Volume Title: New Developments in Productivity Measurement

Volume Author/Editor: John W. Kendrick and Beatrice N. Vaccara, eds.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-43080-4

Volume URL: http://www.nber.org/books/kend80-1

Publication Date: 1980

Chapter Title: Returns to Research and Development Expenditures in the Private Sector

Chapter Author: Zvi Griliches

Chapter URL: http://www.nber.org/chapters/c3919

Chapter pages in book: (p. 419 - 462)

# 8 Returns to Research and Development Expenditures in the Private Sector

Zvi Griliches

## 8.1 Introduction

In late 1965, the Bureau of the Census and the Office of Manpower Studies of the National Science Foundation asked me to consider a project to analyze the available historical data on company research and development expenditures together with other data for the same companies collected in different Census inquiries. During 1966–67, a plan of work was outlined, cut down to size, and agreed upon. The Census undertook to develop a company record, edited for consistency, to produce regressions and related outputs free of disclosures for individual companies, and to pass on the reasonableness of the various series employed. Only Census employees were to have (and have had) access to individual company data, and the treatment of outliers was in accordance with the usual criteria employed by the Census. The process of matching the same companies in different data sets and over time turned out to be quite a difficult and time-consuming task. Because the results were slow in coming, and in the context of severe budgetary cuts, the Office of Manpower Studies of the NSF bowed out as a direct partner in this study in 1968. The rest of the financing for this project

Zvi Griliches is at Harvard University.

A large number of people were essential and helpful in initiating and carrying through this work. I am grateful to, among many others, Max Conklin, Owen Gretton, L. Jack Owen, Walter Heller, Milton Eisen, and Ruth Rynyan at the Bureau of the Census; to Thomas Hogan, Pat Riley, Ken Sanow, and James Blackman at the National Science Foundation; and to Paul Ryan, Ruth Helpman, and Bronwyn Hall at Harvard University, for help, encouragement, and research assistance. I am indebted to the National Science Foundation for the financial support of this project both through the budget of the Office of Economic and Manpower Studies and through grants no. G–1812, GS–712, GS–2762X, GS–39865X, and SOC 73–05374–A01.

still came from the National Science Foundation, but in the form of a direct research grant to me rather than as a continuation of the in-house research partnership. The funding crisis and other workload pressures on the Census delayed the completion of the data match until 1970. During this long gestation period the project was greatly reduced in scope by abandoning the idea of extending the match to such additional company data sources as the IRS and Compustat tapes and by limiting the number and range of variables to be included in the final data base. First regression results for a restricted set of equations and variables became available in early 1971, and final corrected runs were delivered in 1972. This is the first report based on the results of this project. I am solely responsible for the interpretation and analysis of the results and for the delay since mid-1972.

The original universe of this study consists of large (1000-plus employees) R and D–performing U.S. manufacturing companies. There were 1,154 such companies in 1964. Our final sample is based on data for 883 such companies, accounting for about 90% of total sales and over 92% of total R and D expenditures of all firms in this universe (see table 8.1 for more detail). Since large firms account for most of the reported R and D expenditures in industry, our sample accounted for 91% of all the R and D performed in industry in 1963 including the R and D performed outside our universe of large companies. Thus, in spite of quite a few companies for which some or many of the data are missing, the coverage of our sample is rather complete, especially in comparison to other micro-data sets of this kind.

The data base consists of individual company time series on research and development expenditures (company-financed and total), on the number of research scientists and engineers, and on total company employment and sales—all based on the 1957–65 annual NSF–Census R and D surveys—and of additional company data on value added, assets, depreciation, and other economic magnitudes, based on the match with the 1958 and 1963 Census of Manufactures and Enterprise statistics. Because of problems of handling confidential data I received only matrices of correlation coefficients and standard deviations for the various variables in the data base, broken down into six rather broad industry groupings, and never had access to the actual individual observations. The restriction of this study to variables contained in the original data sets and the associated inability to add such things as prices, stock valuations, or concentration ratios, the availability of data only in the form of moment matrices, the relative shortness of the available time series, and the lack of detailed industrial breakdown, all severely limit the range of questions that can be asked and largely predetermine the feasible modes of analysis.

**Table 8.1** **Sample Coverage in 1963: R and D–Performing Companies with 1,000 or More Employees**

| SIC Industry | Number of Companies | | | Total Sales (billions of dollars) | | | Total R and D Expenditures (millions of dollars) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Population | Sample | Coverage Ratio | Population | Sample | Coverage Ratio | Population | Sample | Coverage Ratio |
| 1. 28, 29, 13: Chemicals and petroleum | 134 | 110 | .82 | 52.6 | 48.4 | .92 | 1,556 | 1,294 | .83 |
| 2. 34, 35: Fabricated metal products and machinery | 257 | 187 | .73 | 32.1 | 23.7 | .74 | 1,111 | 958 | .86 |
| 3. 34, 48: Electrical and communication equipment | 134 | 102 | .76 | 28.2 | 23.2 | .82 | 2,866 | 2,579 | .90 |
| 4. 371, 373–9: Motor vehicles and other transport equipment | 55 | 34 | .62 | 32.0 | 29.6 | .92 | 1,090 | 1,062 | .97 |
| 5. 372, 19: Aircraft and missiles | 53 | 31 | .58 | 17.4 | 16.8 | .97 | 4,712 | 4,619 | .98 |
| 6. All others | 521 | 419 | .80 | 97.8 | 90.5 | .93 | 1,137 | 922 | .81 |
| Total | 1,154 | 883 | .77 | 260.1 | 232.2 | .89 | 12,472 | 11,434 | .92 |

SOURCE: Unpublished census tabulations.

When this study was initiated in the mid 1960s, my own interests centered on sources of productivity growth and on estimating the contribution of nonmarket factors to growth using production function models and econometric estimation techniques. The study reported below bears the marks of this interest. It focuses on estimating the coefficient of cumulated R and D expenditures in company-level production functions or its equivalent in company productivity growth equations. Because the data are for individual companies, this study can explore only the magnitude of *private* returns to such expenditures. It cannot deal with the very important issue of externalities—returns that accrue to other firms and to society at large and are not captured by the original investors. In a later report I shall try to deal with this problem by comparing the estimates presented here with those derivable from aggregate industry and economy-wide time series. Here we'll limit ourselves, however, to what direct information can be gleaned from the data at hand.

The next section of this paper outlines the theoretical model used and the statistical problems associated with its estimation. The variables used in this study are described in section 8.3 and the main results are summarized in section 8.4. Section 8.5 digresses to consider the relation of R and D to firm size. Concluding remarks are contained in section 8.6, while more detail on the matching process and data construction can be found in the Appendix.

## 8.2 Models and Problems

Both the theoretical and empirical literature on the relationship between research and productivity have been reviewed recently by several authors (cf. Griliches 1973, 1974; Mansfield 1967, 1972; and Nordhaus 1969, among others) and we shall not go over the same ground again here except to present the simplest possible model of this process which will serve as the framework of our estimation efforts below.

This model, which is common to most analyses of the contribution of research to productivity growth, can be summarized along the following lines:

(1) $$Q = TF(C,L),$$

(2) $$T = G(K,O),$$

(3) $$K = \Sigma w_i R_{t-i},$$

where $Q$ is output (sales, or value added), $C$ and $L$ are measures of capital and labor input, respectively, $T$ is the current level of (average) technological accomplishment (total factor productivity), $K$ is a measure of the accumulated and still productive (social or private) research capital ("knowledge"), $O$ represents other forces affecting productivity,

$R_t$ measures the real gross investment in research in period $t$, and the $w_i$'s connect the levels of past research to the current state of knowledge.[1]

For estimation purposes, the $F$ and $G$ functions are usually specialized to the Cobb-Douglas form and $O$ is approximated by an exponential trend. The whole model then simplifies to

$$Q_t = Ae^{\lambda t} K^{\alpha}{}_t C^{\beta}{}_t L^{1-\beta}{}_t, \tag{4}$$

where $A$ is constant, $\lambda$ is the rate of disembodied "external" technical change, and constant returns to scale have been assumed with respect to the conventional inputs ($C$ and $L$). Equations like this have been estimated by Griliches (1964) from several agricultural cross-sections, and by Evenson (1968) and Minasian (1969) from combinations of time series and cross-section data for agricultural regions and chemical firms, respectively. Alternatively, if one differentiates the above expression with respect to time and assumes that conventional inputs are paid their marginal products, one can rewrite it as

$$f = q - \hat{\beta}c - (1-\hat{\beta})l = \lambda + \alpha k, \tag{5}$$

where $f$ is the rate of growth of total factor productivity, lower-case letters represent relative rates of growth of their respective upper-case counterparts $[x = \dot{X}/X = (dX/dt)/X]$, and $\hat{\beta}$ is the estimated factor share of capital input.[2] Equation (5) is a constrained version of (4). Versions of such an equation were estimated by Evenson (1968) for agriculture and by Mansfield (1965) for manufacturing industries, among others. In either form, the estimates of $\alpha$ have tended to cluster around .05 for public research investments in agriculture (Evenson and Griliches) and around .1 for private research investments in selected manufacturing industries (Mansfield, Minasian, and Terleckyj).

Up to now I have been deliberately vague as to the operational construction of the various variables. The difficulties here are myriad. Perhaps the two most important problems are the measurement of output ($Q$) in a research-intensive industry (where quality changes may be rampant), and the construction of the unobservable research capital measure ($K$). Postponing the first for later consideration, we note that $K_t = \Sigma w_i R_{t-i}$ can be thought of as a measure of the distributed lag effect of past research investments on productivity. There are

1. Note that in writing equations (1) and (2) in this fashion we have implicitly assumed the separability and ultimate neutrality of the research process from the production process. Since theoretical generalization is cheap, we could have extended the model to make the coefficients of $C$ and $L$ also dependent on $K$, but our data could not sustain such complications.

2. To the extent that research inputs are included amcng the conventional input measures, they have already been imputed the average private rate of return.

at least three forces at work here: the lag between investment in research and the actual invention of a new technique or product, the lag between invention and the development and complete market acceptance of the new technique or product, and the disappearance of this technique or product from the currently utilized stock of knowledge due to changes in external circumstances and the development of superior techniques or products by competitors (depreciation and obsolescence). These lags have been largely ignored by most of the investigators. The most common assumption has been one of no or little lag and no depreciation. Thus, Griliches and Minasian have defined $K_t = \Sigma R_{t-i}$ with the summation running over the available range of data, while Mansfield assumed that since $R$ has been growing at a rather rapid rate, so also has $K$ (i.e., $\dot{K}/K \approx \dot{R}/R$). Evenson (1968) has been the only one to investigate this question econometrically, finding that in the aggregate data for U.S. agriculture, an "inverted V" distributed lag form fitted best, with the peak influence coming with a lag of five to eight years and the total effect dying out in about ten to sixteen years. There is some scattered evidence, based largely on questionnaire studies (see Wagner 1968), that such lags are much shorter in industry, where most of research expenditures are spent on development and applied topics.[3]

Because of the difficulties in constructing an unambiguous measure of $K$, many studies have opted for an alternative version of equation (5), utilizing the fact that

$$\alpha = \frac{dQ}{dK}\frac{K}{Q}$$

and

$$\alpha k = \frac{dQ}{dK}\frac{K}{Q}\frac{\dot{K}}{K} = \frac{dQ}{dK}\frac{\dot{K}}{Q},$$

allowing one to rewrite (5) as

$$f = \lambda + \alpha k = \lambda + \rho I_R/Q, \quad (5')$$

where $\rho$ is the rate of return to research expenditures (the marginal product of $K$) while $I_R/Q$ is the net investment in research as a ratio to total output. In practice, to make some connection between gross and net investment in research one needs information about its "depreciation" which, if available, would have allowed us to construct a measure of $K$ in the first place.

While our models are written as if the main point of research expenditures is to increase the physical productivity of the firm's production

3. In the U.S. about three-fourths of all expenditures on R and D in industry have been spent on development and most of the rest on "applied research." Only about 5% of the total R and D expenditure has gone to "basic" research. Thus, one should not expect long lags *on the average*.

process, most of the actual research in industry is devoted to the development of new products or processes to be sold and used outside the firm in question. Assuming that, on average, the outside world pays for these products what they are worth to it, using sales or value added as our dependent variable does in fact capture the private returns to such research endeavors. However, the observed private returns may underestimate the social returns because, given the competitive structure of the particular industry, the market price of the new product or process will be significantly below what consumers might have been willing to pay for it. On the other hand, part of the increase in sales of an individual firm may come at the expense of other firms and not as the result of the expansion of the market as a whole. Also, some of the increase in prices paid for a particular new product may come from changes in the market power of a particular firm induced by the success of the research program. Moreover, some of the gains in productivity or in the sales of new products may be based on the research results of other firms in the same or some other industry. Such factors could result in the observed private returns overestimating the social returns significantly. We cannot say much about the net impact of such forces on the basis of the data at hand. It requires a detailed comparison of the individual firm results with estimates based on industry and economy-wide returns to research, a topic beyond the scope of this paper. But since expected private returns are presumably a determinant of private investment flows into this activity, the estimates presented below may be of some interest even if they cannot answer the social-returns question unequivocally.

Another important problem arises as soon as we write down a system of equations, such as (1)–(3), a problem that will stay with us throughout this paper. Ideally, we would like to distinguish between capital and labor used to produce current "output" and capital and labor used in research (the production of future knowledge and the maintenance of the current stock). In fact, we are usually unable to observe these different input components and are forced to use totals for $C$ and $L$ in our investigations. This leads to a misspecification of equation (4) or (5). Moreover, if components of $L$ and $C$ are weighted in proportion to their current returns, the resulting estimates of the contribution of $K$ (or $R$) represent, errors in timing apart, excess returns above and beyond the "normal" remuneration of such factors of production.

Given the limited range of our time series, we decided early on a *two*-pronged research strategy: (*a*) Concentrate on estimating versions of equation (5) based on average *rates of growth* for the whole 1957–65 period. (*b*) Estimate equation (4) based on the 1963 cross-section levels. Equation (5) has the advantage that, dealing with rates of growth, one essentially differences out permanent efficiency differences across firms and does not allow them to influence the final results. Equation (4) has the advantage that it does *not* ignore the cross-sectional differences in levels, which are a major source of variance in the data and of intrinsic interest themselves. Given our limited data base, additional compromises had to be made in the definition and the choice of variables which are best discussed after we describe, in the next section, the available data and the variables constructed from them.

## 8.3 Data, Variables, and Caveats

Table 8.1 gives some detail on our sample and its coverage. We have data on 883 large R and D–performing companies, divided into six industrial groupings.[4] Unfortunately, the industrial groupings are rather coarse and the number of companies in some of them is rather small, especially in the motor vehicles and aircraft and missiles groups. Most of our attention will be devoted, therefore, to the combined total industry results, though, for comparison purposes, we will also present the individual industry group results and comment on them.

Our data base was limited to the short list of the R and D survey variables on the matched historical R and D tapes (i.e., R and D expenditures—company and total, sales, total employment, and the employment of scientists and engineers) and the limited number of variables that could be matched to them from the 1958 and 1963 Census of Manufacturers and Enterprise Statistics schedules. Moreover, since the original data could not be released except in the form of moment matrices for selected variables, an irreversible decision had to be made about the choice and functional form of the variables to be included in them. The choice was guided by the following research strategy decision: Given the fact that we have only relatively short time series at hand and assuming that much of the individual annual fluctuations in these series are of a transitory nature, our analysis will concentrate on *two* dimensions of these data—average *rates of growth* over the whole observation period (1957–65) and *levels* in 1963.

Thus, a major subset of the variables included in this study are *rates of growth* computed from regressions of the natural logarithms of the annual observations in the historical R and D tapes on a time-trend. They are the estimated slope coefficients ($b$'s) from $\ln X = a + bt$ type equations, fitted to the whole 1957–65 period or to the sub-period of available data, provided that four or more years of data were available to compute such time-trend regressions.

4. See Appendices B and C for details on the criteria for inclusion of companies in the sample and the methods of imputation for missing data. The Standard Industrial Classification code of a company is determined by its main activity, and its entire research and development operations are classified in that industry.

Appendix table 8.A.1 lists the sixty variables for which moment matrices were released by the Census Bureau. These variables can be divided roughly into the following sets: (1) potential dependent variables; (2) various measures of R and D growth and intensity; (3) measures of physical capital and its age composition; (4) measures of total company employment; (5) quality of data measures; (6) other background variables. In what follows we shall discuss only the variables used intensively in this study.

The major dependent variable used in the growth rates section of this study is BPT (number 41 in table 8.A.1), or partial productivity growth, computed as the difference between the estimated rate of growth of total company sales in 1957–65 (31. BS) and the product of the rate of growth of total company employment (32. BE) and the average share of labor (total payroll) in sales (12. ALSS), in 1958 and 1963. That is, BPT = BS — ALSS • BE is a partial approximation to equation (5) with $\beta c$ taken to the right-hand side:

$$q - (1-\beta)\, l \approx \text{BPT} \approx \lambda + \alpha k + \beta c + u, \tag{5''}$$

where ALSS is an approximation to $(1-\beta)$, $\lambda$ is the average exogenous rate of productivity growth, $c$ is the rate of growth of physical capital, and $u$ is a catchall mnemonic for all other systematic and random factors affecting productivity. Because we have no explicit measure of the growth of company physical capital, we could not construct an explicit total factor productivity measure ($f$) and use the direct version of equation (5). The procedure of using each individual firm's labor share as an approximation to its output-labor elasticity has the virtue of allowing this elasticity to differ across firms, adjusting thereby for rather wide differences in vertical integration across firms.

The missing company rate of growth of physical capital is approximated by two variables: the ratio of accumulated depreciation to the total stock of physical capital in 1963 (6. Age $C$ = [gross fixed assets — net fixed assets]/gross fixed assets) and the depreciation rate (7. $D$ = depreciation charged in 1963/gross fixed assets in 1963). These two proxy variables (Age $C$ and $D$) taken together should approximate rather well the unobserved true rate of growth of fixed capital, assuming that it remained reasonably constant over the period in question. Moreover, it can be shown that the estimated coefficient of $D$ should be on the order of $\beta$, the elasticity of output with respect to physical capital.[5]

5. Let $g$ be the rate of growth of fixed investment and $d$ its depreciation rate. If $g$ has been approximately constant and $d$ can be taken as (or approximated by) a fixed declining balance scheme, then

$$\text{Age } C = \frac{\text{Gross Stock} - \text{Net Stock}}{\text{Gross Stock}} = 1 - \frac{g}{g+d-dg} \approx \frac{d}{g+d}.$$

Our major measure of the growth in research capital ($k$) is the estimated rate of growth in total company expenditures on research and development during 1957–65 (34. BTRD). Note that we are approximating the rate of growth in the *stock* of research capital by the rate of growth in gross *investment* in this type of activity. For variables whose initial level is rather low while the rate of growth of investments is rather high, the assumption of proportionality in these rates of growth ($\dot{K}/K \approx \dot{R}/R$) is not a bad one (cf. Mansfield 1965).[6] Other measures of R and D growth include the rate of growth in company-financed (excluding federally supported) R and D expenditures (35. BCRD) and the rate of growth in the *number* of scientists and engineers engaged in research and development (33. BSE). In addition we also use, in various contexts, the average total R and D to sales ratio (28. AR/S, average of 1958 and 1962) as a measure of research intensity, the ratio of company funds to total cumulated R and D expenditures during

Fluctuations in Age $C$ can then be approximated by a second-order Taylor expansion as

$$\text{Age } C \approx \overline{d/(g+d)} - \overline{d/(g+d)} \cdot g + \overline{g/(g+d)} \cdot d,$$

where bars indicate an evaluation at the mean levels of these variables. Now, in the function we need $\beta g$, where $\beta$ is the elasticity of output with respect to fixed capital. Substituting $a_1$ Age $C + a_2 d$ for it, and ignoring constants, we get:

$$a_1 = -\beta \frac{\overline{(g+d)}^{\,2}}{d} \text{ and } a_2 = \beta \bar{g}/\bar{d}.$$

Since $g/d \approx 1$, the estimated coefficient of $d$ should be close to $\beta$, while the estimated coefficient of Age $C$ ($a_1$) should be on the order of a quarter of $\beta$ (assuming $g \approx d \approx .06$). Note that this construction made no allowance for differences in capital utilization among firms or overtime. The available data base contains no information on this topic.

6. Assume no depreciation and let research expenditures $R$ grow at a constant rate $\rho$. Then the rate of growth of $K$, say, $g$, is given by

$$\begin{aligned} g_t = R_t/K_{t-1} &= R_o\,(1+\rho)^t/\Sigma\, R_{t-1-i} \\ &= R_o\,(1+\rho)^t/R_o\,\Sigma\,(1+\rho)^{t-1-i} \\ &= (1+\rho)/\Sigma\,[1/(1+\rho)]^i \\ &= (1+\rho)/1/[1-1/(1+\rho)] = \rho. \end{aligned}$$

Allowing for depreciation and a variable past would make $\rho$ an underestimate or an overestimate of $g$ depending on whether $K_o$, the level of accumulated stock at the beginning of the period, was relatively small or large. For total U.S. industry during this period (1957–65), taking initial level estimates for 1948 from Kendrick (1973), extrapolating the NSF figures back from 1953 to 1948, and assuming a depreciation rate of 10% per year, gives a $g$ of .10 instead of the observed $\rho$ of about .07, or a 30% underestimate of $g$ when using $\rho$. However, an allowance for the rising relative costs of research (deflation of these figures) would bring the two together rather closely.

1957–62 (24. FP62) as a measure of the composition of R and D funds, and the logarithm of total cumulated R and D expenditures over the 1957–62 period (54. LGK62) and the logarithm of the average number of research scientists and engineers during 1957–62 (53. LGANSE) as measures of the absolute size of the company research endeavor.

In the level regressions, the main dependent variable is the logarithm of value added in 1963 (51. LGVA63) and the main independent variables are a measure of capital services in 1963 (46. LGC2 = the logarithm of the sum of depreciation plus rentals plus 8% of net fixed assets and inventories), employment in the manufacturing establishments of the company (47. LGEM63), and the previously described cumulated R and D variable (54. LGK62). Among other variables used we should note the company's (five-digit) specialization ratio in 1963 (18. SPR 63), the fraction of the total company labor force that is employed in establishments classified as manufacturing in 1963 (11. M), and several "quality of data" variables: a dummy variable for no imputations (42. DNI), and the standard errors for the computed trend growth rates for sales (36. SBS) and for total R and D (37. SBTRD). A number of other variables are used occasionally, especially as instruments in the context of allowing for simultaneity. They will be identified as we go along. Of some intrinsic interest, however, is an estimate of the overall company profitability rate in 1963 (20. NRR), computed as value added in 1963 minus total manufacturing payroll, minus equipment rentals, and minus depreciation, all divided by net fixed assets plus inventories.

As these variables are introduced and described, several problems and difficulties immediately come to mind. First, note that in the growth-rate equations the basic data are for the company as a whole and not just for its manufacturing component, and that the dependent variable is based on the growth of sales rather than of value added. In the level equations we try to stick to the manufacturing portion of these companies, but the division of the labor force into these components is far from perfect and no separate data were available on fixed assets for the manufacturing establishments only. All of the variables except employment and the various ratio variables are in undeflated current or historical prices. Since we have no explicit information about the specific product mix of the various companies we could, at best, construct only industry-wide deflators. But then all companies within an industry would be treated alike and additively (given our largely linear-in-the-logarithms framework), affecting only the constants in the various equations. Hence, the whole deflation adjustment can be subsumed and allowed for by including separate industry dummy variables (the *I*'s, 1–5) in the overall regressions.

Another major issue is one of lags, timing, and possible simultaneity. In the growth equations we use the growth in R and D over the whole 1957–65 period as an independent variable. On the whole, we believe that we gain more by averaging over a longer period than we lose by introducing a possible simultaneous-equation bias due to contemporaneous correlation between the disturbances in the output and R and D–determining equations. Given our data base, we did not have enough of a history to experiment with fancier lag structures. We shall attempt to check our results below for robustness with respect to the simultaneity problems by (*a*) using intensity rather than growth measures of R and D, and (*b*) estimating equation (5″) using instrumental variable methods. Similar problems of interpretation and the possibility of bias arise also in the level equations where our measure of accumulated research capital is the simple unweighted sum of total R and D expenditures for the whole 1957–62 period, allowing for little lag and no depreciation.

To recapitulate, we have to use makeshift proxies for the growth in both physical and research capital. We confound price changes with quantity changes in our productivity measures, and our treatment of lags and simultaneity is both crude and cavalier. Nevertheless, it is about the best that we could do with these data. It is our belief that in spite of their shortcomings and in spite of our many simplifications and dubious assumptions, our data are interesting and rich enough, and the underlying relationships are strong enough, to show through and yield valuable insights into the R and D process and its effects on productivity and growth.

## 8.4 The Main Results

The relationship between the rate of growth of partial productivity during the 1957–65 period and measures of growth in fixed capital and in R and D is investigated, for the combined sample, in table 8.2. Under the assumption of relatively constant rates of growth of fixed capital, the ratio of (gross — net)/gross stock and the depreciation rate together act as a proxy for the unobserved rate of growth of fixed capital. Each of the regressions includes five industry dummy variables, allowing for separate industry intercepts and for differential rates of price inflation in these industries. In addition to trying out various R and D variables, some of the regressions also include a set of "quality of data" variables: the estimated standard errors of the rate of growth of sales (SBS) and of R and D (SBRD), and a dummy variable signifying a record with no imputations (DNI).

For all firms combined, both the fixed capital and the R and D growth variables are "highly significant" and of the right sign. Total

**Table 8.2** **All Industries Combined: Growth Rates 1957–65 Dependent Variable BPT = BS – ALSS × BE, Partial Productivity Growth, N = 883**

| Reg. No. | Age C | D | R and D Variables | LGANSE | Other Variables* | $R^2$ | S.E. |
|---|---|---|---|---|---|---|---|
| | Coefficients of (standard errors) | | | | | | |
| **BTRD** | | | | | | | |
| 1 | –.069 (.011) | .334 (.077) | .076 (.013) | | I's | .105 | .0561 |
| 2 | –.074 (.016) | .350 (.064) | .073 (.011) | –.003 (.001) | I's SBTRD–, DNI– | .113 | .0559 |
| 3 | –.052 (.016) | .286 (.061) | .074 (.010) | | I's, SBS+, SBTRD–, DNI– | .402 | .0459 |
| **BCRD** | | | | | | | |
| 4 | –.070 (.019) | .343 (.075) | .063 (.012) | | I's | .096 | .0564 |
| 5 | –.054 (.016) | .301 (.061) | .063 (.010) | –.002 (.001) | I's SBS+, SBTRD–, DNI– | .399 | .0460 |
| **BSE** | | | | | | | |
| 6 | –.072 (.019) | .345 (.076) | .087 (.014) | | I's | .109 | .0560 |
| 7 | –.055 (.015) | .294 (.061) | .087 (.011) | | I's SBS+, SBTRD–, DNI | .409 | .0456 |

*Coefficients that are statistically significant at the conventional .05 level are identified by their respective signs.

Age C = (gross fixed assets – net fixed assets)/gross fixed assets in 1963
D = Depreciation rate, depreciation charged in 1963/gross fixed assets in 1963
BS = Rate of growth of sales, 1957–65
BE = Rate of growth of employment, 1957–65
BTRD = Rate of growth of total R and D expenditures, 1957–65
BCRD = Rate of growth of company R and D expenditures, 1957–65
BSE = Rate of growth in the employment of scientists and engineers, 1957–65
LGANSE = Logarithm of the average number of scientists and engineers, 1957–62
SBS = Standard error of the estimated rate of growth of sales
SBTRD = Standard error of the estimated rate of growth of total R and D expenditures
DNI = Dummy variable = 1 when there were "no imputations" in the data
I's = Industry dummy variables (five)

R and D growth is a somewhat better variable than company R and D growth, while the growth in the number of scientists and engineers is marginally better than either one of the dollar measures. The implied elasticity of output with respect to cumulated R and D is about .07 and there is an indication (in the more detailed results not reported here) of some diminishing returns to the absolute size of the research program (LGANSE) and of a negative impact of variability in it (SBTRD). The overall fit is low and a large fraction of the variance is accounted for by the "quality of data" variables.

Table 8.3 summarizes the results for the individual industry groups. They are roughly similar except that the .07 estimate for the combined cross-section can be seen to be an average of a somewhat higher elasticity (.1) for the research-intensive industries and a somewhat lower coefficient (.04) for the rest (the "other" half of the sample).

A complementary analysis of the problem can be had by looking at the *levels* of productivity and their relationship to the cumulated total of past R and D expenditures (K62). Table 8.4 presents estimates of such 1963 cross-sectional production relationships. They are surprisingly reasonable, and the estimated coefficient of cumulated R and D is rather close to that derived from the time series (growth rates) regressions. At the individual industry level the estimated coefficients are somewhat lower, suggesting that the time series results may be a bit biased upward due to the simultaneity between the growth in research and in sales. But the differences are not statistically significant, as we shall show below. There is no evidence in these data of increasing returns to firm size as such, while both specialization (SPR) and average *plant* (but not firm) size (LSE) are positively related to productivity.

There are interesting consistencies between the estimates given in tables 8.2 and 8.3 and those of table 8.4, though each is based on a very different cut across the data base. We noted before (in footnote 5) that the coefficients of $D$ in table 8.2 are approximate estimates of the physical capital elasticity, and that the coefficients of Age $C$ should be on the order of a quarter of (and of opposite sign to) the coefficients of $D$. Both estimates are of the right order of magnitude (about .33 and .07, respectively). Moreover, they are not too far from the directly estimated coefficients of log $C2$ in table 8.4, which hover around .4. Similarly, the R and D coefficient is about .07 in the growth equations in table 8.2, and about .06 in the level equations in table 8.4, for all industries combined. Since both the dependent and independent variables are quite different, this consistency reinforces our belief that this is the right order of magnitude for this coefficient.

We can check in greater detail whether the data are mutually consistent by estimating a combined multivariate regression, imposing the pairwise equality of the $D$ and log $C2$ and of the BTRD and log K62 coefficients and testing whether these restrictions are rejected by the

**Table 8.3** **Dependent Variable: Partial Productivity Growth BPT = BS − ALSS × BE, 1957–65, by Industry**

| Industry | Coefficients of alternative research variables, standard errors of the coefficients, $R^2$'s and standard errors of the regressions (other variables included: Age C, D, SBS, SBTRD, DNI) | | |
|---|---|---|---|
| | BTRD | BCRD | BSE |
| 1. Chemicals and petroleum | .093 | .090 | .089 |
| $N = 110$ | (.038) | (.038) | (.042) |
| | .230(.042) | .229(.042) | .220(.042) |
| 2. Metals and machinery | .102 | .087 | .123 |
| $N = 187$ | (.022) | (.023) | (.023) |
| | .209(.043) | .179(.044) | .237(.042) |
| 3. Electric Equipment | .106 | .055 | .093 |
| $N = 101$ | (.030) | (.019) | (.029) |
| | .405(.040) | .384(.040) | .393(.040) |
| 4. Motor vehicles | .126 | .143 | .044 |
| $N = 34$ | (.070) | (.055) | (.083) |
| | .491(.036) | .543(.034) | .435(.038) |
| 5. Aircraft | .107 | .034 | .250 |
| $N = 31$ | (.077) | (.050) | (.064) |
| | .229(.042) | .183(.044) | .491(.034) |
| 6. Other | .052 | .051 | .062 |
| $N = 419$ | (.015) | (.015) | (.016) |
| | .556(.047) | .555(.047) | .559(.047) |

See Notes to Table 8.2 for definitions of variables.

sample. Table 8.5 presents the original independent estimates, industry by industry, and the estimated constrained cross-equation coefficients. It also gives the computed chi-square values for the tests of these restrictions. It is clear, at a glance, that except for the two small sample industries (4 and 5), the different estimates are quite close. In no case do the tests reject the hypothesis that the estimates arise from a population having these parameter values in common.

A basic difficulty with the results presented in tables 8.2, 8.3 and 8.5 is the likelihood of simultaneity between the productivity and R and D *growth* measures. One way of guarding against this possibility is to treat BTRD as an endogenous variable and use instrumental variable methods to estimate its coefficient in equation (5″). The results of doing so are given in table 8.6. The instruments used are basically intensity and level variables as of 1957 and 1963, which should be less correlated with the disturbances in the 1957–65 growth equations. On the whole, the results are very encouraging. Except for industries 2 and 6, the TSLS results are similar to the original ones, indicating little simultane-

**Table 8.4** **1963 Cross-sectional Production Functions by Industry**
Dependent Variable—Log VA63

| Industry | Coefficients of (standard errors) | | | Other Variables in Regression* | $R^2$ | S.E. |
|---|---|---|---|---|---|---|
| | Log $C_2$ | Log EM | Log K62 | | | |
| 1. Chemicals and petroleum | .381 (.067) | .538 (.097) | .115 (.040) | Age C−, SPR+, DNI, LSE | .893 | .391 |
| 2. Metals and machinery | .455 (.050) | .282 (.048) | .075 (.022) | Age C, SPR, DNI+, LSE+ | .895 | .305 |
| 3. Electric equipment | .534 (.065) | .439 (.071) | .029 (.020) | Age C, SPR+ | .950 | .272 |
| 4. Motor vehicles | −.048 (.106) | 1.067 (.117) | .063 (.042) | Age C, SPR, DNI | .981 | .233 |
| 5. Aircraft | .176 (.072) | .795 (.090) | .037 (.034) | Age C, SPR, DNI | .987 | .173 |
| 6. Other | .414 (.028) | .542 (.035) | .045 (.012) | Age C, M−, SPR+, DNI | .920 | .299 |
| All industries combined | | | | | | |
| a. | .422 (.018) | .435 (.022) | .069 (.009) | I's | .918 | .330 |
| b. | .376 (.021) | .527 (.026) | .061 (.008) | I's, Age C−, M−, SPR+, LSE+ | .922 | .322 |

ity bias. Only in industry 6 do the TSLS results not yield a significant R and D coefficient. If anything, the overall TSLS results give somewhat higher estimates for the R and D coefficient, indicating that our main problem may not be simultaneity but error (random noise) in the R and D data.

To the extent that the simultaneity problem is the result of too close a contemporaneity of the sales growth and R and D growth variables one could deal with it by either shortening the period over which the R and D growth is estimated or by using intensity variables such as R and D as a percentage of sales, or number of engineers and scientists as a fraction of total employment, instead of the suspect growth rates. While the results of doing so are somewhat more difficult to interpret, on the whole they do support the finding of a significant and apparently nonspurious influence of R and D on productivity growth. For example, in industry 6 (all others) where the instrumental-variables approach did not yield a significant R and D coefficient, if instead of BTRD we use LGK62/LTRD57 we get a significant coefficient on the order of .01 (.004). Assuming a constant rate of growth of R and D between 1957 and 1962, this stock over initial flow variable approximates the rate of growth of R and D times 3 (ignoring constants).[7] Thus, the implied

7. If we assume that $R_t = R_{57}\,(1+\rho)^{t-57}$, then $K62 = \sum^{62} R_t = R_{57} \sum^{5} (1+\rho)^i$ $= R57\,(6 + 15\rho + 20\rho^2 + \ldots)$. Ignoring terms of order $\rho^3$ and higher and assuming that $\rho \approx .1$ and hence $20\rho^2 \approx 2\rho$, gives

$$\begin{aligned} \log K62/R57 &\approx \log 6\,(1 + 17\rho/6 + \ldots) \\ &\approx \log 6 + \log\,(1 + 3\rho \ldots) \\ &\approx \log 6 + 3\rho. \end{aligned}$$

The first term goes into the constant, implying that the estimated coefficient of $\log K62/R57$ should be multiplied by about 3 to convert it into a coefficient of $\rho$.

*Coefficients that are statistically significant at the conventional .05 level are identified by their respective signs.

| | |
|---|---|
| VA 63 | = Value added in 1963 |
| $C_2$ | = Capital services in 1963; depreciation plus rentals plus 8% of net fixed assets and inventories |
| EM | = Total employment in manufacturing establishments |
| SPR | = 1963 company industry (five-digit) specialization ratio |
| M | = Fraction of total company employment in manufacturing establishments |
| LSE | = Logarithm of the average size of establishment in 1963 (total employment / number of establishments) |
| LFP | = Logarithm of the fraction of cumulated research expenditures (by 1963) that were financed by company funds; FP = "fraction private" |

See the notes to table 8.2 for the definition of the other variables. The number of observations is the same as in tables 8.2 and 8.3.

**Table 8.5** **Constrained Multivariate Regression Estimates**
(a) Growth Rates (BPT) and (b) Levels (LVA63) Combined

| | Coefficients of R and D | | | Coefficients of Capital | | | |
|---|---|---|---|---|---|---|---|
| | Unconstrained | | | Unconstrained | | | |
| Industry | (a) BTRD | (b) LK62 | Constrained | (a) Deprec. | (b) LC2 | Constrained | Estimated Chi-square |
| 1 | .122 | .186 | .140 | .303 | .360 | .355 | 1.66 |
| | (.030) | (.041) | (.023) | (.241) | (.047) | (.045) | |
| 2 | .098 | .093 | .085 | .235 | .453 | .449 | 2.42 |
| | (.023) | (.021) | (.014) | (.165) | (.049) | (.042) | |
| 3 | .077 | .031 | .041 | .320 | .507 | .456 | 2.70 |
| | (.033) | (.020) | (.016) | (.111) | (.064) | (.054) | |
| 4 | .025 | .043 | .028 | .284 | .017 | .074 | 3.10 |
| | (.060) | (.038) | (.028) | (.137) | (.072) | (.059) | |
| 5 | .114 | .032 | .048 | −.143 | .196 | .176 | 1.80 |
| | (.063) | (.031) | (.024) | (.363) | (.065) | (.058) | |
| 6 | .054 | .044 | .046 | .535 | .467 | .471 | .20 |
| | (.021) | (.011) | (.009) | (.178) | (.024) | (.024) | |
| 7 Total | .072 | .069 | .067 | .324 | .454 | .422 | 1.60 |
| | (.013) | (.009) | (.007) | (.074) | (.022) | (.017) | |

NOTES: Estimated standard errors are given in parentheses. Other variables in equations: (a) DNI, Age C; (b) LEM63, and industry dummies in the total (industry 7) equation.

Estimated chi-square: Twice the difference in the estimated log likelihood between the unconstrained and constrained multivariate regressions. The expected value of this statistic under the null hypothesis of the validity of the two cross-equations restrictions is 2. The critical value of $\chi^2$ with two degrees of freedom is 6 at the .05 significance level and 4.6 at the .1 level. The estimated chi-squares are thus not even close to the critical values.

**Table 8.6** **Alternative Estimates of the Coefficient of R and D, by Industry**

| Industry | OLS | TSLS |
|---|---|---|
| 1 | .122 | .110 |
| | (.030) | (.048) |
| 2 | .098 | .232 |
| | (.023) | (.069) |
| 3 | .077 | .099 |
| | (.033) | (.072) |
| 4 | .096 | .117 |
| | (.067) | (.094) |
| 5 | .114 | .113 |
| | (.063) | (.072) |
| 6 | .054 | .011 |
| | (.021) | (.035) |
| 7 All (combined) | .072 | .139 |
| | (.013) | (.049) |

Dependent variable: BPT.
Included independent variables in addition to BTRD: Age C, D, DNI. Also SBTRD for industry 4. Industry dummies in the combined (all industries) equation.
Instruments (excluded independent variables): M, AV/S, AI/V; SCE 58, SPR 63, GRR, FP 62, AR/S, SE/E, SBTRD, LGW58, LGANSE, ALVA. In industry 4, SBTRD is not used as an instrument. In industry 6, the instruments were AI/V, GRR, SE/E, SBTRD, LTRD57, LGFP62, K/SC, and LGVA57. (See table 8.A.1 for definitions.)

coefficient of the rate of growth of R and D over the shorter period is about .03, not much less than the earlier estimate of .04. Alternatively, if one substitutes the ratio of research scientists and engineers to total employment (29. SE/E), one gets a coefficient of .38 with a standard error of .21. The intensity variables do a better job for all industries combined, the substitution of the average R and D to sales ratio (20. AR/S) resulting in a coefficient of .07 (.02).

Another way of asking a similar question is to relate profitability rates to past research investments. Assume profits consist of two types of returns $\Pi \approx r_1 C + r_2 K$, where $r_1$ is the rate of return on physical capital and $r_2$ is the rate of return on "knowledge" capital. Then regressing the observed profit rate $\Pi/C = r_1 + r_2 K/C$ on the ratio of cumulated R and D to fixed capital would provide an estimate $r_2$. Unfortunately, because we really don't have the right numbers we can only approximate such an estimate. Since the returns to R and D are distributed over time, we'd like to have a time series in profitability or some estimate of permanent or average profits. Actually, we don't have a perfect measure even for one year. What we do have is gross profits (called by Telser [1972] the "contribution to overhead") in one year (1963) as a ratio to total domestic assets. This variable (19. GRR) is computed by subtracting total payrolls and equipment rentals from value added and dividing the result by total domestic assets. It is an estimate of the gross company rate of return, before depreciation and corporate taxes. Also, bypassing the problems involved in the measurement of the stock of R and D capital ($K$) discussed earlier, we do not have an explicit measure of $K/C$. It was not one of the variables included in our matrices. But we do have log $K/C$ and can use that to approximate it. In addition, there will be a problem in interpreting the resulting $r_2$ estimates, since past and current R and D expenditures are treated as current expense and subtracted from profits rather than capitalized, while the equipment used in the R and D process is already included in the total fixed capital measure ($C$). Thus, the resulting estimates are to some extent a measure of the *excess* rate of return, above and beyond that already imputed to the conventional factors used in the R and D process.

With these reservations out of the way, we can turn to table 8.7, which presents the results of such regressions for the six separate industrial groupings and the total sample. In addition to the log $K/C$ measure, we include also a measure of absolute size (log $C2$)[8] and industry dummies (in the combined regression). The estimated coefficients of log $K/C$ are always positive and significant, except in the case of industry 3. Since we used log $K/C$ instead of $K/C$ as our variable, we have to multiply the resulting coefficient by $C/K$ to get at an estimate of $r_2$. Evaluating it at the approximate arithmetic means of $C$ and $K$, i.e., at $\bar{C}/\bar{K}$, gives the numbers in column (4).[9] Dividing these numbers in turn by the ratio of average company-financed to total R and D (24. FP 62) translates them into rates of return to company-financed R and D. These are listed in the last column of table 8.7. On the whole the estimates appear to be both reasonable and high. The highest rates of return are estimated for the chemical, drugs, and petroleum industry group. Metals and machinery, motor vehicles, and all other industries show a rather high overall rate of return, in excess of 20%. Allowing for a depreciation rate of 10% still would leave an *excess* rate of return above 10%, or about double that earned by physical capital during the same period.

8. We use log $C2$ instead of log $C1$ to reduce the possible spurious relationship between the various measures. But the results of using log $C1$ are very similar to those reported here.

9. Because we were not given the actual means for our samples, but only means rounded to lower class interval boundaries, we cannot really use the supplied geometric means to evaluate anything (since being off by 1 on a natural logarithm is to be off by a factor of 2.7). But since the arithmetic means are very large, rounding introduces little error there.

Two industries, 3 (electrical equipment) and 5 (aircraft and missiles), yield the lowest estimates. These industries have the highest federal involvement in their research activity. The fraction that company-financed R and D is of the total was .65 in industry 3 and only .28 in industry 5 in 1962. The relative specificity of federally supported R and D may explain the estimated low rates of return in these industries. Since together these two industries accounted for over 60% of total R and D in 1963 (see table 8.1), they have a strong depressing effect on the estimated rate of return for the total combined sample. Still, an *excess* gross rate of return of 19% on average company R and D investment is no small matter.

**Table 8.7** **Relationship between Company Profitability (GRR) and Past Research Investment (K62), by Industry**

| Industry | Coefficients of (standard errors) | | $R^2$ and S.E. | Implied Rate of Return to R and D Investments[1] | |
|---|---|---|---|---|---|
| | log K62/C1 (1) | log C2 (2) | (3) | Total (4) | Company[2] (5) |
| 1 | .077<br>(.018) | −.039<br>(.018) | .344<br>.241 | .93 | 1.03 |
| 2 | .055<br>(.013) | −.041<br>(.014) | .112[3]<br>.204 | .25 | .28 |
| 3 | .015<br>(.010) | −.021<br>(.013) | .037<br>.148 | .02 | .03 |
| 4 | .046<br>(.017) | −.014<br>(.015) | .191<br>.121 | .23 | .29 |
| 5 | .104<br>(.036) | −.079<br>(.029) | .332<br>.227 | .05 | .17 |
| 6 | .010<br>(.005) | −.033<br>(.008) | .041<br>.155 | .23 | .26 |
| 7 (combined total) | .033<br>(.007) | −.034<br>(.005) | .136<br>.185 | .17 | .19 |

DEFINITIONS

Log K62/C1: logarithm of cumulated total R and D as a fraction of total domestic assets in 1963; log C2: logarithm of capital services as of 1963; dependent variable: 19; GRR: approximate company gross rate of return in 1963; S.E.: estimated residual standard error.

[1]Evaluated at the ratio of arithmetic means for K62 and C2.

[2]Column (4) divided by the FP62 ratio.

[3]Also contains a significant SPR variable.

## 8.5 R and D and Firm Size

There are a number of important policy issues connected with the question of optimal size of an R and D program which cannot really be dealt with in this study. Nevertheless, we do have some negative results which are worth reporting.

The question of the relationship between firm size and research productivity has been recently analyzed by Fisher and Temin (1973) who show that one can tell very little, a priori, about this relationship, and that one cannot conclude much from an observed relationship between firm size and research *inputs*. Roughly speaking, it may pay a large firm to engage in more research, pushing it to a point where its marginal return is lower than that for a smaller firm. We cannot, then, conclude that just because a firm is doing relatively more research it would be a good idea to transfer additional resources to it from the smaller firms.

Actually, we can also look at the relation of R and D output to firm size, not just R and D input. The results presented earlier, however, are rather negative. There is no indication of significant increasing returns to scale in the productivity *level* results summarized in table 8.4. For most company-level production function regressions the estimated sum of coefficients *including* the coefficient of cumulated R and D is unity or less. There is some evidence that more specialized (i.e., less diversified) companies having plants of larger than average size are more efficient, but there is no evidence of increasing returns to total company size as such (except possibly in industries 1 and 4).

Nor is there any evidence of increasing returns to the relative size of the research program as such. In the productivity growth rate equations, shown in table 8.2, and in comparable estimates (not shown) for individual industry groups, an absolute measure of the level of R and D investments such as LGK62 or LGANSE always has a negative sign, and this negative relation is usually statistically significant. Similarly, the estimated functional form used in the rate-of-return regressions in table 8.7 (GRR on LGK62/LGC1) implies diminishing productivity with respect to the absolute size of the R and D programs.

There are several reasons why these findings should not be taken seriously as a positive proof of diminishing returns to R and D: some of our variables are subject to errors of measurement which could lead to downward biases in our estimates. Also, the use of rates of R and D *investment* growth as measures of R and D *stock* growth may overestimate the latter for large companies with a long R and D history, and the estimated negative coefficients for the cumulated R and D levels may be due to nothing more than an adjustment for such a specification bias. But the point to be made is that we have found no prima facie evidence that the *rate of growth* of productivity is higher in larger companies with larger R and D programs or that the *level* of productivity is proportionately higher in the largest companies.

Nor is it clear that the larger companies invest more than proportionately in R and D.[10] Ours is the first set of data which allows a look at this question at the micro level for a relatively large number of companies (almost all of the universe). In table 8.8 we present regressions which summarize, for the whole sample, the relationship between different measures of R and D and company size. The major measure of company size used is ALVA—average of the logarithm of value added in 1957 and 1963. The first measure examined is the logarithm of total cumulated R and D (LGK65) over the whole available period (1957–65).[11] The crude results, regression 1, indicate that larger companies did spend relatively more, and significantly so, on R and D than smaller companies. But once we allow for data difficulties (DNI) and differences in specialization (SPR), this relationship evaporates. What remains (in regression 3) is a strong indication that fixed capital-intensive firms tend also to be R and D–intensive. There is also some indication that larger plant firms (LSE) are more R and D–intensive, but not larger companies as such.

The other regressions reported in table 8.8 examine in turn the relationship to firm size of cumulated company (as against total, which also includes federally financed) R and D in 1962 (LGCK62), the average R and D investment to value-added ratio in 1957 and 1963 (AR/V), the average company R and D to value-added ratio (CAR/V), and the log of the fraction that cumulated company R and D was of total cumulated R and D in 1962 (LFP62). The conclusion is the same: Overall there is little evidence of anything more than just a proportional relationship between R and D and size. There is some evidence that federally financed R and D is biased towards larger, more diversified companies, and that total R and D investments are not uniformly distributed across industries and companies. Capital-intensive, large-plant companies tend to invest somewhat more in R and D, which may be related to technological differences and the differential profitability of R and D investments across industries. But holding such differences constant, none of

10. While the relationship of R and D inputs to size does not in general imply much about the relationship of R and D output to size (see Fisher and Temin 1973), for the specific model outlined in section 8.2 of this paper which is homogeneous in R and D and non–R and D input, a more than proportionate increase in input would also imply a more than proportionate increase in output.

11. Value added in 1957 was estimated from value added in 1958 using the relative change in total sales between these years.

**Table 8.8** **Relationship of R and D to Company Size, All Industries Combined**

| Dependent Variable | ALVA | LC2/ALVA | Other Variables | $R^2$ | S.E. |
|---|---|---|---|---|---|
| LGK65 | | | | | |
| 1 | 1.203 (.037) | | I's | .656 | 1.202 |
| 2 | 1.024 (.040) | | I's, DNI+, SPR− | .692 | 1.138 |
| 3 | 1.010 (.090) | .248 (.080) | I's, DNI+, SPR−, LSE | .697 | 1.129 |
| LGCK62 | | | | | |
| 1 | 1.149 (.036) | | I's | .615 | 1.173 |
| 2 | .967 (.088) | .202 (.077) | I's, DNI+, SPR−, LSE+ | .661 | 1.104 |
| AR/V | | | | | |
| 1 | .006 (.004) | | I's | .198 | .124 |
| 2 | .004 (.010) | +.035 (.009) | I's, DNI, SPR, LSE+ | .223 | .123 |
| CAR/V | | | | | |
| 1 | −.003 (.003) | | I's | .05 | .093 |
| 2 | −.003 (.007) | +.017 (.006) | I's, LSE+ | .061 | .092 |
| LFP62 | LVA63 | LC2/VA63 | | | |
| | −.062 (.035) | .077 (.031) | I's, DNI+, SPR+ | .400 | .446 |

the measures yields any evidence for the proposition that the largest firms invest more than proportionately in R and D. They do invest *more*, but not relatively to their size.

In table 8.9 we examine the relationship of the R and D to value-added ratio to company size for each of our six industry groupings separately. Again, once capital intensity is controlled for, there is no significant relationship of R and D intensity to size. The results of using only the company-financed R and D ratio as the dependent variable (not shown here) are similar. In short, in our population of already very large companies (1000-plus employees) there is no indication that either the intensity of R and D investments or their productivity is related positively to company size.

## 8.6 Discussion and Suggestions for Further Research

In spite of various reservations, we have found a rather consistent positive relationship between various measures of company productivity and investments in research and development. In particular, Cobb-Douglas–type production function estimates based on both levels (1963) and rates of growth (1957–65) indicate an overall elasticity of output with respect to R and D investments of about .07, which can be thought of as an average of .1 for the more R and D–intensive industries such as chemicals and .05 for the less intensive rest of the universe. These findings are consistent with the earlier findings of Mansfield and Minasian, but are based on a much larger and more recent data base.

It is rather hard to convert the estimated $\alpha = .07$ into an estimate of the rate of return to R and D investments. Accepting our estimates and the validity of our measures, and using the elasticity formula to derive the implied marginal product estimate yields .27 as the overall estimate of the average gross excess rate of return to R and D in 1963. This is an average for 1963 because it is based on a function fitted across all the firms in our sample and because it is evaluated at the average total cumulated R and D to value-added ratio in 1963 in our

DEFINITIONS

LGK65 $= \log \sum_{57}^{65}$ Total R and D; LGCK62 $= \log \sum_{57}^{62}$ Company R and D

AR/V $= \frac{1}{2}\left[\left(\frac{\text{Total R and D}}{\text{Value Added}}\right) 57 + \left(\frac{\text{Total R and D}}{\text{Value Added}}\right) 63\right]$

ALVA $= \frac{1}{2}$ (log VA57 + log VA63), Average of log value added

CAR/V = similar for company R and D

LFP62 = LGCK62 − LGK62

**Table 8.9 Relationship of R and D Intensity to Company Size by Industry**
Dependent Variable—AR/V

| Industry and Regression | | Coefficient of (standard error) | | | |
|---|---|---|---|---|---|
| | | ALVA | LC2/ALVA | $R^2$ | S.E. |
| 1 | a | −.002 | | .004 | .043 |
| | | (.004) | | | |
| | b | −.002 | −.008 | | |
| | | (.007) | (.006) | .021 | .043 |
| 2 | a | .013 | | .017 | .091 |
| | | (.007) | | | |
| | b* | −.002 | .081 | .176 | .083 |
| | | (.018) | (.015) | | |
| 3 | a | .005 | | .000 | .311 |
| | | (.026) | | | |
| | b | .006 | .171 | .047 | .305 |
| | | (.081) | (.078) | | |
| 4 | a | .010 | | .304 | .025 |
| | | (.003) | | | |
| | b | .011 | .006 | .322 | .025 |
| | | (.007) | (.007) | | |
| 5 | a | .064 | | .009 | .239 |
| | | (.032) | | | |
| | b | .057 | .124 | .174 | .236 |
| | | (.102) | (.090) | | |
| 6 | a | −.000 | | .000 | .025 |
| | | (.001) | | | |
| | b | −.000 | −.000 | .000 | .025 |
| | | (.003) | (.003) | | |

*Also includes DNI.

sample ($K/V = .26$).[12] It is "gross" because neither our measures of output or of input allow for any depreciation of past R and D investments, and it is "excess" because the conventional labor and fixed capital measures already include the bulk of the current R and D expenditures once.

12. The average $K$ suffers from conflicting biases. It contains nothing for pre-1957 R and D investments and hence *it* is too low, but *it* allows no depreciation in the past accumulation and hence is too high. The two effects are likely to cancel each other out, at least as of 1963. For total industrial R and D, taking Kendrick's (1973) estimates for 1948 cumulated R and D capital as a benchmark and assuming a 10% annual depreciation rate yields a stock estimate of $K$ as of 1963 only about 6% higher than what we get by just summing from 1957 to 1962.

While our industry groupings differ in the estimated level of this elasticity, they also differ markedly in their R and D intensity, which actually results in much less difference in the estimated rates of return than one might have thought to start out with. Taking tables 8.5 and 8.6 together, one might conclude that $\alpha$ is about .1 or higher for industries 1 and 2, between .05 and .1 for industries 3, 4, and 5, and less than .05 for industry 6. Since the average $K/V$ ratios for these industries are .23, .23, .6, .16, 1.4, and .09, respectively, the implied rates of return are approximately .43, .43, .08, .31, .04, and .44, respectively (taking $\alpha$ as .1 for industries 1 and 2, .05 for industries 3, 4, and 5, and .04 for industry 6). Thus, except for industries 3 and 5, the resulting estimates of the private rates of return to total R and D are on the order of 30 to 40%. These estimates are larger, but not inconsistent with those presented in table 8.7, based on an entirely different dependent variable (GRR). There, too, the two industries with the largest federal involvement in the financing of R and D (3. electrical equipment and 5. aircraft and missiles) yield the lowest rate-of-return estimates.[13]

It is interesting to note that we have stumbled on this impact of federally financed R and D in the interpretation of our results rather than in the econometric analysis itself. In our regressions we were unable to discover any direct evidence of the superiority of company-financed R and D as against federally financed R and D in affecting the growth in productivity. It may well be the case that within any company a dollar is a dollar, irrespective of the source of financing, but that in these two specific industries the externalities created by the large federally financed R and D investments and the constraints on the appropriability of the results of research that may have been associated with such investments

13. In general these estimates are of the same order of magnitude as those reported by Griliches (1973) and Terleckyj (1974) based on regressions of productivity growth on R and D investment ratios for aggregate interindustry data in the U.S. The first study, based on eighty-five manufacturing industries, yielded estimates of 32 to 40% for the rate of return to R and D. The second study, based on twenty manufacturing industries, yielded an estimated rate of return of 37% to company-financed R and D and essentially zero to federally financed R and D. Both studies were based on R and D data for 1958 only. While the results reported above are of the same order of magnitude, I have not been able to replicate this type of equation on these data and get coefficients of the same order of magnitude. The best equation for the combined sample was

$$BPT = \underset{(.028)}{.135}\, AR/V - \underset{(.008)}{.042}\, K/V + (\text{constant}, I\text{'s}, \text{Age } C,D); \quad R^2 = .089, \quad \text{S.E.} = .058,$$

implying a rate of return of about a half of that discussed above and a depreciation rate of 31%, if it were to be believed. Besides pointing to the difference in time periods and the use of aggregate versus micro data, I do not have a satisfactory way of reconciling these results at the moment.

have driven down the realized private rate of return from R and D significantly below its prevailing rate in other industries.[14]

In general, this paper can be viewed as another link in a chain of a rather limited number of investigations supporting the argument that R and D investments have yielded a rather high rate of return in the recent past. In addition, we find no evidence for, and some evidence against, the notion that larger firms either have a higher propensity to invest in R and D or are more effective in deriving benefits from it.

There is little point in reiterating the various reservations outlined earlier. Some of the difficulties are inherent in the attempt to measure and discuss "research" and "productivity" as if they were clear and unequivocal concepts. But many of the problems, particularly those dealing with timing effects, spillovers, and externalities, could yield to more data and better data analysis. It would be very useful to have more detail on the firms at hand, especially information on the distribution of their research expenditures, on other measures of research output such as patents granted and papers published, on income received from royalties, and on money spent on advertising. All of this is feasible; it requires "only" the additional matching of IRS, SEC, and Patent Office and scientific abstracting services data bases. It would also help to know, for tracing out and following up potential externalities, more about the exact industrial structure of individual firms and their product mix. Finally, it should be relatively easy and quite useful to extend this study, as is, to the 1966–74 period. Such an extension would be particularly interesting since it would allow us to observe a period during which R and D growth largely came to an end for many firms (at least in real terms). Besides helping us to find out something about the structure of lags and the rate of depreciation in such data, it would also, for the first time, break sharply the confounding collinearity between growth in R and D and the growth that occurred in almost all of the other economic variables during the 1956–65 period.

Even without new data, we have not yet exhausted what can be learned from the data at hand. Additional analysis of the data on the number of scientists and engineers as against R and D dollar totals should prove illuminating. This distinction between federally and company-financed R and D has not really been explored in depth yet. Finally, a detailed comparison of the individual industry results with

14. This may explain why the aggregate studies cited in the previous footnote found much higher returns to company-financed R and D investments relative to federally financed ones than we did. Another way of looking at it is that in industries with a high rate of federally financed R and D expenditures the rate of depreciation (obsolescence) of the previously accumulated R and D capital is much higher. Again, this would be a difference which wouldn't be observed at the firm level. It is external to the firm but internal to the industry.

industry aggregates, focusing on the potential externalities (external to the firm but internal to the industry), is required before any strong conclusion could be drawn about *social* rates of return from our estimates of *private* rates of return to R and D.

# Appendix A

**Table 8.A.1 Variables in the R and D Study (Total N=883)**

| Variable | | | Overall Sample | |
|---|---|---|---|---|
| Number | Name | Definition | Mean (approximate) | Standard Deviation |
| 1 | ID1 | Industry dummy: Chemicals and petroleum SIC 28, 29, 13 | $N = 110$ | |
| 2 | ID2 | Metals and machinery SIC 34, 35 | $N = 187$ | |
| 3 | ID3 | Electrical equipment and communication SIC 36, 43 | $N = 102$ | |
| 4 | ID4 | Motor vehicles and transportation SIC 371, 373–9 | $N = 34$ | |
| 5 | ID5 | Aircraft and missiles SIC 372, 19 | $N = 31$ | |
| 6 | AGE C | (gross fixed assets – net fixed assets) divided by gross fixed assets (in 1963) | .5 | .105 |
| 7 | D | Depreciation ratio: Depreciation charged in 1963 divided by gross fixed assets in 1963 | .06 | .028 |
| 8 | D/V | Depreciation to value-added ratio, 1963 | .06 | .057 |
| 9 | C3 | Total domestic assets, 1963 | $260 \times 10^6$ | $766 \times 10^6$ |
| 10 | S57 | Sales in 1957 | $200 \times 10^6$ | $62 \times 10^6$ |
| 11 | M | Ratio of employment in manufacturing establishments to total company employment | .80 | .17 |
| 12 | ALSS | Average share of total payroll in sales (average of the ratios for 1958 and 1963) | .30 | .11 |
| 13 | ALSV | Average share of labor in value added (average of payroll to value added for 1958 and 1963) | .50 | .16 |
| 14 | AV/S | Average ratio of value added to sales (1958 and 1963) | .50 | .16 |

**Table 8.A.1** (cont.)

| Variable | | Definition | Overall Sample | |
|---|---|---|---|---|
| Number | Name | | Mean (approximate) | Standard Deviation |
| 15 | AI/V | Average ratio of investment (total capital expenditures) to value added (1958 and 1963) | .07 | .07 |
| 16 | VA63 | Value added in 1963 | $120 \times 10^6$ | $361 \times 10^6$ |
| 17 | SCE58 | Average number of employees per establishment in 1958 | 350 | 751 |
| 18 | SPR63 | 1963 company industry (five-digit) specialization ratio | 60 | 27 |
| 19 | GRR | Gross rate of return in 1963: Value added minus total manufacturing payroll minus equipment rentals divided by gross domestic assets | .26 | .20 |
| 20 | NRR | "Net" rate of return: Value added minus manufacturing payroll, minus equipment rentals, minus depreciation, divided by net *fixed* assets plus inventories | .50 | .62 |
| 21 | LGS63 | Log total sales in 1963 | 10.00 | 1.20 |
| 22 | LGS57 | Log total sales in 1957 | 10.00 | 1.29 |
| 23 | K62 | Cumulated total R and D expenditures, 1957–62 | $50 \times 10^6$ | $272 \times 10^6$ |
| 24 | FP62 | Fraction private 62: Cumulated company R and D expenditures 1957–62 divided by K62 | .90 | .23 |
| 25 | FP65 | Fraction private 65: Cumulated company R and D expenditures 1957–65, divided by K65 | .90 | .23 |
| 26 | AR/V | Average R and D to value-added ratio, 1957 and 1963 | .05 | .14 |
| 27 | K/V | Cumulated R and D in 1962 to value-added in 1963 ratio | .26 | .51 |
| 28 | AR/S | Average (1957 and 1962) R and D to sales ratio | .03 | .09 |
| 29 | SE/E | Average (1957 and 1962) scientists and engineers to total employment ratio | .02 | .04 |
| 30 | CAR/V | Company R and D to value-added ratio 1957 and 1962 average | .03 | .09 |

Rates of growth ($b$'s), computed from regressions of $\log y = a + bt$, for the period 1957–65

**Table 8.A.1** (cont.)

| Variable Number | Variable Name | Definition | Overall Sample: Mean (approximate) | Overall Sample: Standard Deviation |
|---|---|---|---|---|
| 31 | BS | Rate of growth of sales | .06 | .074 |
| 32 | BE | Rate of growth of employment | .023 | .065 |
| 33 | BSE | Rate of growth of scientists and engineers employment | .05 | .14 |
| 34 | BTRD | Rate of growth of total R and D | .08 | .15 |
| 35 | BCRD | Rate of growth of company R and D | .08 | .16 |
| 36 | SBS | Standard error of estimate rate of growth of sales | .014 | .015 |
| 37 | SBTRD | Standard error of estimate rate of growth of total R and D | .035 | .038 |
| 38 | R/V57 | Total R and D to value-added ratio, 1957 | .05 | .21 |
| 39 | LTRD57 | Log total R and D, 1957 | 6.0 | 2.25 |
| 40 | LCRD57 | Log company R and D, 1957 | 6.0 | 2.13 |
| 41 | BPT | Partial productivity growth 1957–65: BS − ALSS × BE | .05 | .06 |
| 42 | DN1 | Dummy variable 1 if no imputations in the data, zero otherwise | .6 | |
| 43 | LGE63 | Log total employment, 1963 | 8.0 | 1.04 |
| 44 | E57 | Total employment, 1957 | 9,000 | 26,358 |
| 45 | LGC1 | Log gross fixed assets 1963 | 10.00 | 1.48 |
| 46 | LGC2 | Log capital services in 1963; capital services: Depreciation and rentals and 8% of net fixed assets and inventories | 8.00 | 1.32 |
| 47 | LGEM63 | Log manufacturing employment, 1963 | 8.00 | 1.04 |
| 48 | LGFM57 | Log manufacturing employment, 1957 | 8.00 | 1.12 |
| 49 | LGW58 | Log average "wage" in 1958 (wage = payroll per employee) | 1.6 | .20 |
| 50 | LGW63 | Log wage rate in 1963 | 1.8 | .20 |
| 51 | LGVA63 | Log value added in 1963 | 10.00 | 1.15 |
| 52 | LGSCE63 | Log average scale of establishments in 1963 | 5.00 | 1.0 |
| 53 | LGANSE | Log average number of scientists and engineers, 1957–62 | 3.5 | 1.90 |
| 54 | LGK62 | Log cumulated R and D through 1962 | 8.0 | 2.1 |
| 55 | LGK65 | Log cumulated R and D through 1965 | 9.0 | 2.0 |

Table 8.A.1 (cont.)

| Variable | | | Overall Sample | |
|---|---|---|---|---|
| Number | Name | Definition | Mean (approximate) | Standard Deviation |
| 56 | LGFP62 | Log 1962 cumulated company R and D as a fraction of total cumulated R and D | −.18 | .57 |
| 57 | T63 | Log absolute total factor productivity level in 1963: LGVA63 − ALSV × LGEM63 − (1 − ALSV) LGC2 | 2.0 | .34 |
| 58 | GVA | Growth in value added, 1957–63: (LGVA63 − LGVA57)/6 | .06 | .074 |
| 59 | GPT | Growth in partial productivity 1957–63: [GVA − ALSV × (LGEM63 − LGEM57)]/6 | .05 | .058 |
| 60 | GSCE | Growth in average scale of establishments: (LGSCE63 − LGSCE58)/5 | −.02 | .09 |

Additional variables constructed from the above set:

K/SC = 54−53, log of cumulated R and D per scientist
LGCK = 56+54, log of cumulated company-financed R and D, 1962
K/C = 54−45, log of the cumulated R and D to fixed capital ratio
LGVA57 = 51−(6)×.58, log of value added, 1957
ALVA = 51−(3)×.58, average of log value added in 1957 and 1963

Value added in 1957 estimated by extrapolating value added in 1958 using the percentage change in sales between 1957 and 1958.

NOTE: Industry group 6 is "All others," $N = 419$. All dollar figures are in thousands.

# Appendix B

*Criteria Used for Inclusion of a Company in the Griliches-NSF-Census Bureau Project*

1. Only companies with 1,000 or more employees in one or more years and filing annual reports on Research and Development (Form RD–1 or RD–2) were included. The list was further limited to companies classified in manufacturing, Petroleum (SIC 13), and Communications (SIC 48). This is the area included under the term "manufacturing" in the Annual Survey of Research and Development in Industry conducted by the Census Bureau for the National Science Foundation.

2. Subsequently, in the final tabulations, only those companies for which we had R and D reports for four or more years during the period 1957–65 were retained.

3. All companies included were matched to the 1963 and 1958 Enterprise Statistics data. Company data in the Griliches-NSF-Census study are combined and classified according to the 1963 enterprise company composition and industry code. A few R and D companies of relatively small size, not matched to the enterprise lists, were dropped.*

4. During the search and edit routine, all cases outside four standard deviations of the various tests were rechecked by clerical and professional staff. A few small cases that could not be explained were dropped from the project.

# Appendix C

Memorandum to Mr. Owen
23 November 1971
Attachment C

*Imputation and Estimation Methods for Griliches-NSF-Census Project*

1. Imputation of R and D data. Our primary data file contained nine years of data, 1957 through 1965, for five items reported in R and D surveys: sales, employment, employment of scientists and engineers assigned to R and D work, total R and D expenditures, and federal R and D expenditures. For each company in the survey, for each of these items, we imputed zero values as follows:

   Let $X$ represent year with a value of 80 for 1957, 90 for 1958, etc.

   Let $Y$ represent one of the R and D variables.

   For each nonzero $Y$, we cumulated $N$, $\Sigma X$, $\Sigma Y$, $\Sigma X^2$, and $\Sigma XY$.

   Then, $A = \Sigma Y/N$ and $B = (N\Sigma XY - \Sigma X \Sigma Y)/(N\Sigma X^2 - (\Sigma X)^2)$.

*Comments on the R and D–Enterprise match: The 1958 enterprise data were placed in the 1963 format. Mergers and acquisitions during the period were reflected by the addition of two or more 1958 enterprise records to equal one 1963 enterprise record. No case came to light where a single 1958 record represented two or more 1963 records. According to the R and D survey instruction, respondents should report for the entire company. However, the results of the instructions have weaknesses that are avoided in the enterprise statistics (1958 and 1963) by a match to lists of related employer identification numbers and associated employment data. The R and D–enterprise match served to update the R and D company composition data, and to establish changes in broad industry classes, Each zero value of $Y$ was imputed from its matching year value by $Y = A + B\ (X - \bar{X})$; and each imputed value was flagged.

Negative imputed values were set to zero.

This is a straight-line imputation procedure; its effects were partially as follows:

(a) items totally not reported were left at zero and flagged as imputed; (b) items reported in only one year had that value imputed for all years.

2. Estimation of regression variables.

a) Federal R and D values were reset to zero if imputed and any federal R and D greater than total R and D was set equal to the total R and D value.

b) In the following description the numbers in parentheses refer to field positions in the primary data record, Attachment A. The variable abbreviation follows Griliches's document of 13 May 1971 as amended by notes of meetings and other conversations. Only those variables whose derivation is not direct from the Griliches definition are described below.

In all cases not explicitly covered below, the calculation of a ratio with a zero value for numerator or denominator would result in a zero value for the ratio.

i. If (50) = 0, Age C = 0.

ii. If R and D sales for 1957 and 1958 were not reported or 0, $S_{57}$ = (21) and $V_{57}$ = (39); i.e., no 1957 to 1958 ratio adjustment.

iii. If R and D employment for 1957 or 1958 was not reported or 0, $EM_{57}$ = (29)xM; i.e., no 57 to 58 adjustment.

iv. For ALSS, ALSV, AV/S, and AI/V, which require an averaging of two ratios, if either ratio was zero, the other ratio is used and not averaged. If both ratios were 0, the variable would be zero and the case listed.

v. For AR/V, CAR/V, and R/V, which require averaging of a ratio involving a 1962 R and D item and a ratio involving a 1957 R and D item, if the 1962 data were missing, we used 1961; if that was also missing we used 1960; and, similarly, for 1957 we substituted 1958 and 1959. If all three early years were missing, the resulting zero ratio would have been averaged.

---

based upon Census company industry codes developed in the processing of the economic census data.

Since any four years of R and D data were sufficient to include a company, it was possible for a company with no R and D reported in 1958 or 1963 to be included in the sample. A few such cases did turn up in the development of the matched R and D–enterprise data.

vi. For AR/S and SE/E, which require an averaging of a ratio of two 1962 R and D items and a ratio of two 1957 R and D items, if either 1962 item were missing we would use 1963, if 1963 were missing we would use 1964, and if both of these were missing we would set that ratio to zero; similarly, we would substitute 1958 and 1959 for 1957. If both ratios were zero, the case would be listed.

vii. Growth rates and standard errors of the growth rates for the following R and D variables were computed: sales, employment, scientists and engineers employment, total R and D expenditures, and company R and D expenditures. For each variable, for the nine-year period, we let $X$ represent year with a value of 1 through 9, and $Y$ represent the log of the variable for nonzero values. For nonzero values of $Y$ we obtained the following counts and sums: $N$, $\Sigma X$, $\Sigma Y$, $\Sigma X^2$, $\Sigma Y^2$, $\Sigma XY$. If $N$ was less than 4 we set the growth rate and the standard error of the growth rate to zero, and set a dummy variable to one; otherwise

the dummy variable = 0;

the growth rate,

$$b = (N\Sigma XY - \Sigma X\Sigma Y)/[N\Sigma X^2 - (\Sigma X)^2],$$

and the standard error of the growth rate

$$= \text{SQRT}\ ([N\Sigma Y^3 - (\Sigma Y)^2 - b(N\Sigma XY - \Sigma X\Sigma Y)]/\{(N-2)[N\Sigma X^3 - (\Sigma X)^2]\}).$$

viii. If BS or BE could not be calculated, BPT = 0.

ix. The log of a variable with a value of zero would be set to zero.

## References

Evenson, R. 1968. The contribution of agricultural research and extension to agricultural production. Ph.D. diss., University of Chicago.

Fisher, F. M., and Temin, P. 1973. Returns to scale in research and development: what does the Schumpeterian hypothesis imply? *Journal of Political Economy* 81:56–70.

Griliches, Z. 1964. Research expenditures, education, and the aggregate agricultural production function. *American Economic Review* 54: 961–74.

———. 1973*a*. Research expenditures and growth accounting. In B. R. Williams, ed., *Science and technology in economic growth*, pp. 59–95. London: Macmillan.

———. 1973*b*. Productivity and research. In National Commission on Productivity, *Conference on an agenda for economic research on productivity*. Washington: Government Printing Office.

Kendrick, J. W. 1976. *The formation and stocks of total capital*. National Bureau of Economic Research, General Series 100. New York: Columbia University Press.

Mansfield, E. 1965. Rates of return from industrial research and development. *American Economic Review* 55: 310–322.

———. 1967. *Econometric studies of industrial research and technological innovation*. New York: W. W. Norton.

———. 1972. The contribution of research and development to economic growth in the United States. In National Science Foundation, *R&D and economic growth productivity*, NSF 72–303. Washington: Government Printing Office.

Minasian, Jora R. 1962. The economics of research and development. In National Bureau of Economic Research, *The rate and direction of inventive activity*. Princeton: Princeton University Press.

———. 1969. Research and development, production functions, and rates of return. *American Economic Review* 59 (Proceedings issue): 80–85.

National Science Foundation. 1959. *Methodology of statistics on research and development*, NSF 59–36. Washington: Government Printing Office.

———. 1970*a*. *Research and development in industry, 1968*, NSF 70–29. Washington: Government Printing Office.

———. 1970*b*. *National Patterns of R&D Resources, 1953–71*, NSF 70–44. Washington: Government Printing Office.

Terleckyj, N. E. 1960. Sources of productivity advance. Ph.D. diss., Columbia University.

———. 1974. *Effects of R&D on the productivity growth of industries: an exploratory study*. Washington: National Planning Association.

Wagner, L. U. 1968. Problems in estimating research and development investment and stock. In ASA, *1968 Proceedings of the business and economic statistics section*, pp. 189–97. Washington: ASA.

## Comment Edwin Mansfield

Zvi Griliches comes to three principal conclusions in this interesting and useful paper. First, he estimates that the elasticity of output with respect to R and D investment is about 0.1 in the more R and D–intensive manufacturing industries, and about 0.05 for the less R and D–intensive manufacturing industries. Second, he estimates that the rate

Edwin Mansfield is at the University of Pennsylvania.

of return from R and D is on the order of 20% in all manufacturing—much lower than this in industries like aircraft and missiles and electrical equipment, where there is great federal involvement in the financing of R and D, and much higher in industries like chemicals. Third, he finds some evidence against the proposition that larger firms either have a higher propensity to invest in R and D or are more effective in deriving benefits from it.

An important advantage of Griliches's study over earlier ones is its inclusiveness. For the first time, we have, thanks to Griliches, the National Science Foundation, and the Census Bureau, results that pertain to almost 900 firms. This is a far larger sample than has been analyzed in previous studies, and Griliches's results are both interesting and welcome. For some purposes, however, it might have been preferable to use finer industrial categories. As matters stand, some results for individual industry groups are difficult to interpret. For example, the chemical industry includes petroleum, chemical, and drug firms. Thus, when R and D intensity and other variables are regressed on firm size, a considerable part of the relationship must be due to the well-known differences among the petroleum, chemical, and drug industries. Perhaps Griliches or his students may be able to extend the results in this way in the future.

The model that Griliches uses is similar to ones used by previous investigators. For example, it assumes that technological change is neutral, and that the Cobb-Douglas form is appropriate. Also, it assumes that the direction of causation runs from R and D expenditures to output, whereas to some extent there may be an identification problem. In other words, high rates of growth of output may lead to high R and D expenditures, or be associated with firm characteristics (such as the nature and quality of management) related to high R and D expenditure. Thus, there may be some bias in the estimated regression coefficients. Recognizing this fact, he computes two-stage least-squares estimates of the coefficient of R and D in table 8.6. Although he concludes that only in industry 6 do the two-stage least-squares results not yield a significant R and D coefficient, it appears that the coefficient is less than 1.6 times its standard error in industries 3, 4, and 5 as well. Thus, in two-thirds of the industries, the two-stage least-squares estimates of the R and D coefficient are not statistically significant.

Also, as Griliches emphasizes, he is forced to use proxies—"makeshift" proxies, as he characterizes them—to represent the growth in physical capital as well as the growth in R and D capital. In particular, he uses two proxies for the rate of growth of physical capital: the ratio of accumulated depreciation to the stock of physical capital in 1963, and the 1963 depreciation rate. Although these proxies may be serviceable if the growth rate (and depreciation rate) of capital is constant over time, this may not be the case. Thus, as he recognizes, these proxies may be troublesome. And the proxy for the growth rate of R and D capital is the rate of growth of R and D expenditures, which can be troublesome too. Combining the use of these proxies with the use of undeflated output data and the fact that the lag in the effects of R and D is ignored, it is clear that the results are very rough.

Yet, despite these approximations, the extent to which his findings agree with previous studies is quite striking. To begin with, recall his conclusion that the elasticity of output with respect to cumulated R and D expenditures is about 0.1 in the more research-intensive industries. This estimate is quite close to my estimate of 0.12 for chemicals and petroleum, and to Minasian's estimate of 0.11 for chemicals (Mansfield 1968; Minasian 1969). Also, consider his estimate at the end of his paper that the rate of return from R and D is perhaps 40% or more in the chemicals group. This estimate agrees quite well with previous studies (although the figure of 93% in table 8.7 is higher than obtained before).

Turning to his conclusion that the estimated rate of return from R and D is lower in industries where there is a great federal involvement in financing R and D, it is worth noting that Terleckyj, in his very interesting study (1974) of effects of R and D on productivity change, comes to a related conclusion. Terleckyj regressed the rate of growth of an industry's total factor productivity in 1948–66 on its privately financed R and D (as a percent of value added) and its government-financed R and D (as a percent of value added), as well as a number of other variables. He found that the regression coefficient of its government-financed R and D was almost precisely zero, and far from statistically significant, whereas the regression coefficient of its privately financed R and D was highly significant. Similarly, Leonard (1971) found that privately financed R and D had a much larger impact than government-financed R and D on both the growth of industry output and the growth of output per worker.

Further, Griliches's conclusion that the largest firms do not invest more, relative to their size, than somewhat smaller firms is in accord with earlier studies. In particular, I found that this was true in the petroleum, drug, steel, and glass industries; and Scherer, in his much more inclusive study, found that it was true (Scherer 1974; see also Mansfield 1968). In both Scherer and my studies, the chemical industry was an exception; and I suspect that the reason why Griliches does not obtain similar findings is that he includes chemicals, drugs, and petroleum in his chemicals group. Also, Griliches's conclusion that there is no evidence that the largest firms are more effective in deriving benefits from R and D is not at all incompatible with the limited amount of data provided by earlier studies.[1]

1. For example, Cooper (1964).

Having summarized his principal findings, discussed some of the problems he faced, and compared his results with previous studies, let me try to present some information that may help to shed light on a couple of the areas untouched by his paper. First, as Griliches points out, we know far too little concerning the lag between investment in research and the appearance of innovations stemming from the research. This lag should, of course, be incorporated into any model of this sort. Based on work carried out by George Beardsley in his doctoral thesis (1974) at the University of Pennsylvania, the estimated probability distribution of this lag in one of the nation's largest firms is as shown in table C.8.1. As you can see, the probability distribution is different for the firm's applied R and D work than for its more basic work. For the applied R and D, the median lag is about two years for work on products and about three years for work on processes; for the more basic research, it is about four years for products and about five years for processes. Also, note that the median lag for more basic work is shorter now than in the 1960s, a tendency which I believe is true of many firms besides this one.

Of course, these figures pertain to only one large firm, and cannot be regarded as typical for all manufacturing. But the results are not very different from those collected about ten years ago from a large electrical equipment firm (Mansfield 1968). Moreover, we collected data of this

**Table C.8.1** **Estimated Percentage of R and D Budget Devoted to Work which, if Successful, was to be Commercialized at Various Lengths of Time after the R and D Expenditure (large manufacturing firm)**

| | | | Basic Research[a] | | |
|---|---|---|---|---|---|
| | Applied R and D | | | 1970–72 | |
| Time Lag (yr.) | Products | Processes | 1960s Total | Products | Processes |
| < 1 | 10 | 15 | ...... | 3 | 1 |
| 1–2 | 45 | 25 | ...... | 7 | 4 |
| 2–3 | 25 | 20 | 5 | 15 | 10 |
| 3–4 | 15 | 15 | 10 | 25 | 10 |
| 4–5 | 5 | 10 | 20 | 25 | 20 |
| 5–6 | ...... | 10 | 20 | 20 | 10 |
| 6–7 | ...... | 5 | 20 | 5 | 15 |
| > 7 | ...... | ...... | 25 | ...... | 5 |
| Total | 100 | 100 | 100 | 100 | 100 |

NOTE: All numbers exceeding 10 % are rounded to the nearest 5%.

[a]The firm's definition of "basic" research does not accord with that of the National Science Foundation. This is a segment of the firm's R and D that is relatively long-range. Its budget was about 30% of that for applied R and D in 1972.

sort from eleven major chemical firms and six major petroleum firms. These data indicated that the median length of time to completion of an R and D project and an effect on firm profits was about two years for some chemical firms and about four years for others, while it was generally about two or three years for the oil firms (Mansfield et al. 1971). Thus, the combined sample of about 19 firms indicates a median lag from R and D to innovation of about three years. And since the diffusion lag must be added to this lag, the average total lag in at least these areas of manufacturing may not be very much shorter than the figures Griliches quotes from Evenson (1968).

Second, it may be worthwhile to describe a case study of the private rate of return from investments in new technology. In his thesis, Beardsley obtained data concerning the returns from the innovative activities of one of the nation's largest firms. This firm has made estimates since 1960 of the benefits obtained from its R and D efforts, these estimates being used for internal planning purposes. This firm is among the largest members of an industry that is neither among the most research-intensive nor among the least research-intensive. In terms of the percent of sales devoted to R and D, this firm is reasonably representative of our nation's largest firms.

For each year since 1960, this firm has put together a careful inventory of the technological innovations arising from its R and D and related activities. Then it has made detailed estimates of the effect of each of these innovations on its profit stream. Specifically, in the case of product innovations, the firm has computed for each new product the expected difference in cash flows over time between the situation with the new product and without it, including the effect of the new product on its profits from displaced products. In the case of process innovations, it has computed the expected difference in cash flow between the situation with the new process and that without it, this difference reflecting, of course, the savings associated with the new process. In addition, the firm has updated these estimates each year. In other words, as time has gone on, the firm has revised its estimates of the returns from past innovations. This, of course, is of crucial importance, since it means that the firm's estimates for innovations occurring in the early and middle 1960s are based on a decade or more of actual experience, not just forecasts. The data we use are taken from the 1973 revision.

Besides these data on the private benefits from the firm's technological innovations, figures are also available concerning the firm's expenditures on R and D and related innovative activities each year. Using these cost figures, as well as the figures concerning the total cash flow of benefits stemming from the new products or new processes that came to fruition each year, we can compute the rate of return from the investment that resulted in each year's crop of innovations. It is worthwhile noting that this rate of return is based on the investment in both commercialized and uncommercialized (and successful and unsuccessful) projects.

The results, shown in table C.8.2, indicate at least two things. First, the average rate of return from this firm's total investments in innovative activities during 1960–72 was about 19%, a figure that is not too different from the average that Griliches gets. In its internal calculations, this firm regards investments with rates of return exceeding 15% as attractive. According to table C.8.2, the average rate of return from the firm's total investment in innovative activities during 1960–72 exceeded this figure. Second, innovation is a risky activity, and this is reflected in the results. Both for processes and products, the estimated private rate of return fell short of this figure of 15% in about 60% of the years. The year-to-year variation in the private rate of return seems greater for processes than for products, which may be related to the fact that the average rate of return is higher for processes than for products.

Although these figures are interesting, their limitations should be stressed. For one thing, the firm does not attempt to include in its calculations any innovation where the discounted value of its benefits

**Table C.8.2 Private Rates of Return from Total Investment in Research and Development and Related Innovative Activities, Major Industrial Firm, Process and Product Innovations, 1960–72**

| Year | Both Products and Processes | Products | Processes |
|---|---|---|---|
| 1960 | 31 | 21% | 34 |
| 1961 | 9 | 0 | 15 |
| 1962 | 7 | 17 | Negative* |
| 1963 | 26 | 13 | 30 |
| 1964 | 15 | 9 | 18 |
| 1965 | 16 | 27 | −1 |
| 1966 | 25 | 22 | 27 |
| 1967 | 11 | 11 | 12 |
| 1968 | 2 | −1 | 5 |
| 1969 | 3 | 13 | −15 |
| 1970 | 6 | 9 | 3 |
| 1971 | 12 | 16 | 10 |
| 1972 | 14 | 14 | 14 |
| 1960–72 | 19 | 14 | 22 |

*No major process innovations occurred in 1962.

(the discount rate being 15%) is less than $200,000. Since the firm's benefit figures omit the benefits from such minor innovations, the rates of return are almost surely underestimates. Also, the estimates for more recent years are not as reliable as those for the early and middle 1960s. Nonetheless, despite these and other defects in the data, the results seem to provide the most detailed description of a firm's returns from its investments in technological innovation that has been published to date.[2]

Third, as Griliches points out, his results pertain to private rates of return from R and D, not social rates of return. How much difference can there be between private and social rates of return? Judging from a study of seventeen industrial innovations that we have done recently (Mansfield et al. 1977), there can be very wide differences between them. As might be expected, the social rate of return tended to be higher than the private rate of return. Specifically, in our sample, the median social rate of return was about 56%, whereas the median private rate of return was about 25%. These estimates were derived from a detailed investigation of each of these seventeen innovations, the basic data being obtained to a considerable extent from the innovating firms, using firms, and relevant government agencies. The model on which these results are based is similar in spirit to the one used by Griliches in his earlier study of hybrid corn, although we have extended it in a number of major directions (Griliches 1958; Mansfield et al. 1977).

In conclusion, what sorts of implications can be derived from Griliches's study? Since his estimates relate to private, not social, rates of return, the fact that they are relatively high would seem to imply that firms should have expanded their R and D programs.[3] In fact, manufacturing firms did expand their investment in R and D during the middle and late 1960s; total annual expenditure by industry on R and D almost doubled between 1963 and 1969 (National Science Foundation 1972). Thus, this implication seems to be consistent with the facts, at least until about 1969. But from 1969 to 1972, there has been no appreciable increase in real terms in industry's annual R and D expenditures (National Science Foundation 1972). Consequently, one might guess that, if Griliches could have used more recent data, he might have found that the marginal private rate of return from R and D was considerably lower than the figures in his paper. Given the great changes that have occurred in the past 10 or 15 years in the attitude of industry toward R and D, it would be very interesting to see what more recent figures of this sort would reveal.

2. For further analysis of these data, see Beardsley and Mansfield (1978).

3. For some discussion of the limitations of estimates of this sort, see Mansfield (1972) and Mansfield et al. (1977).

**References**

Beardsley, G. 1974. Rates of return from investments in innovation. Ph.D. dissertation, University of Pennsylvania.

Beardsley, G., and Mansfield, E. 1978. A note on the accuracy of industrial forecasts of the profitability of new products and processes. *Journal of Business* 51:127–35.

Cooper, A. 1964. R and D is more efficient in small companies. *Harvard Business Review* 42:75–83.

Evenson, R. 1968. The contribution of agricultural research and extension to agricultural production. Ph.D. dissertation, University of Chicago.

Griliches, Z. 1958. Research costs and social returns: hybrid corn and related innovations. *Journal of Political Economy* 66:419–31.

Leonard, W. 1971. Research and development in economic growth. *Journal of Political Economy* 79:232–56.

Mansfield, E. 1968. *Industrial research and technological innovation.* New York: W. W. Norton.

———. 1972. Contribution of R and D to economic growth in the United States. *Science* 175:487–94.

Mansfield, E.; Rapoport, J.; Romeo, A.; Villani, E.; Wagner, S.; and Husic, F. 1977. *The production and application of new industrial technology.* New York: W. W. Norton.

Mansfield, E.; Rapoport, J.; Romeo, A.; Wagner, S.; and Beardsley, G. 1977. Social and private rates of return from industrial innovations. *Quarterly Journal of Economics* 91:221–40.

Mansfield, E.; Rapoport, J.; Schnee, J.; Wagner, S.; and Hamburger, M. 1971. *Research and innovation in the modern corporation.* New York: W. W. Norton.

Minasian, J. 1969. Research and development, production functions, and rates of return. *American Economic Review* 59:80–85.

National Science Foundation. 1972. *Science indicators.* Washington: National Science Foundation.

Scherer, F. M. 1970. *Industrial market structure and economic performance.* Evanston: Rand McNally.

Terleckyj, N. 1974. *Effects of R and D on the productivity growth of industries: an exploratory study.* Washington: National Planning Association.