Asset Returns and Economic Growth

Author(s): Dean Baker, J. Bradford Delong and Paul R. Krugman

Source: *Brookings Papers on Economic Activity*, 2005, Vol. 2005, No. 1 (2005), pp. 289-315

Published by: Brookings Institution Press

Stable URL: https://www.jstor.org/stable/3805089

DEAN BAKER
Center for Economic and Policy Research
J. BRADFORD DELONG
University of California, Berkeley
PAUL R. KRUGMAN
Princeton University

Asset Returns and Economic Growth

IT IS DIFFICULT to see how real U.S. GDP growth can be as rapid in the next half-century as it has been in the last. The baby boom is long past, and no similar explosion of fertility to boost the rate of labor force growth from natural increase has occurred since or is on the horizon. The modern feminist revolution is two generations old: no reservoir of potential female labor remains to be added to the paid labor force. Immigration will doubtless continue—the United States is likely still to have only one-twentieth of the world's population late in this century and to remain vastly richer than the world on average—but can immigration proceed rapidly enough to make the labor force grow as fast in the next fifty years as it did in the past fifty? Productivity growth, the other possible source of faster GDP growth, is a wild card: although we find very attractive the arguments of Robert Gordon for rapid future productivity growth,[1] his is not the consensus view; this is shown most strikingly by the pessimistic projection of the Social Security trustees that very long run labor productivity growth will average 1.6 percent a year.[2]

A slowing of the rate of real economic growth raises challenges for the financing of pay-as-you-go social insurance systems that rely on a rapidly expanding economy to provide generous benefits for the elderly at relatively low tax rates on the young. An alternative way of financing such systems

1. Gordon (2003). Oliner and Sichel (2003) and Kremer (1993) provide additional reasons to be very optimistic about future productivity growth.

2. Board of Trustees of the Federal Old Age and Survivors Insurance and Disability Insurance Trust Funds (2005; all citations from this report are for the intermediate projection). Contrast this with the 2.0 percent average annual rate of economy-wide labor productivity growth from the fourth quarter of 1989 through the first quarter of 2005.

is to prefund them, and for that reason projections of future rates of return on capital play an important role in today's economic policy debates. The solutions to many policy issues depend heavily on whether historical real rates of return—especially the 6.5 percent or so annual average realized rate of return on equities—are likely to persist: the higher are likely future rates of return, the more attractive become policies that, at the margin, shift some additional portion of the burden of financing social insurance onto the present and the near future, thus giving workers' contributions the power to compound over time.

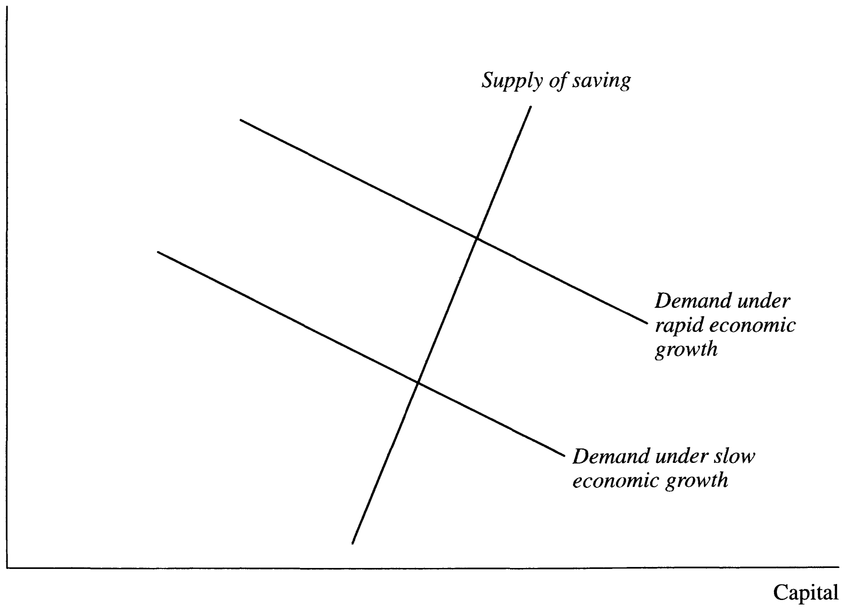
We believe that the argument for prefunding—that slowing economic growth creates a presumption that the burden of financing social insurance should be shifted back in time toward the present—is much shakier than many economists recognize.[3] It is our belief that *if* forecasts of slower real GDP growth come to pass, *then* it is highly likely that future real returns to capital will likewise be significantly below past historical averages. In our view the links between asset returns and economic growth are strong: the algebra of capital accumulation and the production function and the standard macrobehavioral analytical models that economists use as their finger exercises suggest this; arithmetic suggests this as well, for we cannot see any easy way to reconcile current real bond, stock dividend, and stock earnings yields with the twin assumptions that asset markets are making rational forecasts and that rationally expected real rates of return will be as high in the future as they have been in the past half-century.

Our basic argument is very simple. Consider a simple chart of the supply and demand for capital in generational perspective (figure 1). The supply of capital—the amount of investable assets accumulated by savers—presumably follows a standard (if probably steeply sloped) supply curve,[4] with relative quantities of total saving and thus of capital plotted on the horizontal axis, and the price of capital—that is, its rate of return—on the vertical axis. The demand for capital by businesses will, of course, depend

3. An argument challenged, for reasons similar to but not exactly aligned with those we discuss here, in Cutler and others (1990).

4. Supply is likely to be steeply sloped because of opposing income and substitution effects. An increase in the rate of return increases the total lifetime wealth of savers, which presumably increases their consumption when young and so diminishes their saving. An increase in the rate of return also increases the incentive to save, which presumably increases saving. The net effect—which we believe to be positive—is likely to be relatively small.

Figure 1. The Supply and Demand of Capital and the Rate of Return



on the rate of return demanded by the savers who commit their capital to businesses: the higher this required rate of return, the lower will be business demand for capital—and the more eager will businesses be to substitute labor for capital in production. The demand for capital by businesses depends on many other factors as well, from which we single out two:

—*The rate of growth of the labor force.* Labor and capital are complements. A larger labor force for firms to hire from will raise the marginal product of capital for any given level of the capital stock, making businesses more willing to pay higher returns in order to get hold of capital.

—*The rate of improvement in the economy's level of technology.* Better technology—also a complement to capital—will boost business demand.

What is the effect of a slowdown in economic growth—through either a fall in the rate at which the labor force grows, or a fall in the rate at which technology and thus equilibrium labor productivity increase—on this equilibrium? Assume that these changes do not affect the saving behavior of the accumulating generation:[5] then they affect only the demand curve and not the supply curve. Each of these shocks moves the demand curve leftward: having fewer workers reduces the marginal product of capital and hence firm demand for capital; slower productivity growth does the same. The equilibrium capital stock falls, and the rate of return that savers can demand, while still finding businesses willing to invest what they have saved, falls as well. Slower economic growth brings with it lower real rates of return.

We make our case as follows. After first laying out what we see as the major issues to be resolved, we discuss how the algebra of the production function and capital accumulation suggests that rates of return and rates of growth are strongly linked. We then analyze the standard, very simple, macrobehavioral models that economists use to address these issues and find that they, too, lead us to not be surprised by a strong positive relationship between economic growth and asset returns. We then turn to the arithmetic: starting from current bond, stock dividend, and stock earnings yields, we find it arithmetically very difficult to construct scenarios in which asset returns remain at their historic average values when real GDP growth is markedly slowed.

Next we turn to what we regard as the most interesting possibility for escape from this bind. In the late nineteenth century, slower growth in the British economy was accompanied by *no* reduction in returns on British assets, as Britain exported capital on a scale relative to the size of its economy never seen before or since. Could the United States follow the same trajectory? Yes. Is it likely to? Not without a huge boost to national saving.

Before concluding, we turn to a brief analysis of the equity premium. Much argument and some analysis of the dilemmas of the U.S. social insurance system point to the large historical value of the equity premium in America as a potential source of excess returns. We argue, however, that once one has conditioned on the level of the capital-output ratio, returns on

5. As Gregory Mankiw points out in his comment on this paper, and as we discuss below, in the standard Ramsey model a reduction in the rate of natural increase does affect the saving of the accumulating generation—and shifts the saving supply curve inward exactly as much as investment demand shifts inward, keeping the real rate of return unchanged. This is due to the powerful bequest motive behind the assumption of an infinitely lived representative household whose utility for a given level of consumption *per capita* is linear in the size of the household.

balanced portfolios in the long run depend only on the physical return to capital and the margins charged by financial intermediaries. (However, attitudes toward risk *do* affect the long-run capital-output ratio.) They do not depend on the equity premium or the price of risk.

We conclude that *if* economic growth over the next century falls as far as envisioned by forecasts like those in the 2005 Social Security trustees' report, *then* it is not very likely that asset returns will match historical experience. If the stock market today is significantly overvalued and about to come back to earth, if the distribution of income undergoes a significant shift away from labor and toward capital, or if the United States massively boosts its national saving rate and runs surpluses on the relative scale of pre–World War I Britain, for more than twice as long as Britain did—then a real GDP growth slowdown need not entail a significant reduction in asset returns. But these seem to us to be possible, not probable, scenarios, and not the central tendency of the distribution of possible futures that is a real economic forecast.

Issues

The United States is in all likelihood undergoing a minor demographic transition: from a twentieth century in which the population's rate of natural increase was high, to a twenty-first century in which, many suspect, fertility will be at or below levels consistent with zero population growth. This will translate into a slowdown in growth in labor input. From 1958 to 2004, total hours worked in the economy grew at 1.5 percent a year as the entrance of the baby-boomers—male and female—and their successors into the labor force vastly outweighed a decline in average hours worked. The Social Security Administration's 2005 trustees' report projects that hours worked will grow at only 0.3 percent a year from 2015 through 2045.[6]

Meanwhile some economists—although far from all—are projecting a slowdown in productivity growth.[7] The Social Security Administration foresees economy-wide labor productivity growing at only 1.6 percent a year in 2011 and thereafter. In contrast, between 1995 and 2004 economy-wide labor productivity grew at 2.5 percent a year, between 1990 and 2004

6. Board of Trustees (2005).
7. See Oliner and Sichel (2003); Gordon (2003); Nordhaus (2005).

it grew at 2.0 percent a year, and between 1958 and 2004 at 1.9 percent a year.[8] Thus, less than a decade from now, the Social Security forecasters at least see a significant change in both key factors in economic growth: a fall of 1.2 percentage points a year in the rate of growth of labor input, and a fall of between 0.3 and 0.9 percentage point, depending on whether one takes the long 1958–2004 or the short 1995–2004 baseline, in labor productivity growth. The total growth slowdown forecast to hit in a decade or less is thus in the range of 1.6 to 2.2 annual percentage points of real GDP.

What implications will this growth slowdown—if it comes to pass—have for asset values and returns? One position, taken implicitly by the Social Security Administration and explicitly by others,[9] is that there is no reason to expect asset returns to be lower in the future. Whereas U.S. economic growth is determined by productivity growth and labor force growth in the United States, U.S. asset returns are determined by time preference, the intertemporal elasticity of substitution in consumption, and attitudes toward risk, all in a global economy. Why should they be connected? Thus, we hear, past asset performance is still the best guide to future returns.

We take a contrary position. Yes, safe asset returns are equal to the marginal utility of saving, stock market returns equal safe asset returns plus the cost of bearing equity risk, and the United States is part of a world economy. Yes, economic growth is equal to productivity growth plus labor force growth. But only in the case of a small open economy with fixed exchange rates are asset returns determined independently of the rate of economic growth. In a large open economy, they are jointly determined and will be linked.[10]

8. An alternative breakdown would distinguish 1958–73, during which economy-wide labor productivity growth averaged 2.6 percent a year; the productivity slowdown period of 1973–95, when it averaged 1.2 percent a year; and the post-1995 "new economy" period, when it averaged 2.6 percent a year. Much depends on whether one interprets the 1973–95 productivity slowdown period as an anomalous freak disturbance to the economy's normal structure, or as just one of those things one can expect to see every half-century or so.

9. Council of Economic Advisers (2005).

10. Even in a small open economy, real returns on assets and rates of economic growth will be linked unless the real exchange rate is fixed. Even perfect arbitrage by mammoth amounts of risk-neutral foreign capital only equalizes expected rates of return at home and abroad *calculated in foreign currency*. With a flexibly changing real exchange rate, the rate of return in foreign currency is not the same as the rate of return in domestic currency.

Perhaps an analogy will be helpful. In international trade, the trade balance is the difference between what exporters are able to sell abroad and home demand for imports. In international finance, the trade balance is the difference between national saving and national investment. How can this be? Why should a change in exporters' success at marketing abroad change either national saving or national investment? Great confusion has been caused throughout international economics over how, exactly, to think of the connection. We believe that claims that national economic growth is unconnected with asset returns reflect a similar failure to grasp the whole problem.

This is an important issue to get straight now, because the relative attractiveness of pay-as-you-go versus prefunded social insurance systems depends to some degree on the gap between the return on capital r and the rate of real economic growth $n + g$, the sum of the rate of growth of employment n and the rate of growth of labor productivity g. If we are willing to be simple Benthamites, with a social welfare function that shamelessly makes interpersonal comparisons of utility, the argument is straightforward. The higher is the rate of economic growth $n + g$ relative to the return on capital r, the more attractive do pay-as-you-go social insurance systems become. When $n + g$ approaches r, pay-as-you-go systems appear to be very cheap and effective ways of increasing social welfare by passing resources down from the rich and numerous future to the poorer and less numerous present. By contrast, the larger is r relative to $n + g$, the greater are the benefits of prefunding social insurance systems. Prefunded systems can use high rates of return and compound interest to reduce the wedge between productivity and after-contribution real wages. They thus sacrifice the possibility of raising social welfare by moving wealth from the richer distant future to the near future and the present, but in return they gain by reducing the social insurance tax rate and thus its deadweight loss. And whenever we make utilitarian arguments other than those of pure Pareto-preference for why one set of policies is superior to another set, we are all, in our hearts, secret Benthamites.

Thus, to the extent that the political debate over the future of social insurance in America is conducted in the language of rational policy analysis, getting the gap between r on the one hand and $n + g$ on the other hand right is important. Policies predicated on a false belief that r is much larger relative to $n + g$ than it is will unduly burden today's and tomorrow's young people and will leave many disappointed when returns on assets turn out to be less than anticipated and prefunding leaves large unexpected holes in retirement financing. Policies predicated on the belief that $n + g$ is higher relative to r than it is pass up opportunities to lighten the overall tax burden and still provide near-equivalent income security benefits in the long run.

Algebra

Let us begin by distinguishing a number of different rates of return. In this paper we use r to stand for a physical gross marginal product of capital, and we assume that it is the product of a Cobb-Douglas production function:

$$r = \alpha \frac{Y}{K}. \tag{1}$$

We distinguish this physical capital rate of gross profit r from the *net* rate of return on a balanced financial portfolio r_f and from the *net* rate of return on equities r_e.

Only under the assumptions of constant depreciation rates δ, constant financial markups, and a constant price and amount of risk is the mapping among these three straightforward. Toward the end of this paper we briefly consider the equity premium, but otherwise we assume that depreciation rates, financial markups, and other factors that could vary the wedges between r, r_f, and r_e are unimportant. Thus we will move back and forth between these three different rates of return: things that raise or lower the return on stocks will also raise or lower the return on bonds and (after the capital stock has adjusted) the physical marginal product of capital as well.

Robert Solow studied a constant-returns Cobb-Douglas production function with α as the returns-to-capital parameter, and with Y, K, L, and E as aggregate output, the capital stock, the supply of labor, and the level of labor-augmenting technology, respectively:[11]

$$Y = K^{\alpha} (EL)^{1-\alpha}. \tag{2}$$

11. Solow (1956).

Assume constant rates of labor force growth n, of labor-augmenting technical change g, of depreciation δ, and of gross saving s. In the closed-economy case, in which all of domestic capital K is owned by domestic residents and in which all of national saving goes into increasing the domestic capital stock, we know that, along a steady-state growth path of the economy,

$$\frac{K}{Y} = \frac{s}{n + g + \delta}. \tag{3}$$

This tells us that, along such a growth path,

$$r = \alpha\left(\frac{n + g + \delta}{s}\right). \tag{4}$$

If permanent shocks that reduce $n + g$ cause the economy to transit from one steady-state growth path to another, the rate of return on capital falls, with the change in r being

$$\Delta r = (\alpha/s)(\Delta n + \Delta g). \tag{5}$$

As long as α is greater than or equal to s—that is, as long as the economy is not dynamically inefficient[12]—the reduction in r will be greater than one for one. From this algebra we would expect the roughly 1.5-percentage-point reduction in the rate of real GDP growth forecast by the Social Security Administration to carry with it a greater than 1.5-percentage-point reduction in r.

These are steady-state results. How relevant are they for, say, the seventy-five-year standard forecast horizon used in analyses of the Social Security system? In the Solow model the capital-output ratio approaches its steady-state value at an exponential rate of $-(1 - \alpha)(n + g + \delta)$, which, at historical values, is roughly 3.6 percent a year. That closes half the gap to the steady-state capital-output ratio in twenty years. After seventy-five years the capital-output ratio has closed 93 percent of the gap between its initial and its steady-state value.

In this simple Solow setup, only three things can operate to prevent a permanent downward shock to $n + g$ from reducing r. Perhaps the depreciation

12. We have every reason to believe that the economy is dynamically efficient, in that capital in the steady state exceeds the "golden rule" level. See Abel and others (1989).

rate δ could fall. We have been unable to think of a coherent reason why a reduction in labor force growth n or labor productivity growth g should independently carry with it a reduction in δ. (However, the reduction in r could plausibly carry with it an extension of the economic lives of equipment and buildings, and so bring about a partly offsetting fall in δ that would moderate the decline in r.) Or perhaps the production function could shift to increase the capital share of income α.

Last, perhaps a permanent downward shock to $n + g$ could also bring about a reduction in the saving rate s. If it were the case that

$$\text{(6)} \qquad ds = -\left(\frac{s}{n + g + \delta}\right)(dn + dg),$$

then the rate of return r would be constant. There is a reason to think that a fall in n would carry with it a reduction in s: an economy with slower labor force growth is an aging economy with relatively fewer young people and, presumably, if the young do the bulk of the saving, a lower saving rate. (A decline in g, however, would tend to work the other way: the income effect would tend to raise s.)

Analysis

Are the effects just discussed plausibly large enough to keep the rate of return on capital constant at the rate of economic growth? To assess that, we need to model saving decisions, which requires moving from algebra to model-based analysis.

The Ramsey Model

We move now from Solow to Ramsey-Cass-Koopmans.[13] Consider a version of this Ramsey model in which the representative household has the following utility function:

$$\text{(7)} \qquad \sum_{t=0}^{\infty}(1+\beta)^{-t}\left(U(C_t)\right)N_t^{1-\lambda},$$

13. See Romer (2000).

where β is the pure rate of time preference, C_t is consumption per household member, and N_t is the number of members of the representative household, growing according to

$$N_{t+1} = (1 + n)N_t, \tag{8}$$

where n now measures growth in the size of the household. In the standard Ramsey model setup as presented by David Romer,[14] the parameter λ equals zero, so that the household utility function becomes

$$\sum_{t=0}^{\infty}(1 + \beta)^{-t}\left(U(C_t)\right)N_t. \tag{9}$$

This choice drives the result that changes in labor force growth do not have long-run effects on steady-state capital-output ratios or rates of return. But, to us at least, this assumption seems artificial. If it is indeed the case that the utility function is that specified in equation 9, then the more members of the household, the merrier: household utility is linear in the number of people in the household but suffers diminishing returns in consumption per capita. A household with this utility function, provided it has control over its own fertility, would choose to grow as rapidly as possible; that would be the way to make individual units of consumption contribute as much as possible to total household utility. It seems reasonable to allow λ to be greater than zero and so have a utility function with diminishing returns both with respect to household consumption per capita and with respect to household size.

There is yet another reason to be uncomfortable with the assumption that $\lambda = 0$. If the term "golden rule" were not already taken in the growth theory literature, we would use it here, for $\lambda = 0$ requires that those household members making decisions in period t love others (the new household members joining in period $t + 1$) as they love themselves. They assemble the household utility function by treating the personal utility that others receive in the future from their consumption per capita as the equivalent of their own personal utility. Since we cannot call this the "golden rule," we instead call it *perfect familial altruism.* If $1 > \lambda > 0$, there is *imperfect familial altruism:* those making decisions in period t care about the personal utility of extra family members in period $t + 1$, but not as much as they care about their own.

14. Romer (2000).

And if $\lambda = 1$, decisionmakers in period t act as if they care only about their own personal utility. We are comfortable with altruism; we are uncomfortable with perfect familial altruism.

To the extent that changes in population growth are due to changes in rates of international migration, the assumption that $\lambda = 0$ is not defensible. The representative agent in period t would then regard the future-period utility of unrelated strangers of different nationality who migrate into the country on an equal footing as her own utility, or the utility of her direct descendants.[15]

In this version of the Ramsey-Cass-Koopmans model, the first-order condition for the representative household's consumption-saving decision is

$$U'(C_t)\mathrm{d}C_t = \frac{(1+n)^{1-\lambda}}{(1+\beta)} U'(C_{t+1})\mathrm{d}C_{t+1}. \tag{10}$$

If the household faces a net rate of return on financial investments of r_f, then

$$\frac{1+r_f}{1+n}\mathrm{d}C_t = \mathrm{d}C_{t+1}, \tag{11}$$

because resources in period $t+1$ must be split among more members of the expanded household.

For log utility we then have

$$\frac{C_{t+1}}{C_t} = \frac{(1+n)^{-\lambda}(1+r_f)}{(1+\beta)}. \tag{12}$$

Along the economy's steady-state growth path, with consumption per worker growing at the rate of labor augmentation g, this becomes

$$r_f = (1+g)(1+n)^{\lambda}(1+\beta) - 1, \tag{13}$$

and in the continuous-time limit,

$$r_f = \beta + g + \lambda n. \tag{14}$$

15. Approximately 0.3 percentage point a year of the slowdown in labor force growth projected by the Social Security trustees' report (Board of Trustees, 2005) is due to a slowdown in immigration.

Looking across steady-state growth paths, one sees that reductions in the rate of output growth per worker g reduce r_f one for one in the case of log utility. (They reduce r_f by a multiplicative factor γ of the change in g in the case of constant-relative-risk-aversion utility: $U(C_t) = [(C_t)^{1-\gamma}]/[1-\gamma]$.) Reductions in the rate of labor force growth n also reduce r_f except in the case of $\lambda = 0$. If $1 > \lambda > 0$, slower rates of labor force growth reduce r_f, but less than one for one. And if $\lambda = 1$, decisionmakers in period t are not altruistic at all: they act as if they care only about their own personal utility, and reductions in n reduce r_f one for one—the same amount as do reductions in g.

The Ramsey model converges to a balanced-growth path, and this plus the assumption of a representative agent is sufficient to nail down the relationship between economic growth and asset returns. In the steady state, consumption per capita is growing at rate g, and so the relative marginal utility of consumption per capita one period into the future is

$$(1+\beta)^{-1}(1+g)^{-1} \tag{15}$$

in the case of log utility. And the rate at which consumption per capita can be carried forward in time is

$$(1+r_f)(1+n)^{-1}. \tag{16}$$

To drive the rate of return on capital r_f away from

$$r_f = (1+g)(1+n)^{\lambda}(1+\beta) - 1 \tag{17}$$

requires that the consumption of those agents who are marginal in making the consumption-saving decision in period t grow at a rate different from that of growth in consumption per capita. This requires heterogeneous agents. And the simplest suitable model with heterogeneous agents is the Diamond model.

The Diamond Model

In the overlapping-generations model of Peter Diamond,[16] each agent lives for two periods, works and saves when young, and earns returns on

16. Diamond (1965).

capital and spends when old. Thus, for a given generation that is young in period t, their labor income per worker when young w_t, their consumption per worker when young c_{yt}, their consumption per worker when old c_{ot+1}, the *net* rate of return on capital r_{t+1}, and the economy's capital stock per worker in the second period k_{t+1} are all linked:

(18) $$w_t = c_{yt} + k_{t+1}$$

(19) $$c_{ot+1} = (1 + r_{t+1})k_{t+1}.$$

With a Cobb-Douglas production function, output per (young) worker when the period-t generation are young—in period t—is

(20) $$y_t = E_t^{1-\alpha}\left(\frac{k_t}{1+n}\right)^{\alpha},$$

where E is our measure of the efficiency of labor, growing at proportional rate g each period, and where $(1 + n)$ appears in the denominator because n is the rate of population growth per generation. With this production function, labor income is a constant fraction of output per worker,

(21) $$w_t = (1 - \alpha)y_t,$$

and the real return on capital will be the residual, capital income, divided by the capital stock:

(22) $$r_t = \frac{\alpha y_t}{k_t} = \frac{\alpha E_t^{1-\alpha} k_t^{\alpha-1}}{(1+n)^{\alpha-1}}.$$

Once again take time-separable log utility for our utility function,

(23) $$\ln(c_{yt}) + \frac{\ln(c_{ot+1})}{1+\beta}$$

and look for steady states in capital per effective worker by requiring that

(24) $$k_t = E_t k^*.$$

From this we get the following steady-state first-order condition:

(25) $$\frac{1}{c_{yt}} = \frac{(1+r)}{(1+\beta)}\frac{1}{c_{ot+1}}.$$

The model can be solved by substituting in the budget constraint,

(26) $$\frac{1}{\left[(1-\alpha)E_t^{1-\alpha}\left(\frac{k_t}{1+n}\right)^{a} - k_{t+1}\right]} = \frac{(1+r)}{(1+\beta)}\frac{1}{(1+r)k_{t+1}},$$

to get

(27) $$\frac{1}{\left[\frac{1-\alpha}{1+g}\left(\frac{k*}{1+n}\right)^{\alpha} - k*\right]} = \frac{1}{(1+\beta)k*},$$

which leads to

(28) $$k* = \left(\frac{(1-\alpha)}{(1+g)(1+n)^{\alpha}(2+\beta)}\right)^{\left(\frac{1}{1-\alpha}\right)}.$$

Recalling that $r = (\alpha k*^{\alpha-1})/(1+n)^{\alpha-1}$, we have

(29) $$r = \left(\frac{\alpha(1+g)(1+n)(2+\beta)}{(1-\alpha)}\right).$$

In this equation, the lower the rate of productivity growth g, and the lower the rate of labor force and population growth n, the lower is the rate of return on capital r.

Conclusion

Thus, in the Diamond overlapping-generations model as well as in the Ramsey model and the Solow model, slower economic growth comes with lower net returns on capital. In the Ramsey model, there is reason to think that reductions in labor productivity growth have a greater effect on rates of return than do reductions in labor force growth:

—In the basic Solow algebra, the reduction in gross returns r is proportional to (α/s) times the reduction in growth.

—In the Diamond model, the reduction in net returns r_f is equal to $2\alpha/(1 - \alpha)$ times the reduction in labor productivity growth g and, to a first approximation, equal to $2\alpha/(1 - \alpha)$ times the reduction in labor force growth n.

—In the Ramsey model, the reduction in r_f is equal (with log utility) to the reduction in labor productivity growth g and, to a first order, to λ times the reduction in labor force growth n (where λ is the degree to which familial altruism is imperfect).

At some level, the same thing is going on in all three setups. Reductions in economic growth in these setups are all declines in the rate of growth of effective labor relative to the capital stock provided by previous investment. Effective labor becomes relatively scarcer and capital relatively more abundant. The terms of trade move against capital, and so the return to capital falls.

Why, then, does a fall in labor force growth not reduce rates of return in the Ramsey model in the case of perfect familial altruism, $\lambda = 0$? Because a reduction in population growth also reduces the utility value of moving consumption forward in time—an important component of the value of saving in the Ramsey model with perfect familial altruism comes from the possibility of dividing the saving among more people in the future and thus escaping the diminishing marginal utility of consumption. Thus the marginal household utility of saving falls in the Ramsey model when population growth falls. This fall reduces the effective supply of capital by as much as the fall in the rate of population growth reduces the effective supply of labor. To the extent that a slowdown in economic growth is driven by a reduction in the rate of immigration, this representative-agent effect in the Ramsey model is not an effect that we want the model to have: perfect familial altruism is not an assumption that anyone would wish to make.

These models say that there is some economic reason to believe that a slowdown in economic growth would carry a reduction in asset returns with it. These models are the standard models that economics graduate students and their professors use routinely. They are oversimplified. They are abstract. They are ruthlessly narrow in their conceptions of human motivation and institutional detail. Are they relevant to the real world? Are they telling us something that we should hear when we try to forecast the long-run future?

Arithmetic

Is it possible to imagine scenarios in which asset returns remain close to their historical averages even when real GDP growth slows markedly? Yes. Are any such scenarios plausible forecasts in the sense of being the central tendency of a distribution of possible futures? We believe not. In this section we conduct some simple arithmetic exercises to make our case.

Earnings and Returns

Jeremy Siegel believes that stocks are "in the middle range of fair market value" and that therefore the current earnings yield of 5.45 percent is a "good long-term estimate of real returns."[17] The sum of dividend payouts, net buybacks, and investment financed by net retained earnings must add up to 5.45 percent of today's stock values.[18] Returns to investors are payouts—dividends and net buybacks—plus the value of investments financed by net retained earnings.

Firms, which have traditionally paid out, on average, roughly 60 percent of their accounting profits through dividends and buybacks and rely on retained earnings to finance a substantial share of any increase in their capital stock, have little room to boost risk-adjusted returns by massively expanding payouts, unless they can do so without crippling their earnings growth—that is, unless a good deal of today's retained earnings are wasted. Firms similarly have little room to boost risk-adjusted rates of return on their equity by cutting back on payouts, unless there are very large wedges between rates of return on retained and reinvested earnings and rates of return in the market—that is, unless firms have been massively underinvesting. Current earnings yields thus suggest that the stock market is in accord with the logic of our algebra and analysis: it is not anticipating the average real return on the stock market of 6.5 percent a year or so realized over the past half century.

But reported accounting earnings are not true Haig-Simons earnings (that is, equal to the amount that can be consumed from earnings without

17. Siegel (2005, p. 8).

18. There is a wedge of 0.3 percentage point a year between the GDP deflator and the CPI. Siegel's estimated real rate of return becomes 5.15 percent a year in the CPI-basis numbers used by the Social Security Administration.

changing wealth).[19] There is good reason to believe that returns on retained earnings are higher than market returns.[20] And it is at least plausible that the wedge between market returns and returns on retained earnings depends on the rate of economic growth: faster growth means higher demand and greater profits if returns to scale are increasing. So the argument that earnings yields do not support high expected equity returns needs to be shored up by an explicit look ahead at how payouts and values might evolve.[21]

Dividend Yields, Returns, and Growth

Begin with the identity that is the Gordon equation for equity prices:

$$P = \frac{D}{r_e - g}, \tag{30}$$

where D are the dividends paid on a stock or an index of stocks, P is the corresponding price, r_e is the expected real rate of return *on equities,* and g is the expected permanent real growth rate of dividends. This is a standard way to approach the determinants of equity prices as a whole.[22] In this framework the real rate of return on equities is

$$r_e = \frac{D}{P} + g. \tag{31}$$

Returns on an index of stocks differ from the current dividend yield plus the growth rate of economy-wide corporate earnings for two important reasons:

—First, g will be less than the growth rate of economy-wide corporate earnings because those earnings are the earnings of newly created companies that were not in the index last period. Corporate earnings are a return to entrepreneurship as well as capital; hence the rate of growth of economy-wide earnings will in general outstrip that of the earnings of the companies represented in a stock index.

—Second, dividends are not the only way firms pump cash to shareholders. Stock buybacks decrease the equity base and thus push up the

19. Haig (1921).
20. See Hubbard (1998).
21. See Baker (1997) for the first argument along these lines of which we are aware.
22. Campbell and Shiller (1988).

rate of growth of the earnings on the index (as opposed to the earnings of the companies in the index).

It is convenient to think of both of these factors as affecting the payout ratio rather than the growth rate, and to replace equation 31 with

$$r_e = \frac{D + B}{P} + g, \tag{32}$$

where B is net share buybacks (buybacks less initial public offerings), and g is now the growth rate of $D + B$.[23]

The 2005 report of the Social Security trustees projects a long-run real GDP growth rate of 1.8 percent a year on a GDP deflator basis.[24] It projects that labor and capital shares will remain constant in the long run.[25] With a long-run gap of 0.3 percentage point between the consumer price index (CPI) and the GDP deflator,[26] and with an auxiliary assumption that capital structures are in balance, this is an implicit forecast that the variable g in the Gordon equation will be 1.5 percent a year. Current dividend yields on the Standard and Poor's (S&P) 500 index are 1.9 percent a year. Current net stock buybacks are 1.0 percent a year. The sum of these is 4.4 percent a year, which is thus the expected real rate of return r in the Gordon equation. That is significantly lower than the 6.5 percent real rate of return that is the historical experience of the American stock market.

Possible Ways Out

Are there ways to escape from this arithmetic of earnings and payouts? Yes. The U.S. economy is not on a steady-state growth path. Three potential ways out seem most worth exploring:

—Perhaps the stock market is currently overvalued and will decline, significantly raising payout yields.

23. Subtracting initial public offerings ensures that the ratio of total economy-wide earnings to the earnings of companies in the index does not grow. Adding gross buybacks takes account of the antidilution effects of narrowing the equity base of companies currently in the index.

24. Board of Trustees (2005, table V.B2).

25. The assumption of a constant income share follows from the derivation of real wage growth from productivity growth, which is discussed on pages 85–88 of Board of Trustees (2005).

26. Board of Trustees (2005, table V.B1).

—Perhaps payout growth will be unusually rapid in the near term before slowing to its long-term forecast trend rate of 1.5 percent a year.

—Perhaps the distribution of world investment will shift in a way that allows U.S. companies to earn greater and greater shares of their profits abroad.

Diamond argues for the first possibility.[27] A decline in the stock market, relative to the economy's growth trend, of 40 percent would carry payout yields up to the 5.0 percent consistent with a long-run real return of 6.5 percent a year and real profit and dividend growth (on a CPI basis) of 1.5 percent a year. Such a scenario is certainly possible: it was the stock market's experience between the late 1960s and the early 1980s. But we have a hard time seeing it as the central tendency of the distribution of possible futures.[28]

The second possibility requires payouts—both dividends and net stock buybacks—to grow rapidly in the near term to validate a subsequent real growth rate of 1.5 percent a year and a current expected real return of 6.5 percent a year. If such growth were to be concentrated in the next decade, the real payouts of the companies in the S&P index would have to grow at an average of 8.6 percent a year. Over the past fifty years the earnings on the S&P index have grown at an average rate of 2.1 percent a year. It could happen: perhaps we are in the middle of a permanent shift in the distribution of income away from labor and toward capital. But, once again, we regard these as unlikely scenarios, not as the central tendency of the distribution of possible futures that is a rational forecast.

The third way out is the one that we regard as the most interesting possibility. We take it up in the next section.

The Open-Economy Case

In any open economy the steady-state Gordon equity valuation equation is as before, except that the rate of growth is not that of the domestic cor-

27. Diamond (2000).

28. Certainly no investment adviser who anticipates that real equity returns will average −0.6 percent a year over the next decade has any business advising clients to shift their portfolio in the direction of equities today. That is true even when the U.S. government is the adviser, and the relatively young future beneficiaries of Social Security are the clients.

porate capital stock g but that of the capital stock owned by American companies, g_k:

$$r_e = \frac{D + B}{P} + g_k. \tag{33}$$

If foreign companies, on net, invest in America—that is, if the United States on average runs a current account deficit—then the rate of growth of the earnings of American companies in our domestic stock market index will be slower than the rate of growth of economy-wide earnings and of real GDP. The open economy will then deepen rather than resolve the problem of combining slow expected growth with high expected returns. If instead it is American companies that, on net, invest abroad, then the rate of growth of the capital stock, and thus of the earnings of companies in the index, will exceed the rate of growth of the domestic economy g.

How much larger? If we look over spans of time long enough for adjustment costs in investment not to be a major factor, the value of the capital stock will be proportional to the size of the capital stock.[29] If we assume in addition that companies maintain stable debt-equity ratios, we have

$$g_k = g + x\left(\frac{Y}{K}\right), \tag{34}$$

where x is that component of the current account surplus (as a share of GDP) that corresponds to American companies' net investments abroad,[30] and Y/K is the ratio of current output to corporate capital.

29. We here dismiss the possibility that investments overseas might provide higher risk-adjusted rates of return in the long run than domestic investments: Tobin's $q = 1$ both here and abroad. The Bureau of Economic Analysis reports that as of the end of 2003 the market value of foreign-owned assets in the United States is about \$10.5 trillion, compared with foreign assets held by U.S. residents of about \$7.9 trillion, yet the associated income flows are about the same. We attribute this difference to a difference in risk. The experience of nineteenth-century British investors with such landmarks of effective corporate governance as the Erie Railroad suggests that, although there are supernormal returns to be earned in the course of rapid economic development, people with offices separated by oceans are unlikely to be the ones who reap them.

30. The phrase "corresponds to American companies' net investment abroad" is needed to abstract from current account deficits that finance net government consumption or net household consumption.

Here, again, we return to arithmetic. Our rate of return on equities is

$$r_e = \frac{D+B}{P} + g + x\left(\frac{Y}{K}\right). \quad (35)$$

From the previous section this is

$$r_e = 4.4\% + x\left(\frac{Y}{K}\right). \quad (36)$$

Assuming a capital-output ratio of 3, we then have

$$x = 3\left(r_e - 4.4 \text{ percent}\right). \quad (37)$$

In words: for any excess of the rate of return on equities over the closed-economy benchmark case of 4.4 percent a year, three times that figure is the current account surplus associated with net corporate investment overseas needed to produce the higher return.

Note that, for a constant rate of return, the needed surplus grows over time. In equations 34 through 36, *Y*/*K* is not the physical domestic output-to-capital ratio; it is the ratio of domestic output to total capital owned by American companies—including capital overseas. As overseas assets mount, the needed surplus for constant payout yields mounts as well.

Such enormous current account surpluses are possible. Great Britain had them in the quarter-century before World War I, when it ceased to be the workshop of the world and became for a little while its financier.[31] Slowing economic growth in the late Victorian and Edwardian eras and reduced investment relative to national saving were cause (or consequence, or possibly both) of the direction of Britons' saving and of British companies' investment overseas. We see no signs that the United States will undertake a similar trajectory over the next several generations. And we are impressed by the scales involved: to be consistent with current payout yields, and given a forecast real GDP growth rate of 1.8 percent a year, to achieve 6.5 percent annual returns on equity the current account surplus produced by American net corporate investment abroad would have to begin at 6 percent of GDP and grow thereafter.

31. Edelstein (1982).

Could such large outward levels of net corporate investment abroad be consistent with relatively balanced overall trade—in other words, could they be offset by large net portfolio investment inside the United States? Not without additional forces at work. The reason is that the open-economy saving-investment relation,

$$S - NX \equiv I, \tag{38}$$

(where NX is net exports) is an identity. Consider the three uses that such large inward portfolio investments could have:

—They could be used to purchase securities newly issued by American businesses to finance investment in the United States. The flow of inward portfolio investment would add as much to domestic investment as the outward-directed flow of corporate investment would have subtracted. There would be no slowdown in the rate of growth of the domestic capital stock. Thus the rising domestic capital-output ratio would push down rates of return at home. Since foreigners are making these large portfolio investments in the United States, this fall in domestic rates of return would be associated with a similar fall in foreign rates of return as well.

—They could be used to purchase securities newly issued by American businesses to finance investment abroad. In this case, gross foreign direct investment by domestic firms would have to be large enough not only to absorb the difference between domestic investment and domestic saving, and so slow down the rate of growth of the domestic capital stock, but also to neutralize the portfolio capital inflow. We are thus back to square one.

—They could be used to purchase already-existing assets from Americans, who then do not reinvest the proceeds either in expanding the domestic capital stock or in further funding American investment abroad, but instead consume the proceeds.[32] This means massive dissaving on the part of those who sell their assets to foreigners: a large fall in S. Once again, we see a possible scenario but not the central tendency of the distribution of possible futures that would constitute a forecast.

32. This is the possibility that Mankiw stresses in his comment on this paper: that if domestic saving rates fall sharply, the reduction in the rate of growth of the domestic capital stock required to keep rates of return high can be accomplished without a large current account surplus.

The Equity Premium

Economists do not have a good explanation of the equity premium. Rajnish Mehra and Edward Prescott titled their well-known paper "The Equity Premium: A Puzzle," for good reason.[33] Stocks have outperformed fixed-income assets by more than 5 percentage points a year for as far back as records go. As Martin Feldstein, former chairman of the Council of Economic Advisers, has often said, it is as if the market's attitude toward systematic equity risk were that of a rich sixty-five-year-old male with a not-very-healthy lifestyle, whose doctor has told him that he is likely to live less than a decade. Yet we believe that properly structured markets should—and can—mobilize a much deeper set of risk-bearers with a much greater risk tolerance. That they do not appear to have done so is a significant mystery. We find ourselves persuaded by Mehra that the equity premium remains a puzzle, unexplained by rational agents in models that maximize individual utility.[34]

It is quite possible that a substantial part of the equity premium is a thing of the past, not the future.[35] In the distant past the fear of a recurrence of the railroad and other "robber baron" scandals, and in the more recent past the memory of the Great Depression, kept some investors excessively averse to stocks. In addition, the United States has had remarkably good economic luck—a point stressed by Robert Shiller.[36] And, over time, as people realized that their predecessors had been excessively fearful of equity risk, rising price-dividend ratios pushed a further wedge between stock and bond returns. But today our arithmetic projects stock returns of 4.4 percent a year, for an equity premium of perhaps 2.5 percentage points, not 5.

To the extent to which this past behavioral anomaly was the result of an excessive fear of stocks and an excessive attachment to bonds, it is not clear that its erosion should have an impact on the expected return on a balanced portfolio. The simplest, crudest, and most extremely ad hoc model of the equity premium would embed the stock-versus-bond investment decision in the simplest possible Diamond-like overlapping-generations

33. Mehra and Prescott (1985).

34. Mehra (2003).

35. In conversation, Randall Cohen of the Harvard Business School has been an especially forceful advocate of this point of view.

36. Shiller (2005).

model, with the capital stock each period being the wealth accumulated when young by the old, retired generation. Assume that each generation, when it saves, invests a share e_h of its savings in equities and a share $1 - e_h$ in bonds. Firms, however, are unhappy with such a capital structure. Unwilling to run a significant risk of bankruptcy, they are unwilling to commit less than a share e_f, where $e_f > e_h$, of their payouts to equity. A smaller cushion—in the sense that a smaller cyclical decline in relative profits would run the risk of missing bond payments and drawing an appointment with a bankruptcy court—is simply unacceptable to entrenched managers.

If a physical unit of saving when one is young yields returns to physical capital r when one is old, the rates of return on equity and debt, r_e and r_d, respectively, are then calculated as

$$1 + r_e = (1 + r)\left(\frac{e_f}{e_h}\right) \tag{39}$$

$$1 + r_d = (1 + r)\left(\frac{1 - e_f}{1 - e_h}\right), \tag{40}$$

with the equity premium being

$$\frac{1 + r_e}{1 + r_d} = \frac{e_f/(1 - e_f)}{e_h/(1 - e_h)}. \tag{41}$$

In this excessively simple framework, it does seem highly plausible that $(1 + r_e)/(1 + r_d)$ has fallen with greater household willingness to hold equity, because of institutional changes (such as revisions of the "prudent man" rule, the growth of IRAs and 401(k)s, and lower transactions costs associated with stock trading), the fading memory of 1929, two decades of fabulous bull markets, and increased financial sophistication on the part of households.

Thus, even if there were no reasons connected with slowing growth to expect lower returns on capital, one might well anticipate lower returns on equity in the future than in the past. And past decades have seen institutional changes that one would expect, from a behavioral perspective, to boost the share of financial assets channeled to equities.[37]

37. Barberis and Thaler (2003).

A lower rate of return on the assets in a balanced portfolio has powerful implications for economic policy. A lower equity premium seems, to us at least, to have powerful implications for one issue, namely, whether the stock market's apparent failure to mobilize society's risk-bearing capacity is a large-scale market failure, and whether a government-run social insurance scheme can and should attempt to profit from (and thus repair) this failure to mobilize society's risk-bearing resources. The government has the greatest ability of any agent in the economy to manage systematic risk. If other agents are not picking up their share—and if, as a result, there are properly adjusted excess returns to be earned by the government's taking a direct position itself or assuming an indirect position by reinsuring individuals' social insurance accounts—why should the government not do so?

The difference between, broadly speaking, the economists of the coasts and the economists of the interior is that the first specialize in thinking up clever schemes to repair apparent market failures, whereas the second specialize in thinking up clever reasons why apparent market failures are not really so. Even though we are from the coasts, we find enough reasons to believe that the equity premium will be smaller in the future than in the past to prefer that attempts to exploit it be implemented slowly and gingerly.

Conclusion

We see strong reasons to think that, over the long run, rates of return on assets are correlated and causally connected with rates of economic growth. We would expect the reduction in asset returns to be greater for a given reduction in productivity growth than for an equal reduction in labor force growth. We think that reductions in asset returns could be offset and even neutralized by other factors—by capital expropriating some of what has been labor's share of income, by a failure of today's stock market values to soberly reflect likely future returns rather than irrational exuberance, or by the United States cutting its consumption beneath its production for generations and following Britain's pre–World War I trajectory as supplier of capital to the world. But we see these as unlikely (although possible) scenarios. We do not see any of them as the central tendency of the distribution of possible futures that is a proper economic forecast. And although a combination of partial moves in each of the three directions could achieve the result, we see no good reason to presume that such a scenario is likely.

We see the two strands of our argument—our arithmetic demonstration that equity returns as high in the future as in the past are unlikely, and our analytical arguments that rates of return and rates of growth are likely to move together—as reinforcing each other. Returns must be consistent with the saving decisions of households, the investment decisions of firms, and the technologies of production. But returns must also equal payout yields plus capital gains—only in stock market bubbles can capital gains diverge widely from economic growth, and then only for a little while.

Powerful economic forces work to make sure that what the economy's behavioral relationships produce is consistent with its equilibrium flow-of-funds conditions. That is the logic that applies here: if slower economic growth reduces the arena for the profitable deployment of capital, rates of return will fall until less capital is deployed. By how much will they fall? Until—in steady state—payout yields plus retained earnings are equal to profits, and retained earnings are no larger than the sustainable growth of the capital stock permits.