NBER WORKING PAPER SERIES

EXTERNALITIES AND GROWTH ACCOUNTING

Jess Benhabib

Boyan Jovanovic

Working Paper No. 3190

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1989

This paper is part of NBER's research program in Growth. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #3190
December 1989

EXTERNALITIES AND GROWTH ACCOUNTING

ABSTRACT

We reexamine several bodies of data on the growth of output, labor, and capital, within the context of a model that admits the possibility of an externality to the capital input. The model is an augmented version of Paul Romer's (1987) reformulation of the Solow model. Unlike Romer, however, we find no evidence of an externality to capital. This finding implies nothing about the size of possible spillovers in the creation of knowledge because in our model, causality runs exclusively from knowledge to capital.

Jess Benhabib
New York University
Department of Economics
269 Mercer St, 7th Floor
New York, NY 10003

Boyan Jovanovic
New York University
Department of Economics
269 Mercer St, 7th Floor
New York, NY 10003

We reexamine several bodies of data on the growth of output, labor, and capital, within the context of a model that admits the possibility of an externality to the capital input. The model is an augmented version of Paul Romer's (1987) reformulation of the Solow model. Unlike Romer, however, we find no evidence of an externality to capital. This finding implies nothing about the size of possible spillovers in the creation of knowledge because in our model, causality runs exclusively from knowledge to capital.

There seems, on first consideration, to be little reason to expect a firm's investment in capital to have substantial beneficial spillover effects in reducing production costs of other firms. But if firms with more capital also have more productive knowledge, and if this knowledge spreads to other firms, then unless one can somehow measure knowledge and control for it, an increase in the capital stock of one firm will appear to lower the production costs of other firms. In the same vein, if the economy's capital stock in positively rela ed to the availability of specialized intermediate inputs and if these inputs are not measured, the growth of capital will appear to increase aggregate output by more than its private marginal product.

In a pioneering article, Romer (1987) argues that a large positive externality in capital formation is needed to explain the strong positive association in aggregate data (over countries and over epochs) between the "Solow residual" and the growth of the capital stock. Moreover, the size of his externality estimate is staggering: the social marginal product of capital, suggests Romer, is perhaps twice or even three times its private marginal product, and, by implication, the equilibrium level of investment falls far short of its socially optimal level.[1]

Romer's argument implicitly assumes that growth in capital <u>causes</u> a growth in knowledge or a growth in the availability specialized inputs, or both.[2] It is convenient to begin our inquiry by assuming that this assumption is correct. While hard evidence in the external benefits of specialization seems hard to come by, there is, in contrast, a huge body of evidence on the magnitude of spillovers in the process of creation of knowledge. So let us assume that an increase in a firm's capital stock <u>causes</u> the firm's productive knowledge to go up in the same proportion. Under this assumption, one can use estimates from micro data on externalities in R&D as an estimate of the size of the capital externality.

No firm conclusions on this question emerge from the micro data, however, perhaps because of the severe problems of measurement; Griliches (1979) discusses some of these problems. Griliches and Lichtenberg (1982) find only tenuous support for the conjecture that there are beneficial R&D spillovers across industries: Process R&D performed in the industry in question is far more significant than R&D embodied in the products of other industries. On the other hand, Jaffe (1986) identifies a firm's "technological neighbors" and finds that while neighbors' R&D <u>lowers</u> the firm's profits and market value, it does tend to raise the firm's patents per R&D dollar and sales; he interprets the later finding as arising from the presence of spillovers. A handful of studies, however, have found evidence of large spillovers. Mansfield et al. (1977) find very large spillovers in a dozen or so selected innovations, and Bernstein and Nadiri (1988) find that in four industries, the social returns to <u>intra</u>industry spillovers of R&D are very high, ranging from 30% to 123% of the private returns to R&D.[3] Such large estimates of the knowledge externality are, however, an exception: In a summary of the literature on the elasticity of output with respect to the R&D input, Griliches (1988) reports that "while the presence of spillovers would make one expect the industry-level coefficients to be higher than those estimated at the firm level, the econometric estimates do not show this in any convincing fashion."[4] So, if aggregating up to the industry-level makes little difference to the estimates of the R&D coefficient, it would be quite surprising if aggregating to the whole economy would produce a large upward revision (specifically, a tripling) of the R&D coefficient. Yet this is exactly what Romer's (1987) argument implies, and the micro data do not seem to support it.

Our model is a variant of that described by Prescott (1986). In contrast to Romer, causality runs entirely from knowledge to capital. Knowledge evolves exogenously; we do not estimate its external effects, and, indeed, under our assumption about causality, micro evidence in spillovers of knowledge says <u>nothing</u> about spillovers to the capital input. The popular view that some capital investment is needed for the implementation of new ideas favors our causality assumption, since it is natural (as in Shleifer (1986), for instance) to imagine new ideas as proceeding the installation of the capital equipment needed to implement them. Moreover, at the level of the individual firm at least, the data indicate that R&D Granger-causes investment, but that investment does not Granger-cause R&D (Lach and Schankerman (1988)).

While it reverses the assumption about causality between capital and knowledge, our model still admits the possibility of an externality to the capital input, and is in fact almost the same as Romer's. Our conclusions, however, are quite different: we examine a variety of bodies of data, and find no evidence to support the hypothesis that there are beneficial spillovers arising from the capital input. The reason why our conclusions differ from Romer's is roughly this: Romer faces simultaneity problems when he estimates a production function in which capital and labor are endogenous and correlated with the disturbance to the production function. One source of disturbances to the production function is the business cycle, and Romer tries to remove it by filtering out the high frequencies with long run averages. He further recognizes that even in the long-run data, low-frequency movements in technology might create a correlation between the inputs and the production function disturbance, but he argues[5] that the extent of this correlation could not plausibly be so large as to reverse his conclusions. This, however, is where we disagree with his argument. We make explicit assumptions about the way in which the capital and labor inputs evolve in response to changes in the state of technology. These assumptions enable us to calculate the correlation between the inputs and the disturbance. We find that even in the long run data this correlation is plausibly high enough to explain the high empirical elasticity of output with respect to the capital input. No externalities are needed.

The next section presents our model which consists of five structural equations. In Section 2, we present maximum likelihood and least squares estimates for the model using post-war quarterly and annual U.S. data and find no evidence of an externality. In Section 3 we then discuss some of the model's implications about the convergence of GNP among different countries, and interpret the apparent empirical validity of "Gibrat's Law" in the behavior of countries' GNP series over extended periods. In Section 4, we reinterpret Romer's regression results (that use data on growth of inputs and output over long epochs) in terms of the simultaneity biases that we calculate, and we conclude that even those data offer no evidence for the conjectured positive externality to the capital input.

After the empirical evidence discussed in Sections 2, 3, and 4, Section 5 presents two models that give rise to the structural equations first introduced in Section 1. The first is a stochastic Diamond-type of overlapping generations model, the second a Brock-Mirman type of infinite-horizon model. The sixth and final section offers some concluding remarks; some of the computations are included in three appendices.

1. The Augmented Solow Model

In Romer's model, the representative firm produces output Y_t with hired inputs K_t and L_t, taking as given the economy-wide capital stock $\bar{K}_t$ per firm, and the state of knowledge Z_t. The production function is:

$$Y_t = K_t^{\alpha} L_t^{1-\alpha} \bar{K}_t^{\theta} Z_t. \tag{1}$$

The parameter θ measures the external effect of capital, an effect that the firm ignores when making its decisions. Since all firms are the same, $K_t = \bar{K}_t$. Letting lower-case symbols denote logarithms, (1) reads

$$(2) \qquad y_t = (\alpha+\theta)k_t + (1-\alpha)\ell_t + z_t \ .$$

If the firm is a price-taker in its product and factor markets, α is capital's share in output, and $1-\alpha$ is labor's share. This is Romer's reformulation of Solow's model.[6]

To this, we now add assumptions about how knowledge grows, and about how the equilibrium k_t and ℓ_t evolve. Knowledge evolves exogenously, as follows:

$$(3) \qquad z_{t+1} = \mu + \rho z_t + \omega_t, \quad |\rho| \leq 1$$

and

$$(4) \qquad \omega_t = \epsilon_t + \lambda_1\epsilon_{t-1} + \lambda_2\epsilon_{t-2}.$$

Thus the z_t process is an ARMA(1,2). Three features of (3) and (4) deserve mention. First, μ is the rate of exogenous technical change; one expects it to be positive since some knowledge comes for free from abroad, and in addition, some knowledge is generated for free <u>domestically</u> as a byproduct of everyday economic activity. Second, the parameter ρ, or rather $1-\rho$, measures the rate at which knowledge depreciates.[7] From the work on business cycles by Prescott and his coworkers, we expect ρ to be about .99 in the quarterly U.S. data, and .96 in the annual data. Third, the MA(2) specification for ω_t was entirely arbitrary -- it took two moving average terms to remove the autocorrelation from the residuals in the quarterly U.S. data.

Next we specify the behavior of the capital and labor inputs. In Section 6 we shall present two separate micro models[8] that imply the following two equations as governing competitive equilibrium allocations:

$$k_{t+1} = \gamma + y_t, \tag{5}$$

where γ is a constant,[9] and

(6) (ℓ_t) a stochastic process independent of (z_t).

Equations (2)-(6) make up the model. In sum, it is Romer's model with the added assumptions of exogenous knowledge and endogenous capital and labor. The next three sections describe how we have estimated its parameters. Section 2 uses post-war US data, Section 3 uses the Summers-Heston data, and Section 4 uses the longer-run data that Romer compiled from Maddison and elsewhere. None of these bodies of data supports the hypothesis that θ is positive.

2. Estimates From Post-War U.S. Data.

We begin our empirical inquiry by looking at the post-war US data. The reader will not be surprised, however, to learn that the post-war US data offer no support for a positive θ, since (a) Romer himself did not cite these data as supportive of capital externalities, and (b) Prescott (1986) has, with some success, used a model quite similar to ours but with θ set equal to zero to fit de-trended post-war US data.

A problem presented by the capital input is that it is likely to be poorly measured. Our estimation procedure in this subsection begins by treating k_t as an unobservable. This assumption underlies the calculation of the estimates in the first four tables. Tables 5-8, on the other hand, do use capital data.

Substitution of (5) into (2) yields

$$(7) \qquad y_t = (\alpha+\theta)\gamma + (\alpha+\theta)y_{t-1} + (1-\alpha)\ell_t + z_t.$$

This is the equation that formed the basis for the estimation, which used annual data on logGNP for y_t and loghours worked for ℓ_t. The data are not detrended.

We present two sets of estimates. Table 1 reports the unconstrained ML estimates,[10] for the annual and the quarterly data separately. Table 2 reports estimates for the remaining parameters when α is constrained to equal one third, i.e., capital's post-war share in income.

Several points are noteworthy. For the quarterly data, the unconstrained estimates are virtually the same as the constrained estimates, and the likelihood ratio does not significantly differ from one. The estimate of ρ is about the same as Prescott's ((1986), p.15). The estimate of θ is close to zero, and does not significantly differ from zero. When α is freed up, its estimate does not significantly differ from 1/3. All in all, then, the model does pretty well with the quarterly data.

Such is not the case with the annual data. The unconstrained estimated of α is negative, but not significantly different from zero. And, θ is positive, but quite small and insignificant. Thus, the social marginal product of capital appears to be low in these data. This is especially clear in the first panel of Table 2 in which, when α is constrained to 1/3, θ is large but negative, and highly significant. The likelihood ratio test resoundingly rejects the restriction that $\alpha = 1/3$. These results with annual data are quite similar to the regression results that Romer gets in line 2 of his Table 2; in this regression he allows, as we do, for exogenous technical change, and measures, as we do, the labor input by hours worked.

An important source of downward bias on θ deserves mention, however. While our specification (4) does allow for transitory components in y, there may nevertheless be measurement error in y that will cause the coefficient of lagged y in (7), namely $(\alpha+\theta)$, to be underestimated because of errors in variables bias. In Table 2 this will cause θ to be underestimated, while in Table 1, where α is freed, both α and θ may be underestimated. Put differently, measurement error in y will lead to a spurious negative dependence between the quasi-first differences in footnote 10. Such a negative bias could hide a positive θ.

A second set of problems arises because it takes time to build capital, and time for the external benefits of capital accumulation to be felt. That is, it is possible that not only are there significant building-time delays, but externalities affect output with a lag. To test for the presence of such delays, we considered a production function $Y_t = K_{t-p}^{\alpha} L_t^{1-\alpha} \overline{K}_{t-s} Z_t$ where p and s represented lags. Presumably $0 \leq p \leq s$. We derived the corresponding reduced form (the analogue of (7)) for the infinite horizon, representative agent model where y on the right hand side appeared with lags s and p. We estimated this model using quarterly data for various values of s and p and found that the best fits, in terms of likelihood, were for $s = p = 0$. For all values that we checked for $s > 0$ (up to $s = 12$) and $p = 0$, the externality coefficient θ was zero and for values $s = p > 0$, it tended to be negative.

A third set of problems surround our assumption in eq. (5) about the way in which capital evolves. Distinct from the issues (discussed in the previous paragraph) concerning the length of time that it takes to build capital, there is the issue of how long capital remains productive after it has been built; that is, how fast it depreciates. If y_t is measured by GNP, as we have done, then eq. (5) implies that there is 100% depreciation. On the other hand, if Y_t is interpreted as wealth, then we have not measured wealth correctly -- instead we ought to have used net national product plus the entire existing capital stock. But then it is not clear that (1) is the correct production function, and empirical implementation demands an accurate K_t series. For these reasons we did not pursue this route. Instead we took the following alternative. In place of eq. (5) (for which a micro-based justification exists -- see section 3), we posited the ad-hoc Solow-type constant savings rule out of income, along with the conventional assumption that capital depreciates at a constant rate δ. This leads to the following equation for the growth of capital:

$$(5)' \qquad K_{t+1} = sY_t + (1-\delta)K_t$$

Tables 3 and 4 report, respectively, the unconstrained and constrained (by $\alpha = 1/3$) estimates of the parameters when (5)′ is used in place of (5). Only annual data are used since, at the depreciation rates that are commonly used, we did not have a long enough quarterly time series.[11] The first set of estimates in Tables 3 and 4 sets δ at 10%, the second sets it at 8%. The parameter C is the same as before (see note 11) with $\gamma = \ln s$. Further details are in Appendix 3.

All of the estimates imply a significantly negative marginal social product of capital, at magnitudes that are simply incredible. Evidently, (5)' lends even less support than does (5) to the idea that there are positive externalities to the capital input, or to the notion that in the aggregate production function returns to scale are increasing. The first four tables relied on (5) or (5)' to eliminate the capital input from the production function. The next four tables present estimates that use the capital series directly. In Table 5, the low estimate of α is probably due to the high short-run elasticity of output with respect to labor, which is higher than labor's share $1-\alpha$. Given that α is set at zero, θ becomes the output elasticity with respect to capital. Thus Table 5 cannot really be interpreted as supporting the hypothesis that θ is positive. The estimates in Table 6 are more favorable to the hypothesis; α is now constrained to a third, yet θ is still positive and significant. This is the only solid piece of evidence in favor of Romer's hypothesis that we can find in the post-war data. At the same time, the estimate of ρ is surprisingly low. Tables 7 and 8 present OLS estimates of the aggregate production function. If ρ is close to one, and if (5) is appropriate for annual (as opposed to quarterly) data, then according to the model in equations (2)-(6), the OLS estimates of the coefficients in the growth-rates equation (Table 8) are unbiased. On the other hand, the levels equation involves upward bias on the capital coefficient and downward bias on the labor coefficient.

Given the wide diversity of the estimates for θ, α and ρ reported by the eight tables in this section, it seems that the assumptions that we have added to Romer's model do not appear to have substantially improved the ability of his model to rationalize high frequency data. We therefore agree with Romer's view (1987, p. 186) that data from more countries and larger epochs should provide additional and perhaps better information on the model's parameters, and in particular, on θ. We look at cross-country data next.

3. Cross-Country Evidence on the Univariate Representation for y_t.

Under certain assumptions, the Summers-Heston panel data on countries' GNP's will provide additional information on the parameters of the model. This is what we examine next. We assume that all countries have the same production functions and tastes and that the only difference among them is in their initial values for k_t, ℓ_t and z_t. Because we shall be looking only at the y_t process,[12] and because ℓ_t (in addition to k_t and z_t) will also be treated as an unobservable, some further assumptions will now be added. First, we shall assume that $\lambda_1 = \lambda_2 = 0$, so that $\omega_t = \epsilon_t$. This is done for analytical convenience, and it ought not to make much quantitative difference in this subsection because it looks at growth-rates over a period of 25 years, and not at annual or quarterly growth rates, as was done in the previous subsection, so that the two-year moving average induced by the λ's should not matter much, if at all. Second, we shall assume a particular stochastic process for the ℓ_t sequence; in each country ℓ_t is assumed to follow the stochastic process

$$\ell_t = m + r\ell_{t-1} + w_t, \quad |r| \leq 1,$$

where w_t is iid, and independent of ϵ_t. We estimated this equation using U.S. annual data and OLS, and obtained:

$$\ell_t = \underset{(.18)}{-.22} + \underset{(.02)}{1.03}\ell_{t-1} \qquad R^2 = .98 \quad DW = 1.85$$

Our ML results together with this one suggest that at least in the U.S., both ρ and r are quite close to unity. We shall then take the bold step of assuming that this is true in all the countries in the Summers-Heston sample.[13]

Although, even under these additional assumptions, a study of the y_t process on its own will not identify θ, it will nevertheless rule out a great many possible values that the pair of crucial parameters (ρ, θ) can assume. One source of information about the behavior of y_t in 115 countries comes from the Summers-Heston sample (which is now updated to 1985). The regression below represents the relationship between the average 1960-1985 rate in GNP-growth of a country on the one hand, and its 1960 GNP on the other.[14] That is, the growth of countries is regressed on their initial size. The regression results reveal no significant relation between the two:

$$\Delta y_i = \underset{(.015)}{.047} - \underset{(.001)}{.0004}y_i \qquad i = 1,\ldots,115. \quad \text{Residual variance} = .0004 \tag{8}$$

where y_i is the logarithm of 1960 GNP for country i, and Δy_i its growth per year over the 1960-85 period. Standard errors are in parentheses. Thus the updated sample roughly confirms the insignificant relation between growth and initial size, that Romer and others have found,[15] a (non!) relation that is in other contexts often referred to as "Gibrat's Law."

To find out what the seeming absence of a relation between size and growth means for our structural parameters, combine equations (2) and (5) to get

$$y_t = (\alpha+\theta)(\gamma+y_{t-1}) + \eta_t, \tag{9}$$

where $\eta_t = (1-\alpha)\ell_t + z_t$. Repeated substitution for lagged y's leads to the following predicted relation between growth and initial size:

$$y_{t+T} - y_t = [(\alpha+\theta)^T-1]y_t + \sum_{j=0}^{T-1}(\alpha+\theta)^j[(\alpha+\theta)\gamma + \eta_{t+T-j}]. \tag{10}$$

Equation (10) cannot, without further work, be used to interpret the regression results reported in eq. (8), because y_t will be correlated with the disturbance in eq. (10). One can see this by assuming that $r = \rho$, so that $\eta_t = (1-\alpha)m + \mu + \rho\eta_{t-1} + \epsilon_t + (1-\alpha)w_t$, and by recursively substituting for lagged η's in eq. (7) to obtain

$$\eta_{t+T-j} = \rho^{T-j}\eta_t + (T-j)[\mu + (1-\alpha)m] + \sum_{s=0}^{T-1-j} \rho^j v_{t+T-j-s}, \tag{11}$$

where $v_t = \epsilon_t + (1-\alpha)w_t$. As long as $\rho > 0$, innovations in η tend to persist and η_t and y_t will, for each country, be positively correlated. Substituting from (11) into (10) then implies that the least-squares estimate of b in the regression $\Delta y_i = a + by_i$ is <u>identically</u>:

$$(12) \qquad \hat{b} = (\alpha+\theta)^T - 1 + [\mathrm{Cov}_i(\eta_{it}, y_{it})/\mathrm{Var}_i(y_{it})] \sum_{j=0}^{T-1} (\alpha+\theta)^j \rho^{T-j}.$$

The subscript i on Cov_i and Var_i is there to emphasize that it is i that varies while t is held fixed at t = 1960.

To compute the expected value of $\hat{b}$, we invoke our assumption that the parameters of the y_t process are identical for all countries, in which case the empirical bivariate distribution of (y_{it}, η_{it}) over countries i at t approximates the stationary distribution of (y_t, η_t) for a <u>given</u> country when this distribution exists. When either $\rho \to 1$, or $(\alpha+\theta) \to 1$, this distribution blows up,[16] but Appendix 1 shows that the ratio $\mathrm{Cov}(y,\eta)/\mathrm{Var}(y)$ still converges:

$$(13) \quad \lim_{r,\rho\to 1} \mathrm{Cov}(\eta_t, y_t)/\mathrm{Var}(y_t) = 1 - (\alpha+\theta).$$

Substituting from (13) into (12) and noting (once more from Appendix 1) that $\lim_{\rho\to 1} \sum_{j=0}^{T-1} (\alpha+\theta)^j \rho^{T-j} = (1 - (\alpha-\theta)^T)/(1-(\alpha+\theta))$, yields

$$(14) \quad \lim_{\rho\to 1} E(\hat{b}) = (\alpha+\theta)^T - 1 + [1 - (\alpha+\theta)][1 - (\alpha+\theta)^T]/(1-(\alpha+\theta)) = 0.$$

So, if ρ and r are roughly one, Gibrat's Law will hold <u>regardless of the value of θ</u>. This means that the failure of GNP-levels to converge does not identify θ.

And now, a caveat: Our analysis treats countries as closed economies and looks for scale effects or spillover effects <u>within</u> but <u>not across</u> countries. Yet, geographical borders are in some respects an arbitrary division of geographical space and are therefore "noisy" measures of market areas within which, according to our analysis (and Romer's), these scale or spillover effects are assumed to be confined. Nevertheless, in some instances, differences in culture and language, and the presence of trade barriers do support the use of geographical borders to delineate the extent of the market. Of course, if significant cross-country spillovers in knowledge do exist, they surely run mainly from the rich nations toward the poor ones, and if so, they represent a force in support of convergence.[17]

A second set of questions emerges from our assumption that the bivariate distribution of (y, η) among countries at a point in time is the same as the stationary distribution of (y,η) for a given country over time. The first point to note here is that the truth of this hypothesis is completely independent of the length of the epochs (T_{it}); it instead has to do with how long the stochastic processes y_{it} have followed the law of motion (9), and with the speed of convergence to the stationary distribution implied by the parameters $(\alpha+\theta)$, ρ and r.

The assumption that the cross-section distribution coincides with the stationary distribution can also be tested. Let g_{it} be the growth-rate of GNP in country i between periods t and t+1. If the hypothesis is true, the distribution of g_{it} $(t = 1960, \ldots, 1984)$ should, for each fixed i, be roughly the same as the distribution of g_{it} $(i = 1, \ldots, 115)$ for each fixed t. In particular, if σ_i^2 and σ_t^2 be the variances of the two respective distributions, we should have $\sigma_i^2 \approx \sigma_t^2$, at least for most i and most t. In fact, σ_i^2 and σ_t^2 both vary considerably as i and t vary, although <u>on average</u> they are roughly the same:

$$(1/115) \sum_{i=1}^{115} \sigma_i^2 = .0035, \quad \text{and} \quad (1/25) \sum_{t=1960}^{1984} \sigma_t^2 = .0034.$$

The variability of σ_t^2 is documented in Table 9, while the variability of σ_i^2 is described in the histogram in Figure 1 (in which Iraq, whose $\sigma_i^2 \times 1000 = 30$, was omitted).[18] A comparison of Table 9 and Figure 1 reveals that the variability of σ_i^2 is greater than that of σ_t^2, which is what one would expect (if in truth they were equal) on sampling grounds since (a) σ_t^2 averages the variance of g_{it} over 115 countries while σ_i^2 averages it over 25 years only, and (b) the observations g_{it} are, for fixed i, autocorrelated.

While the variability of growth rates <u>among</u> countries does not seem to differ from the growth-rates' variability <u>within</u> countries, one might still wonder if differences in growth rates among countries are too <u>persistent</u> to be consistent with our model. The model asserts that except for initial conditions, the η_t process is the same over countries. One way to pose the question about persistence is to ask about the cross-country variance of the mean-growth rates over the 25-year periods. That is, does the model allow for a reasonable chance that some countries will grow much faster than others over a period as long as 25 years, or is this possibility a remote one?

Since $\Delta k_t = \Delta y_{t-1}$, the steady-state variance of Δy coincides with that of Δk, and this expression is provided in Table A.1. If we hypothesize the truth of the Solow neoclassical model and insert $T = 25$, $\theta = 0$ and $\alpha = 1/3$ in this expression, it reads $a_{kk} = 24\sigma_w^2 + (48.5)\sigma_\epsilon^2$. Now this expression, when divided by $(25)^2$, should be equated to the <u>empirical</u> value of the cross-country variance of the 25-year averaged growth rates. This turns out to be equal to slightly less than the variance of the residual in equation (8), namely .000355. Thus, setting $(25)^{-2}a_{kk}$ equal to this number yields a linear restriction on σ_w^2 and σ_ϵ^2. If we take $\sigma_w^2 = \sigma_\epsilon^2$, this yields $\sigma_\epsilon^2 = .003$. If, on the other hand, we take the least favorable case (for us) where $\sigma_w^2 = 0$, we come up with $\sigma_\epsilon^2 = .0045$.

A striking feature of both of these estimates of σ_ϵ^2 (.003 or .0045) that are obtained from looking at the distribution over countries of 25-year rates of growth, is that they are much larger than the estimates we obtained from the US post-war annual time series, which were equal to .00023 and .00015 for the constrained and unconstrained, respectively. Recall, however, that in this subsection we had assumed that $\lambda_1 = \lambda_2 = 0$; had we assumed this in the estimates reported in Tables 1 and 2, the resulting estimate of σ_ϵ^2 would have been higher. The relevant comparison is with $\sigma_w^2 = [1 + \lambda_1^2 + \lambda_2^2]\sigma_\epsilon^2$, the estimates of which are .00034 and .00017 for the constrained and unconstrained cases, respectively.

Even after this adjustment, the estimate of σ_ϵ^2 from the cross-country data is at least 10 times as large as the US estimate. This should not be altogether surprising since σ_ϵ^2 is identified essentially from the variability over time in the US growth rate, which is far smaller than that of the median country.[19] Nevertheless, the cross-country variability in 25-year growth rates may be up to twice as large as one would have expected if the cross-section distribution coincided with the stationary distribution for a given country. To account for this discrepancy, one or more of the parameters that we have assumed to be the same for all countries, might have to be made country-specific.[20]

While the above discussion leaves some unanswered questions about our model's ability to explain (a) the lack of convergence of GNP <u>levels</u>, and (b) the existence of of persisting differentials in growth <u>rates</u>, we should in all fairness point out that an alternative explanation for (a) and (b) simultaneously, is as yet unavailable. For instance, in Romer's model with a constant savings propensity tacked on, $\alpha+\theta > 1$ ($\alpha+\theta < 1$) implies that the growth rate will be positively (negatively) correlated with the size of the capital stock, while $\alpha+\theta \approx 1$ implies independence. In the latter case, which seems supported by data, differences in growth rates among countries must be due either to differences in technology and savings rates or to shocks. The mere presence of externalities ($\theta > 0$) does not by itself account for differences in growth rates.

4. The Relation Between Inputs and Output Over Longer Epochs.

Consider a regression such as the one that Romer (1987) reports in his eq. (18). In country i, over a period length T_{it}, differences in the growth of inputs and outputs are calculated, so that for instance, $\Delta y_{it} = y_{i,t+T_{it}} - y_{it}$. That is, the regression is:

$$(15) \qquad T_{it}^{-1}\Delta y_{it} = b + b_k T_{it}^{-1}\Delta k_{it} + b_\ell T_{it}^{-1}\Delta \ell_{it} + u_{it}.$$

Romer uses 18 observations that span seven countries (subscript i), and four epochs (subscript t) of at least thirty years in length; the measure of the labor input is hours worked. The least squares regression results that he reports in his eq. (18) are: $\hat{b}_k = .87$ with a standard error of .08, and $\hat{b}_\ell = .04$ with a standard error of .18. Our aim here is to calculate the expectations of $\hat{b}_k$ and $\hat{b}_\ell$ in light of the added assumptions that we have imposed on the evolution of k, ℓ, and z. The least-squares estimates of the coefficients, denoted by hats, are identically equal to[21]

$$\begin{bmatrix} \hat{b} \\ \hat{b}_k \\ \hat{b}_k \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha+\theta \\ 1-\alpha \end{bmatrix} + \begin{bmatrix} n & \bar{k} & \bar{\ell} \\ \bar{k} & a_{kk} & a_{k\ell} \\ \bar{\ell} & a_{k\ell} & a_{\ell\ell} \end{bmatrix}^{-1} \begin{bmatrix} \bar{a} \\ a_{ku} \\ a_{\ell u} \end{bmatrix}$$

from which one can show that since $Ea_{\ell u} = 0$ by eq. (6),

$$E\begin{bmatrix} \hat{b}_k \\ \hat{b}_\ell \end{bmatrix} = \begin{bmatrix} \alpha+\theta \\ 1-\alpha \end{bmatrix} + E\,\frac{1}{a_{kk}a_{\ell\ell}-a_{k\ell}^2}\begin{bmatrix} a_{\ell\ell} & -a_{k\ell} \\ -a_{k\ell} & a_{kk} \end{bmatrix}\begin{bmatrix} a_{ku} \\ 0 \end{bmatrix}$$

(16)

$$= \begin{bmatrix} \alpha+\theta \\ 1-\alpha \end{bmatrix} + E\,\frac{1}{a_{kk}a_{\ell\ell}-a_{k\ell}^2}\begin{bmatrix} a_{\ell\ell}a_{ku} \\ -a_{k\ell}a_{ku} \end{bmatrix}$$

Since the a_{ij} are all positive, $\hat{b}_k$ will be biased upwards while $\hat{b}_\ell$ will be biased downwards, with the bias on $\hat{b}_k$ equalling $-a_{\ell\ell}/a_{k\ell}$ times the bias on $\hat{b}_\ell$. Appendix 2 calculates the a_{ij} on the assumption that all countries are subject to the same stochastic process but that they face different realizations of the ϵ's and w's as well as different initial con ditions. The resulting expressions for the a_{ij}'s are quite messy, but the following limiting results are worth noting:[22]

(17) $\lim_{r,\rho\to 1}(\lim_{T\to\infty}[E(\hat{b}_k)]) = 1,$

(18) $\lim_{r,\rho\to 1}(\lim_{T\to\infty}E(\hat{b}_\ell)) = 0.$

These results are of relevance if the epochs (which are of length T) are long, and if z and ℓ are roughly random walks, as appears to be the case empirically. So, $E(b_\ell)$ is zero regardless of the relative values of σ_ϵ^2 and σ_w^2, while $E(\hat{b}_k) = 1$ if $\sigma_w^2 = 0$. These expected values are of course not far from Romer's actual estimates, $\hat{b}_k = .87$ and $\hat{b}_\ell = .04$.

Table A.1 of the Appendix reports the expressions for the a_{ij} that one can use to calculate the bias in the least squares estimates $\hat{b}_k$ and $\hat{b}_\ell$ for the cases where a) T remains finite but ρ and r tend to unity, and b) for ρ and r less than unity but T going to infinity. In both those cases the bias remains positive, but difficult to represent analytically in a compact way. The main point is that the limiting values expressed in equations (17) and (18) are good approximations for the values that $\hat{b}_k$ and $\hat{b}_\ell$ would be expected to take for large T and for r and ρ reasonably close to one.

Equations (17) and (18) are the same as what Christiano (1987) gets under a different but related set of assumptions. He allows for country-specific fixed effects μ_i in (3) and m_i in the equation governing the evolution of ℓ, while assuming $\rho = r = 1$ and $\sigma_w^2 = 0$. His theoretical results also assume $\sigma_w^2 = 0$, while his simulations allow for $\sigma_w^2 > 0$; both yield the analogues of (17) and (18), and he argues, as we do, that the results that Romer reports in his equation (18) were consistent with θ being zero.

Several additional insights follow from our analysis, however. In explaining these it is worthwhile to elaborate on the differences between our model and Christiano's fixed effects model. These differences are best explained under the assumption that $\sigma_w^2 = 0$ -- i.e., that labor supply is non-random. In the fixed effects model a country's long-run growth rate is μ_i. In our model, it is

$$\lim_{T \to 0} T^{-1} \sum_{t=1}^{T} (z_{t+1} - z_t) = \begin{cases} \mu \text{ if } \rho = 1 \\ 0 \text{ if } \rho < 1. \end{cases}$$

Since μ is the same over countries, countries must, in the long run, all grow at the same rate <u>regardless of the value of ρ</u>. Thus, a fixed effects model delivers a positive variance of long-run growth-rates among countries while ours does not.[23]

In our model the upward bias on the capital coefficient is positively related to ρ. Staying with the case in which labor input is non-random ($\sigma_w^2 = 0$), we find from the second column of Table A.1 that as $T \to \infty$, the bias on the capital coefficient approaches

$$\frac{\rho[1 - (\alpha+\theta)^2]}{1 + \rho(\alpha+\theta)} .$$

As ρ goes to unity, so that (18) obtains, the bias becomes $1 - (\alpha+\theta)$, while as ρ goes to zero, the bias becomes zero.

The conclusion we draw from this exercise is the same as the one Christiano draws: The regression results that use data on long-run movements of output and both inputs also provide no support for the hypothesis that θ is significantly positive.

5. Microfoundations of the Model.

Our remaining task is to provide a firmer analytic foundation for the equations used in our estimation process. The key element driving the results that stem from the structural equations (2)-(6) is the dependence of the capital stock, through savings behavior, on the stochastic shock to production in the previous period. If this shock is serially correlated, current output will also depend on the shock in the previous period. Therefore, the correlation of contemporaneous output and capital does not only reflect the internal and external impact of contemporaneous capital on output but contains an additional component through the joint dependence of output and capital on the previous productivity shock. The ignoring of this element results in exaggeration of the importance of capital in production. In this section we spell out a stochastic overlapping generations (OLG) model as well as a stochastic Brock-Mirman type of growth model to justify the equations of the preceding section, especially equations (5) and (6).

We start with a special OLG model where the representative agent in generation t faces a wage w_t and a stochastic rate of return on his savings, r_{t+1}. Therefore, his consumption in the second period of his life is $c_{t+1} = (w_t - c_t) r_{t+1}$. We assume that the agent has a logarithmic utility function

$$\beta \ell n c_t + (1-\beta) E \ell n (w_t - c_t) r_{t+1},$$

that he maximizes by choosing c_t. The production function is assumed to be of Cobb-Douglas form with a multiplicative productivity shock, and is given by (1).

Population growth is stochastic, so that $L_t = L_{t-1}(1+N_t)$, where N_t is IID with mean zero. The wage rate and interest rate are equal to the marginal products of labor and capital, respectively. Since the agent bases his saving decision on the marginal product of capital in the next period, he faces a stochastic interest rate (on account both of the stochastic productivity shock and the stochastic growth of labor).

The agent's optimal savings do not depend on the interest rate so that $s_t = (w_t - c_t) = (1-\beta)w_t$. Thus, the total savings which determine next period's capital stock are

$$s_t L_t = K_{t+1} = (1-\beta)(1-\alpha)Z_0 K_t^{\alpha+\theta} L_t^{1-\alpha} = (1-\beta)(1-\alpha)Y_t,$$

since the share of labor is the fraction $(1-\alpha)$ of output. Taking logarithms immediately yields equations (5) and (6) of the previous section.

Before moving on the infinite horizon model, we should discuss the role of specific functional forms and assumptions. The logarithmic utility function simplifies matters considerably by eliminating the dependence of savings on the next period rate of return. But its use in this context goes beyond algebraic convenience. Slightly altering the utility function, say to one with a constant relative risk aversion, may yield a savings function that either increases or decreases with the rate of interest, depending on whether the risk aversion coefficient is less than or greater than unity in absolute value. Since an increase in the productivity shock leads to an expected increase in the shock next period and raises the expected interest rate, productivity shocks may, if the direct wealth effect through wages is dominated, in fact decrease savings and next period's capital stock, resulting in a negative correlation between K_{t+1} and Z_{t+1} and contradicting equation (5) in the previous section. (This issue will also arise in the infinite horizon model considered below.) In drawing generalizations from the example, therefore, we should keep in mind that we may need a preference specification for which savings are a non-decreasing function of the interest rate.[24]

We now turn to the specification with an infinitely lived representative agent. Before exploiting specific functional forms we present a general version to again pinpoint the role of the assumptions embodied in our specific functional forms.

The representative agent has an instantaneous, twice differentiable utility function $U(c_t, a_t - L_t)$, defined on feasible consumption and leisure sets, where a_t is a stochastic labor endowment and L_t is the labor supply. a_t may be specified as a multiplicative Markov process to reflect population growth, since the actual supply of labor will be endogenously chosen. The twice differentiable production function is given by $Y_t = Z_t f(K_t, \bar{K}_t, L_t)$ where Z_t is the stochastic shock to the production function and $\bar{K}_t$ $(=K_t)$ enters the production function to reflect an externality. Let δ be the depreciation rate of capital. The agent, facing constraints $K_{t+1} = Z_t f(K_t, \bar{K}_t, L_t) + (1-\delta)K_t - c_t$ and a given K_0, maximizes $E \sum_0^{\infty} \beta^t U(c_t,\ a_t - L_t)$ by choosing each c_t and L_t after observing Z_t and a_t at every t. In dynamic programming form, the problem can be expressed as

$$V(K_0, Z_0, a_0) = \max_{K_1, L_0} U(Z_0 f(K_0, \bar{K}_0, L_0) + (1-\delta)K_0 - K_1,\ a_0 - L_0) + \beta EV(K_1,\ Z_1,\ a_1).$$

To simplify matters, we assume the value function V is twice differentiable in (K,Z). (The twice differentiability of V in certain stochastic cases can be established by the methods of Blume and Easley (1982).) Let the derivatives of $V(K,Z,a)$ with respect to K and Z be denoted as V_k and V_z, and let second order derivatives be denoted in the usual way as V_{kk}, V_{kz} and V_{zz}. Similarly, let U_c and U_L be the derivatives of the utility function with respect to consumption and leisure, with second derivatives U_{cc}, U_{cL} and U_{LL}. Again, for simplicity, we will assume that $U_{cL} = 0$. Finally, let the derivatives of the production function be denoted by f_k, $f_{\bar{k}}$ and f_L. Standard methodology establishes the first order conditions for the representative agent's problem with the usual interpretation:

$$(19) \quad U_c(Z_0 f(K_0,\bar{K}_0,L_0) + (1-\delta)K_0 - K_1,\ a_0 - L_0) = \beta V_k(K_1,\ Z_1,\ a_1),$$

and

$$(20) \quad U_c(Z_0 f(K_0,\bar{K}_0,L_0) + (1-\delta)K_0 - K_1,\ a_0 - L_0) Z_0 f_L(K_0,\bar{K}_0,L_0)$$
$$= U_L(Z_0 f(K_0,\bar{K}_0,L_0) + (1-\delta)K_0 - K_1,\ a_0 - L_0)$$

From equation (20) we can obtain the optimal labor supply function as $L_0 = L(K_1, K_0, \bar{K}_0,\ a_0,\ Z_0)$. Let L_0^z and L_0^k indicate the derivative of L_0 with respect to to Z_0 and K_0.

As discussed earlier, we want to investigate the effect of Z_t on K_{t+1} so as to establish the nature of the covariance between K_{t+1} and Z_{t+1}. Using (19) and (20),

$$dK_1/dK_0 = \left[(U_{LL} + U_cZ_0f_{LL})(U_{cc}Z_0F_k) - U_{cc}Z_0f_LU_cZ_0f_{kL}\right]/D \quad > 0$$

and

$$dK_1/dZ_0 = \left[(U_{LL} + U_cZ_0f_{LL})(U_{cc}f - \beta EV_{kz}dZ_1/dZ_0) - U_{cc}Z_0^2f_L^2\beta EV_{kz}dZ_1/dZ_0\right]/D$$

where $D = (U_{LL} + U_cZ_0f_{LL})(U_{cc} + \beta EV_{kk}) + \beta EU_{cc}V_{kk}Z_0^2f_L^2 > 0$, where $F_k = f_k(K_0, \overline{K}_0, L_0) + f_{\overline{k}}(K_0, \overline{K}_0, L_0) + (1-\delta)$, and where V_{kk} and V_{kz} are evaluated at $(K_1\ Z_1, a_1)$. The policy function $K_1 = h(K_0, Z_0)$ is therefore increasing in K_0.[25] Also $dK_1/dZ_0 > 0$ if $V_{kz} > 0$. To evaluate V_{kz} we first compute

$$V_k(K_0, Z_0, a_0) = U_cZ_0(f_k+f_{\overline{k}} + (1-\delta)),$$

so as to obtain

$$V_{kz}(K_0, Z_0, a_0) = U_c((U_{cc}/U_c)(\partial c_0/\partial Z_0)Z_0+1)(f_k+f_{\overline{k}}+(1-\delta)) + U_cZ_0f_{kL}(dL(K_0, \overline{K}_0, Z_0, a_0)/dZ_0).$$

The sign of V_{kz} and therefore of dK_1/dZ_0 is ambiguous for the same reasons as in the OLG case. First it depends on the degree of relative risk aversion in the term $[(U''/U')(\partial c/\partial Z_0)Z_0+1]$: If this term is sufficiently negative, $\partial K_1/\partial Z_0$ may become negative. Furthermore, unlike our specification in the OLG model, the labor supply is endogenous. An increase in Z_0, through its effect on Z_1, leads to an increase in the expected interest rate and may produce not only lower savings but also a lower labor supply; that is, we may have $dL/dZ < 0$. This also tends to make V_{kz} negative, and, if sufficiently strong, may result in $dK_1/dZ_0 < 0$. In the special case of a logarithmic utility function, coupled with a Cobb-Douglas production function and full depreciation ($\delta=1$), V_{kz} is identically zero, as can be easily computed using the solution of this special case reported below. Therefore, for our purposes, it seems that the main restrictions imposed by a "logarithmic utility, Cobb-Douglas production with full depreciation" model is to eliminate the possibility of a saving policy and a labor supply which decrease in response to increases in the rate of return. To see this consider the policy function for the general case given by $K_1 = h(K_0, Z_0)$ and assume that $\partial h/\partial Z_0 > 0$. We than have

<u>Lemma 1</u>. Let $K_{t+1} = f(K_t, Z_t)$, where f is strictly increasing. If Z_t follows the process described by equations (3) and (4) with $\lambda_i \geq 0$, $(i = 1, 2)$, then k_t is stochastically strictly increasing in z_t.

<u>Proof</u>: Recursive substitution for lagged capital shocks in f yields $k_t = \phi(z^t)$, where $z^t = (z_{t-1}, z_{t-2}, \ldots)$, and where ϕ is strictly increasing. Applying Bayes' rule along with eq. (2) yields that for any vector $\tilde{z} \in R$, $\text{Prob}(z^t \leq \tilde{z} | z_t)$ is stochastically strictly increasing in z_t, and the claim follows. Q.E.D.

A corollary of the lemma is that the steady-state covariance between k_t and z_t is strictly positive, and this is all that is required for an upward bias on the capital coefficient in an OLS context.

The advantage of the specifications with "log utility, and with Cobb-Douglas production with full depreciation", is that we can solve explicitly for the optimal consumption, savings and labor supply policies. Using (10) and (11) and adopting the logarithmic instantaneous utility function

$$\lambda \ell n\, c + (1-\lambda) \ell n(a-L),$$

together with the Cobb-Douglass production function $Z_0 K^{\alpha} \bar{K}^{\theta} L^{1-\alpha}$, it can easily be verified that savings, or next period's capital stock, will be

(13) $$K_1 = \alpha\beta Z_0 K_0^{\alpha+\theta} L_0^{1-\alpha}$$

and that labor supply is given by

(14) $$L_1 = \lambda\theta - a_t/[(1-\lambda)(1-\alpha\beta) + \lambda\theta].$$

If the random endowment follows a multiplicative first order Markov process, (13) and (14), after taking logarithms, correspond exactly to equations (5) and (6). Note that to make labor supply stochastic we could have made the taste parameter stochastic rather than assume a stochastic endowment. Alternatively, if a, λ and other relevant parameters in (14) were constant, labor supply would be as well, and we would run into identification problems in the previous section. (Note that in the general specification of the model, labor would be stochastic even if a and λ were fixed.)

We conclude therefore that the specifications represented by equations (5) and (6) that drive our results in the previous section, and underlie our empirical conclusions, can be obtained under reasonable assumptions in either the OLG or the infinitely-lived agent models of stochastic growth.

6. Conclusions

Given the assumption that knowledge causes capital but not the other way around, our failure to find a positive θ implies nothing whatsoever about externalities in the generation of knowledge. The Solow model with no externalities to either labor or capital but with stochastic shocks to knowledge does not appear to be contradicted by long run data on output and the two inputs, and what is more, it is also consistent with micro evidence on knowledge spillovers. The apparent validity of Gibrat's law in countries' GNP series does not contradict it, nor do the seemingly sizable medium-run differentials in growth rates over countries. The model fits in with the recent business cycle literature that explains properties of cycles with productivity shocks. We have asked if this model can rationalize data other than the post-war US business-cycle, and our findings favor the model.

The realizations of our technology shocks, the z's, are allowed to differ over countries, but the stochastic process forming them is assumed to be the same over all countries, as indeed are all the parameters of our model. That the technology shocks can assume different values in different countries seems reasonable if one interprets these shocks broadly, to also include shifts in institutional and organizational structures, like shifts in the corporate, legal, or bureaucratic structures, or even in attitudes towards work. These elements can greatly enhance or retard the effective use and operation of factors of production. While such changes in institutional or organizational structures may not be permanent, they tend to be quite persistent, so that productivity in different economics can diverge for extended periods of time.

APPENDICES

Appendix 1: The Derivation of eq. (13). Note first that

$$\sum_{j=0}^{T-1} (\alpha+\theta)^j \rho^{T-j} = \rho^T[(1- [(\alpha+\theta)/\rho]^T)/(1- (\alpha+\theta)/\rho)].$$

If $(\alpha+\theta) = \rho$, this expression is equal to $T\rho^T$. Next, note from eqs. (2) and (5) that $k_{t+1} = \gamma + (\alpha+\theta)k_t + \eta_t$, where $\eta_t = (1-\alpha)\ell_t + z_t$. Then,

$$\text{Cov}(\eta_t, k_t) = \text{Cov}(\eta_t, (\alpha+\theta)k_{t-1} + \eta_{t-1}) = (\alpha+\theta)\ \text{Cov}(\eta_t, k_{t-1}) + \text{Cov}(\eta_t, \eta_{t-1}).$$

Expanding further, we obtain:

$$\text{(A.1)} \qquad \text{Cov}(\eta_t, k_t) = \sum_{j=1}^{\infty} (\alpha+\theta)^{j-1}\text{Cov}(\eta_t, \eta_{t-j}).$$

But, since $(1-\alpha)\ \ell_t = (1-\alpha)\ m + (1-\alpha)\ell_{t-1} + (1-\alpha)w_t$, then if $\rho = r$,

$$\eta_t = (1-\alpha)\ m + \mu + \rho\ \eta_{t-1} + (\epsilon_t + (1-\alpha)\ w_t),$$

so that $\text{Cov}(\eta_t, \eta_{t-j}) = \rho^j\sigma_\eta^2$, where $\sigma_\eta^2 = (\sigma_\epsilon^2 + (1-\alpha)^2\sigma_w^2)/(1-\rho^2)$.

Then using (A.1),

$$\text{(A.2)} \qquad \text{Cov}(\eta_t, k_t) = \sigma_\eta^2 \sum_{j=1}^{\infty} \rho^j(\alpha+\theta)^{j-1} = \rho\sigma_\eta^2/(1-\rho(\alpha+\theta)).$$

But from eq. (2), $Cov(\eta_t, y_t) = (\alpha+\theta)Cov(\eta_t, k_t) + \sigma_\eta^2$. Substituting into this expression from (A.2) yields

$$\text{(A.3)} \quad Cov(\eta_t, y_t) = \sigma_\eta^2[1 + \rho(\alpha+\theta)/(1-\rho(\alpha+\theta))] = \sigma_\eta^2/(1-\rho(\alpha+\theta)).$$

Next, we need to compute Var(y). Since $y_t = (\alpha+\theta)k_t + \eta_t$,

$$\text{(A.4)} \quad Var(y_t) = (\alpha+\theta)^2 Var(k_t) + \sigma_\eta^2 + 2(\alpha+\theta)Cov(\eta_t, k_t).$$

Now, since $k_{t+1} = \gamma + y_t$, $Var(y_t) = Var(k_t)$. Using this in (A.4), and substituting from (A.3) into (A.4) for $Cov(\eta_t, k_t)$ yields

$$\text{(A.5)} \quad Var(y_t) = [\sigma_\eta^2/(1- (\alpha+\theta)^2)][1 + 2(\alpha+\theta)/(1 - \rho(\alpha+\theta))]$$

The expressions in (A.4) and (A.5) both explode when $\rho \to 1$ because σ_η^2 goes to infinity, but their ratio does not:

$$\lim_{\rho\to 1} [Cov(\eta, y)/Var(y)] = 1 - \alpha - \theta.$$

This is eq. (13) of the text, since, by assumption, $\rho = r$.

Appendix 2.

This appendix derives expressions for the a_{ij} in eq. (16) under various assumptions. Deterministic components of z and ℓ are ignored. We assume $\lambda_1 = \lambda_2 = 0$, so that $\epsilon_t = w_t$, and so that the w_t are also iid. Repeated substitution in (3) leads to

$$z_{t+T} = \rho^T z_t + \rho^{T-1} \epsilon_t + \rho^{T-2}\epsilon_{t+1} \quad \cdots\cdots + \epsilon_{t+T-1},$$

or

$$\Delta^T z_t = z_{t+T} - z_t = (\rho^T - 1) z_t + \rho^{T-1}\epsilon_t + \rho^{T-2}\epsilon_{t+1} \cdots + \epsilon_{t+T-1}.$$

But

$$z_t = \rho^{j+1} z_{t-j-1} + \rho^j \epsilon_{t-j-1} + \rho^{j-1}\epsilon_{t-j} + \rho^{j-2}\epsilon_{t-j+1} \cdots + \epsilon_{t-1},$$

so that

$$\Delta^T z_t = (\rho^T - 1)(\rho^{j+1} z_{t-j-1} + \rho^j \epsilon_{t-j-1} \cdots + \epsilon_{t-1}) + \rho^{T-1}\epsilon_t + \rho^{T-2}\epsilon_{t+1} \cdots + \epsilon_{t+T-1},$$

and also

$$\Delta^T z_{t-j-1} = (\rho^T - 1) z_{t-j-1} + \rho^{T-1}\epsilon_{t-j-1} + \rho^{T-2}\epsilon_{t-j} + \rho^{T-3}\epsilon_{t-j+1} \cdots + \epsilon_{t-j+T-2}.$$

Note that subscripts on $\Delta^T z_t$ for the ϵ's run from t-j-1 to t-1+T, and that subscripts on $\Delta^T z_{t-j-1}$ for the ϵ's run from t-j-1 to t-1+T-j-1. We shall consider two separate cases: (i) $T-j-1 > 0$, and (ii) $T-j-1 \le 0$, both for $j \ge 0$.

Case (i):

$$\mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1}) = (\sigma_\epsilon^2/(1-\rho^2))((\rho^T-1)^2 \rho^{j+1})$$

$$+ \sigma_\epsilon^2(\rho^T - 1)\sum_{i=1}^{j+1} \rho^{j+1-i}\rho^{T-i} + \sigma_\epsilon^2 \sum_{i=1}^{T-j-1} \rho^{T-i}\rho^{t-j-1-i}$$

$$= (\sigma_\epsilon^2/(1-\rho^2))(\rho^T - 1)^2 \rho^{j+1} + \sigma_\epsilon^2(\rho^T-1)\rho^{T+j-1}[(1-\rho^{-2(j+1)})/(1-\rho^{-2})]$$

$$+ \sigma_\epsilon^2 \rho^{2T-j-3}[(1-\rho^{-2(T-j-1)})/(1-\rho^{-2})].$$

As a check on the algebra, note that $\lim_{\rho \to 1} \mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1}) = \sigma_\epsilon^2(T-j-1)$, because first term goes to zero by L'Hopital's rule and second term is zero. This result is exactly as expected.

Case (ii). $(T-j-1) \leq 0$.

$$\mathrm{Cov}(\Delta^T z_t,\ \Delta^T z_{t-j-1}) = (\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2\rho^{j+1} + (\rho^T-1)(\sum_{i=1}^{T} \rho^{j+1-i}\rho^{T-i})\sigma_\epsilon^2.$$

Again, as a check on the algebra, not that $\lim_{\rho\to 1} \mathrm{Cov}(\Delta^T z_t,\ \Delta^T z_{t-j-1}) = 0$, as it should be. As a further check, note that when $j = 0$, we have

$$\mathrm{Var}(\Delta^T z_t) = (\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2 + ((\rho^{T-2} - \rho^{-2})/(1-\rho^{-2}))\sigma_\epsilon^2,$$

so that $\lim_{\rho\to 1} \mathrm{Var}(\Delta^T z_t) = \sigma_\epsilon^2 T$, as it should be. Moreover, for $\rho < 1$,

$$\lim_{T\to\infty} \mathrm{Var}(\Delta^T z_t) = \sigma_\epsilon^2/(1-\rho^2) + (-\rho^{-2}/(1-\rho^{-2}))\sigma_\epsilon^2 = 2\sigma_\epsilon^2/(1-\rho^2).$$

Now we shall compute $\mathrm{Cov}(\Delta^T k_t, \Delta^T z_t)$, first for arbitrary ρ and T, and then we shall take limits. Combining cases (i) and (ii),

$$\mathrm{Cov}(\Delta^T z_t, \Delta^T k_t) = \sum_{j=0}^{\infty} (\alpha+\theta)^j \mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$

$$= \sum_{j=0}^{T-2} (\alpha+\theta)^j \mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1}) + \sum_{j=T-1}^{\infty} (\alpha+\theta)^j \mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$

$$= \sum_{j=0}^{T-2} (\alpha+\theta)^j [(\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2 \rho^{j+1} + \sigma_\epsilon^2(\rho^T-1) \sum_{i=1}^{j+1} \rho^{j+1-i}\rho^{T-i} + \sigma_\epsilon^2 \sum_{i=1}^{T-j-1} \rho^{T-i} \rho^{t-j-1-i}]$$

$$+ \sum_{j=T-1}^{\infty} (\alpha+\theta)^j [(\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2 \rho^{j+1} + \sigma_\epsilon^2(\rho^T-1) \sum_{i=1}^{T} \rho^{j+1-i}\rho^{T-i}]$$

If we now let $T \to \infty$ so that the second summation on the right goes to zero, we obtain

$$\sum_{j=0}^{\infty} (\alpha+\theta)^j [(\sigma_\epsilon^2/(1-\rho^2))[(\rho^T-1)^2 \rho^{j+1}] + (\rho^T-1)\rho^{T-1}\rho^j((1- \rho^{-2(j+1)})/(1-\rho^{-2}))\sigma_\epsilon^2$$
$$+ \sigma_\epsilon^2\rho^{2T-3}\rho^{-j}((1-\rho^{-2(T-j-1)})/(1-\rho^{-2}))$$
$$= (\rho^T-1)^2 (\sigma_\epsilon^2/(1-\rho^2))\rho(1- (\alpha+\theta)\rho)^{-1} + (\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)\rho^{T-1}[(1- (\alpha+\theta)\rho)^{-1}$$
$$- \rho^{-2}(1 - (\alpha+\theta)\rho^{-1})^{-1}] + \rho^{2T-3}[(1- (\alpha+\theta)\rho-1)^{-1} - \rho^{-1}/(1- (\alpha+\theta)\rho)]\sigma_\epsilon^2/(1-\rho^{-2}).$$

Now we note that as $T \to \infty$, the second term above also goes to zero. The first goes to $\sigma_\epsilon^2\rho/(1-\rho^2)(1 - \rho(\alpha+\theta))$, while the third goes to $\rho^{-1} \sigma_\epsilon^2/(1-\rho^{-2})(1 - \rho(\alpha+\theta))$. Therefore,

(A.5a) $$\lim_{T\to\infty} \mathrm{Cov}(\Delta^T k_t,\Delta^T z_t) = [\sigma_\epsilon^2/(1- \rho(\alpha+\theta))](\rho/(1-\rho^2) - \rho^{-1}/(1-\rho^{-2}))$$
$$= 2\sigma_\epsilon^2\rho/(1-\rho^2)(1 - \rho(\alpha+\theta)).$$

Next we calculate the limit as $\rho \to 1$, for fixed T.

(A.5b) $$\lim_{\rho\to 1} \mathrm{Cov}(\Delta^T z_t,\Delta^T k_t) = \sigma_\epsilon^2 \sum_{j=0}^{T-2} (\alpha+\theta)^j(T-j-1)$$
$$= \sigma_\epsilon^2(T-1 + (\alpha+\theta)(T-2) \ \dots \ (\alpha+\theta)^{T-2})$$
$$= \sigma_\epsilon^2[T(1 - (\alpha+\theta))^{-1} - (1- (\alpha+\theta)^T)(1- (\alpha+\theta))^{-2}]$$
$$= \sigma_\epsilon^2 \ [T(1- (\alpha+\theta)) - (1- (\alpha+\theta)^T)](1- (\alpha+\theta))^{-2}.$$

Next we turn to the computation of $\mathrm{Var}(\Delta^T k_t)$. We have

$$\mathrm{Var}(\Delta^T k_t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} (\alpha+\theta)^{i+j}[(1-\alpha)^2\mathrm{Cov}(\Delta^T \ell_{t-j},\Delta^T \ell_{t-i}) + \mathrm{Cov}(\Delta^T z_{t-j}, \Delta^T z_{t-i})].$$

Let $\hat{A} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} (\alpha+\theta)^{i+j} (1-\alpha)^2\mathrm{Cov}(\Delta^T \ell_{t-j},\Delta^T \ell_{t-i})$.

We will compute $\hat{A}$ later. If we let $\rho \to 1$ for a given T, remembering that $\lim_{\rho\to 1} \mathrm{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1}) = 0$ for $t-j-1 \le 0$, we obtain:

$$\operatorname{Var}(\Delta^T k_t) = \sum_{j=0}^{\infty} \sum_{i=j+1}^{T+j} (\alpha+\theta)^{i+j}(T-(i-j))\sigma_\epsilon^2 + \sum_{i=0}^{\infty} \sum_{j=i+1}^{T+i} (\alpha+\theta)^{i+j}(T-(j-i))\sigma_\epsilon^2 + \sum_{k=0}^{\infty} (\alpha+\theta)^{2k} T\sigma_\epsilon^2 + \hat{A}$$

$$= \{2\sigma_\epsilon^2(\alpha+\theta)[T/(1-(\alpha+\theta)) - (1-(\alpha+\theta)^T)/(1-(\alpha+\theta))^2]$$

$$+ T\sigma_\epsilon^2(1-(\alpha+\theta)^2)^{-1})/(1-(\alpha+\theta)^2) + \hat{A}.$$

We also compute $\operatorname{Var}(\Delta^T k_t)$ for $\rho < 1$ at $T \to \infty$: We have

$$\operatorname{Var}(\Delta^T k_t) = (1-\alpha)^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \operatorname{Cov}(\Delta^T \ell_{t-j}, \Delta^T \ell_{t-i})(\alpha+\theta)^{i+j} + \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \operatorname{Cov}(\Delta^T z_{t-j}, \Delta^T z_{t-i})(\alpha+\theta)^{i+j}.$$

Let $\hat{A}$ again denote the first of these expression. We shall compute it later. Next, observe that for $i > j$ and $T - (i-j) > 0$,

$$\operatorname{Cov}(\Delta^T z_{t-i}, \Delta^T z_{t-j})$$

$$= (\sigma_\epsilon^2/(1-\rho^2))((\rho^j-1)^2 \rho^{i-j} + \sigma_\epsilon^2(\rho^T-1)\rho^{T+(i-j)-2}(1-\rho^{-2(i-j)})/(1-\rho^{-2}))$$

$$+ \sigma_\epsilon^2 \rho^{2T-(i-j)-2}(1-\rho^{-2T+2(i-j)})/(1-\rho^{-2}).$$

If we let $T \to \infty$, note that $T - (i-j) > 0$ and $T - (j-i) > 0$ for all fixed i,j. Now, break the summation for $\operatorname{Var}(\Delta^T k_t)$ into three parts: $i > j$, $i < j$, and $i = j$. But the expressions for $i > j$ and $i < j$ are symmetric. So compute twice the value for $i > j$:

$$\lim_{T\to\infty} \operatorname{Var}(\Delta^T k_t) = \lim_{T\to\infty} \Big\{ 2 \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} (\alpha+\theta)^{i+j}(\sigma_\epsilon^2/(1-\rho^2))[(\rho^T-1)^2 \rho^{i-j}$$

$$+ (\rho^T-1)(\sigma_\epsilon^2/(1-\rho^{-2}))\rho^{T-2}\rho^{i-j}(1-\rho^{-2i}\rho^{2j}) + (\sigma_\epsilon^2/(1-\rho^{-2}))(\rho^{2T-2}\rho^{-i}\rho^{j}(1-\rho^{-2T}\rho^{2i}\rho^{-2j}))]$$

$$+\Sigma\,\Sigma(\alpha+\theta)^{i+j}[(\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2+(\sigma_\epsilon^2/(1-\rho^{-2}))\rho^{T-2}(0)+(\sigma_\epsilon^2/(1-\rho^{-2}))(\rho^{2T-2}(1-\rho^{-2T})) + \hat{A}\Big\}$$

$$= \lim_{T\to\infty} \Big\{2\sigma_\epsilon^2(\sum_{j=0}^{\infty} (\rho^T-1)^2 (\alpha+\theta)^j\rho^{-j}(\alpha+\theta)^{j+1}\rho^{j+1}(1-(\alpha+\theta)\rho)^{-1}(1-\rho)$$

$$+ \rho^{T-1} (1-\rho^{-2})^{-1} (\alpha+\theta)^j \rho^{T-2}\rho^{-j}[\rho^{j+1}(\alpha+\theta)^{j+1}(1 - (\alpha+\theta)\rho)^{-1}$$

$$- \rho^{2j}[(\alpha+\theta)^{j+1}\rho^{-(j+1)}(1- (\alpha+\theta)\rho^{-1})^{-1}]$$

$$+(1-\rho^{-2})\rho^{2T-2} (\alpha+\theta)^j \rho^j[(\alpha+\theta)^{j+1}\rho^{-(j+1)}(1- (\alpha+\theta)\rho^{-1})^{-1}$$

$$- \rho^{-2T}\rho^{-2j} (\alpha+\theta)^{j+1}\rho^{j+1}(1- (\alpha+\theta)\rho)^{-1}]$$

$$+ [\sigma_\epsilon^2(1-\rho^{-2})^{-1}(\rho^T-1)^2 + \sigma_\epsilon^2(1-\rho^{-1})^{-1}\rho^{2T-2} - \sigma_\epsilon^2(1-\rho^{-2})^{-1}\rho^{-2}](1- (\alpha+\theta)^2) + \hat{A}\Big]$$

$$= \lim_{T\to\infty} \Big\{2[(\sigma_\epsilon^2/(1-\rho^2))(\rho^T-1)^2\rho(\alpha+\theta)(1- (\alpha+\theta)^2)(1- (\alpha+\theta)\rho)^{-1}$$

$$+ (\rho^T-1)\rho^{T-2}(\sigma_\epsilon^2/(1-\rho^{-2}))\rho(\alpha+\theta)(1- (\alpha+\theta)^2)^{-1}(1- (\alpha+\theta)\rho)^{-1}$$

$$- (\sigma_\epsilon^2/(1-\rho^{-2}))(\rho^T-1)\rho^{T-2}(\alpha+\theta)\rho^{-1}(1 - (\alpha+\theta)^2)^{-1}(1- (\alpha+\theta)\rho^{-1})^{-1}$$

$$+ (\sigma_\epsilon^2/(1-\rho^{-2}))\rho^{2T-2} (\alpha+\theta)\rho^{-1}(1- (\alpha+\theta)^2)^{-1}(1- (\alpha+\theta)\rho^{-1})^{-1}$$

$$- (\sigma_\epsilon^2/(1-\rho^{-2}))(\alpha+\theta)\rho^{-1}(1- (\alpha+\theta)^2)^{-1}(1- (\alpha+\theta)\rho^{-1})^{-1}$$

$$+ (1- (\alpha+\theta)^2)[\sigma_\epsilon^2(1-\rho^2)^{-1}(\rho^T-1)^2 + \sigma_\epsilon^2(1-\rho^{-2})\rho^{2T-2} - \sigma_\epsilon^2(1-\rho^{-2})\rho^{-2}] + \hat{A}\Big]$$

$$= \lim_{T\to\infty} \Big\{(\sigma_\epsilon^2/(1-\rho^2))(1- (\alpha+\theta)^2)^{-1}[2(\rho^T-1)^2(1-\rho(\alpha+\theta))^{-1}\rho(\alpha+\theta)$$

$$- 2(\rho^T-1)\rho^T\rho(\alpha+\theta)(1- (\alpha+\theta)\rho)^{-1} + 2(\rho^T-1)\rho^T(\alpha+\theta)\rho^{-1}(1- (\alpha+\theta)\rho^{-1})$$

$$- 2\rho^{2T} (\alpha+\theta)\rho^{-1}(1- (\alpha+\theta)\rho^{-1})^{-1} + 2(\alpha+\theta)\rho(1 - (\alpha+\theta)\rho)^{-1}$$

$$+ ((\rho^T-1)^2 - \rho^{2T} + 1) + 2(1-\rho^T) + \hat{A}\Big]$$

$$= \lim_{T\to\infty} \Big\{2(\sigma_\epsilon^2/(1-\rho^2))(1- (\alpha+\theta)^2)\{(\rho^T-1)^2\rho(\alpha+\theta)(1- (\alpha+\theta)\rho)^{-1}$$

$$+ (\rho^T-1)\rho^T[\rho^{-1}(1- (\alpha+\theta)\rho^{-1})^{-1} - \rho(1- (\alpha+\theta)\rho)^{-1}](\alpha+\theta)$$

$$+(\alpha+\theta)[\rho(1- (\alpha+\theta)\rho)^{-1} - \rho^{-1}(1- (\alpha+\theta)\rho^{-1})\rho^{2T}] + (1-\rho^T)\} + \hat{A}\Big]$$

Therefore,

$$\text{(A.6)} \quad \lim_{T\to\infty} \operatorname{Var}(\Delta^T k_t) = 2(\sigma_\epsilon^2/(1-\rho^2))(1- (\alpha+\theta)^2)^{-1}\Big[2(\alpha+\theta)\rho(1- (\alpha+\theta)\rho)^{-1} + 1\Big] + \lim_{T\to\infty} \hat{A}.$$

Finally, we need to compute $\hat{A}$. Since $\ell_t = r\ \ell_{t-1} + w_t$, the process ℓ_t behaves like the z_t process, with r replacing ρ and w replacing ϵ. Therefore, using earlier formulas for z,

$$\text{(A.7)} \qquad \operatorname{Cov}(\Delta^T\ell_{t-i},\Delta^T\ell_{t-j}) = (T - (i-j))\sigma_w^2 \quad \text{for} \quad r \to 1,$$

so that $\operatorname{Var}(\Delta^T \ell_t) \to T\sigma_w^2$ and,

$$\operatorname{Cov}(\Delta^T\ell_{t-i},\Delta^T\ell_{t-j}) = 2(\sigma_w^2/(1-r^2))r^{(i-j)} \quad \text{for} \quad T \to \infty,\ r < 1,$$

so that $\operatorname{Var}(\Delta^T\ell_t) \to 2\sigma_w^2/(1-r^2)$.

Now, for $r = 1$, $\hat{A} = (1-\alpha)^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \operatorname{Cov}(\Delta^T\ell_{t-i},\Delta^T\ell_{t-j})(\alpha+\theta)^{i+j}$

$$= [(1-\alpha)^2/(1- (\alpha+\theta)^2)][2\sigma_w^2(\alpha+\theta)[T/(1- (\alpha+\theta)) - (1- (\alpha+\theta)^T)(1- (\alpha+\theta))^{-2} + T\sigma_w^2].$$

On the other hand, for $r < 1$,

$$\begin{aligned}\lim_{T\to\infty} \hat{A} &= \lim_{T\to\infty} \Big((1-\alpha)^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \operatorname{Cov}(\Delta^T\ell_{t-i},\Delta^T\ell_{t-j})(\alpha+\theta)^{i+j}\Big) \\ &= \lim_{T\to\infty} \Big\{(1-\alpha)^2\Big[(2\sigma_w^2/(1-r^2))(1- (\alpha+\theta)^2) \Big((r^T-1)^2 r(\alpha+\theta)(1- (\alpha+\theta)r)^{-1} \\ &\quad + (r^T-1)r^T[r^{-1}(1- (\alpha+\theta)r^{-1})]^{-1} - r(1-(\alpha+\theta)r)^{-1}(\alpha+\theta) \\ &\quad + (\alpha+\theta)[r(1- (\alpha+\theta)r)^{-1} - r^{-1}(1- (\alpha+\theta)r^{-1})r^{2T}] + (1-r^T)\Big)\Big]\Big\}.\end{aligned}$$

Therefore

$$\text{(A.8)} \qquad \lim_{T\to\infty} \hat{A} = 2(\sigma_w^2/(1-r^2))(1- (\alpha+\theta)^2)^{-1}[2(\alpha+\theta)r(1- (\alpha+\theta)r)^{-1} + 1]$$

Finally

$$\text{(A.9)}\quad \mathrm{Cov}(\Delta^T k_t, \Delta^T \ell_t) = (1-\alpha)\sum_{j=0}^{\infty} \mathrm{Cov}(\Delta^T \ell_t, \Delta^T \ell_{t-j-1})(\alpha+\theta)^j$$

$$= (1-\alpha)\sigma_\epsilon^2\,[T(1-(\alpha+\theta)) - (1-(\alpha+\theta)^T)]/(1-(\alpha+\theta))^2, \quad \text{for } r = 1.$$

In general, for arbitrary r,

$$\lim_{T\to\infty} \mathrm{Cov}(\Delta^T k_t, \Delta^T \ell_t) = \lim_{T\to\infty} \Big\{ (1-\alpha)\sigma_\epsilon^2/(1-r^2)\,\Big[(r^T-1)^2 r(1-(\alpha+\theta)r)^{-1}$$
$$+ (r^T-1)r^{T+1}[r^{-2}(1-(\alpha+\theta)r^{-1})^{-1} - (1-(\alpha+\theta)r)^{-1}]$$
$$- r^{2T-1}(1-(\alpha+\theta)r^{-1})^{-1} + r(1-(\alpha+\theta)r)^{-1}\Big]\Big\}$$

Taking the limit, for $r < 1$,

$$\text{(A.10)}\quad \lim_{T\to\infty} \mathrm{Cov}(\Delta^T \ell_t, \Delta^T k_t) = (1-\alpha)2\sigma_\epsilon^2/(1-r^2)\,(1-(\alpha+\theta)r)^{-1}.$$

The following table summarizes the results of the appendix relevant for the bias described in equation (16) of the text.

Combined cases: $r, \rho \to 1$, $T \to \infty$. We shall how use the expression in the second column, and we shall send r^2 and ρ^2 to unity at the same rate. The resulting expressions are then used in eqs. (16) and (17) of the text. The Landau symbol "O" refers to order of the expression.

$$a_{kk} = O\frac{1}{1-\rho^2}\cdot\frac{2\sigma_\epsilon^2}{1-(\alpha+\theta)^2}\left(\frac{2(\alpha+\theta)}{1-(\alpha+\theta)}+1\right) + O\frac{1}{1-r^2}\cdot\frac{2\sigma_w^2(1-\alpha)^2}{1-(\alpha+\theta)^2}\left(\frac{2(\alpha+\theta)}{1-(\alpha+\theta)}+1\right)$$

$$= 0\frac{1}{1-\rho^2}\cdot\frac{2\sigma_\epsilon^2}{1-(\alpha+\theta)^2}\cdot\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)} + 0\frac{1}{1-r^2}\cdot\frac{2\sigma_w^2(1-\alpha)^2}{1-(\alpha+\theta)^2}\cdot\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)}$$

$$a_{k\ell} = 0\frac{1}{1-r^2}\cdot\frac{2\sigma_w^2}{1-(\alpha+\theta)}(1-\alpha)$$

$$a_{\ell\ell} = 0\frac{1}{1-r^2}\cdot 2\sigma_w^2$$

and

$$a_{ku} = 0\frac{1}{1-\rho^2}\cdot\frac{2\sigma_\epsilon^2}{1-(\alpha+\theta)}$$

Therefore, letting A_{ij} be the constant in the expression for a_{ij},

$$\frac{a_{\ell\ell}a_{ku}}{a_{kk}a_{\ell\ell}-a_{k\ell}^2} = \frac{0\frac{1}{1-r^2}\cdot A_{\ell\ell}\,0\frac{1}{1-\rho^2}\cdot A_{ku}}{\left[0\frac{1}{1-\rho^2}\cdot A_{kk}^1 + 0\frac{1}{1-r^2}\cdot A_{kk}^2\right]0\frac{1}{1-r^2}\cdot A_{\ell\ell} - \left[0\frac{1}{1-r^2}\right]^2 A_{k\ell}^2}$$

where A_{kk}^1 and A_{kk}^2 are the first and second terms in the expression for a_{kk}. Now send $1-\rho^2$ and $1-r^2 \to 0$ at the same rate, to get

$$\frac{a_{\ell\ell}a_{ku}}{a_{kk}a_{\ell\ell}-a_{k\ell}^2} \longrightarrow \frac{A_{\ell\ell}A_{ku}}{(A_{kk}^1 + A_{kk}^2)A_{\ell\ell} - A_{k\ell}^2}$$

$$= \frac{\frac{4\sigma_w^2\sigma_\epsilon^2}{1-(\alpha+\theta)}}{\left[\frac{2\sigma_\epsilon^2}{1-(\alpha+\theta)^2}\cdot\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)} + \frac{2\sigma_w^2(1-\alpha)^2}{1-(\alpha+\theta)^2}\cdot\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)}\right]2\sigma_w^2 - \frac{4\sigma_w^4(1-\alpha)^2}{[1-(\alpha+\theta)]^2}}$$

Observing that $1-(\alpha+\theta)^2 = [1-(\alpha+\theta)](1+\alpha+\theta)$ and making that substitution on the bottom line of the above expression leads to $[1-(\alpha+\theta)]^2$ entering everywhere on the bottom of the denominator. Then multiplying top and bottom by $[1-(\alpha+\theta)]^2/4\sigma_w^2$ leaves us with

$$\frac{\sigma_\epsilon^2[1 - (\alpha+\theta)]}{\sigma_\epsilon^2 + (1-\alpha)\sigma_w^2 - (1-\alpha)\sigma_w^2} = 1 - (\alpha+\theta)$$

Substituting this into (16) leads to (17). We now calculate the bias on $\hat{b}_\ell$:

$$\frac{-a_{k\ell}a_{ku}}{a_{kk}a_{\ell\ell}-a_{k\ell}^2} = \frac{-0\frac{1}{1-r^2}\cdot A_{k\ell}0\frac{1}{1-\rho^2}\cdot A_{ku}}{0\frac{1}{1-\rho^2}\cdot A_{kk}^1 + 0\frac{1}{1-r^2}\cdot A_{kk}^2]0\frac{1}{1-r^2}\cdot A_{\ell\ell} - [0\frac{1}{1-r^2}]^2 A_{k\ell}^2}$$

$$\longrightarrow \quad \frac{-A_{k\ell}A_{ku}}{(A_{k\ell}^1 + A_{kk}^2)A_{\ell\ell} - A_{k\ell}^2}$$

$$= \frac{\dfrac{-4\sigma_w^2\,\sigma_\epsilon^2(1-\alpha)}{[1-(\alpha+\theta)]^2}}{\left(\dfrac{2\sigma_\epsilon^2}{1-(\alpha+\theta)^2}\cdot\dfrac{1+(\alpha+\theta)}{1-(\alpha+\theta)} + \dfrac{2\sigma_w^2(1-\alpha)^2}{1-(\alpha+\theta)^2}\cdot\dfrac{1+(\alpha+\theta)}{1-(\alpha+\theta)}\right) 2\sigma_w^2 - \dfrac{4\sigma_w^4(1-\alpha)^2}{[1-(\alpha+\theta)]^2}}.$$

But

$[1-(\alpha+\theta)^2](1-(\alpha+\theta)) = (1+\alpha+\theta)(1-(\alpha+\theta))^2$. Therefore, the above equals

$$= \frac{-4\sigma_w^2\sigma_\epsilon^2(1-\alpha)}{2\sigma_w^2(2\sigma_\epsilon^2 + (1-\alpha)^2 2\sigma_w^2) - 4\sigma_w^2(1-\alpha)^2} = \frac{-\sigma_\epsilon^2(1-\alpha)}{\sigma_\epsilon^2 + (1-\alpha)^2\sigma_w^2 - \sigma_w^2(1-\alpha)^2}$$

$$= -(1-\alpha).$$

When substituted into (16) this leads to (18).

Appendix 3.

In this appendix, we briefly describe our analysis of the equation $K_{t+1} = sY_t + (1-\delta)K_t$. This analysis led to the estimation reported in Tables 3 and 4. Under this hypothesis,

$$\begin{aligned} y_{t+1} &= z_{t+1} + (1-\alpha)\ell_{t+1} + (\alpha+\theta)\ln(sY_t + (1-\delta)K_t) \\ &= z_{t+1} + (1-\alpha)\ell_{t+1} + (\alpha+\theta)\ln(sY_t + (1-\delta)(sY_{t-1} + (1-\delta)K_{t-1})) \\ &= z_{t+1} + (1-\alpha)\ell_{t+1} + (\alpha+\theta)\ln s + (\alpha+\theta)\ln(\sum_{j=0}^{\infty}(1-\delta)^j Y_{t-j}). \end{aligned}$$

Therefore, the analogue of the equation in footnote 11 is:

$$\begin{aligned} \text{(A.11)} \qquad y_{t+1} - \rho y_t &= \omega_t + (1-\alpha)(\ell_{t+1}-\rho\ell_t) + (\alpha+\theta)(1-\rho)\ln s \\ &\quad + (\alpha+\theta)[\sum_{j=0}^{\infty}(1-\delta)^j Y_{t-j} - \rho\ln\sum_{j=0}^{\infty}(1-\delta)^j Y_{t-j-1}]. \end{aligned}$$

The ω_t process was once again assumed to follow eq. (4). The infinite sums in (A.11) were truncated at $j = 20$. This was possible because only yearly data are used in Tables 3 and 4. Since at least about 20 years of data are needed to construct a reasonable approximation to the infinite sum of past Y's, we could not use quarterly data, as these are available only for the post-war years. See footnote 12, however.

REFERENCES

Baily, Martin Neal, "Comment," NBER Macroeconomic Annual (1987): 205-8.

Baumol, Will and Wolff, Edward, "Productivity Growth, Convergence, and Welfare: Reply," American Economic Review, 78, no. 5 (December 1988): 1155-1159.

Barro, Robert, "The Persistence of Unemployment," American Economic Review, 78, no. 2 (May 1988): 32-37.

Benhabib, Jess and Laroque, Gey, "On Competitive Cycles in Production Economies," Journal of Economic Theory, 45, no. 1 (June 1988): 145-170.

Benhabib, Jess and Nishimura, Kazuo, "Stochastic Equilibrium Oscillations," International Economic Review, 30 (1989):85-102.

Bernstein, Jeremy and Nadiri, Ishaq, "Research and Development and Intraindustry Spillovers: An Empirical Implication of Dynamic Duality," C.V. Starr Paper 88-06, New York University, February, 1988.

Blume, Lawrence and Easley, David, "Characterization of Optimal Plans for Stochastic Dynamic Programs," Journal of Economic Theory, 28, 1982: 221-234.

Campbell, John and Mankiw, Gregg, "Are Output Fluctuations Transitory? "Quarterly Journal of Economics, (November 1987): 857-80.

Christiano, Lawrence, "Comment on Romer's 'Crazy Explanations of the Productivity Slowdown'," unpublished manuscript, Federal Reserve Bank of Minneapolis, 1986.

Cochrane, John, "How Big is the Random Walk in GNP?" Journal of Political Economy, 96, no. 5 (October 1988): 893-920.

DeLong, Bradford, "Productivity Growth, Convergence and Welfare: Comment," American Economic Review, 78, no. 5 (December 1988): 1138-1154.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," Bell Journal of Economics, 10, no. 1 (Spring 1979):92-116.

__________, 'Productivity Puzzles and R&D: Another Nonexplanation," Journal of Economic Perspectives, 2, no. 4 (Fall 1988):9-21.

Griliches, Zvi and Lichtenberg, Frank, "Intraindustry Technical Flows and Productivity Growth: A Reexamination." Review of Economics and Statistics, LXVI, no. 2 (May 1984): 324-28.

Jaffe, Adam, "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," American Economic Review, 76, no. 5 (December 1986): 984-1001.

Jovanovic, Boyan and MacDonald, Glenn, "Competitive Diffusion," unpublished paper, October 1988.

Jovanovic, Boyan and Rob, Rafael, "The Growth and Diffusion of Knowledge," Review of Economic Studies, October 1989.

Lach, Saul and Schankerman, Mark, "Dynamics of R&D and Investment in the Scientific Sector," Journal of Political Economy, 97, no. 4 (August 1989): 880-904.

Mansfield, E., et al., "Social and Private Rates of Return from Industrial Innovations," Quarterly Journal of Economics, 91, no. 2 (1977): 221-240.

Nelson, Charles and Plosser, Charles, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." Journal of Monetary Economics, 10 (September 1982): 139-62.

Pakes, Ariel and Schankerman, Mark, "The Rate of Obsolescence of Patents, Research Gestation Lags and the Private Rate of Return to Research Resources," in Zvi Griliches, ed., R&D, Patents and Productivity, NBER, University of Chicago Press, 1984a.

————————————, "An Exploration into the Determinants of Research Intensity," in Z. Griliches, ed., R&D, Patents and Productivity, NBER, University of Chicago Press, 1984b.

Prescott, Edward C., "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis Quarterly Review, (Fall 1986): 9-22.

Romer, Paul, "Crazy Explanations for the Productivity Slowdown," NBER Macroeconomic Annual, (1987):163-201.

Schmookler, Jacob, Invention and Economic Growth, Cambridge: Harvard University Press, 1966.

Shleifer, Andrei, "Implementation Cycles," Journal of Political Economy, 94, no. 6 (December 1986):1163-90.

Vernon, R. "Technological Development: The Historical Experience," EDI seminar paper 39, The World Bank, Washington D.C. 1989.

FOOTNOTES

* Professors of Economics, New York University, 269 Mercer Street, New York, NY 10003. We thank the C.V. Starr for Applied Economics for financial and technical assistance, Francesco Goletti and Ray Atje for research assistance, and Bob Barro, Will Baumol, Larry Christiano, Stan Fischer, Zvi Griliches, Ned Nadiri, Ariel Pakes and Paul Romer for helpful comments. An earlier version was given at an *NBER* growth conference in Cambridge in October, 1989.

1 Christiano (1987) challenges these conclusions, claiming that a balanced-path outcome for the Solow model is consistent with the data, and that no capital externality is required. Indeed, in the deterministic case, along the balanced-growth path, the externality cannot be identified. Baily (1987) and others have made the same point.

2 See especially his reference to evidence from Schmookler (1966) to the effect that in various industries patenting tends to follow investment.

3 Their results must, however, be viewed with caution. Mansfield et al. look only at *successful* innovations, so that their sample of innovations does not accurately represent the outcome of investments in R&D. On the other hand, while the absolute value of the private and social rates of return is clearly biased upward in their sample, their relative magnitudes are perhaps not biased. Among their 18 innovations, the social rate of return averaged 77%, while that of the private rate averaged 33%. These results do support Romer's claim that the social returns might exceed the private rate by a factor of more than two to one. Bernstein and Nadiri's results are based on a deterministic model, and simultaneity biases are

likely to be present because of left-out time effects and unobservable industry effects. In essence, they evaluate the spillover from the partial correlation between a firm's investment in physical capital and R&D on the one hand, and the industry investment in R&D on the other. These variables will usually be positively correlated because they both will respond to industry shocks, time effects, and so on, and this response will cause an upward bias on any estimate of R&D spillovers that relies on this partial correlation.

4 Pakes and Schankerman (1984b) find a much stronger correlation between industry-wide R&D and lagged industry growth than they do between firm R&D and firm growth. However, they interpret the causality as running from industry growth to R&D, in the spirit of Schmookler's argument that the incentive to do R&D increases as market size grows. It seems crucial, at the industry level, for assumptions be imposed that allow one to distinguish shifts in product-demand from shifts in technological opportunity. A different, and more questionable, source of evidence on spillovers is that on the rate at which the economic value of private knowledge decays. The faster a piece of knowledge spills over to other firms, the faster, presumably, is the loss of economic rent that the firm can extract from that piece of knowledge. Pakes and Schankerman (1984a) find that knowledge depreciates much faster than physical capital, although they do not interpret this as implying a high spillover-rate for knowledge. Unfortunately, the value of private knowledge may, as Griliches (1979) points out, decay not just because it "leaks" to other firms but also because it is superseded by new knowledge generated by other firms. In other words, the economic value of knowledge would depreciate even in a world with no spillovers, and its depreciation rate is thus an unreliable indicator of the extent and

speed of spillovers.

5 Especially on p. 194 with reference to evidence on the persistence of cross-country differentials in growth rates.

6 At least a part of this model is in Griliches (1979, p. 102) who there attributes it to an unpublished note by Grunfeld and Levhari.

7 Either because it is superseded by other knowledge, or because some of the people who possess it die off.

8 One of these is in many respects similar to the model that Prescott (1986) proposes for business-cycle analysis.

9 Equation (5) is an exact specification for the infinite horizon, representative agent model with logarithmic utility, Cobb-Douglas production and 100% depreciation of capital. The details are in section 5. At this stage we should point out, however, that for a model with less than 100% depreciation and general functional forms, the qualitative features of this relationship, that is the positive covariance of k_{t+1} with k_t as well as z_t, that are the critical elements driving our results in the following section, will be preserved under very reasonable assumptions. This issue is explicitly discussed in section 5 (especially see lemma 1 and the surrounding discussion). Moreover, in section 2 we shall also present the estimates for the model's parameters when instead of (5), the evolution of the capital stock obeys $K_{t+1} = sY_t + (1-\delta)K_t$. See Tables 3 and 4. We shall also present estimates in Tables 5-8 that use capital data and hence bypass (5) and (5)'.

10 The likelihood was derived as follows. Multiplying (7) through by ρ, lagging one period and subtracting the result from (7) yields

$$y_t - \rho y_{t-1} = C + (\alpha+\theta)(y_{t-1}-\rho y_{t-2}) + (1-\alpha)(\ell_t-\rho\ell_{t-1}) + \omega_t,$$

where $C = (1-\rho)[(\alpha+\theta)\gamma+\mu]$. The ϵ_t are assumed to be normally distributed. Since ω contains two moving average terms, we used the Box-Jenkins procedure, setting the two pre-sample error terms to their zero mean.

[11] We did experiment with post-war quarterly data, using a 10 year weighted average of past Y_t's to construct the capital stock. The estimate of θ (with α not constrained) was -1.52 and significantly different from zero. Thus, when (5)' is used in place of (5), the annual and quarterly data both yield estimates of θ far below these in Table 1.

[12] The Summers-Heston data set has information on population but not on the labor input. It also has no information on the capital input.

[13] Barro's (1988)cross-country study of the univariate process for log un-employment (again, with annual data) revealed some significant cross-country differences in the degree of persistence in that variable. Nevertheless, at least in the post-war samples, the AR(1) coefficient estimate typically does not differ significantly from unity. There are, however, good reasons to suspect the truth of our assumptions about ℓ_t. First, human capital should respond positively to ϵ, in much the same way as physical capital. This would tend to induce a positive correlation between ℓ_t and ϵ_t. On the other hand, fertility responds negatively to income, and this would tend to induce a negative correlation between ℓ_t and longer lags of ϵ_t.

[14] Kuwait was excluded from the regression, as it is an extreme outlier.

[15] But this finding is for countries as a group, most of whom are small and have little R&D investment. For industrialized countries, the result is somewhat different -- see Baumol and Wolff (1988), and DeLong (1988).

[16] Because y_t acquires a permanent component if either $\rho = 1$, or if $(\alpha+\theta) = 1$. The findings of Nelson and Plosser (1982), Campbell and Mankiw (1987) and Cochrane (1988) that one cannot reject the hypothesis that in the univariate ARMA representation of GNP, innovations to GNP have a permanent component do not, therefore, by themselves tell us whether $\alpha+\theta = 1$, or

whether ρ or r are equal to unity.

[17] In our model, the parameter μ presumably depends inversely on what is known domestically relative to what is known abroad. Two models of learning in situations where different agents know different things are Jovanovic and Rob (1989) and Jovanovic and MacDonald (1988). In both these theoretical models, those who are farther behind learn more (through imitation) than those a who are closer to the leaders, simply because they have more to learn. This argues for a higher μ for the poorer nations. But, such a perspective ignores the constraints on the capacity of people in a developing country to absorb and apply the technologies that the more advanced countries had already created and put in place. See Vernon (1989) for a viewpoint that emphasizes these constraints.

[18] σ_i^2 turns out to be significantly negatively correlated to $y_{i,1960}$. That is, initially larger countries have less variable growth-rates. For instance, for the US, $\sigma_i^2 = .0006$.

[19] For the US, σ_i^2 is .0006, whereas the median country (see Figure 1) has a σ_i^2 that is three or four times that.

[20] For instance, the parameter μ. Country-specific fixed effects are, in this context at least, simply a label for one's ignorance, and the calculations about variances reported in the above paragraph are too rough and tentative to convince us that the country-specific fixed effect is needed here. Our z_{it}'s are, we submit, less objectionable because they at least are stochastically equal among countries, although their particular realizations can vary. Moreover, even if ρ and r are unity, the long-run growth rate of z is just μ for each country, and there can be no long-run differences in growth. We discuss this in greater detail in the next subsection.

[21] This equation follows directly from the application of the least

squares formula. The number of observations (i.e., the number of country-epoch pairs) is n. The a_{ij} are the raw moments. For instance, $a_{k\ell} = \sum_{i,t} T_{it}^{-2}\Delta k_{it}\Delta \ell_{it}$ and so on. The variables with a bar over them are the mean-growth rates over the sample. For instance, $\bar{k} = \sum_{i,t} T_{it}^{-1}\Delta k_{it}$, and so on.

[22] We present the results only for this particular limiting case because the general expressions would be very lengthy. Table (A.1) at the end of Appendix 2 presents results that make it possible to compute $E(\hat{b}_k)$ and $E(\hat{b}_\rho)$ for finite T, or for ρ and r less than unity.

[23] Table A.1 contains information about the speed of convergence to zero of magnitudes such as a_{kk} the variance long-run growth rate of the capital stock). The table shows that this and other variances and covariances go to zero at the rate T^{-1} when $\rho = 1$, while they go to zero at the rate T^{-2} when $\rho < 1$.

[24] Another set of problems that plague the OLG model relates to the continuum of equilibria. While our special specification avoids these problems, multiplicities will arise either if outside money is introduced as an additional asset, or if the labor supply decision is endogenized and the logarithmic specification of utility is dropped. (For a detailed analysis see Benhabib and Laroque, JET 1988.)

[25] This monotonicity property can be established rigorously without assuming the differentiability of the value function. A proof is in Benhabib and Nishimura [1989], in Lemma 1 of the Appendix. Although the model there is slightly different, with very minor modifications the proof goes through.

TABLE 1: Unconstrained ML Estimates (by α = 1/3), Based on (5).

Yearly Data

Observations: 37 Log Likelihood = 101.63 Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.92	0.02	38.84	0.000
θ	0.01	0.09	0.14	0.887
C	-0.36	0.09	-3.72	0.000
λ_1	0.32	0.09	3.24	0.002
λ_2	0.17	0.09	1.80	0.080
α	-0.12	0.09	-1.23	0.225
σ_ϵ^2	0.00014			

Quarterly Data

Observations: 166 Log Likelihood = 547.86 Degrees of freedom: 160

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.002	397.21	0.000
θ	-0.13	0.034	-3.87	0.000
C	-0.01	0.004	-3.72	0.000
λ_1	-0.15	0.039	-3.99	0.000
λ_2	0.19	0.042	4.51	0.000
α	0.35	0.042	8.24	0.000
σ_ϵ^2	0.00007			

TABLE 2: Constrained ML Estimates (by α = 1/3), Based on (5).

Yearly Data

Observations: 37   Log Likelihood = 94.09   Degrees of freedom: 32

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.94	0.02	43.34	0.0[illegible]
θ	-0.53	0.14	-3.66	0.000
C	0.10	0.08	1.23	0.227
λ_1	0.61	0.30	2.03	0.050
λ_2	0.38	0.26	1.45	0.155
σ_ϵ^2	0.00023			

Quarterly Data

Observations: 166   Log Likelihood = 547.82   Degrees of freedom: 161

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.01	51.45	0.000
θ	-0.11	0.04	-2.73	0.006
C	-0.02	0.03	-0.54	0.585
λ_1	-0.16	0.04	-3.81	0.000
λ_2	0.18	0.04	4.39	0.000
σ_ϵ^2	0.00007			

TABLE 3: Unconstrained ML Estimates (by α = 1/3). Yearly Data Only. Based on Ad-Hoc Savings Rule (5)'.

δ = .10

Observations: 37   Log Likelihood = 101.20   Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.00	159.52	0.000
θ	-1.65	0.57	-2.89	0.005
C	0.29	0.08	3.52	0.000
λ_1	0.31	0.17	1.83	0.072
λ_2	0.26	0.16	1.62	0.109
α	0.02	0.10	0.23	0.818

δ = .08

Observations: 37   Log Likelihood = 101.81   Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.00	190.49	0.000
θ	-1.84	0.55	-3.35	0.001
C	0.33	0.07	4.29	0.000
λ_1	0.28	0.17	1.59	0.116
λ_2	0.25	0.16	1.51	0.137
α	0.02	0.09	0.27	0.784

TABLE 4: Constrained ML Estimates (by $\alpha = 1/3$). Yearly Data Only. Based on Ad-Hoc Savings Rule (5)'.

$\delta = .10$

Observations: 37 Log Likelihood = 96.88 Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.00	220.51	0.000
θ	-2.62	0.57	-4.60	0.000
C	0.38	0.08	4.62	0.000
λ_1	0.45	0.19	2.36	0.021
λ_2	0.40	0.18	2.19	0.032

$\delta = .08$

Observations: 37 Log Likelihood = 97.42 Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.98	0.00	245.34	0.000
θ	-2.84	0.60	-4.68	0.000
C	0.41	0.88	4.71	0.000
λ_1	0.41	0.20	2.07	0.042
λ_2	0.38	0.18	2.12	0.038

TABLE 5: ML Estimates Using (Annual) Capital Data: α unconstrained.

Observations: 37   Log Likelihood = 101.68   Degrees of freedom: 31

Parameter	Estimate	Std. Error	t-Stat	P-Value
θ	0.31	0.10	2.93	0.006
α	-0.04	0.12	-0.32	0.749
Const.	-1.16	0.58	-1.97	0.057
λ_1	0.31	0.18	1.70	0.099
λ_2	0.18	0.16	1.09	0.281
ρ	0.83	0.06	12.78	0.000

TABLE 6: ML Estimates Using (Annual) Capital Data: $\alpha = 1/3$ constrained.

Observations: 37   Log Likelihood = 97.47   Degrees of freedom: 32

Parameter	Estimate	Std. Error	t-Stat	P-Value
ρ	0.73	0.07	9.66	0.000
θ	0.23	0.05	4.17	0.000
Const.	-1.44	0.52	-2.75	0.009
λ_1	0.36	0.18	2.03	0.050
λ_2	0.27	0.16	1.68	0.101

TABLE 7: OLS Estimates Using (Annual) Capital Data: Levels

Observations:	37	Degrees of freedom:	34
R-squared:	0.98	Rbar-squared:	0.98
Residual SS:	0.04	Std. error of est.:	0.03
Total SS:	4.14	F(3, 34):	1420.61
P-value:	0.00	Durbin-Watson Stat:	0.60

Parameter	Coeff.	Std. Error	t-Stat	P-Value
Const.	-0.40	1.42	-0.28	0.777
$1-\alpha$	-0.13	0.18	-0.70	0.483
$\alpha+\theta$	1.06	0.09	11.67	0.000

TABLE 8: OLS Estimates Using (Annual) Capital Data: Growth Rates

Observations:	36	Degrees of freedom:	33
R-squared:	0.54	Rbar-squared:	0.51
Residual SS:	0.01	Std. error of est.:	0.01
Total SS:	0.02	F(3, 33):	19.79
P-value:	0.00	Durbin-Watson Stat:	0.78

Parameter	Coeff.	Std. Error	t-Stat	P-Value
Const.	0.02	0.02	1.26	0.215
$1-\alpha$	1.01	0.16	6.13	0.000
$\alpha+\theta$	-0.35	0.75	-0.47	0.641

Table 9: Mean and variance of the cross-country distribution of year-to-year growth rates in GNP*

Year	Mean	Variance (σ_t^2)
1961	0.0509	0.0034
1962	0.0528	0.0038
1963	0.0497	0.0023
1964	0.0527	0.0036
1965	0.0498	0.0032
1966	0.0487	0.0034
1967	0.0427	0.0023
1968	0.0536	0.0024
1969	0.0559	0.0030
1970	0.0543	0.0027
1971	0.0493	0.0027
1972	0.0511	0.0029
1973	0.0577	0.0036
1974	0.0429	0.0036
1975	0.0185	0.0059
1976	0.0526	0.0027
1977	0.0398	0.0028
1978	0.0453	0.0030
1979	0.0323	0.0052
1980	0.0327	0.0037
1981	0.0210	0.0061
1982	0.0023	0.0029
1983	0.0133	0.0033
1984	0.0362	0.0046
1985	0.0263	0.0020

* The Summers-Heston sample. Kuwait excluded. Growth-rates are over the previous year.

Table A.1.

	Case 1	Case 2
	$\lim_{r\to 1}$, $\lim_{\rho\to 1}$; T finite	$\rho, r < 1$, $\lim_{T\to\infty}$
a_{kk}	$\dfrac{2\sigma_\epsilon^2(\alpha+\theta)\left[\dfrac{T}{1-(\alpha+\theta)} - \dfrac{1-(\alpha+\theta)^T}{(1-(\alpha+\theta))^2}\right] + T\sigma_\epsilon^2}{1-(\alpha+\theta)^2} + (1-\alpha)^2\dfrac{2\sigma_w^2(\alpha+\theta)\left[\dfrac{T}{1-(\alpha+\theta)} - \dfrac{1-(\alpha+\theta)^T}{(1-(\alpha+\theta))^2}\right] + T\sigma_w^2}{1-(\alpha+\theta)^2}$	$\dfrac{2\sigma_\epsilon^2}{(1-\rho^2)(1-(\alpha+\theta)^2)}\left[\dfrac{2(\alpha+\theta)\rho}{1-(\alpha+\theta)\rho} + 1\right] + \dfrac{(1-\alpha)^2 2\sigma_w^2}{(1-r^2)(1-(\alpha+\theta)^2)}\left[\dfrac{2(\alpha+\theta)r}{1-(\alpha+\theta)r} + 1\right]$
$a_{k\ell}$	$(1-\alpha)\sigma_w^2 \dfrac{T(1-(\alpha+\theta)) - (1-(\alpha+\theta)^T)}{(1-(\alpha+\theta))^2}$	$2(1-\alpha)\dfrac{\sigma_w^2}{1-r^2}(1-(\alpha+\theta)r)^{-1}$
$a_{\ell\ell}$	$T\sigma_w^2$	$2\sigma_w^2/(1-r^2)$
a_{ku}	$\dfrac{\sigma_w^2[T(1-(\alpha+\theta)) - (1-(\alpha+\theta)^T)]}{(1-(\alpha+\theta))^2}$	$\dfrac{2\sigma_\epsilon^2\rho}{(1-\rho^2)(1-\rho(\alpha+\theta))}$

<u>Note</u>: To obtain the a_{ij}, the expressions in the first column of the table should be divided by T^2. The second column reports $\lim_{T\to\infty} T^2 a_{ij}$.

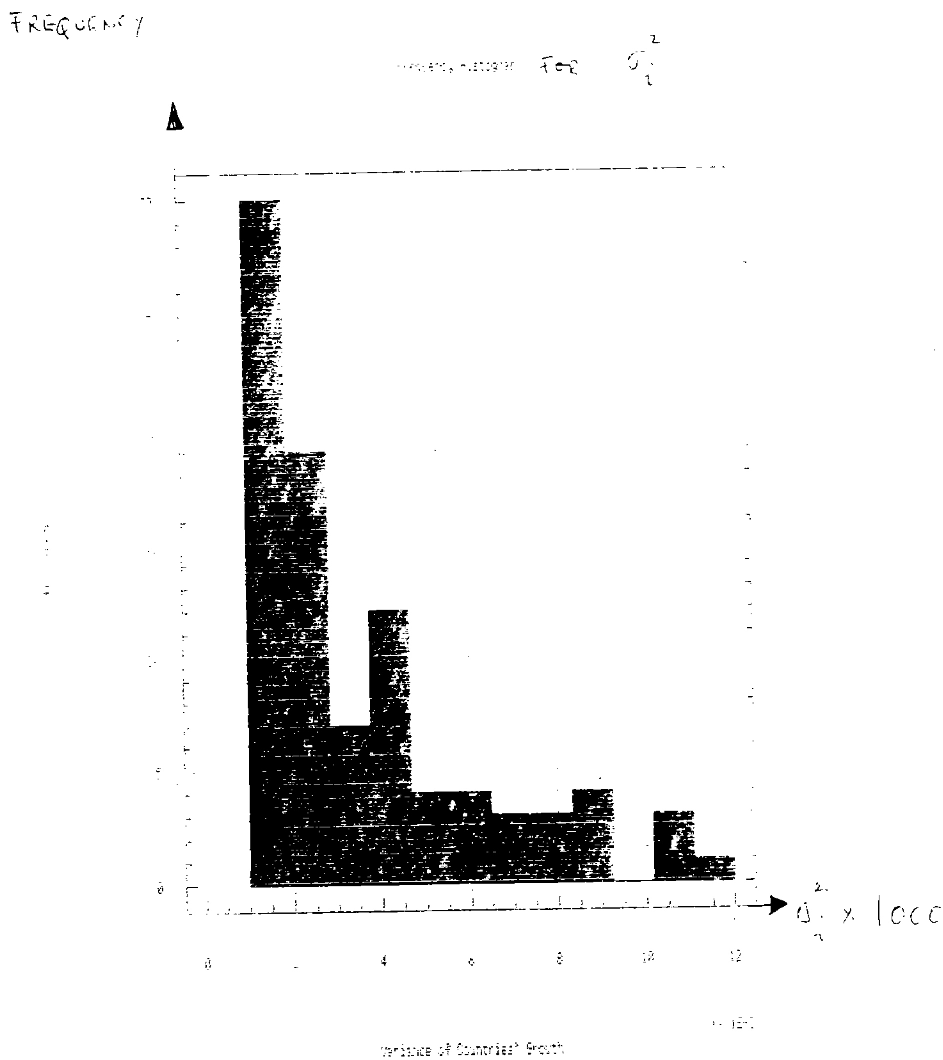
FREQUENCY
FOR
0
4
6
8
10
12
X 1000
Variance of Countries' Growth

FIGURE 1