NBER WORKING PAPER SERIES

PRODUCTIVITY, R&D, AND BASIC RESEARCH AT THE FIRM LEVEL IN THE 1970s

Zvi Griliches

Working Paper No. 1547

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1985

The research reported here is part of the NBER's research program in Productivity and project in Productivity and Industrial change in the World Economy. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #1547
January 1985

Productivity, R&D, and Basic Research at the Firm Level in the 1970s

ABSTRACT

A new data set (the NSF-Census match) containing information on the R&D expenditures, sales, employment, and other detail for approximately 1,000 largest manufacturing firms in the U.S. during 1957-1977 is analyzed using a standard production function framework augmented by the addition of an R&D "capital" and "mix" variables (basic as a fraction of total and privately financed as a fraction of total). The results indicate that R&D continued to contribute to productivity growth in U.S. manufacturing also in the 1970's, with no significant decline in its effectiveness as compared to the 1960's; that the contribution of the basic research component of such expenditures was significantly higher than its nominal ratio would imply; and that while federally financed R&D expenditures did have a positive effect on measured productivity growth of these firms, this effect was significantly smaller than the comparable contribution of privately financed R&D expenditures.

Zvi Griliches
National Bureau of
Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

Productivity, R&D, and Basic Research at the Firm Level in the 1970s

Zvi Griliches*

This paper reports new results on the relationship of R&D expenditures, especially expenditures on basic research, to productivity growth in U.S. manufacturing firms during the 1970's. It is based on a new and unique data set, the NSF R&D-Census match, containing information on R&D expenditures, sales, employment, and other detail for approximately 1000 largest manufacturing firms from 1957 through 1977. It updates the earlier work of Griliches (1980) on the precursor of this data set, replicates some of Mansfield's (1980) work on the contribution of basic research to productivity growth using a larger, more recent, and more representative sample of firms, and complements similar work by Griliches and Mairesse (1983, 1984) based on a publicly accessible but more limited data set.

Two topics are explored in some detail: (1) Is there any evidence of a decline in the returns to industrial R&D expenditures, a decline in their "fecundity" in the 1970's as compared to earlier time periods? And, (2) is there evidence that basic research is a relatively more important component of R&D and that there may have been an underinvestment in this component?

A few background facts are worth stressing at this point. In the U.S. total R&D expenditures in industry peaked (in real terms) around 1968, dropped slightly in the early 1970's and recovered somewhat in the late 1970's. Relative to total sales, R&D expenditures in industry declined from 4.2 percent in 1968 to a trough of 2.6 percent in 1979 and then recovered to 3.7 percent by 1982. This pattern masks a strong divergence between the trends in federally and privately supported industrial R&D. Federally supported R&D fell from 2.1 percent of manufacturing sales in 1967 to 0.7 percent in 1979 and has only recently begun to recover, while company financed R&D stayed essentially constant (relative to industry sales) with almost all of the fluctuation coming from the decline in federal support (NSF 1984). During the same period, the economy experienced one of the sharpest and prolonged recessions of the post-war period. The recession of 1974-75 was triggered by the first OPEC increase in energy prices and resulted both in a large and pervasive productivity slowdown and large amounts of unanticipated inflation. Most hard hit were the primary metals, motor vehicles, and other heavy, energy related industries. On the whole, these were the less R&D intensive industries, resulting in a probably spurious relationship between R&D intensiveness and the productivity slowdown. (See Griliches 1980 and Griliches-Lichtenberg 1984 for more discussion of these issues.)

The remainder of the paper is organized as follows. First I describe the data set with its advantages and limitations and present some overall comparative statistics. Second, I outline briefly the framework that underlies the computations to be performed. The results are presented and discussed next and the paper closes with some caveats and suggestions for further research.

II. The Data

The current project is an extension of work originally begun in the mid 1960s. That work was based on the matching of R&D data collected on behalf of the NSF by the Bureau of the Census during 1957-1965 with additional company data from the 1958 and 1963 Census of Manufactures and Enterprise Statistics. The universe consisted of large (1000 or more employees) R&D performing U.S. manufacturing companies. The final sample of 883 of such companies accounted for over 90 percent of total sales and R&D expenditures of all firms in this universe. Because of the confidentiality of the individual data, the final output was in the form of matrices of correlation coefficients and standard deviations, broken down into six rather broad industrial groupings, with no access by us to the actual individual observations.

The main finding of that work (Griliches, 1980) was a rather consistent and positive relationship between various measures of company productivity and its investments in research and development. Cobb-Douglas type production functions, estimated on both levels (1963) and rates of growth (1957-65) yielded an elasticity of output with respect to R&D investments of about 27 percent (as of 1963), a significantly lower rate of return to federally financed R&D expenditures, and no clear evidence of significant scale effects either in R&D investment policies or the returns from it.

In trying to extend the earlier study to the more recent time period it became clear that the earlier work could not be simply updated, since the earlier project tapes has been blanked inadvertently in the interim. Also, it turned out that the 1958 and 1963 Census of Manufactures summaries could not be retrieved in machine-readable form. Luckily, most of the original R&D schedules could still be found, though they had to be repunched from scratch.

Thus, what started out as a simple update, became an almost entirely new data gathering and matching effort, of significantly larger dimension than originally anticipated. Its basic objective was to create a matched body of data on most of the large R&D performing corporation in the U.S., making it possible to analyze both the determinants and the consequences of R&D spending over time. For this purpose a time series record has been created for each company consisting of the major variables in the annual R&D survey for each of the years 1957-1977, supplementary R&D information for selected years (1962, 1967, 1972, and 1975), data from the Enterprise Statistics (NCK-1) for 1967, 1972, and 1977, and a few additional items from the Census of Manufactures establishment record summaries for 1967 and 1972.

The universe of this data match consists of all "certainty" cases in the 1972 R&D survey. I.e., the basic definition is the population of companies as they existed in 1972 (as against 1962 in the earlier study) and the requirement of "certainty" assures that the Census Bureau tried to collect consistent data for these firms for more than one year. The "certainty" cases correspond closely to the earlier restriction to companies with 1000 or more employees, though it is a bit more inclusive. There were approximately 1100 such companies in 1972. A "complete" record, however, exists only for a much smaller number of companies. A number of different matching efforts were involved: First, a company's R&D schedules had to be matched over time. A company, however, may not have existed over the whole period as an independent entity, or was not in the R&D Survey in some of the years. Second, separate matches had to be made to the Enterprise Statistics (NCK-1) and Census of Manufactures summaries in 1967, 1972, and 1977. Each of these matches could fail individually, both because the relevant records may not have been found, and because the definitions of a company on the different surveys may have been inconsistent (due to different rules of consolidation, treatment of foreign operations, etc.).

Given our interest in the analysis of productivity growth, our data can be reclassified into: 1. Output measures (sales annually from HRD, value added from Census of Manufactures for 1967, 1972, and 1977); 2. Employment mreasures (total employment annually from HRD, manufacturing employment from the Census in Census years); 3. Capital data (from Enterprise Statistics for Census years); and 4. R&D data (annually from HRD, with additional mix detail for 1962, 1967, 1972, and 1975). We have also added to the record price indexes for the delfation of sales and value added, at the 2 1/2 digit NSF recode detail (given in Table 1), derived from the BEA and BLS price indexes tapes by 4-digit and input-output detail, an R&D deflator based on the methodology suggested by S. Jaffe (NSF 1972 and Griliches 1984), and investment and capital stock deflators derived from various NIPA publications.[1]

Table 1 gives detail on the industrial composition of the panel and also some indication of the relative success of the various matching criteria. Roughly speaking, if one requires a good match for at least two Census years, the effective sample size is down to about 500 companies, though for a variety of cross sectional questions significantly larger sample sizes are feasible. Table 2 lists the number of firms with good R&D data by individual year, showing both the growth of the R&D collection effort over time and sample attrition in recent years due to merger activity and sample redefinition. Table 3 lists the means and variances for the major variables as of 1972. But before we look at these number in some detail I need to describe the model and the analytical framework that will be used for their analysis. This is the topic of the next section.

III. The Analytical and Econometric Framework[2]

The work reported here focuses primarily on the analysis of productivity growth for these companies, using a rather simple growth accounting approach which 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.

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 is simplifies to

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

where A is constant, λ 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). Alternatively, if one differentiates the above expression with respect to time and assumes that conventional inputs are paid their marginal products, one can reqrite it as

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

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 $[x = \dot{X}/X = (dX/dt)/X]$. Equation (5) is a constrained version of (4), with $\hat{\beta}$, (the observed factor share of capital) used to estimate the true β .

A number of serious difficulties arise when one turns to the operational construction of the various variables (see Griliches 1979 for more detailed discussion). 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 to the unobservable research capital measure (K). Postponing the first for later consideration, we note that $K_t = \Sigma w_i R_{t-1}$ can be thought of as a measure of the distributed lag effect of past research investments on productivity. There are 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 the 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). There is some scattered evidence, based largely on questionnaire studies that such lags are rather short in industry, where most of research expenditures are spent on development and applied topics, and where the private returns from R&D become obsolete much fasher due to the erosion of a firm's specific monopoly position (Pakes and Schankerman 1984).

Because of the difficulties in constructing an unambigous 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

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

allowing one to rewrite (5) as

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

where ρ 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 one to construct a measure of K in the first place. Note that in estimating (5) or (5') one assumes that either α or ρ are constant espectively across firms or industries. It is not clear, a priori, which is the better assumption.

While our models are written as if the main point of research expenditures is to increase the physical productivity of the firm's production 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 the dependent variable does in fact capture the private returns to such research endeavours. 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 firms 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 its research program. Moreover, some of the gains inproductivity 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. I will not be able to say much about the net impact of such forces on the basis of the data at hand. This would require 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 project. But since expected private returns are a determinant of private investment flows into this acticity, this work is of some interest even if it cannot answer the social returns question unequivocally.

One can also use this framework to ask whether different types of R&D (private versus federal, or basic versus applied) are equally "potent" in generating productivity growth. One way of answering this question is to look at the "mix" of R&D expenditures and ask if it matters for the question at hand. Let there be two types of R&D expenditures, R_1 and R_2, and let us assume that the overall analysis is in terms of the logarithm of total R&D expenditurs but that we believe that R_2 should have been weighted more, given a δ premium (or discount). That is, the right variable is

(6) $$R^* = R_1 + (1+\delta)R_2 = R(1+\delta s),$$

where $s = R_2/R$ is the "share" of R_2 in total $R = R_1+R_2$. Then the $\alpha \log R^*$ term can be approximated by $\alpha \log R^* \simeq \alpha \log R + \alpha\delta s$. The sign and significance of the "mix" term s, will give us some clue about the size and magnitude of the δ term.

A similar argument can be made also in the context of a growth-rate formulation. Let lower case letters denote growth rates. Then $r = (1-s)r_1 + sr_2$ while $r^* = (1-s)r_1 + (1+\delta)sr_2$. If, as is mostly the case in our data, the growth rates of r_1 and r_2 are roughly equal, then $r^* = r(1+\delta s)$, and again, the coefficient of the "mix" term s provides us with some information about the "premium" or "discount" on R_2 since αr^* can be approximated by

(7) $$\alpha r^* \simeq (\alpha+\delta\bar{s})r + (\alpha\bar{r}\delta)s$$

Given the pecularities of our data set -- its unbalanced nature (many missing observations towards the beginning and end of the period) the availability of capital and value added only for Census years, the desire to preserve comparability with the earlier study, and the difficulty of doing elaborate programming inside the Census Bureau, I focus primarily on two major dimensions of the data: levels (in 1967, 1972, and 1977) and growth rates, and eschew any attempt at a complete annual data analysis. The annual data are summarized by computing average growth rates for two subperiods 1957-65 (corresponding to the earlier study period) and 1966-77, based on regressions of the logarithms of the relevant variables on time trends (solving thereby the missing years problem within each of these subperiods).

In implementing such a framework of analysis one has to deal with several serious data problems: missing data, erroneous data and possible erroneous matches, and mergers. Except for R&D data, no special effort was made to replace missing values by various imputation procedures. It was our notion that the basic data set represents what the Census did collect, what we actually know, and that any imputation procedure should be done only in the context of a particular research project where its implications for the final analysis could be interpreted. As far as the R&D data are concerned, the Census used the shuttle nature of the original questionnaires to fill in many of the original blanks. To the extent that there remain missing values which are not due to the fact that the whole company is missing before or after some date, they were interpolated on the basis of the estimated growth rates (which require at least five good data points within each sub-period). For other variables, missing values were not imputed. It was not possible, within the constraints of this project, to develop optimal imputation procedures. This would have required several repeated passes at the original numbers. Instead, the analysis is based either on reduced "clean" samples or on "pairwise present" correlation coefficient matrices.

From an econometric point of view, we have to deal with the problem of firm effects (or firm specific left-out variables) and the possibility that the relationships we are trying to estimate may not stay constant either across firms or across time. The first is handled by analyzing first differences or growth rates, transformations that eliminate any unchanging effects from the model. The second problem, the problem of differences across firms, is handled in part by calculating a measure of "partial" productivity growth [$BPT = y - (1 - \hat{\beta})\ell$], using individual firm data on the share of labor in total costs. One can also estimate separate and different parameters for the various industry groupings and include some of the other variables available in the record which might distinguish one firm's environment and response pattern from another's (such as its specialization ratio, size, or vertical integration). The main hypothesis under investigation, that the returns to R&D investments may have declined over time, is tested both by comparing estimates based on the more recent data with the earlier results, and by allowing and testing for systematic changes in the estimated relationships between the three available cross-sections.

IV. Major Results

Before I present some of the preliminary results derived from this data base it may be useful to review both its structure and the main outlines of what happened in the 1970's. Table 3 is intended to describe three aspects of these data: 1. The general characteristics (means and standard deviations) of the sample as of 1972; 2. average rates of growth of the major variables of interest during the 1967-77 period; and 3. how these measures change when the sample is changed to select observations according to the availability of the requisite information.

Turning to the last topic first, note that we tend to lose smaller and more R&D intensive firms as the sample gets more restrictive. Matrix 1 corresponds to the most liberal criterion: a firm had to exist in 1972 and report positive R&D to be included in this sample. Matrix 2 requires both the ability to compute a growth rate for the 1967-77 period (i.e., at least five good time series observations) and a successful match to the 1972 Census of Enterprise data (NCK-1). Matrix 3 adds to this also the requirement of a successful match to the 1977 Census data while Matrix 4 asks for a match with the 1967 Census instead of the 1977 one. Most of the difference occurs in the transition from Matrix 1 to Matrix 2 where trying to match to the Census we lose a relatively large number of smaller firms for which there are still data on the R&D Survey files. The firms that can be also found in 1977 are slightly larger and have had a somewhat higher rate of growth in employment, R&D, and productivity. The firms that also existed in 1967 are even larger but have on average grown somewhat more slowly than those that exited only in the 1972-77 period. If we look at two of our major variables or interest, partial productivity growth and the ratio of basic to total R&D, there is almost no difference in their means across the relevant matrixes (2, 3, and 4) and hence it is unlikely that subsequent conclusions will be subject to a serious sample selection bias. I will, therefore, ignore this topic for the purposes of further discussion here.

Looking at the levels of the variables in 1972 we see that the average firm in the sample is quite large with 5000+ employees: it employs close to a hundred R&D scientists and engineers, and is making only a relatively modest investment of its own money (about 2.5 percent of sales) in R&D, with very little of that, less than 3 percent, being devoted to Basic R&D.[3] This picture is somewhat misleading, however. The actual distribution of firms is quite skewed, with a small number of larger firms spending much larger amounts on both total and basic R&D. Looking at growth rates one can observe that on average these firms grew only moderately during this period: about 1 percent per year in total employment, about 2.5 percent per year in partial productivity, and almost zero growth in deflated R&D expenditures (though a slightly positive rate of growth in the number of R&D scientists and engineers). Here again, while on average there is little movement, there is a great deal of variability at the individual firm level. The standard deviations of the rates of growth of partial productivity and total R&D are 3.5 and 8 percent per annum respectively, with many firms growing much faster (and also much slower) than the average.

Looking at some of the R&D ratios over time, not reported in Table 3, one cannot see any significant decline in the rate of private investment in R&D. While the total R&D to Sales ratio falls from .042 in 1962 to .035 in 1972 (in matrix 4) and again from .032 in 1972 to .029 in 1977 (matrix 3), the CRS (company financed R&D to sales) ratios are essentially unchanged (.025 in 1962 and 1972 in matrix 4 and .023 in 1972 and 1977 in matrix 3.) On the other hand, while the BR ratio fell only modestly, from .033 to .031 between 1962 and 1972 (in matrix 4) and from .027 to .023 between 1972 and 1977 (in matrix 3), coupled with the decline in the overall total R&D to Sales ratio this implies about a 40 percent reduction in the relative intensity of industrial investment in basic research, relative to industry sales. Almost all of this decline came from the overall decline in federally financed R&D which declined from about 55 percent of total R&D in industry in 1965 to about 35 percent in 1982. The federal government financed about 32 percent of all basic research in industry in 1967 but only 19 percent in 1982 (see NSF, 84-311). The reduction was so steep that basic research in industry declined not only relative (to sales) but also absolutely, from a peak of $813 million in 1966 (in 1977 dollars) to a trough of $581 million in 1975 and did not surpass the 1960's levels until the early 80's. How one interprets the consequences of such declines depends on one's view of the relative productivity of governmentally financed R&D expenditures in industry, a topic I will be exploring below.

Let us look now at the first set of substantive results. Table 4 reports the results of estimating cross-sectional production functions (eq. 4) separately for each Census year, adding to the standard capital and labor variables a measure of total R&D capital accumulated by the firm and two R&D "mix" variables: the fraction of total R&D that was spent on basic research and the fraction of accumulated R&D which had been financed privately. All the reported estimates allow for 18 to 20 (depending on the matrix) separate industry intercepts. The first and last two columns report estimates which are based on the same number of firms and use the same dependent variables, differing only by the year of observation. The two middle columns present additional estimates for 1972 based on different sample and dependent variable definitions with the main intent being to show that the major conclusions are insensitive to such differences. There are three major points to be made about these estimates:

1. The stock of R&D capital contributes significantly to the explanation of

cross-sectional differences in productivity and there is little evidence of a decline in its coefficient over time.[3] There is a minor rise in the estimated coefficient from 1967 to 1972 and a somewhat larger but not really significant decline from 1972 to 1977. Given this particular measure of R&D capital, based on a 15 percent per year declining balance depreciation formula (the results are insensitive to the particular formula used), the implied average (at the geometric mean of the sample) gross rate of return to R&D investment rises in a similar fashion from .51 in 1967 to .62 in 1972 (in matrix 4) and falls from .39 in 1972 to .33 in 1977 (in matrix 3). In either case the estimated rate of return is quite high and there does not appear to be any dramatic fall in it over time.

The second major finding is the significance and rather large size of the basic research coefficient. It seems to be the case that firms that spend a larger fraction of their R&D on basic research are more productive, have a higher level of output relative to their other measured inputs, including R&D capital, and that this effect has been relatively constant over time. If anything, it has risen rather than fallen over time. Using the formulation of equation (6) in Section III implies a very high premium on basic versus the rest, a $\hat{\delta}$ of between 2.5 to 4.5, a several hundred percent premium on basic research. Before I explore the implications of this result, I want to examine some other dimensions of these data and see whether similar effects can be observed there too.

The last major result of interest is the significant positive coefficient on the privately versus federally financed R&D mix variable. This variable is of most import for the older more established firms in matrix 4, but its sign is consistent throughout, indicating a positive premium on privately financed R&D, or equivalently a discount as far as federally financed expenditures are concerned. Here the implied premium is smaller, between 50 and 180 percent, but still quite large.

All the above results were based on cross-sectional level regressions which are subject to a variety of biases, the main one being the possibility that "rich" successful firms are both more productive and can afford to spend more of their own money on such luxuries as R&D and especially the basic variety. One can reduce somewhat the possibility of this type of bias by focusing on firm growth rates, the changes that occurred, rather than on their levels. To the extent that firms have idiosyncratic productivity coefficients which may be also correlated with their accumulated R&D levels, considering growth rates is equivalent to doing a "within" firms analysis, one that eliminates such "fixed effects" from the analysis. The next two tables present, therefore, the results of analyzing the growth in the partial productivity of these same firms during the whole 1966-77 period.

Table 5 presents the results of estimating partial productivity equations in the largest possible sample (matrix 6) and in the sample with a successful 1972 Census match (matrix 2). Here again we find our three main results confirmed: the R&D growth term and the two mix variables: the basic research ratio and the fraction of research financed privately all contribute significantly to the explanation of productivity growth.

On the assumption that the growth rate in the stock of R&D is roughly proportional to the growth in deflated R&D itself, the coefficient of BTRD should be estimating the same number as the coefficient of the R&D stock variable in Table 4. The results are in fact surprisingly close: about .12

in Table 5 as against .09 to .17 in Table 4. Moreover, there seems to have been no decline in this coefficient relative to the earlier 1957-65 period. In the previous study (Griliches 1980) I estimated the same coefficient to be .073. In the current replication and extension of this sample a similar equation for 1957-65 yields a BTRD coefficient of .086. Thus, if anything, the coefficient of R&D went up between the early 60's and the early 70's.

The second major finding of interest is the positive and significant basic research coefficient. It is hard to interpret its magnitude since the approximation outlined in Section III breaks down when the average growth rate of deflated R&D and of basic R&D is close to zero or negative. Consider, however, the following illustrative calculation. Raising the BR ratio by one standard deviation, from .026 to .097 at the mean, would increase the rate of growth of partial productivity by close to half a percent per year (.071 x .059 = .0042). This same increase would raise the growth of total R&D by .107 for one year and would contribute a once and for all increase in the level of productivity of .0125. Discounting the more "permanent" effect of basic research by a real interest rate of .05 yields an "equivalent" one year effect of .084, or a 7 to 1 ratio in favor of basic research! If one allows for industry dummies which in this formulation represent separate industry trends of disembodied technical change, the effect of basic research is cut by about 50 percent, implying perhaps that a significant fraction of the estimated effect comes from spillovers which diffuse throughout the industry. Note that it is the only coefficient that is affected substantively when separate industry dummies are allowed for. Nevertheless, even a 3.2 to 1 ratio is quite high!

The third finding is the significant positive premium on company financed R&D. Here too the impled premia are quite high, but given that the mix variable is defined in terms of stocks rather than flows the calculations are more cumbersome. Consider starting from a zero growth position and a .7 ratio of private to total R&D stock. To move this fraction from .7 to .75 we would need to raise the private stock by 29 percent and the overall stock by 20 percent (without reducing absolutely the stock of federally financed R&D capital). There are different possible investment paths that would achieve this goal and would have somewhat different present value consequences. If one roughly doubled the rate of privately financed R&D expenditures, from the previous replacement level of .105 (.7 x .15) to .205, one could achieve this target in slightly over two years. Ignoring discounting, this would lead to a once and for all growth in productivity of .024, due to the growth in the total stock of R&D <u>and</u> a .0011 permanent increase in the rate of growth due to the shift of the fraction private ratio from .7 to .75. The present value of this second term is about.022, or of the same order of magnitude as the first term. That is, raising the stock of R&D by 20 percent but shifting it all into the private component doubles the effect of such dollars.

There are problems, however, with such an interpretation. If private R&D expenditures contribute more to productivity growth, one might have thought that when they are substituted for the total R&D growth measure they might fit better and also have a higher coefficient. But that is not the case as can be seen from the results presented in columns 2 and 5 of Table 4. The total R&D measure does a little bit better both in terms of fit and in the overall size of its coefficient, implying that the contribution of federal dollars is not zero. That is perhaps what one should expect. Most of the direct output of federal research dollars is "sold" back to the government at "cost plus" and is unlikely to show up as an increase in the firm's own productivity. Thus all that one could expect to measure here are the within firm spillover effects of such expenditures. What we may be detecting is that such effects are indeed present and positive but we should not have expected them to be of the same order of magnitude as would be the case for the firm's own investments in improving its productivity or profitability.

There are a number of eocnometric questions that can be raised about the robustness and sensitivity of such results. I will discuss only a few of these here. The most obvious question arises from the fact that even though I allowed, in the growth rates version, for separate firm intercepts and different industry trends, I am still assuming common R&D and conventional capital coefficients across rather different industries. This is done from necessity rather than as a virtue. Estimating the same models industry by industry reduces the sample sizes drastically and raises greatly the relative noise level, making it rather hard to interpret the resulting estimates. Nevertheless, these estimates, which are summarized in Table 6, are quite consistent with the earlier story: 17 out of the separately estimated 19 coefficients for the R&D growth variable are positive and more than half of them are statistically significant at conventional significance levels. Similarly, the coefficients of the basic research ratio variable are positive in 14 of our 19 industries and significant in over a third of them. The fraction private variable is less robust to the division of the sample into industries, with more than half of the coefficients still positive, but only four of them are statistically significant within particular industries. Two of these industries are indeed the ones where one would expect to find such an effect, aircraft and electronics, industries where the bulk of federal monies is spent. Nevertheless, it seems that the effect that is being caught by the fraction private variable has an important industry component, something that had been already noted in the earlier study (Griliches 1980), as does also the effect associated with the basic research variable, though to a lesser extent.

A number of other versions were computed using the growth in capital services rather than the depreciation and age composition variables which had been used to keep the results comparable to the earlier study, and the growth in R&D "capital" rather than the flow (and also different definitions of such "capital"). I also estimated versions using the "intensity" form for the R&D variable, to make it more comparable to other studies in the literature (Griliches-Lichetenberg 1984, Mansfield 1980, and others). By and large the results of these alternatives were somewhat weaker but not substantively different. Perhaps the most interesting alternative estimate is the intensity version using the growth of capital between 1967 and 72 as its capital measure:

$$\underset{}{BPT6677} = \ldots \quad \underset{(.069)}{.243ACRS} + \underset{(.024)}{.045\ ABR} + \underset{(.130)}{.180DLCS} \qquad \text{See} = .0316 \ \text{(Matrix 4)}$$

where ACRS is the average company R&D to sales ratio, averaged over 1967 and 1972. ABR is a similar average basic to total R&D ratio, and DLCS is the rate of growth in deflated capital services between 1967 and 1972. This version is closest in form to the kind of equations estimated by Mansfield (1980) on much smaller samples. The basic results are similar, however. Basic R&D is a significant contributor to productivity growth with an implied basic to company premium of about 5 to 1 (given an average R&D to sales ratio of .035).

The final set of results to be presented here, in Table 7, relate to the relative profitability of our firms in 1972 and 1977. The dependent variable, GRR, is the ratio of gross profits (value added minus labor costs and plus R&D) to total gross fixed assets. The independent variables include the ratio of R&D capital (undepreciated) to total fixed assets and our ubiquitous R&D mix variables: the basic research and fraction private ratios. Even though the dependent variable is quite different the overall results are rather similar to the earlier ones. The R&D capital variable is positive and almost always statistically significant though its coefficient is a bit low if it is to be interpreted as a rate of return to it. The basic research variable is both large and significant though possibly too large to be credible. Given that the ratio of total R&D capital to total fixed assets is only about .05 on average in 1972 the estimated coefficients imply a δ of about 30 to 60. The fraction private ratio also contributes positively to profitability but its effect largely disappears once industry differences are allowed for. The results for 1977 are weaker than those for 1972, the residual variance is significantly higher, but they too suggest the importance of basic research even in this context. A similar analysis was performed using an estimate of the net rate of return as the dependent variable, subtracting depreciation from the numerator of GRR and using a net stock concept for the denominator and also in the definition of the R&D capital variable. While the fit was significantly worse when using this definition of the dependent variable, the overall results were rather similar. The net return version was also available for 1967 and the results using it indicate a relatively constant and significant coefficient for the basic research ratio while the coefficient of the total R&D stock rises from 1967 to 1972 and than falls again in 1977 (from .11 to .16 and down to .06). It is doubtful whether these fluctuations represent real trends or, more likely, reflect the larger noise level in the 1977 data and the changing composition of these samples. In any case, the profitability regressions are consistent with the productivity level and productivity growth rate based results described earlier (Tables 4, 5, and 6).

V. Discussion and Summary

The three major findings of this paper: that R&D contributed positively to productivity growth and seems to have earned a relatively high rate of return, that basic research appears to be more important as a productivity determinant than other types of R&D, and that privately financed R&D expenditures are more effective, at the firm level, than federally financed ones, are not entirely new or very surprising. The first finding has been documented in a number of earlier studies (see Griliches 1980, Griliches and Mairesse 1984, Link 1981a, and others). What is new in this paper in this regard is a confirmation of this finding on a much larger and more recent data set. It also presents evidence for the view that this effect has not declined significantly in recent years, in spite of the overall slowdown in productivity growth and the general worry about a possible exhaustion of technological opportunities.

The evidence for a "premium" on basic research is much more scarce. The major previous paper suggesting this type of a result is Mansfield (1980) which uses aggregate data for 20 industries for 1948-66 and data for 16 firms during 1960-76 and finds a significant premium on basic research, on the order of 2 to 1 at the industry level and 16 to 1 at the firm level. (See also Link 1981b for similar results for 1973-78 based on data for 55 firms.) In this paper I get similar though somewhat smaller effects at the firm level, using a much larger and more representative sample. I also find that differences in levels of productivity and profitability are related to differences in the basic research intensity of firms.

Such findings are always subject to a variety of econometric and substantive reservations. In this context the two major related issues are simultaneity and the question how can major divergences in private rates of return persist for such long periods. It is possible to argue that it is not R&D, or its basic research component, the casues firm "success" as measured by productivity and profitability but rather that success allows firms to indulge in these types of luxury pursuits. It is difficult to argue about causality on the basis of what are essentially correlational data. It is possible to use simultaneous equation techniques to estimate such models but then the argument shifts to the validty of the exogeneity assumption for the particular instruments. In the context of our specific data set it is hard to think of any valid instruments except for possibly lagged values of the same variables, which raises some problems of its own. The best evidence for the notion that these results are not entirely spurious is provided by the growth rates where the individual firm levels are partialed out of the analysis. But, here too, one could argue about the impact of common unanticipated "luck" elements. Unfortunately, it is unlikely that one could use lagged growth rates as instruments since there is very little correlation in growth rates over time at the firm level. While an attempt will be made in further work with these data to estimate more extended simultaneous equations versions of such models, I am not too optimistic as to what can be accomplished in this regard. The evidence presented here should not be interpreted as "proving" that R&D, and expecially its basic component, are important for productivity growth but rather as presenting some prima facia evidence in support of such an interpretation. In this sense it is an exercise in economic rhetoric (McCloskey 1983).

It is even more difficult to respond to the theoretical a priori argument that such results cannot be true since they imply widely differing rates of return to different activities under the control of the same firms. One's response to this depends on one's views as to the prevalence of equilibria in the economy. While it is likely that major divergences in rates of return are eliminated or reduced in the long run, the relevant runs can be quite long. R&D as a major component of firm activity was undergoing a diffusion process in the 1950's and 1960's and may not have reached full equilibrium even by the end of our period. This may be especially true of the basic research component where the risks are much greater and the uncertainty introduced by changing government policies and the changing economic environment make it quite difficult to decide what is the right level for it.

A somewhat different version of this argument would claim that the world is indeed in approximate equilibrium but that different firms face different opportunities for doing research, basic or otherwise, are in different ecological niches, and hence have different coefficients in their "production functions." This would explain why different firms are observed to spend different amounts on R&D while actually earning about the same rate of return on it. When a constant coefficients production function is fit to such data it will fit because it is approximating a market equilibrium relation. If the level of R&D invested were independent of the coefficient then such a function would just reproduce its average share and not produce any evidence of excess returns. But if, as is reasonable, R&D is invested optimally with firms which have better opportunities, higher coefficients, investing more, this will induce a positive correlation between R&D and its individual coefficient and lead to an upward bias in the estimated "average" coefficient.[4] The resulting "larger" coefficient, larger than the observed factor share, will be interpreted, wrongly, as implying a higher rate of return than is actually prevailing at the individual level.

This argument may be recognized as a version of the earlier attacks on the Cobb-Douglas production function combined with a random coefficients interpretation of the same phenomenon. In its extreme form it is testable. Since there are time series data available for individual firms one could try to estimate individual firm parameters and check whether they are in fact distributed as is predicted by this particular argument. While individual parameters are unlikely to be well estimated, given the relative shortness of the available times series, the parameters of the distribution of such coefficients might be estimable with more precision. I intend to pursue this possibility in further work.

To restate again the major points of the paper: A newly available body of data on all the major R&D doing firms in the U.S. has been examined and evidence has been presented for the proposition that R&D contributes significantly to productivity growth, that the basic research component of it does so even more strongly, and that privately financed R&D expenditures have a significantly larger effect on private productivity and profitability than federally financed R&D. These findings are open to a number of reservations. Nevertheless, they do raise the issue that the overall slowdown in the growth of R&D and the absolute decline in basic research in industry which occurred in the 1970's may turn out to have been very costly to the economy in terms of foregone growth opportunities.

Table 1

CENSUS-NSF-Griliches R&D Sample by Industry

Industry	SIC	Number of Firms in Each Matrix* 1	6	5	2	3	4
Food	20	69	68	53	52	38	
Textiles	22,23	42	40	in Misc.	29	26	
Petroleum	29	27	26	18	20	-	
Stone, Clay, & Glass	32	24	23	19	20		45
Non-manufacturing	1-19,40-63,73,80-89	94	89	41	21	50	
Miscellaneous Mfg.	21,27,31,39	73	50	49	30		
Lumber, Wood & Paper	24,25,26	59	56	41	44	31	31
Industrial Chemicals	281-2	48	45	39	49	33	37
Drugs	283	31	29	23			
Other Chem. (Agric., Paint)	284-9	40	39	29	28	23	33
Rubber	30	27	26	18	22	18	
Primary Metals	33	62	59	51	39	29	30
Fabricated Metals	34	56	54	35	36	29	21
Machinery	35 except 357	138	136	106	85	68	55
Electronics	365,366,367	81	92	65	46	35	32
Computers	357	18					
Electrical Equipment	361-4,369	67	63	52	45	31	37
Motor Vehicles	371,373-5,379	36	36	31	28	37	34
Aircraft	372,376	41	34	27	21		
Scientific Instruments	38	72	69	48	46	36	31
Total		1105	991	745	652	491	386

* The samples in each of the six matrices are defined as follows:

1. All firms in our universe.
6. Growth rates of the R&D variables computable for 1966-1977.
5. Growth rates of the R&D variables computable for 1957-1965.
2. Growth rates for 1966 to 1977 and a successful match to the 1972 NCK-1 survey.
3. Growth rates for 1966 to 1977 and successful matches to both the 1972 and the 1977 NCK-1 surveys.
4. Growth rates for 1957 to 65 and 1966 to 1977, and successful matches to the 1967 and 1972 NCK-1 surveys.

Table 2: Number of Firms in Panel Reporting Total R&D Expenditures, by year, 1957-1977

YEAR	NUMBER
1957	671
1958	727
1959	750
1960	745
1961	800
1962	846
1963	834
1964	858
1965	868
1966	871
1967	1000
1968	1002
1969	1013
1970	1063
1971	1076
1972	1079
1973	1060
1974	1030
1975	876
1976	875
1977	801

Table 3: Major Variables in 1972 and 1966-77 Growth Rates by Subsample. Means and Standard Deviations (in parentheses)*

Variable	Matrix Number and Sample Size 1 1105	2 652	3 491	4 386
A. Levels in 1972				
Sales in million dollars	146 (1.61)	205 (1.43)	223 (1.40)	236 (1.44)
Total Employment	4038 (1.48)	5570 (1.27)	6212 (1.30)	6698 (1.31)
R&D Scientists and Engineers	89 (1.66)	74 (1.70)	82 (1.71)	106 (1.72)
R&D in million dollars	2.3 (1.74)	3.0 (1.78)	3.4 (1.77)	4.3 (1.83)
R&D to Sales Ratio	.051 (.131)	.033 (.064)	.032 (.051)	.035 (.048)
Company R&D/Sales Ratio	.028 (.069)	.022 (.026)	.023 (.026)	.025 (.026)
Basic to Total R&D Ratio	.025 (.074)	.026 (.071)	.026 (.075)	.027 (.073)
Value Added, million dollars		100 (1.32)	113 (1.31)	121 (1.34)
Gross Fixed Assets million dollars		115 (1.67)	124 (1.59)	147 (1.65)
B. Growth Rates 1966-77				
Employment growth		.012 (.046)	.015 (.041)	.006 (.040)
Partial Productivity growth		.025 (.036)	.026 (.034)	.025 (.035)
Total R&D growth		-.001 (.079)	.003 (.074)	-.007 (.070)
Scientists and Engineers growth		.008 (.087)	.012 (.084)	.004 (.078)
Company R&D growth		.004 (.081)	.008 (.076)	-.000 (.071)

*Geometric means and standard deviations of the logarithms (approximate coefficient of variation) except for growth rates or ratios.

Table 4: NSF-Census Study: Cross-Sectional Production Functions Log Value Added Dependent,[1] U.S. Firms, 1967, 1972 and 1977. Coefficients (standard errors)

Variables	1967	1972	1972	1972[2]	1972[2]	1977[2]
L. Empl (nr)	.604 (.045)	.622 (.046)	.623 (.035)	.586 (.038)	.578 (.038)	.611 (.039)
L Capit. Serv.	.224 (.041)	.199 (.044)	.161 (.032)	.234 (.036)	.254 (.036)	.291 (.035)
L T R&D Stk. (db)	.113 (.023)	.135 (.026)	.165 (.019)	.126 (.019)	.115 (.018)	.089 (.017)
B/R	.396 (.240)	.340 (.261)	.274 (.215)	.499 (.191)	.517 (.189)	.401 (.189)
FP	.190 (.097)	.247 (.106)	.068 (.100)	.133 (.088)	.138 (.088)	.044 (.084)
N	386	386	652	491	491	491
SEE	.312	.336	.390	.312	.309	.290

[1]Value added and materials used in research in 1967 and 1972.

[2]Value added only.

L Employment - log (total employment--employment of scientists and engineers)
L Capit. Serv. - log of (depreciation plus interest on net assets plus machinery and equipment rentals)
L T R&D (db) - log of the "stock" of total R&D expenditures based on a 15 percent per year declining balance depreciation assumption.
B/R - basic research as a fraction of total R&D. 1972 in the 1977 equation, 1967 in 1967 and 1972.
FP - fraction of R&D stock "private", company financed R&D stock as a ratio to the total R&D stock, as of t.

All equations include also a constant term and industry dummies. The number of industry dummies used depends on matrix. In 1967-72, Matrix 4, 18 industries. In 1977, Matrix 3, 19 industries.

Table 5: Growth Rate of Partial Productivity, 1966-77

Different Estimates

Variables	Matrix 6 N=911		Matrix 2 N=652		
Const.	.019	.009	.012	with industry dummies	
BTRD 6677	.107 (.014)		.117 (.017)	.119 (.016)	
BCRD 6677		.095 (.014)			.106 (.015)
BR 72	.056 (.017)	.056 (.017)	.059 (.019)	.035 (.018)	.034 (.018)
FP 72	.011 (.005)	.019 (.005)	.017 (.006)	.022 (.007)	.030 (.007)
SEE	.0383	.0384	.0337	.0305	.0307

Dep. variable: BPT 6677 = trend growth rate of deflated sales minus the trend growth of total employment multiplied by the share of payroll in total sales.

BTRD - trend growth of deflated total R&D expenditures
BCRD - same for company financed R&D expenditures
B/R - basic research expenditures as of fraction of total research expenditures.
FP - ratio of company financed R&D stock to total.
SEE - residual standard error.

All equations contain also a term reflecting the variance of R&D and terms representing the growth of physical capital: age composition and depreciation as of 1972.

Table 6: Growth Rate of Partial Productivity,

by industry, 1966-77. Matrix 6, Total N = 991

Coefficients of	Coefficients by the estimated t-ratio			
	<-1.5	-1.5-0	0-1.5	1.5+
BTRD		2	7	10:Misc, Ind. Chem, Drugs, St. & Gl., Mach., Electron., Elect. Eq., Transp. Eq., Sc. Inst., Non Mfg.
BR72		5	8	6:Wood & P, Oth. Chem, Oil, Mach., Aircraft, Non Mfg.
FP72	2	6	7	4:Oil, Rubber, Electron., Aircraft

Other variables in equation: Const. ACOMP72, DEPR72, STRD6677.

Table 7: Gross Profit Rate Regressions:

GPR = (Value added - Payrolls + R&D)/(Gross Assets)

Dept. Variable matrix and sample size		Const.	Coefficients of R&D capital to Total Fixed Assets Ratio	Basic R&D Ratio	Fraction Private	SEE
<u>GRR 72</u>						
2: N = 652	a	.144 (.049)	.088 (.012)	.344 (.144)	.107 (.048)	.262
	b		.060 (.013)	.187 (.138)	-.012 (.052)	.237
3: N = 491	a	.117 (.052)	.080 (.013)	.514 (.139)	.154 (.051)	.264
	b		.061 (.015)	.366 (.138)	.074 (.057)	.227
<u>GRR 77</u>						
3: N = 491	a	.341 (.064)	.031 (.019)	.402 (.187)	.033 (.068)	.313
	b		.004 (.022)	.261 (.187)	-.028 (.077)	.292

b regressions contain industry dummies.

FOOTNOTES

*Harvard University and the NBER. I am indebted to NSF grant No. SES82-08006 for the support of this work, to Doug Dobos and Bronwyn H. Hall for making the data gathering effort possible and to David Body for research assistance. I have also benefited from comments of seminar participants at the NBER and Yale.

1. See Griliches-Hall, 1982, and Hall 1984 for more details.
2. This section borrows heavily from Griliches 1974.
3. Here and subsequently, all statements about statistical "significance" should not be taken literally. Besides the usual issue of data mining clouding their interpretation, the "samples" analyzed come close to covering completely the relevant population. Tests of significance are used here as a metric for discussing the relative fit of different versions of the model. In each case, the actual magnitude of the estimated coefficients is of more interest than their precise "statistical significance."
4. A positive correlation is not enough, but itself, for a positive bias. The weight of an individual firm slope coefficient in the cross-sectional estimate is proportional to the square of the deviation of R&D stock from its mean. A positive correlation between levels does not translate itself directly into a positive correlation between the level of one variable and the square of the other, except for certain skewed distributions. Since we do not observe the individual coefficients directly, it is rather difficult to check out this conjecture.

References

Bound, J., C. Cummins, Z. Griliches, B. Hall, A. Jaffe, 1984, "Who Does R&D and Who Patents?", in Z. Griliches (ed.), R&D, Patents, and Productivity, Chicago: University of Chicago Press, pp. 21-54

Griliches, Zvi, 1979, "Issues in Assessing the Contribution of R&D to Productivity Growth," The Bell Journal of Economics, 10(1), Spring, pp. 92-116.

______________, 1980a,"Returns to Research and Development Expenditures in the Private Sector," In J.W. Kendrick and B. Vaccara (eds.), New Developments in Productivity Measurement, NBER, Studies in Income and Wealth, 44, pp. 419-454, Chicago: University of Chicago Press.

______________, 1980b,"R&D and The Productivity Slowdown," American Economic Review, 70(2), pp. 343-348.

______________, 1984, "Comment" on Edwin Mansfield, "R and D and Innovation: Some Empirical Findings," in Zvi Griliches (ed.), R&D, Patents and Productivity, Chicago: University of Chicago Press.

Griliches, Zvi and B.H. Hall, 1982, " Census-NSF R&D Data Match Project: A Progress Report," in Development and Use of Longitudinal Establishment Data, U.S. Dept. of Commerce, Bureau of the Census, Economic Research Report, ER-4, pp. 51-68.

Griliches, Zvi and F. Lichtenberg, 1984, "R&D and Productivity Growth at the Industry Level: Is There Still A Relationship?", in Zvi Griliches (ed.), R&D, Patents and Productivity, Chicago: University of Chicago Press, 465-496.

Griliches, Zvi and J. Mairesse, 1983, "Comparing Productivity Growth: An Exploration of French and U.S. Industrial and Firm Data," European Economic Review, 21, pp. 89-119.

Griliches, Zvi and J. Mairesse, 1984, "Productivity and R&D at the Firm Level," in Zvi Griliches (ed.) R&D, Patents and Productivity, Chicago: University of Chicago Press, pp. 339-374.

Hall, Bronwyn H., "Historical R&D Panel: 1957-77. Public Use Correlation, Matrices-Tape Documentation," September 1984.

Link, A.N., 1981a, Research and Development Activity in U.S. Manufacturing, New York: Proeger.

__________, 1981b, "Basic Research and Productivity Increase in Manufacturing: Additional Evidence," American Economic Review, 71(5), December, pp. 1111-1112.

Mansfield, Edwin, 1980, "Basic Research and Productivity Increase in Manufacturing," American Economic Review 70(5), December, pp. 863-873.

McCloskey, Donald N., 1983. "The Rethoric of Economics," Journal of Economic Literature, XXI(22), pp. 481-517.

National Science Foundation, 1972, Jaffe, S.A., "A Price Index for Deflation of Academic R&D Expenditures," 72-130, Government Printing Office.

____________________________, 1974, National Patterns of Science and Technology Resources, NSF 84-311, Washington, D.C.: Government Printing Office.

Pakes, Ariel and M. Schankerman, 1984, "The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources," in Zvi Griliches (ed.), R&D, Patents and Productivity, Chicago: University of Chicago Press., pp. 209-232.