Externalities and Growth Accounting
Benhabib, Jess;Jovanovic, Boyan
The American Economic Review; Mar 1991; 81, 1; ProQuest Central
pg. 82

Externalities and Growth Accounting

By JESS BENHABIB AND BOYAN JOVANOVIC*

This paper tackles two puzzles: the high empirical elasticity of aggregate output with respect to the measured capital input and the seemingly high variability of growth rates over countries in the medium run. We find that one need not invoke increasing returns or externalities to capital to explain these two puzzles. Rather, they are consistent with a constant-returns-to-scale aggregate production function, so long as the exogenous Solow residual process has enough persistence in it. In our model, causality runs exclusively from knowledge to capital, and therefore the apparent absence of an external effect to the capital input says nothing about the importance of spillovers in the creation of knowledge. (*JEL* 110)

This paper addresses the question of how to explain the variation in the levels and rates of growth of output across countries. It focuses in particular on the question of whether or not this cross-country variation offers support for the suggestion that there are aggregate increasing returns to capital and labor caused either by external effects associated with capital investment or by a secular increase in the variety of intermediate inputs. By looking again at the evidence considered by Paul Romer (1987) and Lawrence Christiano (1987), we show that, under plausible assumptions about the behavior of the economy, there is no support for the assertion that capital-related externalities are present. We show instead that the variation in countries' growth rates is consistent with each country having the same constant-returns-to-scale production function and with a stochastic process for technological change that is the same across different countries but starts from different initial positions.

*Professors of Economics, New York University, 269 Mercer Street, New York, NY 10003. We thank The C. V. Starr Center 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, Massachusetts, in October 1989.

A. *The Issues*

Two issues are at hand. The first concerns the prediction of Robert Solow's (1957) model that the elasticity of output with respect to capital should equal capital's share in output, which is roughly one-third. Yet when one looks at data over longer periods of time and in several different countries, as Romer (1987) has, one finds an elasticity that is closer to unity. There are three extensions of Solow's model that can generate this. One is to make the identifying assumption that Romer makes, at least implicitly, that the fundamental exogenous variation is in the rate of savings and investment. In this case, one must change the elasticity of output with respect to capital by invoking capital-related external effects or increasing returns stemming from input variety. The second possibility is to make the identifying assumption that Christiano (1987) makes and that we discuss in Section IV, namely that the exogenous variation across countries is in the underlying exogenous process of technical change. The third alternative, suggested here, is to assume that the fundamental stochastic process for technology is the same but that in the sample the realized paths differ across countries. In this case, something like Christiano's explanation can be constructed (and no capital externalities are needed), but this can be done without invoking his unattractive "fixed effects" as-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

sumption that there are permanent, exogenous differences in the technology faced by different countries. Rather, the difference across countries lies in the initial conditions and implicitly in the sequence of historical accidents that lead to those conditions. We deal with this first issue in Section IV.

The second issue is whether (our version of) Solow's model can be reconciled with the seemingly large variation in countries' growth rates since World War II. We shall argue that this variation is roughly consistent with Solow's model. Part of the new evidence that we bring to bear on this assertion is the finding that, for the population of countries as a whole, the time-series variation in output growth in a representative country is consistent with the cross-country variation in output growth measured over 25 years. This bears on the question of the inherent differences, if any, that must be invoked to explain the disparate behavior of the different countries. We deal with this second issue in Section III.

B. *Capital and Knowledge*

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. However, 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 is positively related 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 his pioneering article, Romer (1987) offers two separate models each of which can generate a capital elasticity of output larger than one-third. Within the context of his first model, he 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 second model introduces knowledge explicitly in the form of ideas for intermediate goods. In its implications for comovements between aggregate output, capital, and labor, this model is observationally equivalent to a version of his first model, as is evident in his equation 11. This model need not have any *direct* spillovers of knowledge to yield aggregate increasing returns, although in later versions Romer introduces direct spillovers in the research sector. The "externality" that the final goods sector enjoys is, as Romer points out, a pecuniary one, and a divergence between equilibrium and social optimum could arise solely because of the monopoly power introduced into the intermediate-goods sector so as to provide incentives for inventions.

Romer's two models have a common feature: growth in capital *causes* a growth in knowledge or a growth in the availability of specialized inputs, or both.[2] Evidence of direct spillovers of knowledge then constitutes support for Romer's first model. On the other hand, evidence that *pecuniary* spillovers are present (and that such spillovers increase with the capital stock) would provide support for the second model, so long as some increasing returns (as inputs *and* variety vary) are also present. Un-

[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. Martin N. Baily (1987) and others have made the same point. We return to Christiano's argument in Section IV.

[2]See especially Romer's reference to evidence from Jacob Schmookler (1966) to the effect that in various industries patenting tends to *follow* investment.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

fortunately, the available evidence often mixes the two types of externality (pecuniary and nonpecuniary) so that their relative (as well as absolute, it turns out) importance is hard to pin down.

We begin with the assumption, implicit in Romer's first model, that an increase in a firm's capital stock causes the firm's productive knowledge to go up in the same proportion, so that we can use estimates from micro data on externalities in R&D as an estimate of the size of the capital externality. If direct spillovers to R&D do exist, they are likely to be largest among firms in the same industry, since those firms are likely to be using similar technologies. The largest *intraindustry* spillover estimates were obtained by Jeffrey Bernstein and Ishaq M. Nadiri (1989), who find that, in four industries, the social returns to intraindustry spillovers of R&D ranged from 30 percent to 123 percent of the private returns to R&D.[3] Such large estimates are, however, exceptions: in a summary of the literature on the elasticity of output with respect to the R&D input, Zvi Griliches (1988 p. 15) 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] 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. However, this is exactly what Romer's first argument implies, and the micro data do not seem to support it.

The micro evidence does not seem to favor Romer's second argument either. Frederick M. Scherer (1982) finds that productivity growth in an industry is strongly correlated with the extent to which it purchases R&D-embodying products. Additionally, a wealth of evidence points to sustained cost reductions in a whole range of intermediate-inputs services. It is beyond dispute, therefore, that final-goods producers have benefitted from sustained improve-

[3]Edwin Mansfield et al. (1977) also find large spillovers for a select group of innovations, but since these were all *successful* innovations, their sample 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 percent, while that of the private rate averaged 33 percent. These results do support Romer's claim that the social returns might exceed the private rate by a factor of more than two. Bernstein and Nadiri's results must also be viewed with caution, because they are based on a deterministic model, and simultaneity biases are likely to be present because of omitted 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. Similarly, Adam Jaffe's (1986) finding that there were significant spillovers in a cross section is questionable on grounds that his assumption (on p. 992) about an absence of correlation between his instrumental variables and the unobserved "technological opportunity" parameter that each firm faces is unlikely to be met.

[4]Ariel Pakes and Mark 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, to impose assumptions 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 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, as Griliches (1979) points out, the value of private knowledge may 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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

ments in input quality. This does not mean however, that there are increasing returns in the aggregate. In fact, we know of no evidence that (a) aggregate returns to measured inputs *and* variety increase or (b) the provision of variety is fueled by a larger stock of capital.

In short, the micro data are so far not conclusive on Romer's hypotheses. Our aim here is to show that *aggregate* data are consistent with a view that neither direct nor pecuniary spillovers are fueled by physical capital. In doing this, we leave open the possibility that there are increasing returns due to something else—human capital, perhaps.

C. *Our Argument*

In our model, 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 *nothing* 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 Andrei Shleifer [1986], for instance) to imagine that new ideas precede the installation of the capital equipment needed to implement them. Moreover, at the level of the individual firm at least, micro data indicate that R&D Granger-causes investment, but that investment does not Granger-cause R&D (Saul Lach and Mark Schankerman, 1989).

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 as follows. 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. Moreover, the positive association between knowledge shocks and capital investments also seems to explain most of the variance in countries' growth rates in the postwar period. No externalities or increasing returns are needed.

Section I presents our model, which consists of five structural equations. In Section II, we present maximum-likelihood and least-squares estimates for the model using postwar quarterly and annual U.S. data and find no evidence of an externality. In Section III, 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 IV, we reinterpret Romer's regression results (which 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

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

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

evidence for the conjectured positive externality to the capital input.

After the empirical evidence discussed in Sections II–IV, Section V presents two models that give rise to the structural equations first introduced in Section I. 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.

I. The Augmented Solow 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

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

In the first version of Romer's model, the parameter θ measures the external effect of capital, an effect that the firm ignores when making its decisions. In the second version, θ represents increasing returns in the variety of intermediate inputs whose quality is (in an auxiliary equation) linked to the economy-wide capital stock. Since all firms are the same, $K_t = \bar{K}_t$. Letting lowercase symbols denote logarithms, (1) reads

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

When $\alpha+\theta$ is unity, this equation is the same as Romer's equation 11. 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:

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

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

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

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 domestically as a by-product of everyday economic activity. At any rate, the economy's long-run growth rate in GNP per capita will be $\mu(1-\alpha)$ if ρ is unity and 0 if ρ is less than 1. On these grounds, we expect ρ to be roughly equal to unity. Second, the parameter ρ, or rather $1-\rho$, measures the rate at which knowledge depreciates.[7] From the work on business cycles by Edward C. Prescott and his coworkers (e.g., Prescott, 1986), we expect ρ to be about 0.99 in the quarterly U.S. data and 0.96 in the annual data. An important omission here is a purely transitory component to z. Its inclusion would almost certainly raise the estimates of ρ that we present in the next section and would work in our favor. In other words, the inclusion of transitory monetary and other policy effects on the measured Solow residual and even of errors in measuring z_t would, as will become clear below, help rather than hurt our case, which hinges on ρ being close to 1. 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 VI, we shall present two separate micro models[8] that imply the following equation as governing competitive equilibrium allocations:

$$(5)\qquad k_{t+1} = \gamma + y_t$$

where γ is a constant.[9] For some of the

[7]Because it is superseded by other knowledge. Actually, it is not knowledge but rather its economic value that depreciates.

[8]One of these is in many respects similar to the model that Edward 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-percent depreciation of capital. The details are in Section V.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

results in Sections II and IV, we also assume that

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

This assumption, which is also given a micro justification in Section VI, is not used in Section III, however, where we interpret long-run growth differentials in countries' growth experience.

Equations (2)–(5) and assumption (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 II uses postwar U.S. data, Section III uses Alan Heston and Robert Summers's (1984) data, and Section IV uses the longer-run data that Romer compiled from Angus Maddison (1982) and elsewhere. None of these bodies of data supports the hypothesis that θ is positive.

II. Estimates From Postwar U.S. Data

We begin our empirical inquiry by looking at the postwar U.S. data. The reader will not be surprised to learn these data offer no support for a positive θ, since (a) Romer himself did not cite these data as supportive of capital externalities, (b) Prescott (1986) has, with some success, used a model quite similar to ours but with θ set equal to 0 to fit detrended postwar U.S. data, and (c) the short-run data are notorious in that output fluctuations are explained almost entirely by variations in hours worked and hardly at all by the measured capital input.

A problem presented by the capital input is that it is likely to be poorly measured, at least at high frequencies, because of variation in its utilization rate. Our estimation procedure in this section begins by treating k_t as 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

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

This is the equation that formed the basis for the estimates reported in the first two tables, which used data on log(GNP) for y_t and log(hours worked) for l_t. The data are *not* detrended. We present two sets of estimates. Table 1 reports the unconstrained maximum-likelihood estimates[10] for the annual and the quarterly data separately. Table 2 reports estimates for the remaining parameters when α is constrained to equal $\frac{1}{3}$ (i.e., capital's postwar 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 that of Prescott (1986 p. 15). The estimate of θ is close to 0 and does not differ significantly from 0. When α is freed up, its estimate does not

However, for a model with less than 100-percent 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, which 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 V (especially see Lemma 1 and the surrounding discussion). Moreover, in Section II we shall also present the estimates for the model's parameters when the evolution of the capital stock obeys $K_{t+1}=sY_t+(1-\delta)K_t$ instead of (5) (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

$$y_t-\rho y_{t-1}=C+(\alpha+\theta)(y_{t-1}-\rho y_{t-2})+(1-\alpha)(l_t-\rho l_{t-1})+\omega_t$$

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 presample error terms to their zero means. For y, we use log(GNP) in 1982 dollars; for l, we use log of total hours worked; and for k, we use the log of total (private fixed plus government) capital, all for the period 1947–1985, or 1986.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

TABLE 1—UNCONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES, BASED ON (5)

Data	Parameter	Estimate	SE	t	P
Yearly[a]	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[b]	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			

Note: The estimates are unconstrained in that α need not equal $\frac{1}{3}$.
[a]Log likelihood = 101.63 (37 observations, 31 degrees of freedom).
[b]Log likelihood = 547.86 (166 observations, 160 degrees of freedom).

TABLE 2—CONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES, BASED ON (5)

Data	Parameter	Estimate	SE	t	P
Yearly[a]	ρ	0.94	0.02	43.34	0.000
	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[b]	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			

Note: The estimates we constrained in that $\alpha = \frac{1}{3}$.
[a]Log likelihood = 94.09 (37 observations, 32 degrees of freedom).
[b]Log likelihood = 547.82 (166 observations, 161 degrees of freedom).

differ significantly from $\frac{1}{3}$. All in all, then, the model does pretty well with the quarterly data.

Such is not the case with the annual data. The unconstrained estimate of α is negative, but not significant; and, θ is positive, but small and nonsignificant. 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 $\frac{1}{3}$, θ is large but negative and highly significant. The likelihood-ratio test resoundingly rejects the restriction that $\alpha \equiv \frac{1}{3}$. These results with annual data are quite similar to the regression results that Romer reports 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

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

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 it also takes time for the external benefits of capital accumulation to be felt. That is, it is possible not only that there are significant building-time delays, but that 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 represent 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 0, and for values $s \equiv p > 0$, it tended to be negative.

A third set of problems surround our assumption in equation (5) about the way in which capital evolves. Distinct from the issues discussed above 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 equation (5) implies that there is 100-percent depreciation. On the other hand, if Y_t is interpreted as wealth, then we have not measured wealth correctly; instead we should have used net national product plus the entire existing capital stock. However, 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 equation (5) (for which a micro-based justification exists [see Section III]), 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 $\alpha = \frac{1}{3}$) estimates of the parameters when (5′) is used in place of (5). Only annual data are used, since we did not have a long enough quarterly time series at the depreciation rates that are commonly used.[11] The first set of estimates in Tables 3 and 4 sets δ at 10 percent; the second sets it at 8 percent. The parameter C is the same as before (see footnote 10) with $\gamma = \ln s$. Further detail is 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 0, 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 equal $\frac{1}{3}$, yet θ is

[11] We did experiment with postwar 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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

TABLE 3—UNCONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES, YEARLY DATA ONLY, BASED ON AD HOC SAVINGS RULE (5′)

δ	Parameter	Estimate	SE	t	P
0.10[a]	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
0.08[b]	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

Note: The estimates are unconstrained in that α need not equal $\frac{1}{3}$.
[a]Log likelihood = 101.20 (37 observations, 31 degrees of freedom).
[b]Log likelihood = 101.81 (37 observations, 31 degrees of freedom).

TABLE 4—CONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES, YEARLY DATA ONLY, BASED ON AD HOC SAVINGS RULE (5′)

δ	Parameter	Estimate	SE	t	P
0.10[a]	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
0.08[b]	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

Note: The estimates are constrained in that $\alpha = \frac{1}{3}$.
[a]Log likelihood = 96.88 (37 observations, 31 degrees of freedom).
[b]Log likelihood = 97.42 (37 observations, 31 degrees of freedom).

TABLE 5–UNCONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES USING ANNUAL CAPITAL DATA

Parameter	Estimate	SE	t	P
θ	0.31	0.10	2.93	0.006
α	−0.04	0.12	−0.32	0.749
C	−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

Notes: The estimates are unconstrained in that α need not equal $\frac{1}{3}$. Log likelihood = 101.68 (37 observations, 31 degrees of freedom).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

TABLE 6—CONSTRAINED MAXIMUM-LIKELIHOOD ESTIMATES USING ANNUAL CAPITAL DATA

Parameter	Estimate	SE	t	P
ρ	0.73	0.07	9.66	0.000
θ	0.23	0.05	4.17	0.000
C	−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

Notes: The estimates are constrained in that $\alpha = \frac{1}{3}$. Log likelihood = 97.47 (37 observations, 32 degrees of freedom).

TABLE 7—OLS ESTIMATES USING ANNUAL CAPITAL DATA: LEVELS

Parameter	Coefficient	SE	t	P
C	−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

Notes: $R^2 = 0.98$, $\bar{R}^2 = 0.98$, residual SS = 0.04, SE of estimate = 0.03, total SS = 4.14, $F_{[3,34]} = 1{,}420.61$, $P \simeq 0.00$, Durbin-Watson statistic = 0.60 (37 observations, 34 degrees of freedom).

TABLE 8—OLS ESTIMATES USING ANNUAL CAPITAL DATA: GROWTH RATES

Parameter	Coefficient	SE	t	P
C	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

Notes: $R^2 = 0.54$, $\bar{R}^2 = 0.51$, residual SS = 0.01, SE of estimate = 0.01, total SS = 0.02, $F_{[3,33]} = 19.79$, $P \simeq 0.00$, Durbin-Watson statistic = 0.78 (36 observations, 33 degrees of freedom).

still positive and significant. This is the only solid piece of evidence in favor of Romer's hypothesis that we can find in the postwar data. At the same time, the estimate of ρ is surprisingly low. Tables 7 and 8 present ordinary least-squares (OLS) estimates of the aggregate production function. If ρ is close to 1, and if (5) is appropriate for annual (as opposed to quarterly) data, then according to the model in equations (2)–(5) and assumption (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 we have added to Romer's model do not substantially improve the ability of his model to rationalize high-frequency data. We therefore agree with Romer's (1987 p. 186) view that data from more countries and longer 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.

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

Under certain assumptions, the Heston-Summers panel data on countries' GNP's will provide additional information on the parameters of the model. This is what we

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

examine next. We assume that all countries have the same production functions and tastes and that the only differences among them are their initial values for k_t, l_t, and z_t. In particular, z_t obeys the same process in all countries, although its realizations can vary. 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 unobservable, some further assumptions will now be added. First, we shall assume that $\lambda_1 = \lambda_2 = 0$ in equation (4), so that $\omega_t = \varepsilon_t$. This is done for analytical convenience, and it ought not to make much quantitative difference in this section, where we examine growth rates over a period of 25 years (and not at annual or quarterly growth rates, as was done in the previous section), 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

$$l_t = m + rl_{t-1} + w_t \qquad |r| \leq 1$$

where w_t is independently and identically distributed and independent of ε_t. We estimated this equation using U.S. annual data and OLS and obtained

$$\underset{}{l_t} = \underset{(0.18)}{-0.22} + \underset{(0.02)}{1.03 l_{t-1}} \qquad R^2 = 0.98 \quad \text{DW} = 1.85.$$

Our maximum-likelihood results together with this 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 Heston-Summers sample.[13]

A. *Gibrat's Law in Growth of GNP*

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 Heston-Summers 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:

$$(8) \quad \Delta y_i = \underset{(0.015)}{0.047} - \underset{(0.001)}{0.0004 y_i} \qquad i = 1, \ldots, 115$$
$$\text{residual variance} = 0.0004$$

where y_i is the logarithm of 1960 GNP for country i and Δy_i is its growth per year over the 1960–1985 period (standard errors are in parentheses). Thus, the updated sample roughly confirms the nonsignificant relationship between growth and initial size that 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

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

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

[13] Robert Barro's (1988) cross-country study of the univariate process for log(unemployment) (again, with annual data) revealed some significant cross-country differences in the degree of persistence in that variable. Nevertheless, at least in the postwar 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] This finding is for countries as a group, most of which are small and have little R&D investment. For industrialized countries, the result is somewhat different (see William Baumol and Edward Wolff, 1988; Bradford DeLong, 1988).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

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

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

The form of equations (9) and (10) depends, of course, on the savings rule in equation (5), and the theoretical justification that we provide for this savings rule rests on the assumption that capital depreciates fully each period. However, this section and the next both look at data at a frequency no greater than once every 25 years, and so this assumed depreciation rate may approximate reality quite well in this context.

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

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

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

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

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 given 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:

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

Substituting from (13) into (12) and noting (once more from Appendix 1) that

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

yields

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

Therefore, if ρ and r are roughly 1, Gibrat's Law will hold *regardless of the value of* θ. This means that the failure of GNP levels to converge does not identify θ.

[16]Because y_t acquires a permanent component if $\rho = 1$ *or* if $\alpha + \theta = 1$. Therefore, the findings of Charles Nelson and Charles Plosser (1982), John Campbell and Gregg Mankiw (1987), and John 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, by themselves tell us whether $\alpha + \theta = 1$, or whether ρ or r is equal to unity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

As a caveat, we point out that our analysis treats countries as closed economies and looks for scale effects or spillover effects within but not across 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, differences in culture and language and the presence of capital controls and other trade barriers do support the use of geographical borders to delineate the extent of the market.[17]

B. *Growth in the Cross Section and in the Time Series*

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}); instead, it 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

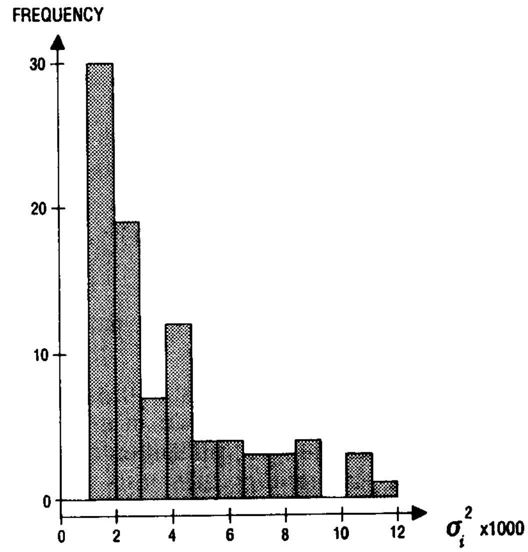

FIGURE 1. THE EMPIRICAL DISTRIBUTION OF COUNTRY-SPECIFIC VARIANCES OF GROWTH

the growth rate of GNP in country i between periods t and $t+1$. If the hypothesis is true, the distribution of g_{it} $(t = 1960,\dots,1984)$ for each fixed i should be roughly the same as the distribution of g_{it} $(t=1,\dots,115)$ for each fixed t.[18] In particular, if σ_i^2 and σ_t^2 are 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 *on average* they are roughly the same:

$$\left(\frac{1}{115}\right)\sum_{i=1}^{115}\sigma_i^2 = 0.0035$$

$$\left(\frac{1}{25}\right)\sum_{t=1960}^{1984}\sigma_t^2 = 0.0034.$$

The variability of σ_t^2 is documented in Table 9, and the variability of σ_i^2 is described in

[17]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. 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 presented in Jovanovic and Rafael Rob (1989) and Jovanovic and Glenn MacDonald (1988). In both of these theoretical models, those who are further behind learn more (through imitation) than those who are closer to the leaders, simply because they have more to learn. This argues for a higher μ for the poorer nations. However, 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 have already created and put in place (see Raymond Vernon [1989] for a viewpoint that emphasizes these constraints).

[18]Actually, this is true only if g_{it} and g_{jt} are independent random variables for $i \neq j$. In fact, Table 9 reveals significant time effects in the world's mean growth rates, especially after the first oil shock. This seems to imply a fair degree of contemporaneous covariation in countries' growth rates from 1974 on. If so, the variance of growth rates over countries at a given date should in fact be less than the variance of growth rates over time in a single country.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

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

Notes: Data are from the Heston-Summers sample, with Kuwait excluded. Growth rates are over the previous year.

the histogram in Figure 1 (in which Iraq, whose $\sigma_i^2 \times 1{,}000 = 30$, was omitted).[19] 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) for fixed i, the observations g_{it} are autocorrelated.

C. *Medium-Run Differences in Growth Among Countries*

While the variability of growth rates among countries does not seem to differ from the variability of growth rates within countries, one might still wonder whether differences in growth rates among countries are too *persistent* 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?

The variance of growth within a country over time ultimately depends on the variance of the process $\eta_t = \varepsilon_t + (1-\alpha)w_t$. The first step is to see what values of σ_ε^2 and σ_w^2 are compatible with the variance of 25-year annualized growth rates. In Appendix 1, we have calculated the variance of the steady-state distribution of Δy as a function of σ_w^2 and σ_ε^2. 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 A1 (Appendix 2), where it is denoted by a_{kk}. If we hypothesize the truth of the Solow neoclassical model and insert $T = 25$, $\theta = 0$, and $\alpha = \frac{1}{3}$ in this expression, it reads $a_{kk} = 24\sigma_w^2 + (48.5)\sigma_\varepsilon^2$. This expression, when divided by 25^2, should equal the *empirical* value of the cross-country variance of the 25-year averaged growth rates. This value turns out to be equal to slightly less than the variance of the residual in equation (8), namely 0.000355. Thus, setting $a_{kk}/25^2$ equal to this number yields a linear restriction on σ_w^2 and σ_ε^2, namely

$$\left(24\sigma_w^2 + 48.5\sigma_\varepsilon^2\right)/25^2 = 0.000355.$$

This is the restriction on σ_w^2 and σ_ε^2 that will generate the medium-run growth differentials across countries that we observe.

If one were to substitute the U.S. values for σ_w^2 and σ_ε^2 into the above expression, one would obtain an expression for its left-hand side that is much smaller than the right-hand side. If ρ and r are both unity, one can estimate σ_w^2 and σ_ε^2 by the variance of Δl and Δz, respectively. Doing this for the postwar annual U.S. data yields $\sigma_w^2 =$

[19] It turns out that σ_i^2 is significantly negatively correlated with $y_{i,1960}$. That is, initially larger countries have less-variable growth rates. For instance, for the United States, $\sigma_i^2 = 0.0006$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

0.00044 and $\sigma_\varepsilon^2 = 0.00037$.[20] Substituting these values into the above expression yields just 0.000046, which is too small by a factor of 7.7.

This is not the end of it, however, because of the heteroskedasticity of growth rates over countries. The variance of the U.S. growth rate is 0.0006, whereas the growth-rate variance for the median country is about 0.003, and its mean is 0.0035. Therefore, the variability of growth in the "average" country exceeds that of the United States by a factor of about 5 or 6. Since the U.S. variability underestimates the right-hand side of the above equation by a factor of 7.7, it is likely that, if we had estimates of σ_w^2 and σ_ε^2 from the *average* country, we would have explained roughly 65–78 percent of the cross-country variability in growth rates. The discrepancy is therefore far smaller than one would have thought; to account for it, one or more of the parameters that we have assumed to be the same for all countries might have to be made country-specific.[21]

The above arguments suggest that the augmented Solow model with $\theta = 0$ is consistent with the bulk of cross-country growth differentials in the medium run. However, it is clearly not consistent with the tremendous heteroskedasticity in yearly growth rates that Figure 1 highlights, although such heteroskedasticity could also have been produced by measurement errors with country-specific variances.

While the above discussion leaves some unanswered questions about our model's ability to explain (a) the lack of convergence of GNP *levels* and (b) the existence of persisting differentials in growth *rates*, 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 (1987) 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. Under independence, 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.

IV. The Relation Between Inputs and Output Over Longer Epochs

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

$$(15) \quad T_{it}^{-1}\Delta y_{it} = b + b_k T_{it}^{-1}\Delta k_{it} + b_l T_{it}^{-1}\Delta l_{it} + u_{it}.$$

Romer uses 18 observations that span seven countries (subscript i), and four epochs (subscript t) of at least 30 years in length; the measure of the labor input is hours worked. The least-squares regression results that he reports in his equation 18 are: $\hat{b}_k = 0.87$ with a standard error of 0.08, and $\hat{b}_l = 0.04$ with a standard error of 0.18. Our aim here is to calculate the expectations of $\hat{b}_k$ and $\hat{b}_l$ 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 carets, are identi-

[20] The covariance between Δl and Δz is 0.00012. Ignoring it, as we do here, hurts rather than helps our case. Note also that the estimate of σ_ε^2 that we have calculated here is actually consistent with its estimates for the annual data in Tables 1 and 2, because, since we are assuming no moving-average terms in this section, the relevant comparison is with $(1+\lambda_1^2+\lambda_2^2)\sigma_\varepsilon^2$ in those tables.

[21] For instance, the parameter μ might have to be made country-specific. 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 section.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

cally equal to[22]

$$\begin{bmatrix}\hat{b}\\ \hat{b}_k\\ \hat{b}_l\end{bmatrix}=\begin{bmatrix}0\\ \alpha+\theta\\ 1-\alpha\end{bmatrix}$$

$$+\begin{bmatrix}n & \bar{k} & \bar{l}\\ \bar{k} & a_{kk} & a_{kl}\\ \bar{l} & a_{kl} & a_{ll}\end{bmatrix}^{-1}\begin{bmatrix}\bar{a}\\ a_{ku}\\ a_{lu}\end{bmatrix}$$

from which one can show that, since $Ea_{lu}=0$ by equation (6),

$$(16)\quad E\begin{bmatrix}\hat{b}_k\\ \hat{b}_l\end{bmatrix}=\begin{bmatrix}\alpha+\theta\\ 1-\alpha\end{bmatrix}+\left(E\frac{1}{a_{kk}a_{ll}-a_{kl}^2}\times\begin{bmatrix}a_{ll} & -a_{kl}\\ -a_{kl} & a_{kk}\end{bmatrix}\begin{bmatrix}a_{ku}\\ 0\end{bmatrix}\right)$$

$$=\begin{bmatrix}\alpha+\theta\\ 1-\alpha\end{bmatrix}+\left(E\frac{1}{a_{kk}a_{ll}-a_{kl}^2}\times\begin{bmatrix}a_{ll}a_{ku}\\ -a_{kl}a_{ku}\end{bmatrix}\right).$$

Since the a_{ij} are all positive, $\hat{b}_k$ will be biased upward while $\hat{b}_l$ will be biased downward, with the bias on $\hat{b}_k$ equalling $-a_{ll}/a_{kl}$ times the bias on $\hat{b}_l$. Appendix 2 calculates the a_{ij} on the assumption that all countries are subject to the same stochastic process but face different realizations of the ε's and w's, as well as different initial conditions. The resulting expressions for the a_{ij}'s are quite messy, but the following limiting results are worth noting:[23]

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

$$(18)\quad \lim_{r,\rho\to 1}\left\{\lim_{T\to\infty}\left[E(\hat{b}_l)\right]\right\}=0.$$

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. Therefore, $E(b_l)$ 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, $\hat{b}_k=0.87$ and $\hat{b}_l=0.04$.

Table A1 of Appendix 2 reports the expressions for the a_{ij} that one can use to calculate the bias in the least-squares estimates $\hat{b}_k$ and $\hat{b}_l$ for the cases in which (a) T remains finite but ρ and r tend to unity and (b) ρ and r are less than unity but T goes to infinity. In both cases, the bias remains positive but is 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 $\hat{b}_l$ would be expected to take for large T and for r and ρ reasonably close to 1.

Equations (17) and (18) are the same as what Christiano (1987) obtains 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

[22]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, $a_{kl}=\Sigma_{i,t}T_{it}^{-2}\Delta k_{it}\Delta l_{it}$, and so on. The variables with overbars are the mean growth rates over the sample. For instance, $\bar{k}=\Sigma_{i,t}T_{it}^{-1}\Delta k_{it}$, and so on.

[23]We present the results only for this particular limiting case because the general expressions would be very lengthy. Table A1 in 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.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

that Romer reports in his equation (18) are consistent with θ being 0.

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 (1987) fixed-effects model. These differences are best explained under the assumption that $\sigma_w^2 = 0$ (i.e., that labor supply is nonrandom). In the fixed-effects model, a country's long-run growth rate is $\mu_i/(1-\alpha)$. In our model, the long-run growth rate for z is

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

Since μ is the same over countries, countries must, in the long run, all grow at the same rate regardless of the value of ρ. Thus, a fixed-effects model delivers a positive variance of long-run growth rates among countries, while ours does not.[24] Our view gets further support from recent empirical work by Danny Quah (1990), who argues that, in the Heston-Summers data, income growth rates show convergence to a common value.

In our model, the upward bias on the capital coefficient is positively related to ρ. Continuing with the case in which labor input is nonrandom ($\sigma_w^2 = 0$), we find from the second column of Table A1 that, as T approaches ∞, the bias on the capital coefficient approaches

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

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.

Aside from the hypotheses that Romer, Christiano, and we have advanced, yet another argument can be advanced to rationalize the high estimated coefficient of Δk and the low estimated coefficient of Δl: the growth of the physical capital stock may be strongly positively correlated with growth of labor quality, while growth in hours worked may be negatively related to the growth of labor quality. The latter association is well documented at high frequencies in the micro data, and Romer's (1987) figure 1 provides some indirect evidence for it at lower frequencies for the United States. Since the empirical estimates of equation (15) use hours worked unadjusted for quality, this would tend to cause an upward bias on capital's coefficient and a downward bias on labor's coefficient. This hypothesis, which does not revolve around either external effects or increasing returns, still awaits a careful theoretical treatment, but whatever its eventual fate, it is important to bear in mind that it is independent of the other hypotheses discussed here. Of course, if one were to insist that increases in the capital stock *cause* measured equiproportionate improvements in unmeasured labor quality, then one is back in a framework captured by equation (1), with, say $\bar{K}^{\theta}$ denoting unmeasured labor quality, and this is a framework that the data do not seem to support.

V. 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)–(5) and assumption (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 not only reflects the internal and external impact of contemporaneous capital on output but contains an additional component

[24] Table A1 contains information about the speed of convergence to zero of variables such as a_{kk}, the variance of the long-run growth rate of 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 $\rho < 1$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

through the joint dependence of output and capital on the previous productivity shock. Ignoring 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 previous section, especially equation (5) and assumption (6).

We start with a special OLG model in which 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

$$\beta \ln c_t + (1-\beta) E \ln(w_t - c_t) r_{t+1}$$

which 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 independently and identically distributed with mean zero. The wage rate and the 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 equation (5) and assumption (6) of the previous section.

Before moving on to 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 in fact decrease savings and next period's capital stock if the direct wealth effect through wages is dominated, 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 nondecreasing function of the interest rate.[25]

We now turn to the specification with an infinitely lived representative agent. Before exploiting specific functional forms, we present a general version to pinpoint again 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. We may specify a_t 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

[25]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 Guy Laroque [1988]).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

$Y_t = Z_t f(K_t, \bar{K}_t, L_t)$, where Z_t is the stochastic shock to the production function and $\bar{K}_t$ ($\equiv K_t$) 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 $E\Sigma_0^\infty \beta^t U(c_t, a_t - L_t)$ 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

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

To simplify matters, we assume that 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 Lawrence Blume and David Easley [1982]). Let the derivatives of $V(K, Z, a)$ with respect to K and Z be denoted by V_k and V_z and let the second-order derivatives be denoted by 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:

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

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

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 derivatives of L_0 with respect to Z_0 and K_0.

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

$$dK_1/dK_0 = \big[(U_{LL} + U_c Z_0 f_{LL})(U_{cc} Z_0 F_k) - U_{cc} Z_0 f_L U_c Z_0 f_{kL}\big]/D > 0$$

and

$$dK_1/dZ_0 = \big[(U_{LL} + U_c Z_0 f_{LL}) \times (U_{cc} f - \beta E V_{kz}\, dZ_1/dZ_0) - U_{cc} Z_0^2 f_L^2 \beta E V_{kz}\, dZ_1/dZ_0\big]/D$$

where

$$D \equiv (U_{LL} + U_c Z_0 f_{LL})(U_{cc} + \beta E V_{kk}) + \beta E U_{cc} V_{kk} Z_0^2 f_L^2 > 0$$

$$F_k \equiv f_k(K_0, \bar{K}_0, L_0) + f_{\bar{k}}(K_0, \bar{K}_0, L_0) + (1-\delta)$$

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.[26] Also, $dK_1/dZ_0 > 0$ if $V_{kz} > 0$. To evaluate V_{kz} we first compute

$$V_k(K_0, Z_0, a_0) = U_c Z_0\big[f_k + f_{\bar{k}} + (1-\delta)\big]$$

[26]This monotonicity property can be established rigorously without assuming the differentiability of the value function. A proof is in Benhabib and Kazuo Nishimura (1989), in lemma 1 of that paper's appendix. Although the model there is slightly different, with very minor modifications the proof applies to the present case.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

to obtain

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

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 model with logarithmic utility and Cobb-Douglas production with full depreciation are to eliminate the possibility of a saving policy and a labor supply which both 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 the following lemma.

LEMMA 1: *Let $K_{t+1} = h(K_t, Z_t)$, where h 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.*

PROOF:

Recursive substitution for lagged capital shocks in h yields $k_t = \phi(\mathbf{z}^t)$, where $\mathbf{z}^t \equiv (z_{t-1}, z_{t-2}, \ldots)$ and where ϕ is strictly increasing. Applying Bayes' rule along with equation (2) yields that for any vector $\tilde{\mathbf{z}} \in \mathbb{R}$, $\Pr\{\mathbf{z}^t \leq \tilde{\mathbf{z}} | z_t\}$ is stochastically strictly increasing in z_t, and the claim follows.

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 ordinary-least-squares context.

The advantage of specifying log utility and Cobb-Douglas production with full depreciation is that we can solve explicitly for the optimal consumption, savings, and labor-supply policies. Using (19) and (20) and adopting the logarithmic instantaneous utility function

$$\lambda \ln c + (1-\lambda)\ln(a - L)$$

together with the Cobb-Douglas 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

$$K_1 = \alpha\beta Z_0 K_0^{\alpha+\theta} L_0^{1-\alpha} \tag{21}$$

and that labor supply is given by

$$L_1 = \frac{\lambda(1-\alpha)a_t}{(1-\lambda)(1-\alpha\beta) + \lambda(1-\alpha)}. \tag{22}$$

If the random endowment follows a multiplicative first-order Markov process, then after taking logarithms, (21) and (22) correspond exactly to equation (5) and assumption (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 (22) 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 equation (5) and assumption (6), which drive our results in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

previous section and which underlie our empirical conclusions, can be obtained under reasonable assumptions in either the OLG or the infinitely-lived-agent models of stochastic growth.

VI. Conclusions

Given our 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; furthermore, 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. Moreover, the model fits in with the recent business-cycle literature that explains properties of cycles with productivity shocks.

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, as indeed are all the parameters of our model. That technology shocks can assume different values over countries seems reasonable if one interprets these shocks broadly to include shifts in institutional and organizational structures, such as shifts in the corporate, legal, or bureaucratic structures, or even in attitudes toward 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 economies can diverge for extended periods of time.

No doubt, a quantum leap in our understanding of growth will occur only when the engine of growth, namely the z_t process, is successfully endogenized. What we think we have shown here, however, is that this engine is fueled primarily by something other than physical capital.

Appendix

Appendix 1: The Derivation of Equation (13)

Note first that

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

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

$$\begin{aligned}\mathrm{Cov}(\eta_t,k_t)&=\mathrm{Cov}(\eta_t,(\alpha+\theta)k_{t-1}+\eta_{t-1})\\&=(\alpha+\theta)\mathrm{Cov}(\eta_t,k_{t-1})\\&\quad+\mathrm{Cov}(\eta_t,\eta_{t-1}).\end{aligned}$$

Expanding further, we obtain

$$\text{(A1)}\quad \mathrm{Cov}(\eta_t,k_t)=\sum_{j=1}^{\infty}(\alpha+\theta)^{j-1}\mathrm{Cov}(\eta_t,\eta_{t-j}).$$

Since $(1-\alpha)l_t=(1-\alpha)m+(1-\alpha)l_{t-1}+(1-\alpha)w_t$, then if $\rho=r$,

$$\begin{aligned}\eta_t&=(1-\alpha)m+\mu+\rho\eta_{t-1}\\&\quad+(\varepsilon_t+(1-\alpha)w_t)\end{aligned}$$

so that $\mathrm{Cov}(\eta_t,\eta_{t-j})=\rho^j\sigma_\eta^2$, where $\sigma_\eta^2=[\sigma_\varepsilon^2+(1-\alpha)^2\sigma_w^2]/(1-\rho^2)$. Then, using (A1),

$$\begin{aligned}\text{(A2)}\quad \mathrm{Cov}(\eta_t,k_t)&=\sigma_\eta^2\sum_{j=1}^{\infty}\rho^j(\alpha+\theta)^{j-1}\\&=\rho\sigma_\eta^2/[1-\rho(\alpha+\theta)].\end{aligned}$$

From equation (2),

$$\mathrm{Cov}(\eta_t,y_t)=(\alpha+\theta)\mathrm{Cov}(\eta_t,k_t)+\sigma_\eta^2.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

Substituting into this expression from (A2) yields

$$\text{(A3)}\quad \mathrm{Cov}(\eta_t, y_t) = \sigma_\eta^2\{1+\rho(\alpha+\theta)/[1-\rho(\alpha+\theta)]\} = \sigma_\eta^2/[1-\rho(\alpha+\theta)].$$

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

$$\text{(A4)}\quad \mathrm{Var}(y_t) = (\alpha+\theta)^2\mathrm{Var}(k_t) + \sigma_\eta^2 + 2(\alpha+\theta)\mathrm{Cov}(\eta_t, k_t).$$

Now, since $k_{t+1} = \gamma + y_t, \mathrm{Var}(y_t) = \mathrm{Var}(k_t)$. Using this in (A4) and substituting from (A3) into (A4) for $\mathrm{Cov}(\eta_t, k_t)$ yields

$$\text{(A5)}\quad \mathrm{Var}(y_t) = \left(\frac{\sigma_\eta^2}{1-(\alpha+\theta)^2}\right)\left(1+\frac{2(\alpha+\theta)\rho}{1-\rho(\alpha+\theta)}\right).$$

The expressions in (A4) and (A5) both explode when ρ approaches 1, because σ_η^2 goes to infinity, but their ratio does not:

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

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

Appendix 2

Here, we derive expressions for the a_{ij} in equation (16) under various assumptions. Deterministic components of z and l are ignored. We assume in equation (4) that $\lambda_1 = \lambda_2 = 0$, so that $\varepsilon_t = w_t$ and so that the w_t are also independently and identically distributed. Repeated substitution in (3) leads to

$$z_{t+T} = \rho^T z_t + \rho^{T-1}\varepsilon_t + \rho^{T-2}\varepsilon_{t+1}\dots + \varepsilon_{t+T-1}$$

or

$$\Delta^T z_t \equiv z_{t+T} - z_t = (\rho^T - 1)z_t + \rho^{T-1}\varepsilon_t + \rho^{T-2}\varepsilon_{t+1}\dots + \varepsilon_{t+T-1}.$$

However,

$$z_t = \rho^{j+1}z_{t-j-1} + \rho^j\varepsilon_{t-j-1} + \rho^{j-1}\varepsilon_{t-j} + \rho^{j-2}\varepsilon_{t-j+1}\dots + \varepsilon_{t-1}$$

so that

$$\Delta^T z_t = (\rho^T - 1) \times \left(\rho^{j+1}z_{t-j-1} + \rho^j\varepsilon_{t-j-1}\dots + \varepsilon_{t-1}\right) + \rho^{T-1}\varepsilon_t + \rho^{T-2}\varepsilon_{t+1}\dots + \varepsilon_{t+T-1}$$

and also

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

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*). For this case,

$$\mathrm{Cov}\left(\Delta^T z_t, \Delta^T z_{t-j-1}\right) = [\sigma_\varepsilon^2/(1-\rho^2)]\left[(\rho^T-1)^2\rho^{j+1}\right] + \sigma_\varepsilon^2(\rho^T-1)\sum_{i=1}^{j+1}\rho^{j+1-i}\rho^{T-i} + \sigma_\varepsilon^2\sum_{i=1}^{T-j-1}\rho^{T-i}\rho^{t-j-1-i}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

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

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

Case (ii). For $(T-j-1)\le 0$,

$$\text{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$
$$= [\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2\rho^{j+1}$$
$$+(\rho^T-1)\left(\sum_{i=1}^{T}\rho^{j+1-i}\rho^{T-i}\right)\sigma_\varepsilon^2.$$

Again, as a check on the algebra, note that $\lim_{\rho\to 1}[\text{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})] = 0$, as it should. As a further check, note that when $j=0$, we have

$$\text{Var}(\Delta^T z_t) = [\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2$$
$$+[(\rho^{T-2}-\rho^{-2})/(1-\rho^{-2})]\sigma_\varepsilon^2$$

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

$$\lim_{T\to\infty}[\text{Var}(\Delta^T z_t)]$$
$$=\sigma_\varepsilon^2/(1-\rho^2)$$
$$+[-\rho^{-2}/(1-\rho^{-2})]\sigma_\varepsilon^2$$
$$=2\sigma_\varepsilon^2/(1-\rho^2).$$

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

$$\text{Cov}(\Delta^T z_t, \Delta^T k_t)$$
$$=\sum_{j=0}^{\infty}(\alpha+\theta)^j\,\text{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$
$$=\sum_{j=0}^{T-2}(\alpha+\theta)^j\,\text{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$
$$+\sum_{j=T-1}^{\infty}(\alpha+\theta)^j$$
$$\times\text{Cov}(\Delta^T z_t, \Delta^T z_{t-j-1})$$
$$=\sum_{j=0}^{T-2}(\alpha+\theta)^j$$
$$\times\Bigg\{[\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2\rho^{j+1}$$
$$+\sigma_\varepsilon^2(\rho^T-1)\sum_{i=1}^{j+1}\rho^{j+1-i}\rho^{T-i}$$
$$+\sigma_\varepsilon^2\sum_{i=1}^{T-j-1}\rho^{T-i}\rho^{t-j-1-i}\Bigg\}$$
$$+\sum_{j=T-1}^{\infty}(\alpha+\theta)^j$$
$$\times\Bigg\{[\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2\rho^{j+1}$$
$$+\sigma_\varepsilon^2(\rho^T-1)\sum_{i=1}^{T}\rho^{j+1-i}\rho^{T-i}\Bigg\}.$$

If we now let T approach ∞, so that the second summation on the right goes to zero,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

we obtain

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

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

$$\text{(A5a)}\quad \lim_{T\to\infty}\big[\text{Cov}(\Delta^T k_t,\Delta^T z_t)\big]$$
$$=\{\sigma_\varepsilon^2/[1-\rho(\alpha+\theta)]\}$$
$$\times\big[\rho/(1-\rho^2)$$
$$-\rho^{-1}/(1-\rho^{-2})\big]$$
$$=2\sigma_\varepsilon^2\rho/(1-\rho^2)[1-\rho(\alpha+\theta)].$$

Next we calculate the limit as ρ approaches 1, for fixed T.

$$\text{(A5b)}\quad \lim_{\rho\to1}\big[\text{Cov}(\Delta^T z_t,\Delta^T k_t)\big]$$
$$=\sigma_\varepsilon^2\sum_{j=0}^{T-2}(\alpha+\theta)^j(T-j-1)$$
$$=\sigma_\varepsilon^2\big[T-1+(\alpha+\theta)$$
$$\times(T+2)\ldots(\alpha+\theta)^{T-2}\big]$$
$$=\sigma_\varepsilon^2\{T[1-(\alpha+\theta)]^{-1}$$
$$-\big[1-(\alpha+\theta)^T\big]$$
$$\times[1-(\alpha+\theta)]^{-2}\}$$
$$=\sigma_\varepsilon^2\{T[1-(\alpha+\theta)]$$
$$-\big[1-(\alpha+\theta)^T\big]\}$$
$$\times[1-(\alpha+\theta)]^{-2}.$$

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

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

Let

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

We will compute $\hat{A}$ later. If we let ρ approach 1 for a given T, remembering that

$$\lim_{\rho\to1}\big[\text{Cov}(\Delta^T z_t,\Delta^T z_{t-j-1})\big]=0$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

for $t-j-1\le 0$, we obtain

$$\begin{aligned}\operatorname{Var}(\Delta^T k_t) &= \sum_{j=0}^{\infty}\sum_{i=j+1}^{T+j}(\alpha+\theta)^{i+j}\times[T-(i-j)]\sigma_\varepsilon^2\\ &\quad+\sum_{i=0}^{\infty}\sum_{j=i+1}^{T+i}(\alpha+\theta)^{i+j}\times[T-(j-i)]\sigma_\varepsilon^2\\ &\quad+\sum_{k=0}^{\infty}(\alpha+\theta)^{2k}T\sigma_\varepsilon^2+\hat{A}\\ &=\Big(2\sigma_\varepsilon^2(\alpha+\theta)\big\{T/[1-(\alpha+\theta)]\\ &\qquad-\big[1-(\alpha+\theta)^T\big]/[1-(\alpha+\theta)]^2\big\}\\ &\quad+T\sigma_\varepsilon^2\big[1-(\alpha+\theta)^2\big]^{-1}\Big)\\ &\quad\times\big[1-(\alpha+\theta)^2\big]^{-1}+\hat{A}.\end{aligned}$$

We also compute $\operatorname{Var}(\Delta^T k_t)$ for $\rho<1$ as T approaches ∞: we have

$$\begin{aligned}\operatorname{Var}(\Delta^T k_t) &= (1-\alpha)^2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\operatorname{Cov}\big(\Delta^T l_{t-j},\Delta^T l_{t-i}\big)\times(\alpha+\theta)^{i+j}\\ &\quad+\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\operatorname{Cov}\big(\Delta^T z_{t-j},\Delta^T z_{t-i}\big)\times(\alpha+\theta)^{i+j}.\end{aligned}$$

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

$$\begin{aligned}\operatorname{Cov}\big(\Delta^T z_{t-i},\Delta^T z_{t-j}\big) &= \big[\sigma_\varepsilon^2/(1-\rho^2)\big]\\ &\quad\times\Big[(\rho^j-1)^2\rho^{i-j}\\ &\qquad+\sigma_\varepsilon^2(\rho^T-1)\rho^{T+(i-j)-2}\\ &\qquad\times\big(1-\rho^{-2(i-j)}\big)/\big(1-\rho^{-2}\big)\Big]\\ &\quad+\sigma_\varepsilon^2\rho^{2T-(i-j)-2}\\ &\qquad\times\big(1-\rho^{-2T+2(i-j)}\big)/\big(1-\rho^{-2}\big).\end{aligned}$$

If we let T approach ∞, 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 $i=j$. The expressions for $i>j$ and $i<j$ are symmetric, so compute twice the value for $i>j$:

$$\begin{aligned}\lim_{T\to\infty}[\operatorname{Var}(\Delta^T k_t)] &= \lim_{T\to\infty}\Big(2\sum_{j=0}^{\infty}\sum_{i=j+1}^{\infty}(\alpha+\theta)^{i+j}\\ &\quad\times[\sigma_\varepsilon^2/(1-\rho^2)]\\ &\quad\times\{(\rho^T-1)^2\rho^{i-j}\\ &\qquad+(\rho^T-1)[\sigma_\varepsilon^2/(1-\rho^{-2})]\\ &\qquad\times\rho^{T-2}\rho^{i-j}(1-\rho^{-2i}\rho^{2j})\\ &\qquad+[\sigma_\varepsilon^2/(1-\rho^{-2})]\\ &\qquad\times[\rho^{2T-2}\rho^{-i}\rho^{j}(1-\rho^{-2T}\rho^{2i}\rho^{-2j})]\}\\ &\quad+\sum_{j=0}^{\infty}(\alpha+\theta)^{2j}\\ &\quad\times\{[\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2\\ &\qquad+[\sigma_\varepsilon^2/(1-\rho^{-2})]\\ &\qquad\times(\rho^T-1)\rho^{T-2}(0)\\ &\qquad+[\sigma_\varepsilon^2/(1-\rho^{-2})]\\ &\qquad\times[\rho^{2T-2}(1-\rho^{-2T})]\}+\hat{A}\Big)\end{aligned}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

$$
\begin{aligned}
= \lim_{T\to\infty} \Big(2\sigma_\varepsilon^2 \Big[\sum_{j=0}^{\infty} & (\rho^T-1)^2(\alpha+\theta)^j \\
& \times \rho^{-j}(\alpha+\theta)^{j+1}\rho^{j+1} \\
& \times [1-(\alpha+\theta)\rho]^{-1}(1-\rho) \\
& + \rho^{T-1}(1-\rho^{-2})^{-1} \\
& \times (\alpha+\theta)^j\rho^{T-2}\rho^{-j} \\
& \times \{\rho^{j+1}(\alpha+\theta)^{j+1} \\
& \quad \times [1-(\alpha+\theta)\rho]^{-1} \\
& \quad - \rho^{2j}(\alpha+\theta)^{j+1}\rho^{-(j+1)} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1}\} \\
& + (1-\rho^{-2})\rho^{2T-2}(\alpha+\theta)^j \\
& \times \rho^j\{(\alpha+\theta)^{j+1}\rho^{-(j+1)} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1} \\
& \quad - \rho^{-2T}\rho^{-2j}(\alpha+\theta)^{j+1} \\
& \quad \times \rho^{j+1}[1-(\alpha+\theta)\rho]^{-1}\} \\
& + \Big[\sigma_\varepsilon^2(1-\rho^{-2})^{-1}(\rho^T-1)^2 \\
& \quad + \sigma_\varepsilon^2(1-\rho^{-1})^{-1}\rho^{2T-2} \\
& \quad - \sigma_\varepsilon^2(1-\rho^{-2})^{-1}\rho^{-2}\Big] \\
& \quad \times [1-(\alpha+\theta)^2]^{-1}\Big] + \hat{A}\Big)
\end{aligned}
$$

$$
\begin{aligned}
= \lim_{T\to\infty} \Big(2\{ & [\sigma_\varepsilon^2/(1-\rho^2)](\rho^T-1)^2 \\
& \times \rho(\alpha+\theta)[1-(\alpha+\theta)^2] \\
& \times [1-(\alpha+\theta)\rho]^{-1} \\
& + (\rho^T-1)\rho^{T-2}[\sigma_\varepsilon^2/(1-\rho^{-2})] \\
& \times \rho(\alpha+\theta) \\
& \times [1-(\alpha+\theta)^2]^{-1}[1-(\alpha+\theta)\rho]^{-1} \\
& - [\sigma_\varepsilon^2/(1-\rho^{-2})](\rho^T-1) \\
& \quad \times \rho^{T-2}(\alpha+\theta)\rho^{-1}[1-(\alpha+\theta)^2]^{-1} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1} \\
& + [\sigma_\varepsilon^2/(1-\rho^{-2})]\rho^{2T-2} \\
& \quad \times (\alpha+\theta)\rho^{-1}[1-(\alpha+\theta)^2]^{-1} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1} \\
& - [\sigma_\varepsilon^2/(1-\rho^{-2})](\alpha+\theta) \\
& \quad \times \rho^{-1}[1-(\alpha+\theta)^2]^{-1} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1}\} \\
& + [1-(\alpha+\theta)^2]^{-1} \\
& \times \Big[\sigma_\varepsilon^2(1-\rho^2)^{-1}(\rho^T-1)^2 \\
& + \sigma_\varepsilon^2(1-\rho^{-2})^{-1}\rho^{2T-2} \\
& - \sigma_\varepsilon^2(1-\rho^{-2})^{-1}\rho^{-2}\Big] + \hat{A}\Big)
\end{aligned}
$$

$$
\begin{aligned}
= \lim_{T\to\infty} \Big(& [\sigma_\varepsilon^2/(1-\rho^2)][1-(\alpha+\theta)^2]^{-1} \\
& \times \{2(\rho^T-1)^2[1-\rho(\alpha+\theta)]^{-1} \\
& \quad \times \rho(\alpha+\theta) \\
& \quad - 2(\rho^T-1)\rho^T\rho \\
& \quad \times (\alpha+\theta)[1-(\alpha+\theta)\rho]^{-1} \\
& \quad + 2(\rho^T-1)\rho^T(\alpha+\theta) \\
& \quad \times \rho^{-1}[1-(\alpha+\theta)\rho^{-1}] \\
& \quad - 2\rho^{2T}(\alpha+\theta)\rho^{-1} \\
& \quad \times [1-(\alpha+\theta)\rho^{-1}]^{-1} \\
& \quad + 2(\alpha+\theta)\rho \\
& \quad \times [1-(\alpha+\theta)\rho]^{-1} \\
& \quad + \Big[(\rho^T-1)^2 - \rho^{2T} + 1\Big] \\
& \quad + 2(1-\rho^T)\} + \hat{A}\Big)
\end{aligned}
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

$$= \lim_{T\to\infty}\Big(2[\sigma_\varepsilon^2/(1-\rho^2)] \times[1-(\alpha+\theta)^2] \times\Big[(\rho^T-1)^2\rho(\alpha+\theta) \times[1-(\alpha+\theta)\rho]^{-1}+(\rho^T-1)\rho^T \times\{\rho^{-1}[1-(\alpha+\theta)\rho^{-1}]^{-1} -\rho[1-(\alpha+\theta)\rho]^{-1}\}(\alpha+\theta) +(\alpha+\theta)\{\rho[1-(\alpha+\theta)\rho]^{-1} -\rho^{-1}[1-(\alpha+\theta)\rho^{-1}]\rho^{2T}\} +1-\rho^T\Big]+\hat{A}\Big).$$

Therefore,

$$\text{(A6)}\quad \lim_{T\to\infty}\left[\mathrm{Var}(\Delta^T k_t)\right] = 2\left[\sigma_\varepsilon^2/(1-\rho^2)\right] \times\left[1-(\alpha+\theta)^2\right]^{-1} \times\{2(\alpha+\theta)\rho \times[1-(\alpha+\theta)\rho]^{-1}+1\} + \lim_{T\to\infty}\hat{A}.$$

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

$$\text{(A7)}\quad \mathrm{Cov}(\Delta^T l_{t-i}, \Delta^T l_{t-j}) = [T-(i-j)]\sigma_w^2 \qquad \text{for } r\to 1$$

so that $\mathrm{Var}(\Delta^T l_t)\to T\sigma_w^2$ and

$$\mathrm{Cov}(\Delta^T l_{t-i}, \Delta^T l_{t-j}) = 2\left[\sigma_w^2/(1-r^2)\right]r^{i-j} \quad \text{for } T\to\infty,\ r<1$$

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

Now, for $r=1$,

$$\hat{A} = (1-\alpha)^2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\mathrm{Cov}(\Delta^T l_{t-i}, \Delta^T l_{t-j}) \times(\alpha+\theta)^{i+j} = \{(1-\alpha)^2/[1-(\alpha+\theta)^2]\} \times\Big(2\sigma_w^2(\alpha+\theta)\{T/[1-(\alpha+\theta)] -[1-(\alpha+\theta)^T] \times[1-(\alpha+\theta)]^{-2}\} + T\sigma_w^2\Big).$$

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

$$\lim_{T\to\infty}\hat{A} = \lim_{T\to\infty}\Big((1-\alpha)^2 \times\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\mathrm{Cov}(\Delta^T l_{t-i}, \Delta^T l_{t-j}) \times(\alpha+\theta)^{i+j}\Big)$$

$$= \lim_{T\to\infty}\Big((1-\alpha)^2[2\sigma_w^2/(1-r^2)] \times\left[1-(\alpha+\theta)^2\right] \times\Big[(r^T-1)^2 r(\alpha+\theta) \times[1-(\alpha+\theta)r]^{-1} +(r^T-1)r^T \times\{r^{-1}[1-(\alpha+\theta)r^{-1}]\}^{-1} -r[1-(\alpha+\theta)r]^{-1}(\alpha+\theta) +(\alpha+\theta) \times\{r[1-(\alpha+\theta)r]^{-1} -r^{-1}[1-(\alpha+\theta)r^{-1}]r^{2T}\} +1-r^T\Big]\Big).$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

Therefore,

$$\text{(A8)}\quad \lim_{T\to\infty} \hat{A} = 2\left[\sigma_w^2/(1-r^2)\right] \times\left[1-(\alpha+\theta)^2\right]^{-1} \times\left\{2(\alpha+\theta)r \times[1-(\alpha+\theta)r]^{-1}+1\right\}.$$

Finally,

$$\text{(A9)}\quad \operatorname{Cov}(\Delta^T k_t,\Delta^T l_t) = (1-\alpha) \times\sum_{j=0}^{\infty}\operatorname{Cov}(\Delta^T l_t,\Delta^T l_{t-j-1}) \times(\alpha+\theta)^j$$
$$=(1-\alpha)\sigma_\varepsilon^2\{T[1-(\alpha+\theta)] -[1-(\alpha+\theta)^T]\}/[1-(\alpha+\theta)]^2 \quad \text{for } r=1.$$

In general, for arbitrary r,

$$\lim_{T\to\infty}\left[\operatorname{Cov}(\Delta^T k_t,\Delta^T l_t)\right] = \lim_{T\to\infty}\left\{\left[(1-\alpha)\sigma_\varepsilon^2/(1-r^2)\right] \times\left((r^T-1)^2 r[1-(\alpha+\theta)r]^{-1} +(r^T-1)r^{T+1} \times\left\{r^{-2}[1-(\alpha+\theta)r^{-1}]^{-1} -[1-(\alpha+\theta)r]^{-1}\right\} -r^{2T-1}[1-(\alpha+\theta)r^{-1}]^{-1} +r[1-(\alpha+\theta)r]^{-1}\right)\right\}.$$

Taking the limit, for $r<1$,

$$\text{(A10)}\quad \lim_{T\to\infty}\left[\operatorname{Cov}(\Delta^T l_t,\Delta^T k_t)\right] =\left[(1-\alpha)2\sigma_\varepsilon^2/(1-r^2)\right] \times[1-(\alpha+\theta)r]^{-1}.$$

Table A1 summarizes the results of this appendix that are relevant for the bias described in equation (16) of the text.

Combined cases: $r,\rho\to 1$, $T\to\infty$. We shall now 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 equations (16) and (17) of the text (the Landau symbol "O" refers to the order of the expression):

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

$$a_{kl} = O\left(\frac{1}{1-r^2}\right)\left(\frac{2\sigma_w^2}{1-(\alpha+\theta)}\right)(1-\alpha)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

TABLE A1—EXPRESSIONS FOR THE a_{ij} FROM WHICH ONE MAY CALCULATE THE BIAS IN $\hat{b}_k$ AND $\hat{b}_l$

Parameter	Case (a): $\lim_{r\to 1}$, $\lim_{\rho\to 1}$, T finite	Case (b): $\rho, r<1$, $\lim_{T\to\infty}$
a_{kk}:	$\dfrac{2\sigma_\varepsilon^2(\alpha+\theta)\left(\dfrac{T}{1-(\alpha+\theta)}-\dfrac{1-(\alpha+\theta)^T}{[1-(\alpha+\theta)]^2}\right)+T\sigma_\varepsilon^2}{1-(\alpha+\theta)^2}$	$\left(\dfrac{2\sigma_\varepsilon^2}{(1-\rho^2)[1-(\alpha+\theta)^2]}\right)\left(\dfrac{2(\alpha+\theta)\rho}{1-(\alpha+\theta)\rho}+1\right)$
	$+(1-\alpha)^2\left[\dfrac{2\sigma_w^2(\alpha+\theta)\left(\dfrac{T}{1-(\alpha+\theta)}-\dfrac{1-(\alpha+\theta)^T}{[1-(\alpha+\theta)]^2}\right)+T\sigma_w^2}{1-(\alpha+\theta)^2}\right]$	$+\left(\dfrac{(1-\alpha)^2 2\sigma_w^2}{(1-r^2)[1-(\alpha+\theta)^2]}\right)\left(\dfrac{2(\alpha+\theta)r}{1-(\alpha+\theta)r}+1\right)$
a_{kl}:	$(1-\alpha)\sigma_w^2\left(\dfrac{T[1-(\alpha+\theta)]-[1-(\alpha+\theta)^T]}{[1-(\alpha+\theta)]^2}\right)$	$2(1-\alpha)\left(\dfrac{\sigma_w^2}{1-r^2}\right)[1-(\alpha+\theta)r]^{-1}$
a_{ll}:	$T\sigma_w^2$	$2\sigma_w^2/(1-r^2)$
a_{ku}:	$\dfrac{\sigma_w^2\{T[1-(\alpha+\theta)]-[1-(\alpha+\theta)^T]\}}{[1-(\alpha+\theta)]^2}$	$\dfrac{2\sigma_\varepsilon^2\rho}{(1-\rho^2)[1-\rho(\alpha+\theta)]}$

Note: To obtain the a_{ij}, the expressions in the first column of the table [case (a)] should be divided by T^2; the second column reports $\lim_{T\to\infty} T^2 a_{ij}$ for case (b).

$$a_{ll}=O\left(\frac{1}{1-r^2}\right)(2\sigma_w^2)$$

$$a_{ku}=O\left(\frac{1}{1-\rho^2}\right)\left(\frac{2\sigma_\varepsilon^2}{1-(\alpha+\theta)}\right).$$

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

$$\frac{a_{ll}a_{ku}}{a_{kk}a_{ll}-a_{kl}^2}=O\left(\frac{1}{1-r^2}\right)(A_{ll})\,O\left(\frac{1}{1-\rho^2}\right)(A_{ku})\times\left\{\left[O\left(\frac{1}{1-\rho^2}\right)(A_{kk}^1)+O\left(\frac{1}{1-r^2}\right)(A_{kk}^2)\right]\times O\left(\frac{1}{1-r^2}\right)(A_{ll})-\left[O\left(\frac{1}{1-r^2}\right)\right]^2(A_{kl}^2)\right\}^{-1}$$

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

$$\frac{a_{ll}a_{ku}}{a_{kk}a_{ll}-a_{kl}^2}\to\frac{A_{ll}A_{ku}}{(A_{kk}^1+A_{kk}^2)A_{ll}-A_{kl}^2}$$

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

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

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 in the bottom of the denominator (i.e., the expression in large braces). Then, multiplying top and bottom by $[1-(\alpha+\theta)]^2/4\sigma_w^2$ leaves us with

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

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

$$\begin{aligned}\frac{-a_{kl}a_{ku}}{a_{kk}a_{ll}-a_{kl}^2} &= -\mathrm{O}\left(\frac{1}{1-r^2}\right)(A_{kl}) \\ &\quad\times\mathrm{O}\left(\frac{1}{1-\rho^2}\right)(A_{ku}) \\ &\quad\times\left\{\left[\mathrm{O}\left(\frac{1}{1-\rho^2}\right)(A_{kk}^1)\right.\right. \\ &\qquad\left.+\mathrm{O}\left(\frac{1}{1-r^2}\right)(A_{kk}^2)\right] \\ &\qquad\times\mathrm{O}\left(\frac{1}{1-r^2}\right)(A_{ll}) \\ &\qquad\left.-\left[\mathrm{O}\left(\frac{1}{1-r^2}\right)\right]^2A_{kl}^2\right\}^{-1} \\ &\to\frac{-A_{kl}A_{ku}}{(A_{kk}^1+A_{kk}^2)A_{ll}-A_{kl}^2} \\ &=\frac{-4\sigma_w^2\sigma_\varepsilon^2(1-\alpha)}{[1-(\alpha+\theta)]^2} \\ &\quad\times\left\{\left[\left(\frac{2\sigma_\varepsilon^2}{1-(\alpha+\theta)^2}\right)\left(\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)}\right)\right.\right. \\ &\qquad\left.+\left(\frac{2\sigma_w^2(1-\alpha)^2}{1-(\alpha+\theta)^2}\right)\left(\frac{1+(\alpha+\theta)}{1-(\alpha+\theta)}\right)\right] \\ &\qquad\left.\times2\sigma_w^2-\frac{4\sigma_w^2(1-\alpha)^2}{[1-(\alpha+\theta)]^2}\right\}^{-1}.\end{aligned}$$

We note that

$$\begin{aligned}&\left[1-(\alpha+\theta)^2\right][1-(\alpha+\theta)] \\ &\quad=(1+\alpha+\theta)[1-(\alpha+\theta)]^2.\end{aligned}$$

Therefore, the above equals

$$\begin{aligned}&\frac{-4\sigma_w^2\sigma_\varepsilon^2(1-\alpha)}{2\sigma_w^2\left[2\sigma_\varepsilon^2+(1-\alpha)^2 2\sigma_w^2\right]-4\sigma_w^2(1-\alpha)^2} \\ &\quad=\frac{-\sigma_\varepsilon^2(1-\alpha)}{\sigma_\varepsilon^2+(1-\alpha)^2\sigma_w^2-\sigma_w^2(1-\alpha)^2} \\ &\quad=-(1-\alpha).\end{aligned}$$

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

Appendix 3

Here, 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,

$$\begin{aligned}y_{t+1}&=z_{t+1}+(1-\alpha)l_{t+1} \\ &\quad+(\alpha+\theta)\ln[sY_t+(1-\delta)K_t] \\ &=z_{t+1}+(1-\alpha)l_{t+1} \\ &\quad+(\alpha+\theta)\ln\{sY_t+(1-\delta) \\ &\quad\times[sY_{t-1}+(1-\delta)K_{t-1}]\} \\ &=z_{t+1}+(1-\alpha)l_{t+1} \\ &\quad+(\alpha+\theta)\ln s \\ &\quad+(\alpha+\theta)\ln\left[\sum_{j=0}^{\infty}(1-\delta)^jY_{t-j}\right].\end{aligned}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

Therefore, the analogue of the equation in footnote 11 is

$$\text{(A11)}\quad y_{t+1}-\rho y_t = \omega_t + (1-\alpha)(l_{t+1}-\rho l_t) + (\alpha+\theta)(1-\rho)\ln s + (\alpha+\theta)\left[\sum_{j=0}^{\infty}(1-\delta)^j Y_{t-j} - \rho \ln \sum_{j=0}^{\infty}(1-\delta)^j Y_{t-j-1}\right].$$

The ω_t process was once again assumed to follow equation (4). The infinite sums in (A11) 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 postwar years (see footnote 11, however).

REFERENCES

Baily, Martin Neal, "Comment," *NBER Macroeconomic Annual*, 1987, *1*, 205–8.

Barro, Robert, "The Persistence of Unemployment," *American Economic Review*, May 1988, *78*, 32–7.

Baumol, William and Wolff, Edward, "Productivity Growth, Convergence, and Welfare: Reply," *American Economic Review*, December 1988, *78*, 1155–9.

Benhabib, Jess and Laroque, Guy, "On Competitive Cycles in Production Economies," *Journal of Economic Theory*, June 1988, *45*, 145–70.

______ **and Nishimura, Kazuo,** "Stochastic Equilibrium Oscillations," *International Economic Review*, February 1989, *30*, 85–102.

Bernstein, Jeffrey and Nadiri, Ishaq, "Research and Development and Intraindustry Spillovers: An Empirical Implication of Dynamic Duality," *Review of Economic Studies*, April 1989, *56*, 249–68.

Blume, Lawrence and Easley, David, "Characterization of Optimal Plans for Stochastic Dynamic Programs," *Journal of Economic Theory*, April 1982, *28*, 221–34.

Campbell, John and Mankiw, Gregg, "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, November 1987, *102*, 857–80.

Christiano, Lawrence, "Comment on Romer's 'Crazy Explanations of the Productivity Slowdown'," unpublished manuscript, Federal Reserve Bank of Minneapolis, 1987.

Cochrane, John, "How Big is the Random Walk in GNP?" *Journal of Political Economy*, October 1988, *96*, 893–920.

DeLong, Bradford, "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, December 1988, *78*, 1138–54.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, Spring 1979, *10*, 92–116.

______, "Productivity Puzzles and R&D: Another Nonexplanation," *Journal of Economic Perspectives*, Fall 1988, *2*, 9–21.

Heston, Alan and Summers, Robert, "Improved International Comparisons of Real Product and Its Composition," *Review of Income and Wealth*, June 1984, *30*, 207–26.

Jaffe, Adam, "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, December 1986, *76*, 984–1001.

Jovanovic, Boyan and MacDonald, Glenn, "Competitive Diffusion," unpublished manuscript, New York University, October 1988.

______ **and Rob, Rafael,** "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, October 1989, *56*, 569–82.

Lach, Saul and Schankerman, Mark, "Dynamics of R&D and Investment in the Scientific Sector," *Journal of Political Economy*, August 1989, *97*, 880–904.

Maddison, Angus, *Phases of Capitalist Development*, New York: Oxford University Press, 1982.

Mansfield, Edwin, Rapoport, John, Romeo, Anthony, Wagner, Samuel and Beardsley, George, "Social and Private Rates of Re-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com

turn from Industrial Innovations," *Quarterly Journal of Economics*, May 1977, *91*, 221–40.

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

Pakes, Ariel and Schankerman, Mark, (1984a) "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*, Chicago: University of Chicago Press, 1984, 73–88.

______ **and** ______, (1984b) "An Exploration into the Determinants of Research Intensity," in Zvi Griliches, ed., *R&D, Patents and Productivity*, Chicago: University of Chicago Press, 1984, 209–32.

Prescott, Edward C., "Theory Ahead of Measurement in Business-Cycle Research," *Carnegie-Rochester Conference on Public Policy*, Autumn 1986, *25*, 11–44.

Quah, Danny, "International Patterns of Growth: Persistence in Cross-Country Disparities," unpublished manuscript, Massachusetts Institute of Technology, January 1990.

Romer, Paul, "Crazy Explanations for the Productivity Slowdown," *NBER Macroeconomics Annual*, 1987, *1*, 163–201.

Scherer, Frederick M., "Inter-Industry Technology Flows and Productivity Growth," *Review of Economics and Statistics*, November 1982, *64*, 627–34.

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

Shleifer, Andrei, "Implementation Cycles," *Journal of Political Economy*, December 1986, *94*, 1163–90.

Solow, Robert, "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, August 1957, *39*, 312–20.

Vernon, Raymond, "Technological Development: The Historical Experience," EDI Seminar Paper 39, The World Bank, Washington, DC, 1989.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.manaraa.com