MEASURING THE SOCIAL RETURN TO R&D*

CHARLES I. JONES AND JOHN C. WILLIAMS

Is there too much or too little research and development (R&D)? In this paper we bridge the gap between the recent growth literature and the empirical productivity literature. We derive in a growth model the relationship between the social rate of return to R&D and the coefficient estimates of the empirical literature and show that these estimates represent a lower bound. Furthermore, our analytic framework provides a direct mapping from the rate of return to the degree of underinvestment in research. Conservative estimates suggest that optimal R&D investment is at least two to four times actual investment.

I. INTRODUCTION

Do advanced economies engage in too much or too little R&D?[1] By how much does private investment in research differ from optimal investment? Given the central role of R&D as an engine of growth, these questions have spawned a large theoretical and empirical literature. Theory has emphasized the importance of market failures such as imperfect competition and externalities in determining outcomes in the market for new goods and ideas.[2] However, because there are incentives working to promote both over- and underinvestment in R&D, theory alone is unable to provide an unambiguous answer to the sign, much less the magnitude, of the net distortion to R&D. The empirical literature attempts to resolve this ambiguity by estimating directly the rate of return to R&D in regressions of productivity growth on measures of R&D.[3] The findings of this literature are summarized by Griliches [1992, p. S43]: "In spite of [many] difficulties, there has been a significant number of reasonably

* A previous version of this paper was circulated under the title "Too Much of a Good Thing? The Economics of Investment in R&D." We would like to thank Eric Bartelsman, Roland Bénabou, Kenneth Judd, Michael Horvath, Samuel Kortum, Ariel Pakes, Scott Stern, Alwyn Young, and participants in numerous seminars for helpful comments. Financial support from the National Science Foundation (SBR-9510916) is gratefully acknowledged. The opinions expressed here are not necessarily shared by the Board of Governors of the Federal Reserve System or its staff.

1. We should emphasize from the beginning that this paper is not about basic science but rather about applied R&D undertaken by profit-maximizing firms. Of course, we recognize that the distinction is sometimes difficult to make in practice.

2. The theoretical literature includes contributions from the IO approach, as reviewed by Tirole [1988], as well as the general equilibrium approach exemplified by Romer [1990], Grossman and Helpman [1991], and Aghion and Howitt [1992].

3. Recent summaries of this literature include Cohen and Levin [1991], Griliches [1992], and Nadiri [1993].

The Quarterly Journal of Economics, November 1998

well-done studies all pointing in the same direction: R&D spillovers are present, their magnitude may be quite large, and social rates of return remain significantly above private rates." The empirical approach seems to provide a clear answer to the question of whether there is too much or too little private R&D; however, it does not indicate by how much R&D investment needs to be increased.

In fact, theory provides some reason to question the findings of the empirical productivity literature. The results of this literature are nearly all based on a neoclassical theory of growth in which R&D is simply an alternative form of capital investment. This simple capital-based approach ignores many of the distortions associated with research that are formalized by new growth theory, including monopoly pricing, intertemporal knowledge spillovers, congestion externalities, and creative destruction. Because these considerations are omitted in the empirical literature, we may in fact have very little information about the true social rate of return to R&D.

The main contribution of this paper is to link the theoretical models of new growth theory to the empirical results in the productivity literature. In doing so, we provide a general method for computing social rates of return. We then derive analytically the relationship between the true social rate of return to R&D and the coefficient estimates from regressions of total factor productivity growth on R&D investment. In the process, we provide an intuitive explanation for the various components that make up the social return to R&D. We also derive the relationship between the magnitude of over- or underinvestment in R&D and the estimated rate of return.

The results are rather surprising. Despite the methodological limitations of the productivity literature—in particular its omission of distortions that might lead to overinvestment—we show that the estimates in this literature represent *lower* bounds on the social rate of return to R&D. Thus, estimates of the rate of return to R&D from the productivity literature of 30 percent or higher imply that advanced economies like the United States substantially underinvest in R&D. Based on results from new growth theory, one might be inclined to question the broad conclusions of the productivity literature; in contrast to this intuition, we show that the findings of the productivity literature are extremely robust.

With an estimate of the social return to R&D in hand, a lower bound on the degree of underinvestment in R&D can be computed directly. Using a conservative estimate of the social return of 30 percent and a private rate of return to capital of 7 to 14 percent, optimal R&D spending as a share of GDP is more than two to four times larger than actual spending.

The methodology developed in this paper allows one to derive these results directly from the production possibilities of the economy: the production function for new ideas and the production function for the consumption/output good. It does not rely on any particular assumptions regarding market structure, the patent system, or distortionary taxes.[4] More generally, this approach can be applied to a wide variety of models.

The remainder of this paper is organized as follows. We begin in Section II with a general derivation of the social rate of return to research. Section III reviews the methodology and results of the empirical productivity literature and relates the true social rate of return to R&D to the estimates in this literature. In Section IV the magnitude of over- or underinvestment in R&D is derived and related to the estimated social return, and Section V concludes.

II. The Social Rate of Return to R&D

What is the rate of return to society from performing an additional unit of R&D? To answer this question, we consider the return associated with the following variational argument. Suppose that we reallocate one unit of output from consumption to R&D today, and then consume the proceeds tomorrow. In particular, we reduce R&D tomorrow to leave the subsequent stock of ideas unchanged. In the market equilibrium an individual agent is indifferent to undertaking this deviation herself, but in the presence of distortions and externalities, society as a whole generally will not be. We define the social rate of return to R&D to be the gain in consumption associated with this variation. This particular definition turns out to have a number of useful properties that we will now explore.

A. General Derivation

We begin with a general derivation of the social rate of return. The first useful result related to our definition is that the social

4. In this sense, it is interesting to compare this approach with Stokey [1995]. Stokey, and our own earlier work, address the issue of investment in R&D by calibrating an R&D-based growth model. The results in this approach depend critically on how one characterizes the market economy.

rate of return can be derived solely from the production possibilities of the economy. In the case of R&D considered here, only two equations are needed: the production function for ideas and the production function for final output. Let A denote the stock of ideas in the economy. New ideas, the change in A, are produced by forgoing consumption of the final output good Y, according to some production function G:

$$A_{t+1} - A_t = G(R_t, A_t), \quad (1)$$

where R represents resources devoted to research. We assume that G is increasing in its first argument: more research leads to more ideas. G might be increasing or decreasing in its second argument, depending on the way past ideas affect the current productivity of research. If $\partial G/\partial A > 0$, then past inventions raise the productivity of research today, a case that corresponds to "knowledge spillovers" in research. On the other hand, if the best ideas are discovered first, G might be decreasing in A.

The consumption/final output good is produced using ideas and a collection of private inputs X according to the production function F:

$$Y_t = F(A_t, X_t). \quad (2)$$

We assume that F is increasing in each of its arguments. Following Romer [1990], one would expect F to exhibit constant returns to X and therefore increasing returns to scale overall.

We will assume the existence of a balanced growth path in which all variables are growing at constant rates over time. This may entail some restrictions on the shapes of G and F; we will specialize to the Cobb-Douglas functional forms shortly. Our use of a more general notation is not necessarily intended to suggest generality. Rather, it illuminates the source of each term in the social rate of return.

The social rate of return to R&D is computed using the following discrete time variational argument. Suppose that we reallocate one unit of output from consumption to R&D at time t, and then consume the proceeds in the next period, $t + 1$. Moreover, we reduce R&D at time $t + 1$ so as to leave the stock of knowledge unchanged from time $t + 2$ onward. The total gain in consumption at time $t + 1$ associated with this variation is the social rate of return to R&D.

The increase in A_{t+1} associated with a small change in R_t is

$$\nabla A_{t+1} = \left(\frac{\partial G}{\partial R}\right)_t,$$

where ∇ is used to denote the change relative to the steady state path. The additional knowledge ∇A_{t+1} increases output at time $t+1$ by $(\partial Y/\partial A)_{t+1}$. An additional increase in consumption at time $t+1$ occurs because R_{t+1} can be reduced to leave the path of knowledge unchanged. To determine how much consumption is gained from reducing R&D, note that

$$A_{t+2} = A_{t+1} + G(R_{t+1}, A_{t+1}).$$

Considering the deviation from the balanced growth path,

$$\nabla A_{t+2} = \nabla A_{t+1} + \left(\frac{\partial G}{\partial R}\right)_{t+1} \nabla R_{t+1} + \left(\frac{\partial G}{\partial A}\right)_{t+1} \nabla A_{t+1}.$$

The deviation in R&D, ∇R_{t+1}, that will return the stock of knowledge to its original path is found by setting $\nabla A_{t+2} = 0$:

$$\nabla R_{t+1} = -\frac{(\partial G/\partial R)_t}{(\partial G/\partial R)_{t+1}} \left(1 + \left(\frac{\partial G}{\partial A}\right)_{t+1}\right).$$

The total gain to consumption in period $t+1$ is the sum of the additional output produced and the reduction in R&D that is made possible. The social rate of return, $\tilde{r}$, is thus given by

$$1 + \tilde{r} = \left(\frac{\partial G}{\partial R}\right)_t \left(\frac{\partial Y}{\partial A}\right)_{t+1} + \frac{(\partial G/\partial R)_t}{(\partial G/\partial R)_{t+1}} \left(1 + \left(\frac{\partial G}{\partial A}\right)_{t+1}\right). \tag{3}$$

The intuition behind this equation becomes more transparent if one thinks of knowledge as an asset "purchased" by society, held for a short period of time in order to reap a dividend, and then sold. The return can then be thought of as the sum of a dividend and a capital gain (or loss). Let $P_{A,t}$ denote the cost to society of a new idea in units of consumption (the numeraire). Then, because a small change in R&D leads to $\partial G/\partial R$ new ideas, P_A is given by

$$P_{A,t} = \left(\frac{\partial G}{\partial R}\right)_t^{-1}.$$

The rate of change in the cost of producing new ideas, denoted $g_{P_{A,t}}$

equals

$$g_{P_{A,t}} = \frac{(\partial G/\partial R)_{t-1}}{(\partial G/\partial R)_t} - 1,$$

which is constant along a balanced growth path.

After rearrangement and substitution, the social rate of return satisfies

$$\tilde{r} = d/P_A + g_{P_A}, \tag{4}$$

where

$$d = \frac{\partial Y}{\partial A} + \frac{\partial G}{\partial A} P_A. \tag{5}$$

In equation (4), d is the "dividend" to society and g_{P_A} is the "capital gain."[5] The dividend associated with an additional idea consists of two components. First, the additional knowledge directly raises the productivity of capital and labor in the economy. Second, the additional knowledge changes the productivity of future R&D investment because of either knowledge spillovers or because subsequent ideas are more difficult to discover. Finally, there is a capital gain or loss associated with any change in the cost of producing new ideas, denoted g_{P_A}.

B. A Specific Model

The preceding derivation is purposefully abstract. To make the ideas concrete, we now derive the social rate of return to R&D using Cobb-Douglas specifications for the final goods and research technologies. For ease of presentation, we switch to continuous time. In this generalized version of Romer's [1990] variety-based endogenous growth model, the final goods technology is given by

$$Y = A^{\sigma} K^{\alpha} L^{1-\alpha}, \tag{6}$$

where L is labor input, and K is the (aggregated) capital stock. We assume that L grows exogenously at rate $n > 0$, and capital is accumulated in the standard way, by forgoing consumption.

The production function for new ideas takes the form,

$$(1 + \psi)\dot{A} = \tilde{\delta} R = \delta R^{\lambda} A^{\phi}. \tag{7}$$

Individual researchers take the productivity of research $\tilde{\delta}$ as

5. The second-order cross term $g_{P_A}(\partial G/\partial A)$ has been suppressed.

given. Because they are small relative to the total number of researchers, they view the production of new ideas as taking place with constant returns to research effort R. However, economywide production of ideas need not be characterized by constant returns. For example, the presence of $0 < \lambda \leq 1$ may reflect duplication of effort in the research process: the social marginal product of R may be less than the private marginal product, a classic congestion externality. The parameter ϕ measures the net effect of knowledge spillovers and "fishing out" effects in research, both external to atomistic research firms. If the net effect is such that $\phi > 0$, we might call this the *standing on shoulders effect*. The duplication externality associated with $\lambda < 1$ might be called the *stepping on toes effect*.

A third distortion in the research process, highlighted by Grossman and Helpman [1991] and Aghion and Howitt [1992], is associated with creative destruction. That is, new ideas may replace old ideas. Creative destruction can provide an incentive for overinvestment in research in that some innovators earn rents on ideas that are not entirely new. In the market economy, creative destruction affects who gets compensated for which idea, and one has to be careful in describing how this process works. However, in terms of the production possibilities of the model (which are relevant for calculating the social return), introducing creative destruction is straightforward. Creative destruction drives a wedge between the production of new ideas and the total output from research. In equation (7) we assume that for every new idea created, ψ existing ideas get "repackaged," with the most recent inventor entitled to all subsequent rents.

With these functional form assumptions, the social rate of return to R&D implied by equation (4) is

$$\tilde{r}(s) = \lambda\sigma g_A/s + \phi g_A + (g_Y - g_A), \tag{8}$$

where $\tilde{r}(s)$ represents the steady state social rate of return to R&D, evaluated at a given steady state R&D share of total output, s. The notation g_x is used to indicate the steady state growth rate of the placeholder x. The first term on the right-hand side of equation (8) is the dividend associated with extra output, the second term is the dividend associated with knowledge spillovers, and the last term is the capital gain associated with the changing relative value of ideas.

Implicit in this presentation is that steady state growth rates are not affected by the allocation of resources. This is true because the model considered here is a semi-endogenous growth model like that of Jones [1995]. In this model, g_A is given by $\lambda n/(1 - \phi - \lambda\sigma/(1 - \alpha))$. More generally, however, the analysis extends to the case in which the steady-state growth rates also depend on the allocation of resources.

Equation (8) identifies the functional relationship between the social return to research and the share of output invested in research by the economy. This functional relationship is plotted in Figure I. This figure, together with the analysis in the previous section, motivates our first key result concerning social rates of return:

RESULT 1. The functional relationship between the social rate of return and the share of resources devoted to research depends only on the production possibilities of the economy. Features of the market economy affect the allocation of resources, which determines the point on the social return function.

The attractiveness of this result is that one does not need to

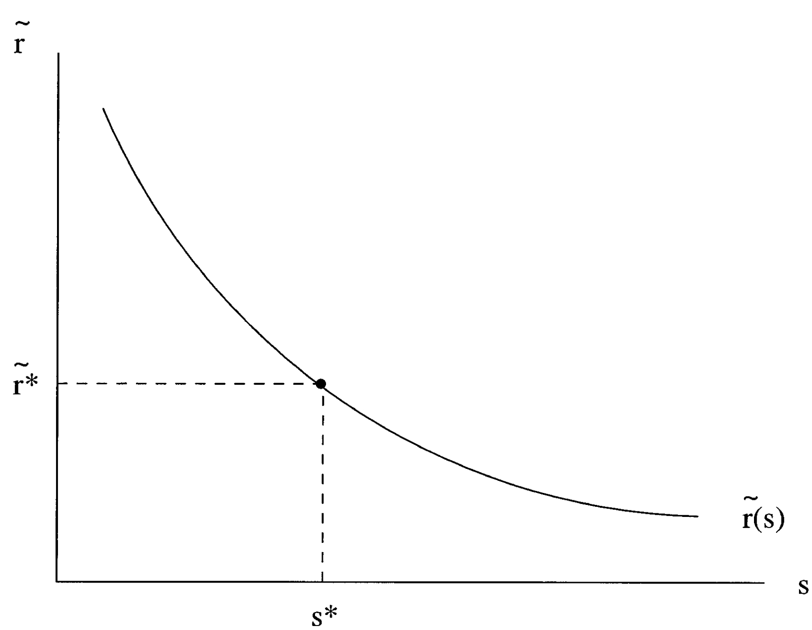

FIGURE I
The Social Return to Research Function

make additional assumptions about the nature of the market economy (market structure, patent arrangements, taxes, etc.) to determine the social return to R&D. Provided that one can write down an accurate representation of the economy's production possibilities, including knowledge of the parameters, one knows the function plotted in Figure I. Then, one can look at the allocation of resources—the s—actually chosen in the economy and "read off" the social rate of return from the figure.

Why is the functional relationship between the social rate of return and the allocation of resources independent of any market features of the economy? Intuitively, the answer is that no allocative decisions are involved in computing $\tilde{r}$: we force the economy to do one more unit of R&D and then calculate the total amount of output that can be consumed with this variation. In this sense, the production possibilities of the model (e.g., the equations describing the social planner's problem) determine the social rate of return function. The market economy and whatever distortions are present determine the allocation of resources, i.e., the point on the social rate of return function. With this definition of the social rate of return, our first key result is almost immediately evident from this insight. Moreover, this insight indicates that the result is more general than our specific application.[6]

III. Estimating the Social Rate of Return

Now consider the following question: how can the available data on productivity and R&D expenditures be used to estimate the social return to R&D? One widely used approach found in the literature is to treat R&D investment simply as an alternative capital investment in a standard neoclassical model.[7] The R&D "stock" is included in the production function, and the partial

6. With this in mind, it is interesting to compare our measure of the social rate of return with an alternative calculation, the change in the "value function" of the decentralized economy. This alternative provides the change in welfare associated with additional R&D taking into account the dynamic response of agents to the variation. However, calculating this alternative requires substantially more structure and effort. First, the value function approach depends critically on the assumptions one makes about market structure and the distortions present in the decentralized economy. As already emphasized, one advantage of our approach is that we do not need this additional structure. Second, analytical solutions are not available with this approach. Finally, in the context of this paper, our calculation is extremely relevant because it is directly related to the estimates in the productivity literature.

7. A second approach, pursued by Bernstein and Nadiri [1989], is to compute the return to R&D using estimated cost functions. These two approaches yield similar results for the return to R&D.

derivative of output with respect to that stock is treated as the rate of return to R&D.[8] The analogy to the marginal product of physical capital is clear. This basic relationship is described by the following two equations:

$$Y = e^{\mu t} Z^{\xi} K^{\alpha} L^{1-\alpha}, \tag{9}$$

$$\dot{Z} = R, \tag{10}$$

where Z is the measured R&D stock and we assume no depreciation of R&D capital.[9]

In this approach, the marginal product of the R&D stock, $\partial Y/\partial Z$, is interpreted as the rate of return to R&D; let us call this marginal product $\tilde{r}^{PL}$. By standard growth accounting logic, estimated TFP growth accounted for by R&D is then $\tilde{r}^{PL} R/Y$. This motivates the following empirical specification for estimating the rate of return to R&D:

$$\Delta \log TFP = \mu + \tilde{r}^{PL}(R/Y) + \epsilon. \tag{11}$$

That is, total factor productivity growth is regressed on the R&D share of output (and perhaps other control variables as well).

The empirical literature distinguishes between the private return and the social return to R&D. The former refers to the estimate of $\tilde{r}^{PL}$ using a firm's own R&D share as the explanatory variable.[10] The latter attempts to mitigate measurement problems and to capture interfirm technology spillovers by focusing on the industry level.[11] Table I provides a partial review of estimates of so-defined "social" rates of return from the productivity literature. In these studies, each observation corresponds to an indus-

8. This basic approach is extended in several directions in the productivity literature. For example, Jaffe [1986] makes progress by incorporating R&D from other industries into the R&D stock to estimate the gains from interindustry spillovers.

9. This assumption of zero depreciation is somewhat standard in the productivity literature. Griliches and Lichtenberg [1984b] and Hall and Mairesse [1995] find that point estimates rise with the assumed rate of depreciation but that the specification's fit is best with a zero rate of depreciation.

10. In regressions with firm-level data, private rates of return cluster around 10 to 15 percent, but can range to more than 30 percent [Hall 1995]. Large estimates could reflect risk premiums or endogeneity problems and suggest that some caution is warranted in interpreting the estimates from this literature.

11. Measurement issues are particularly acute. For example, the development of a new high-speed computer may not be reflected in the developing firm's total factor productivity; some of the measured productivity gain may show up downstream. To the extent that product innovations are created and used in the same industry, aggregation to the industry level helps mitigate these problems.

TABLE I
ESTIMATED RATES OF RETURN TO R&D

Study	(1) $\tilde{r}$ (own)	(2) $\tilde{r}$ (used)	(1) + (2)	# obs.	Years
Sveikauskas [1981]	0.17	—	—	144	59–69
	(.06)	—	—		
Griliches [1994]	0.30	—	—	143	78–89
	(.07)	—	—		
Griliches and Lichtenberg [1984b]	0.34	—	—	27	69–73
	(.08)	—	—		
Terleckyj [1980]	0.25	0.82	1.07	20	48–66
	(.08)	(.21)			
Scherer [1982]	0.29	0.74	1.03	87	73–78
	(.14)	(.39)			
Griliches and Lichtenberg [1984a]	0.30	0.41	0.71	193	69–78
	(.09)	(.20)			

The dependent variable is average TFP growth in an industry over the years indicated, except for Scherer [1982], who uses labor productivity growth and includes growth in the capital-labor ratio as a regressor. Column (1) reports representative point estimates and associated standard errors (in parentheses) of the coefficient on R&D intensity (typically privately financed R&D/Sales). Column (2) reports estimates of the coefficient on used R&D, equal to the sum of own process R&D and imputed purchases of R&D from other industries, where the imputation is based on a technology flow matrix constructed from patent data or industries of origin and use or input-output flows between industries, as indicated. When used R&D is included in the regression, "own" R&D only includes product R&D used in the industry of origin.

try in the manufacturing sector.[12] TFP growth is averaged over a range of years to reduce the effects of the business cycle on measured productivity. R&D intensity is typically measured as the average ratio of privately financed R&D spending to sales during some period before or at the start of the TFP measurement period.

Estimates of the social return average about 27 percent when only R&D from one's own industry is included and average nearly 100 percent when the broadest concept of return (the sum of the two columns in the table) is employed. The return to "used" R&D reported in the table attempts to capture the effect of product R&D in one industry on measured productivity in other industries. For each industry, R&D inputs from other industries are inputed through the use of a technology flow matrix based on information on industries of origin and use of inventions culled from patent data or input-output flows between industries.

The framework used in the empirical approach outlined above places two important restrictions on the R&D stock accumu-

12. Scherer [1982] is the exception in that 6 nonmanufacturing industries are included along with 81 manufacturing industries in the analysis.

lation process. First, no explicit allowance is made for congestion effects. Second, this approach does not explicitly allow for intertemporal knowledge spillovers or diminishing technological opportunities. Assuming that these restrictions on the R&D technology are violated, the model is misspecified. In this case, it is not possible to relate exactly the parameters estimated in the productivity literature to our model parameters. It is possible, however, to obtain a linear approximation to the relationship, accurate in the vicinity of the steady state equilibrium.

Suppose that the economy consists of a number of industries, each described by the production possibilities outlined in Section II. Consider running the regression of the productivity literature in this economy. To determine what this regression will produce, we log-linearize the production function for ideas given in equation (7) around the balanced growth path. The linear approximation is

$$\frac{\dot{A}_t}{A_t} \simeq c + \lambda g_A \frac{s_t}{\bar{s}} + \lambda g_A \ln\left(\frac{Y_t}{\bar{Y}_t}\right) + (\phi - 1) g_A \ln\left(\frac{A_t}{\bar{A}_t}\right), \tag{12}$$

where c is a constant. Multiplying by σ,

$$\frac{d \ln TFP_t}{dt} \simeq \sigma c + \frac{\lambda g_{TFP}}{\bar{s}} s_t + \lambda g_{TFP} \ln\left(\frac{Y_t}{\bar{Y}_t}\right) + (\phi - 1) g_A \ln\left(\frac{TFP_t}{\overline{TFP}_t}\right). \tag{13}$$

Regression of the TFP growth rate on the R&D share of output should yield a coefficient given by

$$\tilde{r}^{PL} = \lambda g_{TFP}/\bar{s}. \tag{14}$$

Our derivation suggests that the regressions in the productivity literature should also include as regressors the percent deviations of output and TFP from their respective steady state levels. Therefore, omitted variable bias may be a problem in interpreting the estimates from the productivity literature. To investigate the potential magnitude of this bias, we mimicked the methodology of the studies reported in the table. We constructed seven consecutive four-year samples, covering the years 1961–1989, of TFP growth and R&D intensity for twelve manufacturing industries at the two-digit level.[13] We then estimated the standard

13. The R&D data were provided by Carol Moylan and are described in Carson, Grimm, and Moylan [1994]; the data on manufacturing TFP and output were provided by William Gullickson and are described in Gullickson [1992].

regression imposing equality of coefficients across industries and over time. The only difference in methodology from that used in the studies reported in the table is that multiple time periods are included. In this pooled regression of TFP growth on a constant and lagged privately financed R&D intensity, the slope coefficient is estimated to be 0.37 (s.e. .09); when the deviations from industries' trend output and TFP are included, the coefficient on R&D intensity is 0.35 (s.e. .07). In this exercise, the bias from excluding the levels of output and TFP is positive but quite small. This analysis is not meant to represent a serious attempt to estimate the social return to R&D but instead to suggest the sign and magnitude of the bias in the existing literature.[14]

Now compare the coefficient estimated in the productivity literature and reported in equation (14) with the true social rate of return given by equation (8). The productivity literature captures only the basic output dividend and ignores the dynamic effects associated with the intertemporal knowledge spillover and the capital gain or loss.[15] Mathematically,

$$\tilde{r}(\bar{s}) = \tilde{r}^{PL} + (\phi g_A + g_Y - g_A). \tag{15}$$

As written, it appears that the term determining the difference between the true social return and the estimate from the productivity literature could be either positive or negative. However, this term can be rewritten to reveal that it is always positive, at least in steady state. Along the balanced growth path, $\lambda g_Y = (1 - \phi) g_A$. Therefore,

$$\tilde{r}(\bar{s}) = \tilde{r}^{PL} + (1 - \lambda) g_Y. \tag{16}$$

This equation reveals a rather surprising result:

RESULT 2. $\tilde{r}^{PL}$ represents an underestimate of the true social rate of return to R&D with a maximum downward bias equal to the rate of growth of output.

The general conclusion from this literature that the social rate of return to R&D is very large evidently survives rigorous analysis in the context of new growth theory.

14. It should be noted that most of the identification of the coefficient on R&D intensity comes from the cross-sectional or "between" dimension; i.e., inclusion of industry dummies negates the statistical significance of the coefficient on R&D intensity, while adding time dummies actually raises the point estimate by 0.03 and sharpens the precision of the estimate.

15. In practice, the estimates in the productivity literature may implicitly capture some of the intertemporal spillovers because R&D shares are highly persistent.

How does the productivity literature nearly get the right answer? The explanation involves two different errors that almost offset. First, the productivity literature focuses on $\partial Y/\partial Z$ as the rate of return to R&D. This focus captures the basic output effect but ignores two dynamic factors that determine the social rate of return to R&D in equation (8): intertemporal knowledge spillovers and the "capital gain" (or loss) due to changes in the relative value of knowledge creation over time. The empirical productivity literature implicitly assumes that these terms equal zero. In fact, both terms may be large in magnitude, but their sum is limited to $(1 - \lambda)g_Y$.[16]

The intuition for why the sum of the knowledge spillover and the capital gain terms is bounded is seen by noting that the capital gain reflects the change in the value of ideas. This value equals the cost in terms of consumption goods of producing a new idea, $R/\dot{A}$. From the production function for ideas, one sees that this cost is proportional to $R^{1-\lambda}A^{-\phi}$. The return to society due to the knowledge spillover, ϕg_A, exactly offsets the capital loss due to the fall in value of ideas as ideas become less costly to generate over time due to the accumulation of knowledge. What remains is the capital gain due to the increase in the value of designs resulting from the growth in R&D and $\lambda < 1$, reflected by the term $(1 - \lambda)g_Y$.

One might expect that the method used in the productivity literature would not correctly incorporate the distortions associated with creative destruction and the monopoly pricing of capital goods. However, the results indicate that these factors enter the rate of return calculation directly through s. More generally, distortions associated with the market economy that do not affect the production possibilities of the economy do not affect either the rate of return calculation or the optimal amount of R&D. Thus, adjustments to estimates of $\tilde{r}^{PL}$ to reflect monopoly pricing, imitation, or creative destruction, as sometimes suggested in the literature, are unnecessary and inappropriate.

IV. The Extent of Underinvestment in R&D

One drawback to discussing underinvestment in terms of social rates of return is that the extent of underinvestment is not

16. We have maintained the assumption that $\lambda \leq 1$; i.e., there are congestion externalities. If instead $\lambda > 1$, indicating complementarity between research today apart from knowledge spillovers, then the productivity literature would underestimate the rate of return. However, notice that the magnitude of the error is small because of the multiplication by g_Y.

readily apparent. Fortunately, the analytic framework we have used to interpret the estimates from the productivity literature provides this translation. This is illustrated in Figure I: intuitively, in order to find the optimal rate of R&D investment, all we need to do is invert the social rate of return function.

First, notice that the actual rate of investment in research by the economy, s^{actual}, satisfies the relation,

$$\tilde{r}^{PL} = \lambda g_{TFP}/s^{actual}. \tag{17}$$

Second, the optimal amount of research is given by the condition that the social rate of return is equal to the real interest rate r. Using this condition and equation (11), the optimal rate of investment in R&D along a balanced growth path is

$$s^{optimal} = \lambda g_{TFP}/(r - (1 - \lambda)g_Y). \tag{18}$$

Combining this equation with equation (17) gives the ratio of optimal investment to actual investment in research:

$$\frac{s^{optimal}}{s^{actual}} = \tilde{r}^{PL}/(r - (1 - \lambda)g_Y). \tag{19}$$

With estimates of $\tilde{r}^{PL}$ in mind, we can compute a conservative "lower bound" on this ratio. First, notice that the denominator is no greater than the real rate of return for the economy. Hence, it is no larger than a number like 7 percent, the average real return on the stock market for the last century [Mehra and Prescott 1985]. Picking a value for $\tilde{r}^{PL}$ of 28 or 30 percent, toward the lower end of the estimates in Table I, equation (19) implies a conservative estimate of $s^{optimal}/s^{actual}$ of about 4. Even if one doubles the private rate of return to 14 percent, the ratio remains high at 2. That is, the optimal share of resources to invest in research is conservatively estimated to be two to four times larger than the actual amount invested by the U. S. economy. The extent of underinvestment is substantial, and could well be much larger.

V. Conclusion

Recent endogenous growth models have emphasized the importance of R&D and the production of knowledge for understanding long-run growth. A key issue is whether the economy undertakes too little or too much R&D, and by how much. In exploring these questions, we uncover several findings. First, we provide a methodological contribution in showing how to compute social rates of return. For the case of R&D, we establish that the functional relationship between the social rate of return and the share of resources devoted to R&D depends only on the production possibilities of the economy. Market distortions such as patents, taxes, and monopoly power affect the allocation of resources to R&D, but not the functional relationship itself. Everything we need to know about the market economy is summarized in the observed allocation of resources.

Second, we examine the answer to these questions provided by the empirical productivity literature. A number of studies in that literature purport to find large rates of return to R&D, suggesting substantial underinvestment. We show that these estimates should be interpreted as a lower bound on the true social rate of return, even in light of the distortions to R&D highlighted by the theoretical literature.

Finally, the approach developed here allows us to go beyond measuring rates of return. Knowledge of the social rate of return function provides a ready mapping between social rates of return and the extent of underinvestment. A conservative estimate indicates that optimal investment in research is more than two to four times actual investment.

STANFORD UNIVERSITY
BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM

REFERENCES

Aghion, Philippe, and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, LX (1992), 323–351.

Bernstein, Jeffrey I., and M. Ishaq Nadiri, "Research and Development and Intra-industry Spillovers: An Empirical Application of Dynamic Duality," *Review of Economic Studies*, LVI (1989), 249–269.

Carson, Carol S., Bruce T. Grimm, and Carol E. Moylan, "A Satellite Account for Research and Development," *Survey of Current Business*, LXXIV (1994), 37–71.

Cohen, Wesley M., and Richard C. Levin, "Empirical Studies of Innovation and Market Structure," in Richard Schmalensee and Robert D. Willig, eds., *Handbook of Industrial Organization* (New York: Elsevier Science Publishers, 1991).

Griliches, Zvi, "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, XCIV (1992), 29–47.

——, "Productivity, R&D and the Data Constraint," *American Economic Review*, LXXXIV (1994), 1–23.

Griliches, Zvi, and Frank Lichtenberg, "Interindustry Technology Flows and Productivity Growth: A Reexamination," *Review of Economics and Statistics*, LXVI (1984a), 324–329.

Griliches, Zvi, and Frank Lichtenberg, "R&D and Productivity Growth at the Industry Level: Is there Still a Relationship?" in Zvi Griliches, ed., *R&D, Patents and Productivity* (Chicago: University of Chicago Press, 1984b).

Grossman, Gene M., and Elhanan Helpman, *Innovation and Growth in the Global Economy* (Cambridge, MA: MIT Press, 1991).

Gullickson, William, "Multifactor Productivity in Manufacturing Industries," *Monthly Labor Review* (1992), 20–32.
Hall, Bronwyn H., "The Private and Social Returns to Research and Development: What Have We Learned?" manuscript, University of California at Berkeley, 1995.
Hall, Bronwyn H., and Jacques Mairesse, "Exploring the Relationship between R&D and Productivity in French Manufacturing Firms," *Journal of Econometrics,* LXV (1995), 263–293.
Jaffe, Adam B., "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value," *American Economic Review,* LXXVI (1986), 984–1001.
Jones, Charles I., "R&D-Based Models of Economic Growth," *Journal of Political Economy,* CIII (1995), 759–784.
Mehra, Rajnish, and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* XV (1985), 145–161.
Nadiri, M. Ishaq, "Innovations and Technological Spillovers," NBER Working Paper No. 4423, 1993.
Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy,* XCVIII (1990), S71–S102.
Scherer, Frederic M., "Inter-Industry Technology Flows and Productivity Growth," *Review of Economics and Statistics,* LXIV (1982), 627–634.
Stokey, Nancy L., "R&D and Economic Growth," *Review of Economic Studies,* LXII (1995), 469–489.
Sveikauskas, Leo, "Technological Inputs and Multifactor Productivity Growth," *Review of Economics and Statistics,* LXIII (1981), 275–282.
Terleckyj, Nestor E., "Direct and Indirect Effects of Industrial Research and Development on the Productivity Growth of Industries," in John W. Kendrick and Beatrice N. Vaccara, eds., *New Developments in Productivity Measurement and Analysis* (Chicago: University of Chicago Press, 1980).
Tirole, Jean, *The Theory of Industrial Organization* (Cambridge, MA: MIT Press, 1988).