries over the period 1964-1990 reveals that ordinary least squares is strongly biased. The consistent estimates are about twice as large in magnitude as the inconsistent estimates. This finding suggests that many inferences reached in the empirical growth literature are invalid and may be seriously misleading. In particular, the inference that economies converge slowly to parallel balanced-growth paths appears to result from biased estimates.

JEL Code: O40

Keywords: Exogenous Growth, Endogenous Growth, Consistency

---

*Mailing Address: Department of Economics, Ohio State University, Columbus, Ohio 43210-1172. Telephone: 614-292-0072. Fax: 614-292-3906. Email: evans.21@osu.edu.

I. INTRODUCTION

The seminal articles of Romer (1986) and Lucas (1988) have sparked an enormous resurgence of interest in formulating growth theories and testing them empirically. Much of the recent empirical growth literature has evaluated growth theories by fitting cross-sectional regressions that relate the average growth rate of per capita output over some period for a sample of countries to the initial value of per capita output and country characteristics. Standard methods of inference are then applied to the estimated coefficients.[1] The method of estimation for these growth regressions has virtually always been ordinary least squares. This paper shows that ordinary least squares consistently estimates growth regressions only under highly implausible conditions and that plausible departures can produce large biases. The paper also formulates a simple alternative method that is consistent under much weaker conditions. Applying ordinary least squares and this alternative method to data on 85 countries over the period 1964-1990 reveals that ordinary least squares produces strongly biased estimates. The consistent estimates produced by the alternative method are about twice as large in magnitude as the inconsistent estimates produced by ordinary least squares. This result suggests that many of the inferences reached in the empirical growth literature are invalid and may be seriously misleading. In particular, inferences that economies converge slowly appear to result from biased estimates.

[1] A far from exhaustive list of such articles is Kormendi and Meguire (1985); Baumol (1986); De Long (1988); Barro (1991); De Long and Lawrence Summers (1991); Dollar (1991); Barro and Sala-i-Martin (1992); Mankiw, Romer, and Weil (1992); Levine and Renelt (1992); King and Levine (1993); De Gregorio (1993); Easterly (1993); Easterly and Rebelo (1993); Easterly, Kremer, Pritchett, and Summers (1993); Persson and Tabellini (1994); and Alesina and Rodrik (1994).

The rest of the paper is organized as follows. Section II evaluates the use of ordinary least squares to estimate growth regressions and formulates a simple alternative method for estimating them. Section III empirically documents that ordinary least squares produces strongly biased estimates. Section IV interprets the estimates obtained. Finally, section V draws a few conclusions.

II. ECONOMETRIC DISCUSSION

Let y_{nt} be the logarithm of per capita output for country *n* during year *t*. Growth regressions are obtained by fitting equations of the form

$$g_n = \alpha + \beta y_{n0} + \gamma' x_n + v_n \qquad (1)$$

to cross-sectional data for a sample of countries 1, ···, *N*. In equation (1), g_n is the average growth rate of per capita output for country *n* between years 0 and *T* [i.e., $(y_{nT} - y_{n0})/T$], y_{n0} is the initial logarithm of per capita output,[2] x_n is a vector of observations for country *n* on variables that control for cross-country heterogeneity, α and β are parameters, γ is a vector of parameters, and v_n is an error term with a zero mean and finite variance. If $\beta < 0$, countries that are initially rich after controlling for permanent differences associated with their *x*s and *v*s grow more slowly than countries that are initially poor on the same basis. The countries therefore converge toward parallel balanced growth paths whose heights depend in part on the *x*s. The more negative β is, the faster this convergence takes place. Because the parameter vector γ determines how the variables in x_n affect the height of country *n*'s balanced growth path, it can be used together with the magnitude of β in order to distinguish among exogenous growth theories.[3] If $b = 0$, γ determines how the variables in x_n affect country *n*'s trend growth rate. In this case, γ can be used to distinguish among endogenous growth theories.

By far the most common approach to evaluating growth theories in the literature has been to apply ordinary least squares to equation (1) and then to use standard methods to infer what β and γ are. Specifically, the *t*-ratio for the resulting estimator $\hat{b}$ is used to infer whether growth is exogenous or endogenous; and *t*-ratios on the entries of the resulting estimator $\hat{g}$ can be used to infer whether specific exogenous (endogenous) growth theories are invalid conditional on $b < 0\,(b = 0)$. Furthermore, confidence intervals for β and the entries of γ can be constructed and used to evaluate the quantitative predictions of growth theories. This procedure, however, yields correct inferences only if these estimators are consistent.

In order to determine whether they are indeed consistent, the data-generating process for y_{nt} must first be specified. I assume that

$$y_{nt} - a_t = d_n + l_n (y_{n,t-1} - a_{t-1}) + \sum_{i=0}^{q} q_{ni} e_{n,t-i} \tag{2}$$

with

$$d_n = k + x'_n x_n + w_n , \tag{3}$$

where e_{nt} is a zero-mean, covariance-stationary error term that is independently distributed both over time and across countries and uncorrelated with x_n. In equation (2),

[2]In lieu of y_{n0}, some studies use per capita output itself or per capita output normalized by the U.S. level. The point made here applies equally well to those studies.

[3] I use the term *exogenous growth theories* to refer to those in which no country can affect its trend growth rate without affecting every other country's trend growth rate equally. Thus, the term comprises not only

the autoregressive parameter l_n lies on (0,1] and the moving-average parameters $q_{n0}, \cdots, q_{nq}$ satisfy the restrictions that $\theta_{n0} = 1$ and the moduli of the roots of the polynomial $\det\left(\sum_i q_{ni} L^i\right)$ exceed $1/\, l_n$. As a result, $y_{nt} - a_t$ has an autoregressive representation with a dominant root of l_n, is covariance stationary if $l_n < 1$, and is difference stationary if $l_n = 1$. For notational convenience, the moving average in equation (2) is assumed to be of order $q < T$. If T is sufficiently large, nothing material is lost by doing so since any covariance-stationary process can be approximated arbitrarily well as a long finite moving average.

The term a_t is the common time-specific effect that every country experiences in year t. I assume that Δa_t is covariance stationary and independent of e_{nt}. If all of the λs are less than one, a_t is the common trend for all the *y*s and is thus the sole impetus for growth in all countries. In that case, growth is exogenous and countries follow parallel balanced-growth paths. By contrast, if $l_n = 1$, country n grows endogenously since y_{nt} diverges from a_t and the *y*s of all other countries. In this case, a_t enters equation (2) as the covariance-stationary shock Δa_t and serves merely to increase the plausibility of the assumption that e_{nt} is independently distributed across countries.[4]

Each country has its own intercept d_n in equation (2). As a result, the countries can have parallel rather than coincident balanced growth paths if all λs are less than one and can grow at different rates if any of the λs is one. The parameter d_n controls the

theories like those of Solow and Cass but also theories like that of Barro and Sala-i-Martin (1995, Ch. 8) in which all useful technical knowledge ultimately diffuses across countries.

[4] Even though a time-specific effect is included in the model, this assumption is problematical. Unfortunately, relaxing it is fraught with difficulties since no general method analogous to that of Newey and West (1987) for time series is available for cross sections.

relative height of country *n*'s balanced growth path if all of the λs are less than one and its relative growth rate if $l_n = 1$. According to equation (3), d_n is linearly related to the vector of country characteristics included in the growth regression (1) with an intercept parameter κ and a vector x_n of slope parameters. The error term w_n, which is the part of d_n that cannot be explained by x_n, is assumed to be uncorrelated with x_n and to have a finite cross-sectional variance.[5]

Solving equation (2) backward from year *T* to year 0, dividing both members of the resulting equation by *T*, substituting from equation (3), and rearranging produces

$$g_n = a_n + b_n y_{n0} + g'_n x_n - b_n w_n / (1 - g_n)$$

$$+ \frac{1}{T}\sum_{i=0}^{T-1} l_n^i \left(\sum_{j=0}^{\min[i,q]} l_n^{-j} q_{nj} \right) e_{n,T-i} + \left(\frac{l_n^T}{T} \right) \sum_{i=0}^{q-1} l_n^j \left(\sum_{j=i+1}^{q} l_n^{-j} q_{nj} \right) e_{n,-i} \; , \qquad (4)$$

where $b_n \equiv (l_n^T - 1)/T$, $g_n \equiv -b_n x_n / (1 - l_n)$, and $a_n \equiv (a_T - a_0)/T - b_n[a_0 + k/(1 - l_n)]$. In equation (4) and the definitions of γ_n and α_n, $-b_n/(1 - l_n)$ should be interpreted as equaling its limit of one if $l_n = 1$. Note that $b_n < 0$ if country *n* grows exogenously ($l_n < 1$) and $b_n = 0$ if it grows endogenously ($l_n = 1$).

In order to investigate the consistency of ordinary least squares in estimating equation (1), I first consider the special case in which every intercept d_n is completely explained by the country characteristics included in x_n ($w_n = 0 \; \forall_n$) and every series $y_{nt} - a_t$ is a first-order autoregression ($q = 0$). In this case, equation (4) reduces to

$$g_n = a_n + b_n y_{n0} + g'_n x_n + \tfrac{1}{T}\sum_{i=0}^{T-1} l_n^i e_{n,T-i} \; . \qquad (5)$$

[5] Assuming that w_n is uncorrelated with x_n involves no essential loss in generality at this point. This assumption is relaxed later when it becomes important.

The estimator $\hat{b}$ can be obtained by regressing y_{n0} on an intercept and x_n to obtain the residual r_n and then regressing g_n on r_n. Because each term in $\frac{1}{T}\sum_{i=0}^{T-1} l_n^i e_{n,T-i}$ is uncorrelated with the intercept, y_{n0}, x_n, and hence with the residual r_n, one has

$$\underset{N\to\infty}{\text{plim}}\,\hat{b} = \left(\underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} a_n r_n + \underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} b_n r_n y_n + \underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} g_n' r_n x_n\right) \Big/ \left(\underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right). \quad (6)$$

On the further assumption that a_n is uncorrelated with r_n, b_n is uncorrelated with $r_n y_n$, and γ_n is uncorrelated with $r_n x_n$, equation (6) reduces to

$$\underset{N\to\infty}{\text{plim}}\,\hat{b} = \left(\underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} b_n r_n^2\right) \Big/ \left(\underset{N\to\infty}{\text{plim}}\,\tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right) \quad (7)$$

since the residual r_n is orthogonal to the regressors in its regression and to the fitted part of y_{n0} by construction.

According to equation (7), $\hat{b}$ converges in probability to a weighted average of the βs. Therefore, unless the possibility that growth is endogenous in some countries and exogenous in others can be ruled out *a priori*, obtaining a statistically significant negative estimate for β in equation (1) does not indicate that growth is exogenous in all countries. One can infer only that growth is exogenous in a positive fraction of all countries.[6]

Ideally, $\hat{b}$ should be useful not only in testing whether growth is exogenous but also in estimating how fast countries converge toward their balanced-growth paths on average. A natural candidate for estimating this rate is $\hat{c} \equiv 1-(1+T\hat{b})^{1/T}$ since the convergence rate for each country n is given by

$$c_n = 1-(1+Tb_n)^{1/T}. \quad (8)$$

[6] Durlauf and Johnson (1995) make a related point.

Jensen's inequality, however, implies that unless the βs are identical for all countries,

$$\operatorname*{plim}_{N\to\infty} \hat{c} < \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N} \sum_{n=1}^{N} c_n r_n^2 \right) / \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N} \sum_{n=1}^{N} r_n^2 \right) \tag{9}$$

since the function defined in equation (8) is strictly convex.

This bias may be substantial. For example, using the median unbiased estimates that I obtained for 48 countries in my 1997 article, the right-hand member of the inequality (9) is .0781 and the left-hand member is only .0137.[7] There is no way to avoid this bias short of entirely abandoning the use of growth regressions. Indeed, this is exactly what I have advocated in a series of papers.[8] Here, I sweep this problem under the rug, assuming henceforth that the λs and thus the βs are identical across all countries.

A similar analysis can be used to show that the entries of $\hat{g}$ converge in probability to weighted averages of the entries of $-bx_n/(1-l)$, $n = 1, 2, \cdots,$ where β and λ are the common values of $b_1, b_2, \cdots$ and $l_1, l_2, \cdots$, respectively. Unfortunately, the weights differ across the entries. As a result, unless the ξs are identical across all countries, no straightforward interpretation attaches to the estimator $\hat{x} \equiv -(1-\hat{l})\hat{g}/\hat{b}$ with $\hat{l} \equiv (1+\hat{b}T)^{1/T}$. I therefore assume that the ξs are indeed identical.

I now relax the restrictions that every intercept d_n is completely explained by x_n and that every series $y_{nt} - a_t$ is a first-order autoregression but impose the additional

[7] The calculation assumes equal weights for the countries and constrains the estimates to be nonnegative.

[8] Evans and Karras (1996b) use panel methods that allow for cross-country heterogeneity in intercepts. Evans and Karras (1996a) and Evans (1996b) use panel methods that allow for some cross-country heterogeneity in dynamics as well as in the intercepts. Evans (1996a) and Evans (1997) allow for still more cross-country heterogeneity. These papers obtain empirical results similar to those of this paper. Islam (1995) and Caselli, Esquivel, and Lefort (1996) also use panel methods to allow for cross-country heterogeneity in intercepts with similar results. The contribution of this paper relative to the others is that it spells out exactly why many of the results in the empirical growth literature are suspect, quantifies the size

restrictions that the λs and ξs and thus the βs and γs are identical across all countries. Equation (4) then takes the form

$$g_n = a + by_{n0} + gx_n - bw_n / (1- g)$$

$$+\frac{1}{T}\sum_{i=0}^{T-1} l^i \left(\sum_{j=0}^{\min[i,q]} l^{-j} q_{nj} \right) e_{n,T-i} + \left(\frac{l^T}{T} \right) \sum_{i=0}^{q-1} l^i \left(\sum_{j=i+1}^{q} l^{-j} q_{nj} \right) e_{n,-i} \,, \tag{10}$$

where $b \equiv (l^T - 1)/T$, $g \equiv -bx/(1- l)$, and $a \equiv (a_T - a_0)/T - b[a_0 + k/(1- l)]$.

Analysis similar to that leading to equation (7) yields

$$\operatorname*{plim}_{N\to\infty} \hat{b} = b + (\Phi + \Psi) \Big/ \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N} \sum_{n=1}^{N} r_n^2 \right) \tag{11}$$

with

$$\Phi \equiv \frac{l^T}{T} \operatorname*{plim}_{N\to\infty} \tfrac{1}{N} \sum_{n=1}^{N} \left[\sum_{i=0}^{q-1} l^i \left(\sum_{j=i+1}^{q} l^{-j} q_{n,j+i+1} \right) r_n e_{n,-i} \right] \tag{12}$$

and

$$\Psi \equiv -\frac{b}{1- l} \operatorname*{plim}_{N\to\infty} \tfrac{1}{N} \sum_{n=1}^{N} r_n w_n \,. \tag{13}$$

Recall that r_n is the residual obtained from regressing y_{n0} on the intercept and x_n.

Equation (11) implies that $\operatorname{plim}_{N\to\infty} \hat{b}$ differs from β if either $q > 0$ or the cross-sectional variance of w_n is positive. Whether the departure of $\operatorname{plim}_{N\to\infty} \hat{b}$ from β should be regarded as *inconsistency* depends on one's purpose.[9] If one wishes only to test whether growth is exogenous, it creates little problem. The reason is that if the data-generating process began operation sufficiently far in the past, $\operatorname{plim}_{N\to\infty} \hat{b}$ is negative if

of the problem, and shows how to overcome the problem if one is willing to entertain the assumption of identical slope parameters.

[9] I owe this point to the referee.

$\lambda < 1$ and is negligibly different from zero if $\lambda = 1$; see Appendix A for proof. As a result, $\hat{b}$ can be used to infer whether economies are exogenous even if $q > 0$ and the cross-sectional variance of w_n is positive.[10] Similarly, $\hat{g}$ can be used to infer whether ξ differs from zero.

By contrast, if one wishes to construct confidence intervals for the asymptotic convergence rate 1-λ or the asymptotic half life –ln2/ln(1-λ), then one should indeed regard $\hat{b}$ as inconsistent.[11] Furthermore, if one wishes to construct confidence intervals for the entries of ξ, one requires consistent estimators for both β and λ since $x = -(1- l)\, g/ b$. Because pinning down the magnitudes of β and the entries of γ enables one to distinguish among growth theories, this paper regards $\hat{b}$ and $\hat{g}$ as inconsistent unless they converge in probability to β and γ.

The bias resulting from $q > 0$ is likely to be negligible in practice. For example, suppose that the cross-sectional variance of w_n is positive and $q = 1$ with $\theta_{11} = \theta_{21} = \cdots \equiv \theta$. In that case, Appendix B shows that

$$\operatorname*{plim}_{N\to\infty} \hat{b} = b + \frac{q\, l^{T-1}(1- l^2)}{T(1+2\, lq + q^2)} \tag{14}$$

if $\lambda < 1$ and the data-generating process (2) has always operated. For $T = 26$, equation (14) implies the following values for $\operatorname{plim}_{N\to\infty} \hat{b}$:[12]

---

[10] Of course, using the t-ratio of $\hat{b}$ to test $\beta = 0$ against $\beta < 0$ may result in tests with little power if y_{nt}-a_t differs appreciably from a first-order autoregression or if the cross-sectional variance of w_n is appreciable.

[11] The *asymptotic convergence rate* is the annual rate at which convergence would be predicted at a forecast horizon infinitely far in the future. For the data-generating process (2), it is obtained at all forecasts horizons of q or more years. The *asymptotic half life* is defined similarly.

[12] The bias is likely to be much smaller than the figures below suggest since θ is estimated to be .088 (.020) if λ = .98, .112 (.021) if λ = .94, and .161 (.021) if λ = .90.

λ	β	θ = -.4	θ = +.4
.98	-.0157	-.0160	-.0156
.94	-.0308	-.0311	-.0307
.90	-.0360	-.0362	-.0359

By contrast, the bias resulting from a positive cross-sectional variance for w_n can be substantial. Appendix C shows that

$$\operatorname*{plim}_{N\to\infty} \hat{b} = \left[\frac{\mathrm{var}(y|x,w)}{\mathrm{var}(y|x)}\right] b \tag{15}$$

and

$$\operatorname*{plim}_{N\to\infty} \hat{g} = \left[\frac{\mathrm{var}(y|x,w)}{\mathrm{var}(y|x)}\right] g \tag{16}$$

if $\lambda < 1$, $q = 0$, and the data-generating process (2) has always operated. The bracketed term in equations (15) and (16) is the ratio of the cross-sectional variance of y_{n0} conditional on both x_n and w_n to the cross-sectional variance of y_{n0} conditional on only x_n. Thus, unless the xs can control for a large fraction of the permanent cross-country variation in the ys, $\hat{b}$ and $\hat{g}$ can be appreciably biased toward zero.

Obtaining consistent estimates of β and γ is straightforward even if $q > 0$ and the cross-sectional variance of w_n is positive. The component w_n of the error term in equation (10) can be eliminated simply by differencing both of its members. The result is

$$\Delta g_n = w + b\Delta y_{n0} + u_n , \tag{17}$$

where Δ is the difference operator [i.e., $\Delta y_{n0} \equiv y_{n0} - y_{n,-1}$], $w \equiv \Delta a = (\Delta a_T - \Delta a_0)/T -$ $b\Delta a_0$, and

$$u_n \equiv \frac{1}{T}\sum_{i=0}^{T-1} l^i \left(\sum_{j=0}^{\min[i,q]} l^{-j} q_{nj} \right) \Delta e_{n,T-i} + \left(\frac{l^T}{T} \right) \sum_{i=0}^{q-1} l^i \left(\sum_{j=i+1}^{q} l^{-j} q_{nj} \right) \Delta e_{n,-i} \; . \qquad (18)$$

The parameter β then can be estimated consistently using instrumental variables correlated with Δy_{n0} but uncorrelated with $e_{nT}, \cdots, e_{n,-q}$. Obvious candidates are the intercept and $y_{n,-q-1}, y_{n,-q-2}, \cdots$. Unfortunately, if λ is not appreciably less than one, $y_{n,-q-1}, y_{n,-q-2}, \cdots$ may be only weakly correlated with Δy_{n0}. For this reason, it may be desirable to supplement them with additional instrumental variables.[13]

Let $\tilde{b}$ be the estimator obtained from the procedure described above. The parameter vector γ can be estimated consistently by using ordinary least squares to regress $g_n - \tilde{b} y_{n0}$ on an intercept and x_n. The resulting estimators for α and γ are consistent because $g_n - \tilde{b} y_{n0}$ converges in probability to g_n-βy_{n0} and w_n and $\varepsilon_{nT}, \cdots, \varepsilon_{n,-q}$ are uncorrelated with the intercept and x_n; see equation (10).[14]

III. AN EMPIRICAL EXAMPLE

Although the analysis of the previous section suggests that ordinary least squares can produce strongly biased estimates of growth regressions, it does not establish that the bias is in fact large for the growth regressions reported in the literature. In this section, I employ ordinary least squares to fit a growth regression similar to many that have

[13] I performed a small-scale Monte-Carlo evaluation of the properties of the instrumental-variables estimator described in the text for the case $q = 0$. I found that the estimator is essentially unbiased and produced reasonably powerful tests even for λs as large as .98 when the instrumental variables are the intercept and $y_{n,-q-1}$, $y_{n,-q-2}$, $y_{n,-q-3}$, $y_{n,-q-4}$. The inclusion of additional instruments may improve the properties of the estimator for $\lambda < 1$ and should enable it to retain decent properties even for $\lambda = 1$.

appeared in the literature and then use the consistent method formulated in the previous section to show that ordinary least squares does indeed produce much different estimates. My starting point is the growth regression reported in the first column of Table V in Mankiw, Romer, and Weil (MRW). According to this regression, the growth in real GDP per worker for 98 countries over the period 1960-1985 is highly significantly related to the initial logarithm of real GDP per worker (y_{n0}), the logarithm of the average ratio of real gross domestic investment to real GDP over the period ($\ln \bar{i}_n$), the logarithm of a measure of schooling ($\ln s_n$), and the logarithm of .05 plus the average growth rate of the labor force over the period [$\ln(.05+\bar{\ell}_n)$]. I fitted a similar regression to data over the period 1964-1990 for a sample of 85 countries consisting of every country used by MRW for which the requisite data are available.[15] The specification differs from theirs only in making the dependent variable the average growth rate (g_n) rather than total growth ($y_{nT} - y_{n0}$). I define years 0 and T to be 1964 and 1990 in order to allow moving-average orders as large as three years in the instrumental-variable estimates reported below. The Penn World Tables of Summers and Heston (1991) as updated in 1995 provide the data on real GDP per worker, $\bar{i}_n$, and $\bar{\ell}_n$; the appendix to MRW's paper provides the data on s_n. I obtained

---

[14] If w_n is correlated with x_n, instrumental-variables estimation should be used in the second step. Although the simultaneity bias resulting from the endogeneity of x_n may be important in practice, I do not investigate it here. See Cho (1997) for an analysis of simultaneity bias in growth regressions.

[15] The countries are Algeria, Argentina, Australia, Austria, Bangladesh, Belgium, Benin, Bolivia, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic, Chad, Chile, Colombia, Congo, Costa Rica, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Germany, Ghana, Greece, Guatemala, Honduras, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Ivory Coast, Jamaica, Japan, Jordan, Kenya, Korea, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Singapore, Spain, South Africa, Sri Lanka, Sweden, Switzerland, Syria, Togo, Trinidad and Tobago, Thailand, Tunisia, Turkey, Uganda, United States, United Kingdom, Uruguay, Venezuela, Zambia, and Zimbabwe.

$$\hat{g}_n = \underset{(.0278)}{.0502} - \underset{(.0022)}{.0125}y_{n0} + \underset{(.0026)}{.0112}\ln(\bar{i}_n) + \underset{(.0028)}{.0110}\ln s_n - \underset{(.100)}{.277}\ln(.05+\bar{\ell}_n) \ , \quad (19)$$

where the figure in parentheses under each estimate is its heteroskedasticity-consistent standard error. Clearly, all four variables included in the regression are highly significant. If the estimates were consistent, it would be appropriate to reject endogenous growth theories on the basis of the coefficient on y_{n0} and to reject those exogenous growth theories in which investment in physical and human capital and labor force growth are unimportant determinants of how high balanced growth paths are.

According to the analysis of the previous section, however, these estimates are unlikely to be consistent. In order to investigate how important the resulting biases in equation (19) are, I fitted equation (17) for q = 0, 1, 2, and 3 using the intercept, $y_{n,-q-1}, \cdots, y_{n,-4}$, and $i_{n,-q-1}, \cdots, i_{n,-4}$ as instrumental variables. (The variable i_{nt} is the investment rate of country n in year t.[16]) Table 1 reports the resulting estimates of β and their standard errors. The estimates are negative, statistically significant at any reasonable level, and over twice as large in magnitude as those reported in the growth regression (19). Because the estimates of β are not sensitive to the instrumental variables employed, serial correlation appears to be unimportant as expected. Table 1 therefore provides strong evidence that ordinary least squares produces seriously biased estimates.

Regressing $g_n - \tilde{b}y_{n0}$ with $\tilde{b} = -.0349$ on the intercept and x_n yields

$$\tilde{g}_n + \underset{(.0050)}{.0349}y_{n0} = \underset{(.0384)}{.1106} + \underset{(.0034)}{.0165}\ln(\bar{i}_n) + \underset{(.0029)}{.0249}\ln s_n - \underset{(.133)}{.746}\ln(.05+\bar{\ell}_n) \ . \quad (20)$$

[16] I also experimented with using lagged labor-force growth rates and enrollment rates in 1960 as instrumental variables. They have little, if any, explanatory power for Δy_{n0} beyond that provided by $y_{n,-q-1}, \cdots, y_{n,-4}$ and $i_{n,-q-1}, \cdots, i_{n,-4}$.

The parameter estimates in equation (20) are about twice as large in magnitude as those in equation (19). Therefore, the latter estimates are not only inconsistent but also seriously biased relative to the consistent estimates of equation (20).

Table 1. Instrumental-Variables Estimates of b and their Standard Errors

q	*Estimate of b*	*Standard Error*
0	-.0349	.0050
1	-.0340	.0055
2	-.0326	.0064
3	-.0338	.0066

IV. INTERPRETATION

This section compares the invalid inferences implied by the growth regression (19) with the valid inferences implied by equation (20). My purpose is to show that the inferences are qualitatively similar but quantitatively very different.

The estimates of β in the growth regressions (19) and (20) are statistically significant at any reasonable level. One would therefore reject the null hypothesis $\beta = 0$ in favor of the alternative hypothesis $\beta < 0$ using either estimate. Under the maintained assumption that the countries in the sample have identical growth dynamics, one can then infer that they converge toward parallel balanced growth paths and thus have identical trend growth rates. For this reason, both estimates lead to the resounding rejection of those growth theories that predict cross-country variation in trend growth rates.

The estimated coefficients on $\ln \bar{i}_n$ and $\ln s_n$ in the growth regressions (19) and (20) are significantly positive at any reasonable level, and that on $\ln(.05+\bar{\ell}_n)$ is significantly negative at any reasonable level. Both sets of estimated coefficients therefore indicate that the height of a country's balanced growth path is increasing in the rates at which it invests in physical and human capital and decreasing in the growth rate of its labor force.

The quantitative implications of equations (19) and (20) differ in three important ways. First, they imply very different asymptotic convergence rates. Substituting the estimates of β from equations (19) and (20) into the formula

$$c = 1-(1+Tb)^{1/T} \tag{21}$$

yields point estimates of .0162 and .0874 for the asymptotic convergence rate. The .95 confidence intervals implied by the point estimates of β and their standard errors are (.0098,.0239) and (.0401,1), respectively.[17] The consistent point estimate is more than five times larger than the inconsistent estimate. Moreover, the two .95 confidence intervals do not even overlap. The consistent convergence rate is near that predicted by David Cass's exogenous growth model if the elasticity of output with respect to reproducible factors of production is near the share of output paid to physical capital in developed countries.[18] This inference contrasts sharply with those that have been drawn in the literature from the low convergence rates produced by ordinary least squares. For

[17] The calculation of these confidence intervals involves two steps. First, the end points of each confidence interval for β are obtained by adding to each estimate ±1.96 times its standard error. Second, these end points are plugged into equation (21). If the bottom end of β's confidence interval is less than -1/*T*, the top end of *c*'s confidence interval is equated to one. Note that for β near –1/*T*, the function relating *c* to β becomes highly convex, implying that large convergence rates are not readily distinguishable. This is similar to my 1997 finding that the interval from the median to the top of each country's .90 confidence band is much longer than the interval from the median to the bottom.

example, Barro and Sala-i-Martin and Barro, Mankiw, and Sala-i-Martin (1995) infer that the elasticity of output with respect to reproducible factors must be about .8.

Second, the large convergence rate estimated here implies that even if the distribution of output per worker was far from its stationary distribution in 1964, it should have been quite close by 1990. For example, if "accidents of history" resulted in 90 percent of the cross-country variance of output per worker in 1964, they would have accounted for only 7 percent by 1990.[19] As a result, most observed cross-country differences at least in recent years are permanent.

Third, the estimated impact effects of the observed country characteristics are substantially larger in the consistent equation (20) than in the inconsistent equation (19). Therefore, a country that somehow permanently alters its characteristics can obtain much larger initial effects as well as a much faster transition to its new balanced-growth path. For example, equation (20) implies that doubling a country's investment rates in both physical and human capital can be expected to raise its output per worker by $26\times(.0165+.0249)\times\ln 2$, or 75 percent, over the next 26 years. The analogous figure is only 41 percent for equation (19). Thus, reform may be more palatable than the inconsistent estimates would suggest since more of the benefits are received early on.

---

[18] According to King and Rebelo (1993), the convergence rate is about 13 percent a year if the elasticity of output with respect to reproducible factors is 1/3 and the intertemporal elasticity of substitution is 1.

[19] Let the initial variance be 1, comprised of .9 resulting from accidents of history and .1 from the stationary distribution. In the 26 years between 1964 and 1990, the former is predicted to decline to $.9\times(1-.0874)^{2\times 26}$, or .00774. The new variance is then .10774. The figure in the text is .00774/.10774, rounded to the nearest percent. In fact, my 1996a paper shows that the cross-country variance of output per worker rose over the period 1964-1990 for 54 countries, suggesting that the accidents of history made a negative contribution in 1960.

V. CONCLUSIONS

This paper has established a theoretical presumption that ordinary least squares inconsistently estimates growth regressions. The resulting biases in estimated coefficients are large unless the econometrician can control for a large fraction of permanent cross-country differences in per capita output. The paper also shows how to estimate growth regressions consistently. Applying ordinary least squares and the consistent method to data for 85 countries over the period 1964-1990, the paper shows that the two methods produce much different estimates. Specifically, the consistent estimates produced by the alternative method are about twice the inconsistent estimates produced by ordinary least squares.

In the empirical growth literature, many authors have used growth regressions estimated with ordinary least squares to make inferences about the validity of competing growth theories. An important implication of this paper is that many of their inferences are likely to be seriously misleading.

APPENDIX A

Suppose that the data-generating process (2) began operation in year $-q\text{-}M$ for some nonnegative integer M. $1+T\hat{b}$ can be estimated by regressing R_n on r_n, where R_n and r_n are the residuals obtained by first regressing y_{nT} and y_{n0} on intercepts and x_n. The triangle inequality implies that

$$1+T\left(\operatorname*{plim}_{N\to\infty}\hat{b}\right)=\left(\operatorname*{plim}_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}r_nR_n\right)/\left(\operatorname*{plim}_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}r_n^2\right)$$

$$< \sqrt{\left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right)\left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} R_n^2\right)} \Big/ \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right)$$

$$< \sqrt{\left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} R_n^2\right) \Big/ \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right)} \qquad \text{(A.1)}$$

since R_n and r_n are not proportional (equal) to each other except by happenstance. If $\lambda < 1$, the right-hand member of the inequality approaches one as M approaches infinity since then $\operatorname{plim}_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} R_n^2 = \operatorname{plim}_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} r_n^2$. In that case, rearranging the inequality (A.1) yields $\operatorname{plim}_{N\to\infty} \hat{b} < 1$.[20]

Now, consider the case $\lambda = 1$. Solving equation (2) backward from year 0 to year $-q$-M and substituting from equation (2) yields

$$y_{n0} = a_0 + (q+M)(k + x'x_n + w_n) + \sum_{i=0}^{q-1}\left(\sum_{j=0}^{i} q_{nj}\right) e_{n,-i}$$

$$+ \left(\sum_{j=0}^{q} q_{nj}\right)\sum_{i=0}^{M} l^i e_{n,-i-q} + (y_{n,-q-M} - a_{-q-M})\,. \qquad \text{(A.2)}$$

Hence, the residual r_n is asymptotically equivalent to

$$\sum_{i=0}^{q-1}\left(\sum_{j=0}^{i} q_{nj}\right) e_{n,-i} + \left(\sum_{j=0}^{q} l^j q_{nj}\right)\sum_{i=0}^{M} e_{n,-i-q} + r_n^* + (q+M) w_n\,, \qquad \text{(A.3)}$$

where r_n^* is the residual obtained by regressing $y_{n,-q-M}$-a_{-q-M} on the intercept and x_n. Substituting the expression (A.3) into equations (11)-(13) in place of r_n yields

[20] This proof is adapted from Quah (1997).

$$\operatorname*{plim}_{N\to\infty}\hat{b}-b=\left\{\tfrac{1}{T}\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}\left[\sum_{i=0}^{q-1}\left(\sum_{j=i+1}^{q}q_{n,j+i+1}\right)\left(\sum_{j=0}^{i}q_{nj}\right)\right]s_{n}^{2}-\frac{b}{1-l}(q+M)\left(\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}w_{n}^{2}\right)\right\}/$$

$$\left[\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}\left(\sum_{j=0}^{i}q_{nj}\right)^{2}s_{n}^{2}+M\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}\left(\sum_{j=1}^{q}q_{nj}\right)^{2}s_{n}^{2}+\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}r_{n}^{*2}+(q+M)^{2}\operatorname*{plim}_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}w_{n}^{2}\right]$$

where s_n^2 is the variance of ε_{nt}. The right-hand member of this equation is negligible if M is sufficiently large.

APPENDIX B

Suppose that $\lambda < 1$, $q = 1$, and $\theta_{11} = \theta_{21} = \cdots = \theta$. Solving the equation (2) backward from period zero to period $-M-1$, substituting from equation (3) with $w_n = 0$, and letting M approach infinity yields

$$y_{n0}=a_{0}+(k+x'x_{n})/(1-l)+e_{n0}+(l+q)\sum_{i=0}^{\infty}l^{i}e_{n,-i-1}\,. \tag{A.4}$$

Hence, the residual r_n is asymptotically equivalent to

$$e_{n0}+(l+q)\sum_{i=0}^{\infty}l^{i}e_{n,-i-1}\,. \tag{A.5}$$

Substituting the expression (A.5) in place of r_n in equations (11)-(13) with $w_n = 0$ gives

$$\operatorname*{plim}_{N\to\infty}\hat{b}=b+\left(\frac{q\,l^{T-1}}{T}\right)\left(\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}s_{n}^{2}\right)/\left[1+\frac{(l+q)^{2}}{1-l^{2}}\right]\left(\lim_{N\to\infty}\tfrac{1}{N}\sum_{n=1}^{N}s_{n}^{2}\right) \tag{A.6}$$

from which equation (14) follows immediately.

APPENDIX C

Suppose that λ < 1 and q = 0. Solving the equation (2) backward from period zero to period –M, substituting from equation (3), and letting M approach infinity yields

$$y_{n0} = a_0 + (k + x' x_n + w_n) / (1 - l) + \sum_{i=0}^{\infty} l^i e_{n,-i} \ . \tag{A.7}$$

Hence, the residual r_n is asymptotically equivalent to

$$w_n / (1 - l) + \sum_{i=0}^{\infty} l^i e_{n,-i} \ . \tag{A.8}$$

Substituting the expression (A.8) in place of r_n in equation (11) results in

$$\operatorname*{plim}_{N\to\infty} \hat{b} = b\left[\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N}\{r_n^2 - w_n^2 / (1 - l)^2\}\right] / \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right). \tag{A.9}$$

Equation (A.7) and expression (A.8) imply that the denominator in equation (A.9) is $\operatorname{cov}(y|x)$ and the numerator is $b\operatorname{cov}(y|x, w)$.

The quantity $g_n - \hat{b} y_{n0}$ is asymptotically equivalent to

$$a + g' x_n + \frac{b}{(1 - l)^2}\left[\left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} w_n^2\right) / \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right)\right] y_{n0} + \tfrac{1}{T}\sum_{i=0}^{T-1} l^i e_{n,T-i} \ . \tag{A.10}$$

Equation (A.7) implies that regressing the quantity (A.10) on the intercept and x_n yields an estimator for γ with the probability limit

$$g + \frac{b}{(1 - l)^2}\left[\left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} w_n^2\right) / \left(\operatorname*{plim}_{N\to\infty} \tfrac{1}{N}\sum_{n=1}^{N} r_n^2\right)\right] \frac{x}{1 - l} \tag{A.11}$$

which should also equal $\operatorname{plim}_{N\to\infty} \hat{g}$. The expression (A.11) equals the right-hand member of equation (16) because βξ/(1-λ) = -γ.

REFERENCES

Alesina, Alberto, and Rodrik, Dani. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics* 109 (May 1994): 465-490.

Barro, Robert J. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* 106 (May 1991): 407-443.

Barro, Robert J.; Mankiw, N. Gregory; and Sala-i-Martin, Xavier. "Capital Mobility in Neoclassical Models of Growth." *American Economic Review* 85 (March 1995): 103-115.

Barro, Robert J., and Sala-i-Martin, Xavier. "Convergence." *Journal of Political Economy* 100 (April 1992): 223-251.

Barro, Robert J., and Sala-i-Martin, Xavier. *Economic Growth*. New York: McGraw-Hill, 1995.

Baumol, William J. "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show." *American Economic Review* 76 (December 1986): 1072-1085.

Campbell, John Y. "Inspecting the Mechanism: An Analytical Approach to the Stochastic Growth Model." *Journal of Monetary Economics* 33 (June 1994): 463-508.

Caselli, Francesco; Esquivel, Gerardo; and Lefort, Fernando. "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics." *Journal of Economic Growth* 1 (September 1996): 363-390.

Cass, David M. "Optimal Growth in an Aggregate Model of Capital Accumulation." *Review of Economic Studies* 32 (July 1965): 233-240.

Cho, Dongchul. " An Alternative Interpretation of Conditional Convergence Results." Manuscript, Korea Development Institute, 1997. Forthcoming in the *Journal of Money, Credit and Banking*.

De Long, Bradford. "Productivity Growth, Convergence, and Welfare. Comment." *American Economic Review* 78 (July 1965): 233-240.

De Long, Bradford, and Summers, Lawrence H. "Equipment Investment and Economic Growth." *Quarterly Journal of Economics* 106 (May 1991): 407-444.

De Gregorio, Jose. "Inflation, Taxation, and Long-Run Growth." *Journal of Monetary Economics* 31 (June 1993): 271-298.

Dollar, David. "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985." *Economic Development and Cultural Change* 40 (April 1991): 523-544.

Durlauf, Steven N., and Johnson, Paul A. "Multiple Regimes and Cross-Country Growth Behaviour." *Journal of Applied Econometrics* 10 (October-December, 1995): 365-384.

Easterly, William. "How Much Do Distortions Affect Growth?" *Journal of Monetary Economics* 32 (November 1993): 187-212.

Easterly, William, and Rebelo, Sergio. "Fiscal Policy and Growth: An Empirical Investigation." *Journal of Monetary Economics* 32 (December 1993): 417-458.

Easterly, William; Kremer, Michael; Pritchett, Lant; and Summers, Lawrence H. "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks." *Journal of Monetary Economics* 32 (December 1993): 459-483.

Evans, Paul. "Using Cross-Country Variances to Evaluate Growth Theories." *Journal of Economic Dynamics and Control* 20 (1996a): 1027-1049.

Evans, Paul. "Using Panel Date to Evaluate Growth Theories." Manuscript, Ohio State University, December 1996b. Forthcoming in *International Economic Review.*

Evans, Paul. "How Fast Do Economies Converge?" *Review of Economics and Statistics* 36 (May 1997): 219-225.

Evans, Paul, and Karras, George. "Convergence Revisited," *Journal of Monetary Economics* 37 (April 1996a): 249-266.

Evans, Paul, and Karras, George. "Do Economies Converge? Evidence from a Panel of U.S. States." *Review of Economics and Statistics* 78 (August 1996b): 384-388.

Islam, N. "Growth Dynamics: A Panel Data Approach." *Quarterly Journal of Economics* 110 (November 1995): 1127-1170.

King, Robert G., and Levine, Ross. "Finance and Growth: Schumpeter Might Be Right." *Quarterly Journal of Economics* 108 (August 1993): 717-738.

King, Robert G., and Rebelo, Sergio. "Transitional Dynamics and Economic Growth in the Neoclassical Model." *American Economic Review* 83 (September 1993): 908-931.

Kormendi, Roger, and Meguire, Philip. "Macroeconomic Determinants of Growth: Cross-Country Evidence." *Journal of Monetary Economics* 16 (September 1985): 141-163.

Levine, Ross, and Renelt, David. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82 (September 1992): 942-963.

Lucas, Robert E., Jr. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July 1988): 3-42.

Mankiw, N. Gregory; Romer, David; and Weil, David N. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107 (May 1992): 407-438.

Newey, Whitney K., and West, Kenneth D. "A Simple-Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (May 1987): 708-708.

Persson, Torsten, and Tabellini, Guido. "Is Inequality Harmful?" *American Economic R*eview 84 (June 1994): 600-621.

Quah, Danny. "Galton's Fallacy and Tests of the Convergence Hypothesis." *Scandinavian Journal of Economics* 95 (4, 1993): 427-443.

Romer, Paul M. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94 (October 1986): 1002-1037.

Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70 (February 1956): 65-94.

Summers, Robert, and Heston, Alan. "The Penn World Table (Mark 5): An Expanded Set of 106 International Comparisons, 1950-1988." *Quarterly Journal of Economic* (May 1991): 327-368.