NBER WORKING PAPER SERIES

THE IMPACT OF R&D INVESTMENT ON PRODUCTIVITY - NEW EVIDENCE USING LINKED R&D-LRD DATA

Frank R. Lichtenberg

Donald Siegel

Working Paper No. 2901

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1989

Lichtenberg is an Associate Professor at Columbia University Graduate School of Business and a Research Associate of the National Bureau of Economic Research. Siegel is a Research Economist of the National Bureau of Economic Research. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #2901
March 1989

THE IMPACT OF R&D INVESTMENT ON PRODUCTIVITY - NEW EVIDENCE USING LINKED R&D-LRD DATA

ABSTRACT

This paper uses confidential Census longitudinal microdata to examine the association between R&D and productivity for the period 1972-1985. These data allow for significant improvements in measurement and model specification, yielding more precise estimates of the returns to R&D. Our results confirm the findings of existing studies:

1) positive returns to R&D investment
2) higher returns to company-financed research
3) a productivity "premium" on basic research

These results are robust to our attempts to adjust for "influential" outliers. Also, it appears that the return to company-financed R&D (but not total R&D) is an increasing function of firm size.

Frank R. Lichtenberg
Graduate School of Business
Columbia University
726 Uris Hall
New York, NY 10027

Donald Siegel
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

The Impact of R&D Investment on Productivity-

New Evidence Using Linked R&D-LRD Data

Introduction

Economists have identified R&D investment as an important determinant of productivity growth. The objective of this paper is to examine the association between R&D and productivity growth using the most comprehensive and accurate longitudinal microdata yet available for productivity analysis. These data allow for improvements in measurement and model specification, yielding more efficient estimates of the effects of R&D investment on productivity.

Our empirical investigation of the strength of the R&D-productivity connection is based on estimation of reduced-form equations derived from the R&D Capital Stock Model (Griliches 1979), which asserts that the stock of a firm's technical knowledge or its knowledge capital is itself a factor of production. If the rate of depreciation of knowledge capital is assumed to be negligible, TFP growth is a function of the "intensity" of R&D investment, which is usually measured as the ratio of R&D expenditure to sales:

$$(1)\ DTFP_t = \alpha + \beta RDINT_t + u_t$$

where u_t is a classical disturbance term. The coefficient on R&D intensity (β) is interpreted as a marginal product or "rate of return" to R&D investment. Point estimates at the company or line of business level are interpreted as measures of the private rate of return to R&D, or those that accrue to the firm or its investors. Due to incomplete appropriability or other factors, this return may not equal the social return to this activity.

This paper discusses and adjusts for the limitations of existing micro-level empirical studies, which have provided estimates of the private rate of return to R&D. While these studies have been useful, productivity estimates at the firm or line of business level contained therein have been based on crude and incomplete measures of output and inputs. The most serious measurement problem has been an inability to control for diversification when calculating firms' productivity. More specifically, productivity estimation in these studies has often been based on the assumption that firms operate in only one line of business (4-digit SIC industry). To the extent that the relative prices of firms' outputs and inputs vary across industries, this approach introduces an element of noise into estimation of TFP growth (DTFP). Even if it is uncorrelated with R&D intensity, this measurement error will reduce the efficiency of estimates of β. We demonstrate that measures of TFP growth based on linked Census R&D-LRD Data are more precise, because we can calculate firms' real output and input at the 4-digit SIC industry level and aggregate to the firm level. As expected from the standard errors-in-variables model, this reduction in measurement error from improved deflation yields more efficient estimates of the rate of return to R&D.

An empirical example is presented that illustrates the extent to which more precise TFP measurement leads to more efficient estimates of rates of return to R&D. This example is derived from a pilot study performed by the authors (Lichtenberg-Siegel (1988)), based on linked R&D-LRD data for a sample of 115 firms for the years 1972-1980. Having demonstrated the desirability and feasibility of using linked Census R&D-LRD data to examine the R&D-productivity connection, a comprehensive analysis of rates of return to R&D is undertaken, using a full sample of over 2,000 firms. Given our ability to disaggregate R&D investment, we can discriminate between the returns to R&D by source of funds (company-funded vs. federally-funded R&D) and by character of use (basic research vs. applied research and development). Also, the panel structure of our data allows us to explore the time series properties of these rates of return throughout the sample period. The sensitivity of our results to outliers (influential observations) is also examined. Finally, we address Schumpeterian issues -- are the returns to R&D an increasing function of firm size? The last section is devoted to a summary of the conclusions that can be drawn from our empirical analysis.

I. Description of Linked Census R&D-LRD Data

Our examination of these issues is based on two confidential longitudinal data sets that were made available to us as participants in the ASA/NSF/Census Research Program. The Longitudinal Research Database (LRD), which brings together data from the Annual Survey and Census of Manufactures, will be used to measure productivity at the firm level (based on plant-level data). The LRD file is the richest source of annual data collected from manufacturing establishments, containing detailed information on their output and inputs. Comprehensive information on the characteristics of this file is presented in McGuckin and Pascoe (1988).

To study the relationship between R&D and productivity, we linked the LRD file to the NSF/Census firm-level Annual Survey of Industrial R&D (RD-1 Survey). The RD-1 Survey contains comprehensive data on firms' R&D investment and its distribution by source of funds (company vs. federally-funded R&D), character of use (basic vs. applied, product field), and many other classifications. The importance of the RD-1 Survey is demonstrated by the fact that it serves as the basis for the official United States R&D statistics, as published by NSF.[1]

A previous study, Lichtenberg-Siegel (1988), was based on linked R&D-LRD data for a sample of 115 large firms. In this

[1] See Lichtenberg (1989) for a discussion of the RD-1 data, including a comparison of them with other (e.g. Compustat) R&D data.

paper, we analyze the R&D-productivity connection for a substantially larger and more representative sample of firms -- the complete universe of firms in the linked R&D-LRD data set-- over 2,000 companies. Definitions of key variables appear in Table 1. Table 2 contains descriptive statistics for these variables. R&D intensity values are derived completely from information contained in the NSF R&D Survey. That is, R&D expenditure, employment, and sales figures reported are consolidated, domestic, firm-level values. We have computed average annual values of each variable for three periods, 1973-1976, 1977-1980, and 1981-1985.[2]

Productivity performance appears to have improved in the later periods. Average annual TFP growth declined 1.4 percentage points during period one and remained virtually constant in periods two and three. This result is consistent with the general improvement in aggregate economic performance during these years. A second stylized fact is that there is only a small degree of variation across periods in all measures of R&D intensity. For example, the average R&D intensity during periods one, two, and three were 2.4%, 2.4%, and 2.8%, respectively.[3]

While our measures of the intensity of R&D investment are based on consolidated, company-wide numbers, our estimates of

---

[2]Explicit information on the construction of these variables is contained in the data appendix.

[3]The unbalanced nature of reporting is due to the fact that firms are <u>not</u> required to report R&D expenditure by character of use (basic vs. applied research and development) and R&D scientists and engineers by source of funds.

firm productivity are based on data collected from the LRD file. This file contains information only for manufacturing establishments that were sampled continuously in the Annual Survey of Manufactures (ASM) for the years 1972-1981. The mean value of the "coverage ratio" COVRAT (the ratio of a firm's total LRD shipments to its consolidated sales) suggests that on average, we are capturing a substantial proportion (approximately 82%) of each corporation's domestic sales. In the next section, we discuss how these linked R&D-LRD Data can be used to estimate the effects of R&D on productivity growth.

II. The Relationship Between R&D Investment and Total Factor Productivity-Review and Critique of Existing Studies

Previous studies have demonstrated that productivity growth is positively correlated with the intensity of R&D investment (usually measured as R&D expenditure per dollar of sales). Our reservations concerning attempts to assess the impact of R&D on productivity are not grounded in doubts relating to the theory that is used to explain this relationship.[4] Our concern, however, is directed towards the poor quality of productivity measurement inherent to previous empirical

[4]In contrast, Nelson (1987) criticizes the interpretation of the empirical results contained in these studies. He argues that R&D intensity is not exogenous; in fact, it is determined by "technological opportunity" and appropriability conditions in specific industries. To some extent, we will control for these factors by measuring each firm's R&D intensity as a deviation from the average R&D intensity in its home industry (2-digit SIC).

implementation of the basic theoretical model.

This model (see Griliches (1979)), common to most existing empirical studies, is based on a Cobb-Douglas production function, including the stock of technical knowledge as a factor of production:

$$(1)\quad Q(t) = A_0 \exp(\lambda t) \prod_{i=1}^{3} X_i(t)^{\beta_i} R(t)^{\alpha}$$

with

A_0 = a constant

λ = a disembodied Hicks-neutral technical change parameter

$X_1(t) = K(t)$ = the stock of physical capital

$X_2(t) = L(t)$ = labor input

$X_3(t) = M(t)$ = materials (inlcuding energy)

$R(t)$ = the (unobservable) stock of technical knowledge

β_i = output elasticity of factor i

α = output elasticity of the stock of R&D

An index of TFP is defined as:

$$(2)\quad TFP(t) = \frac{Q(t)}{\prod_{i=1}^{3} X_i(t)^{\beta_i}}$$

Taking logs, differentiating with respect to time, assuming constant returns to scale, imposing the condition that the output elasticities of the conventional inputs equal their respective shares in total cost (factors are paid their marginal products), and reparameterization of the output elasticity of R&D (see Terleckyj (1974)), yields:

$$(3)\quad DTFP_t = \lambda + \rho\frac{\dot{R}}{Q} + \mu$$

where ρ is the marginal product of research capital, or the "rate of return" to R&D investment, λ is the rate of disembodied "external" technical change, and μ is a classical disturbance term. Assuming zero or negligible depreciation of R&D, $\frac{\dot{R}}{Q}$ is measured as the "intensity," of R&D investment, or the R&D to sales ratio. As discussed in the data appendix, the dependent variable, total factor productivity growth ($DTFP_t$), is calculated based on a Tornqvist approximation to the continuous Divisia index of output and inputs. Equation (3) therefore consititutes the "intensity" version of the R&D Capital Stock Model.

Recent examinations of the relationship between R&D and productivity have yielded somewhat contradictory results. Papers based on data collected from the 1950's and 1960's have found positive effects of R&D on productivity. However, in several studies using 1970's data, including Agnew and Wise (1978), Griliches (1980b), Link (1981a), and Scherer (1981), the R&D coefficient failed to achieve statistical significance. On the other hand, Mansfield (1980), Griliches-Lichtenberg (1984), Griliches-Mairesse (1984), Griliches (1986), and Lichtenberg-Siegel (1988) have found that the association between R&D and productivity did not collapse during the 1970's.

An important consideration in analyzing the effects of R&D on productivity is the level of aggregation of the data. Early studies used economy-wide or industry time series, leading to parameter estimates that have been interpreted as measures of the social return to R&D. Several R&D-productivity studies have also been undertaken at the firm-level (Minasian (1969), Mansfield (1980), Link (1981a), Griliches (1980a, 1986), and Griliches-Mairesse (1984)) and the business segment level (Clark-Griliches (1984)).[5] A micro level analysis is desirable because of greater degrees of freedom and the ability to evaluate private returns to R&D.

While these papers were useful, they were subject to some major restrictions. One problem is that many studies were based only on data collected from firms whose shares are publicly traded.[6] Also, public access files such as Standard and Poor's Compustat file, which is derived from firms' 10-K reports to the SEC, contain information only on company-funded R&D expenditure. This lack of detailed R&D data is bothersome because recent empirical work by Mansfield (1980), Griliches (1986), Griliches-Lichtenberg (1984), and Link (1981b) has focused on differences in productivity returns to specific components of R&D investment (e.g. basic research).[7]

---

[5]Mansfield (1980) used industry-level and firm-level data to estimate the productivity impact of basic research. His results were similar in both cases. The company-level estimates were based on data provided by 10 petroleum firms and 6 chemical firms.

[6]The Clark and Griliches (1984) study is based on the proprietary PIMS file collected by the Strategic Planning Institute (SPI).

[7]Mansfield (1980), Link (1981b), and Griliches (1986) find evidence of a productivity "premium" on basic research. Griliches (1986) and Griliches-Lichtenberg (1984) conclude that privately-financed R&D (but not federally-financed R&D) has a strong effect on productivity growth.

Productivity estimates in existing micro level studies are based on crude and incomplete measures of output and inputs. Public access files include no information on the number of hours worked and energy and materials data are often missing. The major problem associated with estimates of productivity in previous papers has been an inability to control for firm diversification. When computing real values of output and inputs, industry deflators must be applied to nominal variables. These deflators may vary greatly across different lines-of-business in a given firm. However, public information concerning a given firm's activity in different industries is remarkably limited.[8] As a result, in most firm-level studies, productivity estimation has been based on the assumption that firms operate in only one 4-digit SIC industry with real variables calculated according to a single set of price deflators.

In Lichtenberg-Siegel (1988), we illustrated the problems associated with the single-industry classification process by presenting this simple example: Assume that a firm operates in two 4-digit SIC industries.

Let $VQ1_t$ = The firm's output in current dollars, in industry 1 at time t

$VQ2_t$ = The firm's output in current dollars, in industry 2 at time t

---

[8]Compustat, for example, provides only one or several 4-digit SICs for the corporations it samples. Most importantly, the distributions of the variables needed to construct total or partial productivity are unavailable at the line-of-business level.

$PQ1_t$ - The price deflator for industry 1 at time t

$PQ2_t$ - The price deflator for industry 2 at time t

If $VQ1_t > VQ2_t$ then industry 1 is considered to be the firm's major line of business at time t, and industry 1's price deflator is applied to the firm's nominal output, to calculate real output. Thus, the conventional (using publicly-available information) methodology yields:

$$\text{(M1)} \quad Q'_t = \frac{VQ1_t + VQ2_t}{PQ1_t}$$

We believe that the proper way to measure real output is to take account of diversification,[9] by deflating the firm's nominal output in each industry, and then aggregating real output to the firm level. The preferred measure of real output is:

$$\text{(M2)} \quad Q_t = \frac{VQ1_t}{PQ1_t} + \frac{VQ2_t}{PQ2_t}$$

It is clear that Q_t and Q'_t will grow at different rates if there are changes in the relative price of the two industries' outputs. Similar issues are associated with the measurement of real input of diversified firms.

An errors-in-variables argument may be invoked to explain the deviations from the basic regression model that are inherent to each method of TFP estimation, the "conventional" (M1) and the "preferred" (M2) methodologies. Equation (3) is restated as:

$$\text{(4)} \quad DTFP^*_t = \beta RDINT_t + u_t$$

[9] Gollop and Monahan (1984), using quinquennial Census of Manufactures data for 1963-1977, present evidence that firms are becoming increasingly diversified.

where

$RDINT_t$ = the firm's R&D intensity in year t

$DTFP^*_t$ = the firm's true TFP growth in year t

u_t = a classical disturbance term

Using the preferred methodology (M2), we observe:

$$(5) \quad DTFP_t = DTFP^*_t + \epsilon_{1t}$$

where ϵ_{1t} is due to incomplete coverage of firm's establishments, an inability to completely measure the firm's "true" level of diversification, and errors in the industry deflators.[10] In estimating productivity growth according to the conventional methodology (M1), we believe that the following is observed:

$$(6) \quad DTFP'_t = DTFP^*_t + \epsilon_{1t} + \epsilon_{2t}$$

The additional error term, ϵ_{2t}, is due to a failure to control for diversification beyond a single 4-digit SIC industry.

The dependent variable is measured with error in both cases. We invoke the usual errors-in-variables assumptions:

$$E(u_t) = E(\epsilon_{1t}) = E(\epsilon_{2t}) = cov(RDINT_t, u_t) = cov(RDINT_t, \epsilon_{1t}) = cov(RDINT_t, \epsilon_{2t}) = cov(u_t, \epsilon_{1t}) = cov(u_t, \epsilon_{2t}) = cov(\epsilon_{1t}, \epsilon_{2t}) = 0$$

The error terms have zero mean, are pairwise uncorrelated, and are uncorrelated with both the independent variable and the classical disturbance term.

Under these assumptions, both the conventional and preferred methodologies yield unbiased parameter estimates. However, the conventional methodology (M1) produces "noisier" estimates:

---

[10] See Griliches and Lichtenberg (1989) for a description of errors in industry deflators.

$$\mathrm{var}(DTFP'_t) = \sigma_1^2 + \sigma_2^2 + \sigma_*^2 > \mathrm{var}(DTFP_t) = \sigma_1^2 + \sigma_*^2$$

where

σ_*^2 = the "true" variance of total factor productivity growth

The point estimate of the rate of return is not affected by this additional measurement error, but the preferred methodology provides more efficient estimates:

$$\mathrm{var}(\hat{\beta}_{(M1)}) = \frac{\sigma_u^2+\sigma_1^2+\sigma_2^2}{\Sigma RDINT^2} > \mathrm{var}(\hat{\beta}_{(M2)}) = \frac{\sigma_u^2+\sigma_1^2}{\Sigma RDINT^2}$$

In order to assess the importance of these alleged gains in the precision of TFP estimation generated by the "preferred" methodology, we calculated variants of our regression model using (M1) and (M2) in Lichtenberg-Siegel (1988). For a sample of 115 of the largest R&D performing companies, the standard deviation of the conventional measure of TFP growth was 29% higher.

Analysis of the regression results from the pilot study reveals that the additional variance associated with the conventional methodology is also uncorrelated with R&D. Tables 3 and 4 (from Lichtenberg-Siegel (1988)) are based on regressions of average annual productivity growth on average annual R&D intensity (for two periods: 1973-1976 and 1977-1980), using the standard (M1) and improved (M2) methods, respectively, to calculate TFP. In columns (b) and (e), R&D is disaggregated by source of funds, while in column (c), it is classified by character of use.[11] The point estimates in both tables are quite similar, as expected. However, LRD-based TFP measures provide more efficient estimates, as evidenced by the substantial increases in R^2 that arise when we calculate TFP under (M2).[12] Across columns (a)-(e), R^2 increases 30.8%, 46.3%, 69.2%, 88.2%, and 38.1% when we control for firm diversification (Table 4). Due to the overall improvement in goodness of fit when we adopt (M2) rather than (M1), the virtues of using the LRD file to estimate TFP have been firmly established.

III. Rates of Return to R&D Using Linked Census R&D-LRD Data

Having developed a procedure to measure productivity with greater precision in Lichtenberg-Siegel (1988), this section consists of a comprehensive analysis of rates of return to R&D investment using the complete set of over 2,000 firms in the linked Census R&D-LRD data set. Our sample, which includes many small, private firms, is a highly representative sample of companies performing industrial R&D. Table 5 demonstrates that the companies in our sample accounted for 84% of R&D performed (by industrial firms) in the United States in 1976. The sample coverage ratios are also quite high for sales, employment, and

---

[11] In Tables 3 and 4, columns a,b,d, and e are weighted least squares results, to adjust for heteroskedasticity due to incomplete firm "coverage" in the LRD file. That is, we believe that estimates of firm productivity based on only a small fraction of its manufacturing plants are somewhat less reliable, or "noisy." Still, on average, we captured 85% of a firm's sales.

[12] In comparing parameter estimates from these two tables, we note that corresponding regressions (columns (a)-(e)) are based on the same number of observations and R&D intensity values.

all measures of R&D.

In Section II, we discussed how use of linked R&D-LRD data allows us to measure productivity more precisely at the firm level. This is due to our ability to observe more complete measures of firms' output and inputs across industries. However, LRD estimates of a firm's productivity are based only on its continuously operating (between 1972 and 1981) manufacturing plants. We hypothesize that companies with a high percentage of their output in LRD establishments have their productivity measured more accurately than firms with low percentages of LRD activity. We found that the inability to measure firms' total output and inputs introduced an element of heteroskedasticity into OLS estimation of equation (3).

Thus, weighted least squares estimates for variants of equation (3) are presented in Table 6. The weight for these regressions is $(\text{COVRAT})^{1/2}$, where COVRAT = the ratio of the firm's LRD shipments to its consolidated domestic net sales. These results are based on regressions of average annual TFP growth on average R&D intensity values for three periods, 1973-1976, 1977-1980, and 1981-1985. In order to control (imperfectly) for inter-industry differences in R&D intensity due to differential "technological opportunity" and appropriability conditions, we measure each firm's R&D intensity as a deviation from the average R&D intensity in its home industry (2-digit SIC).[13] The parameter estimates of Table 6 will be compared with

[13] All of our regressions control for these industry effects

those from existing macro/industry and firm/line of business empirical studies in Table 7.

As discussed in Section II, it is generally believed that parameter estimates from industry-level studies measure social returns to R&D, while estimates derived from micro studies capture private returns to R&D. Therefore, we do not expect parameter estimates based on models estimated at different levels of aggregation to be equal.[14] In the studies that are not based on the "intensity" model, the parameter of interest is α, the output elasticity of R&D. For each $\hat{\alpha}$, we have imputed a value of $\hat{\rho}$ (the estimated marginal product), dividing $\hat{\alpha}$ by the mean of R/Q. We are especially interested in comparing our results to those of previous studies based on the "intensity" model, using firm or line-of-business data.

The results of Table 6 confirm the existence of a positive relationship between the intensity of R&D investment and average annual productivity growth. Under competitive assumptions, our results imply private rates of return to R&D investment of 13.2% and 9.7% using expenditure and employment measures, respectively. A 13.2% rate of return is lower than comparable estimates of 20% by Clark-Griliches (1984) using business level data, and 27% by Griliches (1980a) and 39% by Griliches (1986), using firm-level

[14] Several authors have avoided imposing the condition that factors are paid their marginal products and instead, estimated Cobb-Douglas production functions directly, assuming only constant returns to scale with respect to the conventional inputs. The "level" and "growth rate" equations estimated in Griliches (1980a, 1980b, and 1986) are based on constant returns to scale and competitive markets for these inputs.

data.

A disaggregation of R&D by source of funds reveals that while the intensity of company-funded R&D investment (expenditure or employment) is a significant determinant of productivity growth, the intensity of federally-funded R&D investment is not. In addition, using both expenditure and employment measures, the hypothesis of homogeneity of returns to company and federally-funded R&D is rejected at the 1% level of significance.[15] The potency of privately-financed R&D (but not federally-financed R&D) has also been documented in Terleckyj (1974), Griliches-Lichtenberg (1984), and Griliches (1986). Our estimate of 35.3% rate of return to company-funded R&D is substantially higher than similar estimates of 24.5% by Griliches (1986) and 27.5% by Mansfield (1980).

A disaggregation of R&D by character of use indicates that the intensity of investment in basic research has a powerful impact on productivity growth. An estimated rate of return of 133.8% to investment in basic research is substantially lower than Mansfield's (1980) estimate of 178.0% and Link's (1981b) estimate of 231.0%. The hypothesis of equality of returns to basic research and other types of R&D was rejected at the 1% level of significance.[16] This evidence of a "premium" on basic

---

[15]The t-statistics for tests of the expenditure and employment measures are 48.53 and 25.03, respectively.

[16]The t-statistic for the test of equality of returns to basic research and applied research is 61.07. The t-statistic for the same test involving basic research and development is 65.78.

research confirms the findings of Griliches (1986), Mansfield (1980), and Link (1981b).[17][18]

The point estimates displayed in Table 6 are based on pooled regressions, imposing a common slope for each period (1973-1976, 1977-1980, and 1981-1985). To investigate the hypothesis that the returns to R&D varied across these two periods, we re-estimated the regression model, allowing for different slopes in all three periods.[19] These results are presented in Table 8. Although we could not reject the hypothesis of equality of rates of return to R&D across periods, the results certainly suggest that the impact of R&D investment on TFP increased substantially in the later periods. In fact, the point estimate on total R&D (expenditure) is almost twice as high in period two. The rate or

[17]To make our results more comparable to the findings of Griliches (1986), we also regressed TFP growth on company-funded R&D intensity and the ratio of basic research to total R&D expenditure. The estimated coefficients and t-statistics (in parenthesis) are, respectively, .313 (11.86) and .096 (8.86). These estimates are quite similar to the findings contained in the aforementioned study, although the "premium" on basic research is somewhat lower.

[18]The same pattern of results emerges when we restrict our sample to include only firms reporting basic research. When we estimate variants of the regression model for the same set of companies, the estimated coefficients and corresponding t-statistics (in parenthesis) for columns a, b, d, and e are-a: .137 (4.66), b: .422 (6.17) .016 (0.41), d: .135 (3.07), e: .160 (3.37) .044 (0.62).

[19]Scherer (1983) argues that specifications such as ours may not capture "true" changes in private returns to R&D over time. He believes that it is important to follow R&D from industry of origin to industry of use, as many firms "purchase" R&D from other firms implicitly when buying certain products and services. Only by identifying these "interindustry technology flows" can we truly measure the impact of R&D on productivity.

return to company-funded R&D, using expenditure or employment measures of R&D investment, is also dramatically higher in period two. Turning to the character of use results, we observe that the "premium" on basic research increased slightly in period two, while the rates of return to applied research and development were quite unstable across periods. In the next section, we examine the sensitivity of the full sample regression results (Table 6) to outlying, influential observations.

IV. Identification of Influential Outliers

Following Kuh, and Welsch (1980) and Neter, Wasserman, and Kutner (1985), our research design is to detect outliers in the dependent and independent variables and determine whether these outliers are influential in the least squares regression fit. Cook (1977, 1979) has proposed an influence statistic that measures the change in the estimated parameter vector that results if the ith observation is deleted. It is well known (see Maddala (1977)) that the confidence region for β is expressed as:

$$(7)\quad \frac{(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)}{kMSE} = F(1-\alpha;k,n-k)$$

where

X is a nxk matrix of independent variables

β is a kx1 column vector of regression parameters

$\hat{\beta}$ is a kx1 column vector of estimated regression parameters

MSE = mean square error

Cook's influence statistic is defined in a similar manner, except it is based on a measure of the difference in the estimated parameter vector including and excluding the ith observation.

$$(8) \quad D_i = \frac{(\hat{\beta}-\hat{\beta}_{(i)})'X'X(\hat{\beta}-\hat{\beta}_{(i)})}{kMSE}$$

where $\hat{\beta}_i$ is the estimate of β without the ith data point. An equivalent expression for D_i involves the residuals and leverage values:

$$(9) \quad D_i = \frac{e_i^2}{kMSE}\left[\frac{h_{ii}}{(1-h_{ii})^2}\right]$$

Relating values of D_i to the F-distribution with k and n-k degrees of freedom, those observations having percentile values (of the F-distribution) of 50% or more are considered influential observations. For each variant of the regression model (equation (3)), we detected observations that are influential <u>and</u> outliers in X and/or y.

The discovery of influential outliers compels us to examine the sensitivity of our regression results to the following remedial measures:

1) reduction of the impact that influential observations have on the fitted regression function
2) deletion of influential outliers

Deleting outlying influential observations is a somewhat drastic approach, unless the researcher is certain that the extreme values are due to what Belsley, Kuh, and Welsch (1980) call "gross measurement error"-keypunch errors or incorrect reporting. If these values are correct, then their deletion eliminates vital information. In our final sample, we believe that we have already discarded a non-negligible percentage of values reflecting gross errors in measurement. An alternative to the least squares estimator that minimizes the effect of egregious errors or outliers is the method of least absolute deviations. One of a class of robust estimators, this "bounded-influence" estimator minimizes the sum of the absolute deviations of the observations from their means. The loss function is:

$$(10) \quad \sum_{i=1}^{n} |y_i - (\beta_0 + \beta_{i1} X_{i1} + \ldots + \beta_{k-1} X_{i,k-1})|$$

These estimates are less sensitive to outliers because the sum of absolute, rather than squared, deviations, is minimized.[20]

In Table 9, we compare our earlier estimates of rates of return to R&D investment to estimates of these same parameters after deleting influential outliers (DEL) and those derived from the method of least absolute deviations (LAD). The LAD estimates of the rates of return to total R&D, company-funded R&D, and basic research are 29.5%, 26.6%, and 31.5% lower, respectively, than our original estimates. Our point estimates of these same three parameters decline 3.8%, 17.3%, and 24.4%, respectively, when influential outliers are discarded. These results are somewhat surprising. We expected the deletion of influential outliers to have a stronger impact on the parameter estimates than attempts to reduce their influence. Instead, the opposite pattern emerged, as the LAD estimates of these three rates of

---

[20]As demonstrated by Charnes et al. (1955), the method of least absolute deviations estimates can be derived from the solution to a linear programming problem. However, standard errors for the parameter estimates are unknown due to the fact that the statistical properties of the sampling distribution of this estimator are not well-defined.

return are always lower than corresponding DEL estimates. Still, our original estimates of the effect of R&D investment on TFP growth are not dramatically sensitive to adjustments aimed at reducing the impact of influential outliers.[21] In the next section, we explore the relationship between rates of return to R&D and firm size.

V. Firm Size and Rates of Return to R&D

The ideas of Joseph Schumpeter (1950) figure prominently in the belief that large firms are especially likely both to undertake, and be successful in, research activities. Recent papers by Griliches (1980a), Scherer (1984), Bound et al. (1984), and Cohen, Levin, and Mowery (1987) have found little evidence to support the position that firm size is positively correlated with R&D intensity. Link (1981a) suggested an alternative approach to empirical investigation of the Schumpeterian hypothesis. He examined and found evidence of a systematic relationship between firm size and the impact of R&D on productivity -- large firms earned higher returns to R&D than small firms.

Estimation of variants of equation (3) was contingent on the assumption of a common rate of return, ρ, among firms. In this section, we test whether large firms R&D investment by large firms is more successful than that undertaken by small firms, by investigating the possibility that ρ varies across firms

---

[21] With influential outliers deleted, we still reject the hypothesis of homogeneity of returns to company and federally-funded R&D. The same is true for the hypothesis of equality of returns to basic research and other types of R&D.

according to size.[22] A test for the structural stability of regression parameters has been developed by Brown, Durbin, and Evans (1975). The null hypothesis of this test is that the regression coefficients are constant over an index of firm size. An analysis of the cumulative sum of squared residuals determines where, if at all, a structural "break" or shift occurs. An attractive property of the Brown-Durbin-Evans "cusum" test is that it does <u>not</u> require prior information concerning the true point of structural change, unlike the Chow test. The use of prior information concerning structural shifts is often quite plausible when analyzing time series data. However, in our context, the implementation of this type of strategy would be arbitrary.[23] Link (1981a), having estimated rates of return to R&D investment based on the reduced form version of the R&D capital stock model (using firm-level data), employed the Brown-Durbin-Evans test to examine the structural stability of these returns with respect to firm size.

Our findings, based on the same research design applied to a larger and more representative sample of firms, do not substantiate the Schumpeterian hypothesis. Using sales and employment as proxies for firm size, we were unable to reject the

---

[22] A strict Schumpeterian interpretation of the heterogeneity of ρ across companies might be that ρ is a function of firms' monopoly power. Large size is neither a necessary nor a sufficient condition for firms' ability to gain or maintain monopoly power.

[23] In fact, cusum test have been applied on macroeconomic time series data in Khan (1974), Heller and Khan (1979), and Stern, Baum, and Greene (1979).

null hypothesis of structural stability for the rate of return to R&D.[24] Although we were unable to determine from the data whether different size "regimes" exist, with respect to the rate of return to R&D, we ranked companies by size and divided the sample into three groups. Table 10 contains estimates of rates of return to R&D, productivity growth, and R&D investment for these three size categories. The small degree of variation in the rate of return to total R&D across groups provides non-parametric evidence in support of structural stability. Mean total factor productivity growth and R&D intensity values are also quite similar.

The returns to company-funded R&D, however, do appear to be higher for large firms. In fact, we observe highly statistically-significant differences (at the 1% level of significance) in rates of return to company-funded R&D between the two groups of large firms and the smallest companies.[25] The apparent instability of this regression parameter compels us to implement the cusum test on this coefficient. One reason for this finding may be that appropriability conditions are more favorable for large firms. This is a subject for future research. We also note that federally-funded R&D has a stronger

---

[24] That is, the test statistic, s_r, based on the normalized cumulative sum of squared residuals, always falls within the 5% confidence intervals above and below the mean value line.

[25] The t-statistics for these tests are 2.75 (largest vs. smallest) and 2.97 (middle vs. smallest). The difference in rates of return to company-funded R&D between the largest and middle groups of firms is statistically insignificant.

impact on the productivity growth of small firms. This result, which has important implications for federal contracting policy, also bears further examination.

Still, the evidence presented in Table 10 demonstrates that our earlier findings concerning rates of return are consistent for each size classification of firms:

a) positive and significant coefficient on R&D intensity

b) higher returns to company-funded R&D

c) a productivity "premium" on basic research.

VI. Conclusions

Given the limited information contained in publicly-available and proprietary data sets, estimates of company productivity used in previous studies have been subject to substantial measurement error. The main source of inaccuracy in these studies has been an inability to adequately control for the diversified activities of corporations. Use of the NSF/Census R&D-LRD Panel allows us to develop more precise estimates of TFP, because the LRD file contains detailed data on the output and input of firms at the 4-digit SIC industry level.

Our results suggest that R&D investment was a significant determinant of productivity growth during the 1970's, as documented in previous studies, and also during the 1980's. However, the estimated private rate of return to R&D expenditure, 13.2%, is substantially lower than previous estimates. As in Griliches-Lichtenberg (1984) and Griliches (1986), we observe statistically-significant differences in the rates of return to company-funded and federally-funded R&D. We find a strong positive correlation between the rate of growth of TFP and the firm's privately-financed R&D, while federally-financed R&D does not appear to be a significant determinant of TFP growth.[26] We also find that a firm's investment in basic research has a strong effect on productivity growth, while investment in other types of R&D apparently has either a small impact on TFP growth, or none at all. This result is consistent with the findings of Mansfield (1980), Link (1981b) and Griliches (1986).

The point estimates we have discussed were based on pooled regressions, assuming constant slopes for each three-year period, 1973-1976, 1977-1980, and 1981-1985. Although we could not formally reject this constraint, we do find evidence to support the hypothesis that the impact of R&D on productivity was stronger in the later periods.[27] Concern about a recent decline in the impact of R&D on total factor productivity, due to a diminution of technological opportunities (Nordhaus (1980)), does not appear to be well-founded.

Next, we analyzed the sensitivity of our regression results to attempts to reduce the influence of "outlying" observations.

---

[26] As Griliches (1979) points out, this result does not necessarily imply that federally-funded R&D is "unproductive." In industries with relatively high levels of publicly-financed R&D, such as the defense or space sectors, output is poorly measured and price indices do not accurately reflect improvements in quality.

[27] Despite considerable variation across periods, the pattern of results in <u>each</u> period is consistent with our overall findings and those of most existing studies.

Two strategies were employed to address this problem:

a) deletion of influential outliers

b) estimation by the method of least absolute deviations a "bounded-influence" estimator

These adjustments had only a small impact on our key empirical findings. Therefore, our regression results are not due primarily to values that may be anomalous and/or erroneous.

The relationship between rates of return to R&D and firm size was also examined. Inherent to our estimation of the reduced form version of the R&D Capital Stock Model is the assumption of a common rate of return among firms. As in Link (1981a), we explored the validity of this assumption by testing for the structural stability of the regression parameter (only the rate of return to total R&D was examined), with respect to firm size. We could not reject structural stability, and an analysis of three groups of firms, ranked in ascending size, seemed to confirm this result. However, the returns to company-funded R&D appear to be an increasing function of firm size. At the same time, the rate of return to federally-funded R&D is higher for small firms.[28] These results indicate that appropriability conditions are more favorable for large firms funding their own research. This non-parametric evidence, which supports the Schumpeterian hypothesis, bears further examination.

---

[28] This may explain why our estimate of the rate of return to company-funded R&D was substantially higher in the pilot study (Lichtenberg-Siegel (1988)), which was based on data collected from large firms.

Data Appendix

Data Sources for Construction of TFP estimates:

To construct estimates of levels and growth rates of TFP based on 3 factors of production, we need estimates of real values of output (Q), capital (K), labor (L), materials (including energy) (M), and factor shares. The LRD (Longitudinal Research Database) file provides data on nominal values of output (VQ), capital (VK), labor (VL), materials (VM), energy, and inventories for 20,493 establishments that were sampled in the Annual Survey of Manufactures for the years 1972-1985. Price deflators were imported from 3 separate files:

1) Bureau of Industrial Economics Output Data Base
   This file contains deflators for shipments, raw materials, work-in-process, and finished goods inventories, at the 4-digit SIC level for the years 1972-1980. Subsequent to 1980, we used the National Bureau of Economic Research (NBER) output deflators, which are derived from the Bureau of Economic Analysis.

2) Bureau of Industrial Economics Capital Stocks Data Base
   Includes data for the years 1972-1982 on the net stock of capital in constant (1972) dollars and the gross stock of capital in historical dollars at the 3/4 digit SIC level. For each industry, we evaluated the ratio of these two numbers for plant and equipment separately. As in Lichtenberg-Siegel (1987), these ratios were applied to the gross plant and gross

equipment figures to evaluate net "benchmark" estimates of plant and equipment in 1972. Using these estimates of the initial capital stock, along with industry estimates of the average rate of capital depreciation and plant-specific deflated capital expenditures, a "perpetual inventory" algorithm is used to generate estimates of the net stock of plant and equipment. The real net stock of capital is then defined as the sum of net plant and net equipment.[29]

3) NBER R&D and Productivity Project File

The NBER has constructed materials and energy deflators at the 4-digit SIC level.

The remainder of this section is devoted to an explanation of how key economic measures of output, labor, capital, and materials (including energy) were defined using current dollar values of inputs and output, and industry deflators (3 or 4 digit SIC level) for the years 1972-1985.

Output:

Output in current dollars is defined as the value of shipments, with adjustments for the net (annual) change in finished goods and work-in-process inventories. Real output is computed by dividing each term by its corresponding industry price deflator.

---

[29]The BIE data set has not been updated beyond 1982. As a result, two-digit industry deflators for producers' durable plant and equipment were used in later years (see March 1987 Survey of Current Business).

<u>Labor</u>:

Current dollar labor input is measured as the sum of salaries and wages and total supplemental labor costs. Real labor input is defined as the ratio of total salaries and wages (TSW) to production worker wages (PWW), multiplied by total production worker hours (PWH).

(3) $L = (TSW \div PWW) \cdot PWH$

This "production worker equivalent" measure of labor input (data on hours of work of nonproduction workers is unavailable) is based on the assumption that the relative wages of production and nonproduction workers are equal to their marginal productivity.

<u>Capital</u>:

Nominal capital is constructed using the assumption of constant returns to scale. We define current dollar capital as current dollar output minus the current dollar costs of materials (including energy) and labor, plus an adjustment for the net change in materials inventories. Our construction of the real net stock of plant and equipment, based on a perpetual inventory algorithm, was discussed on the previous page.

<u>Materials and Energy</u>:

The current dollar values of materials (including energy) is defined as cost of materials (CM), plus an adjustment for the net change in materials inventories. Constant dollar values of materials (including energy) are evaluated by dividing current dollar values of materials and energy by the NBER 4-digit SIC price deflators for materials and energy.

We also compute factor shares, which are used in constructing growth rates of TFP. Using the methodology employed in Griliches-Lichtenberg (1984) and in many other studies, we calculate a Tornqvist index of three inputs:

$$\ell n\left(\frac{IN_t}{IN_{t-6}}\right) = .5(SL_t+SL_{t-6}) * \ell n\left(\frac{L_t}{L_{t-6}}\right)$$

$$+ .5(SK_t+SK_{t-6}) * \ell n\left(\frac{K_t}{K_{t-6}}\right)$$

$$+ .5(SM_t+SM_{t-6}) * \ell n\left(\frac{M_t}{M_{t-6}}\right)$$

$$= \sum_{i=1}^{3} [.5*(S_{it}+S_{i,t-6})] \ell n\left(\frac{X_t}{X_{i,t-6}}\right)$$

where

IN_t = Index of total input at time t

S_{it} = Share of factor i in the total cost of output at time t, factors i = K,L,M (including energy)

X_{it} = Quantity of factor i at time t (in real terms).

Our explicit formula for TFP growth is:

$$(8)\quad DTFP_t = \frac{\dot{TFP}}{TFP} = \ell n\left[\frac{Qt}{Q_{t-1}}\right] - \sum_{1}^{3} [(.5*(S_{it} + S_{i,t-1})] \ell n\left[\frac{Xit}{X_{i,t-1}}\right]$$

Table 1

Variable Definitions

VARIABLE	DEFINITION
AVDTFP	Average Annual TFP growth rate calculated for 3 periods: 1973-1976, 1977-1980, and 1981-1985
RDINT	Total R&D expenditure/Total Domestic Net Sales
COMPRDINT	Total Company-financed R&D expenditure/Total Domestic Net Sales
FEDRDINT	Total Federally-financed R&D expenditure/Total Domestic Net Sales
RDINTSE	Total R&D Scientists and Engineers/Total Domestic Employment
COMPRDINTSE	Total Company-financed R&D Scientists and Engineers/Total Domestic Employment
FEDRDINTSE	Total Federally-financed R&D Scientists and Engineers/Total Domestic Employment
BASICINT	Total Basic Research Expenditure/Total Domestic Net Sales
APDEVINT	Total Applied Research and Development Expenditure/Total Domestic Net Sales
APPLINT	Total Applied Research Expenditure/Total Domestic Net Sales
DEVINT	Total Development Expenditure/Total Domestic Net Sales
COVRAT	Total Domestic LRD Shipments/Total Domestic Net Sales

Table 2

Characteristics of the Distribution of Variables Used in our Analysis

Period 1: 1973-1976
Period 2: 1977-1980
Period 3: 1981-1985

Variable	N			Mean			Median			Std. Deviation		
	Per 1	Per 2	Per 3	Period 1	Period 2	Period 3	Period 1	Period 2	Period 3	Period 1	Period 2	Period 3
(in %) AVDTFP	2207	1522	1512	-1.40	-.015	-.004	-1.22	-0.05	-0.03	8.07	7.64	1.37
RDINT	2207	1522	1512	.024	.024	.028	.008	.010	.011	.056	.044	0.48
COMPRDINT	2207	1522	1512	.018	.019	.024	.008	.010	.011	.034	.026	.036
FEDRDINT	2207	1522	1512	.005	.006	.005	0	0	0	.037	.032	.029
RDINTSE	1953	1341	1425	.024	.027	.029	.010	.013	.014	.043	.045	.043
COMPRDINTSE	1783	963	1424	.056	.026	.026	.015	.015	.013	.108	.031	.034
FEDRDINTSE	1613	962	1426	.014	.005	.003	0	0	0	.052	.043	.029
BASICINT	859	942	1507	.002	.002	.001	0	0	0	.009	.009	.009
APDEVINT	859	952	1509	.031	.029	.028	.013	.014	.011	.055	.047	.047
APPLINT	858	923	1482	.009	.008	.009	.002	.002	.002	.031	.021	.019
DEVINT	859	923	1482	.025	.023	.020	.010	.010	.007	.051	.039	.037
COVRAT	2207	1522	1512	.837	.809	.783	.857	.828	.804	389	386	370

Table 3

Regressions of $DTFP'_t$ on Various Measures of R&D
(from pilot study-Lichtenberg-Siegel (1988))

Type of R&D Intensity	R&D Expenditure / Sales (a)	R&D Expenditure / Sales (b)	R&D Expenditure / Sales (c)	Number of FTE R&D Scientists & Engineers / Employment (d)	Number of FTE R&D Scientists & Engineers / Employment (e)
TOTAL R&D	.079 (1.70)			.150 (1.58)	
COMPANY-FUNDED R&D		.402* (3.48)			.397* (2.51)
FEDERALLY-FUNDED R&D		.015 (.29)			.028 (.25)
BASIC RESEARCH			.535 (.73)		
APPLIED RESEARCH			-.303 ([illegible].11)		
DEVELOPMENT			.097 (1.50)		
INTERCEPT	-.020 (1.71)	-.028* (2.42)	-.024** (2.17)	-.021 (1.79)	-.025** (2.15)
R^2	.026	.067	.039	.017	.042
PERIOD DUMMY	.012 (1.66)	.011 (1.60)	.015** (2.32)	.011 (1.61)	.011 (1.60)
DFE	214	213	201	218	216

* = significant at .01 level
** = significant at .05 level
(t-statistics in parentheses)

Table 4

Regressions of $DTFP_t$ on Various Measures of R&D
(from pilot study-Lichtenberg-Siegel (1988))

Type of R&D Intensity	R&D Expenditure / Sales			Number of FTE R&D Scientists & Engineers / Employment	
	(a)	(b)	(c)	(d)	(e)
TOTAL R&D	.086** (2.28)			.159** (2.07)	
COMPANY-FUNDED R&D		.412* (4.49)			.403* (3.16)
FEDERALLY-FUNDED R&D		.020 (.50)			.041 (.45)
BASIC RESEARCH			1.198** (2.07)		
APPLIED RESEARCH			-.261 (1.21)		
DEVELOPMENT			.108** (2.10)		
INTERCEPT	-.016 (1.71)	-.025* (2.64)	-.023* (2.59)	-.017 (1.85)	-.022** (2.30)
R^2	.034	.098	.066	.032	.058
PERIOD DUMMY	.009 (1.54)	.008 (1.48)	.013 (2.53)	.009 (1.59)	.009 (1.55)
DFE	214	213	201	218	216

* = significant at .01 level
** = significant at .05 level
(t-statistics in parentheses)

Table 5

Sample Coverage of Linked R&D-LED Data Set:
Population-R&D Performing Companies in US-1976*

	(1) Sample	(2) Population	(1)/(2) Percent
Number of Companies with > 1000 employees	1092	1398	78.1
Sales (millions of $)	701877	870871	80.6
Employment (thousands)	11857	15166	78.2
Total R&D (millions of $)	22654	26997	83.9
Company-funded R&D (million of $)	14628	17436	83.9
Federally-funded R&D (million of $)	8026	9561	83.9
Basic Research (millions of $)	751	837	89.7
Applied Research (millions of $)	4162	5102	81.6
Development (millions of $)	17741	21058	84.2
R&D Scientists and Engineers (thousands)	325	364	89.3

*Population values are derived from NSF (1980)

Table 6

Regressions of $DTFP_t$ on Various Measures of R&D

(full sample-controls for industry effects in R&D)

Type of R&D Intensity	R&D Expenditure / Sales			Number of FTE R&D Scientists & Engineers / Employment	
	(a)	(b)	(c)	(d)	(e)
TOTAL R&D	.132* (6.40)			.097* (4.70)	
COMPANY-FUNDED R&D		.353* (13.09)			.172* (8.88)
FEDERALLY-FUNDED R&D		.026 (0.81)			.028 (1.10)
BASIC RESEARCH			1.338* (13.06)		
APPLIED RESEARCH			.108 (1.14)		
DEVELOPMENT			.014 (0.24)		
INTERCEPT PERIOD ONE	-.003 (0.56)	-.009 (1.51)	.023* (2.40)	-.001 (0.08)	.002 (0.28)
INTERCEPT PERIOD TWO	.001 (0.21)	-.051 (0.46)	.021** (2.20)	.003 (0.47)	.009 (1.08)
INTERCEPT PERIOD THREE	-.003 (0.53)	-.012 (1.80)	.015 (1.55)	.000 (0.06)	.009 (1.09)
R^2	.026	.051	.075	.025	.041
DFE	5218	5217	3236	4697	4140

* = significant at .01 level
** = significant at .05 level
(t-statistics in parentheses)

Table 7

Summary of Previous Studies Estimating Rates of Return to R&D

Intensity Study	Period	Level of Aggregation	N	Type of Specification	Additional Covariates	Estimates of Rates of Return to R&D Investment (%) Total R&D	Company-funded R&D
Terleckyj (1974)	1948-1966	Industry	20	Intensity	% of sales to the private sector, % of unionized workers, index of cyclical instability of output	12.0	37.0
Link (1978)	1958	"	45	"	--	18.8	--
Griliches (1980b)	1959-1968	"	39	"within" industries	Age of capital, Man-hours	42.0	--
Griliches (1980b)	1969-1977	"	"	"	"	0	--
Griliches-Lichtenberg (1984)	1959-1963	"	27	Intensity	--	2.7	9.2
Griliches-Lichtenberg (1984)	1964-1969	"	"	"	--	0	20.3
Griliches-Lichtenberg (1984)	1969-1973	"	"	"	--	5.2	33.4
Minasian (1969)	1948-1957	Firm	17	Production Function	--	54.0	--
Griliches (1980a)	1957-1965	"	883	" (on levels and rates of growth)	5 industry dummies, Std Error of sales growth rate, Dummy for "no imputations"	27.0	36.5
Griliches (1986)	1966-1977	" "	491 and 911	" and Intensity	Basic Research/Total R&D, Co. R&D stock/ R&D stock, industry dummies	(production Function) 39.0	(Intensity) 24.3
Mansfield (1980)	1960-1976	"	16	Intensity	--	--	27.5
Griliches-Mairesse (1984)	1966-1977	"	133	Production Function	--	--	62.1* 41.5**
Link (1981a)	1971-1976	"	174	Intensity	--	--	0
Clark-Griliches (1984)	1971-1980	Line of Business	924	"	Industry dummies, rate of capacity utilization, net plant/gross plant, $ of unionized employees	20.0	--

* constant returns to scale imposed
** constant returns to scale <u>not</u> imposed

Table 8

Regressions of $DTFP_t$ on various measures of R&D - pooled, and allowing for different slopes across periods

Type of R&D Intensity	R&D Expenditure / Sales (a) Pooled	(a) Period 1	(a) Period 2	(a) Period 3	(b) Pooled	(b) Period 1	(b) Period 2	(b) Period 3	(c) Pooled	(c) Period 1	(c) Period 2	(c) Period 3
TOTAL R&D	.132*	.095*	.189*	.213*								
	(6.40)	(3.93)	(3.54)	(4.17)								
COMPANY-FUNDED R&D					.353*	.254*	.415*	.513*				
					(13.09)	(7.52)	(5.34)	(9.59)				
FEDERALLY-FUNDED R&D					.026	-.001	.096	-.013				
					(0.81)	(0.02)	(1.24)	(0.14)				
BASIC R&D									1.338*	1.203*	1.409*	1.360*
									(13.06)	(4.26)	(2.83)	(11.78)
APPLIED RESEARCH									.108	.153	.036	-.120
									(1.14)	(1.07)	(0.18)	(0.66)
DEVELOPMENT									.014	-.005	.154	.153
									(0.24)	(0.06)	(1.61)	(1.51)
INTERCEPT-PERIOD ONE	-.003*	-.002	---	---	-.009	-.005	---	---	.023*	.023*	---	---
	(0.56)	(0.25)			(1.51)	(0.78)			(2.40)	(2.34)		
INTERCEPT-PERIOD TWO	.001	---	.000	---	-.003	---	-.003	---	.021**	---	.018	---
	(0.21)		(0.05)		(0.46)		(0.53)		(2.20)		(1.85)	
INTERCEPT-PERIOD THREE	-.003	---	---	-.006	-.012	---	---	-.015**	.015	---	---	.014
	(0.53)			(0.87)	(1.80)			(2.38)	(1.55)			(1.45)
R^2	.026	.027	---	---	.051	.053	---	---	.075	.075	---	---
DFE	5218	5215	---	---	5217	5214	---	---	3236	3233	---	---

* = significant at .01 level
** = significant at .05 level
(t-statistics in parentheses)

Table 9

Regressions of $DTFP_t$ on various measures of R&D expenditure - with and without adjustments for <u>influential outlying</u> observations (NOADJ = no adjustments, DEL = deletion of influential outliers, LAD = method of least absolute deviaitons)

Type of R&D Intensity	(a) NOADJ	(a) DEL	(a) LAD#	(b) NOADJ	(b) DEL	(b) LAD#	(c) NOADJ	(c) DEL	(c) LAD#
	R&D Expenditure / Sales								
TOTAL R&D	.132* (6.40)	.127* (9.43)	.093						
COMPANY-FUNDED R&D				.353* (13.09)	.292* (11.96)	.259			
FEDERALLY FUNDED R&D				.026 (0.81)	.073 (1.81)	.071			
BASIC RESEARCH							1.338* (13.06)	1.012* (9.98)	.917
APPLIED RESEARCH							.108 (1.14)	.043 (1.36)	-.085
DEVELOPMENT							.014 (0.24)	.045 (1.13)	.054
INTERCEPT PERIOD ONE	-.003 (0.56)	-.005 (0.81)	-.005	-.009 (1.51)	-.010 (1.60)	-.007	.023* (2.40)	.024* (2.45)	.014
INTERCEPT PERIOD TWO	.001 (0.21)	.007 (0.28)	.006	-.051 (0.46)	-.055 (0.00)	-.040	.021** (2.20)	.023* (2.33)	.013
INTERCEPT PERIOD THREE	-.003 (0.53)	.004 (0.62)	.001	-.012 (1.80)	.000 (0.30)	.005	.015 (1.55)	.018 (1.70)	.009
R^2	.026	.032		.051	.054		.075	.052	
DFE	5218	4983	5218	5217	4982	5217	3236	3201	3236

* = significant at .01 level

Table 10

Rates of Return to R&D Expenditure and Mean Values of Productivity Growth and R&D Intensity for 3 Size Classes of Firms

FIRM SIZE DISTRIBUTION

Rates of Return to R&D Expenditure (t-statistics in parentheses)	ALL FIRMS	<224 EMPL	224-779 EMPL	>780 EMPL	ALL FIRMS	<224 EMPL	224-779 EMPL	>780 EMPL	ALL FIRMS	<224 EMPL	224-779 EMPL	>780 EMPL
TOTAL R&D	.132* (6.40)	.128* (4.58)	.127* (3.95)	.145* (3.47)								
COMPANY-FUNDED R&D					.353* (13.09)	.238* (7.71)	.397* (7.63)	.412* (7.77)				
FEDERALLY-FUNDED R&D					.026 (0.81)	.073 (1.28)	.017 (0.39)	-.031 (0.38)				
BASIC RESEARCH									1.338* (13.06)	1.313* (8.65)	1.563* (5.30)	1.301* (5.90)
APPLIED RESEARCH									.108 (1.14)	.036 (0.66)	-.038 (0.26)	.162 (1.04)
DEVELOPMENT									.014 (0.24)	.086 (1.60)	.175 (2.22)	-.046 (1.19)
MEAN VALUES (in %)												
AVERAGE ANNUAL TFP GROWTH	-0.7	-0.7	-0.8	-0.7								
R&D INTENSITY	2.6	2.7	2.3	2.8								
COMPANY-FUNDED R&D INTENSITY	2.0	2.1	1.9	2.1								
FEDERALLY-FUNDED R&D INTENSITY	0.5	0.6	0.4	0.7								
R&D EMPLOYMENT AS A % OF TOTAL EMPLOYMENT	2.7	2.7	2.5	2.8								
N	5242	1747	1747	1748								

* = significant at .01 level
** = significant at .05 level
(t - statistics in parentheses)

References

Agnew, Carson E., and Donald E. Wise. 1978. "The Impact of R&D on Productivity: A Preliminary Report." Paper presented at the Southern Economic Association meetings.

Belsley, David A., Edwin Kuh, and Roy E. Welsch. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity New York: John Wiley & Sons.

Brown, R.L., J. Durbin, and J.M. Evans. 1975. "Techniques for Testing the Constancy of Regression Relationships Over Time," Journal of the Royal Statistical Society (series B) 37:149-163.

Charnes, Abe, W.W. Cooper, and R.O. Ferguson. 1955. "Optimal Estimation of Executive Computing by Linear Programming," Management Science 138-151.

Clark, Kim B. and Zvi Griliches. 1984. "Productivity Growth and R&D at the Business Level: Results from the PIMS Data Base, in Z. Griliches, ed., R&D, Patents, and Productivity, Chicago: University of Chicago Press, 1984, 393-416.

Cohen, Wesley M., R.C. Levin, and D.C. Mowery. 1987. "Firm Size and R&D Intensity: A Re-Examination," NBER Working Paper #2205.

Cook, R. Dennis. 1977. "Detection of Influential Observations in Linear Regression," Technometrics 19:15-18.

___________. 1979. "Influential Observations in Linear Regression," Journal of the American Statistical Association 74:169-174.

Gollop, Frank M., and James L. Monahan. 1984. "From Homogeneity to Heterogeneity: A Proper Index of Diversification," Bureau of the Census Technical Notes, mimeo.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth," Bell Journal of Economics 10, no. 1:92-116.

___________. 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 No. 44, Chicago: University of Chicago Press, 1980, 419-454.

___________. 1980b. "R&D and the Productivity Slowdown," American Economic Review 70, no. 2:343-348.

___________. 1986. "Productivity, R&D and Basic Research at the Firm Level in the 1970's," American Economic Review 76, no. 1:141-154.

___________ and Frank R. Lichtenberg. 1984. "R&D and Productivity Growth at the Industry Level. Is There Still a Relationship?" in Z. Griliches, ed., R&D, Patents, and Productivity, Chicago: University of Chicago Press, 1984, 465-496.

___________. 1989. "Errors of Measurement in Output Deflators," NBER Working Paper No. 2000, forthcoming Journal of Business and Economic Statistics

___________ and Jacques Mairesse. 1984. "Productivity and R&D at the Firm Level," in Z. Griliches, ed., R&D, Patents, and Productivity Chicago: University of Chicago Press, 339-374.

Heller, Robert H., and Moshin S. Khan. 1979. "The Demand for Money and the Term Structure of Interest Rates," Journal of Political Economy 87:109-129.

Khan, Moshin S. 1974. "The Stability of the Demand-for-Money Function in the United States 1901-1905," Journal of Political Economy 82:1205-1219.

Lichtenberg, Frank R. 1986. "The Duration and Intensity of Investment in 'Independent Research and Development' Projects," Journal of Economic and Social Measurement 14:207-218.

___________. 1988. "Using Longitudinal Establishment Data to Estimate the Internal Adjustments Costs Model," Review of Economics and Statistics 70:421-430.

___________. 1989. "Issues in Measuring Industrial R&D," Research Policy forthcoming.

___________ and Donald Siegel 1987. "Productivity and Changes in Ownership of Manufacturing Plants," Brookings Papers on Economic Activity 3:1987:

___________. 1988. "Using Linked Census R&D-LRD Data to Analyze the Effect of R&D Investment on Total Factor Productivity Growth," Center for Economic Studies Discussion Paper, U.S. Census Bureau.

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

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

Maddala, G.S. 1977. Econometrics, New York: McGraw Hill.

Mansfield, Edwin. 1980. "Basic Research and Productivity Increase in Manufacturing," American Economic Review 70:863-873.

McGuckin, Robert, and George Pascoe. 1988. "The Longitudinal Research Data Base-Research Possibilities," paper presented at NBER Conferenece on Income and Wealth.

Minasian, Jora R. 1969. "Research and Development, Production Functions and Rates of Return," American Economic Review 59:80-85.

Nelson, Richard R. 1988. "Modeling the Connections in the Cross Section Between Technical Progress and R&D Intensity," mimeo.

Neter, J., W. Wasserman, and Kutner. 1985. Applied Linear Statistical Models Homewood, Il: Irwin.

Nordhaus, William D. 1980. "Policy Responses to the Productivity Slowdown," in The Decline in Productivity Growth, series no. 22. Federal Reserve Bank of Boston.

Scherer, F.M. 1981. "Research and Development, Patenting, and the Microstructure of Productivity." Final Report, National Science Foundation, grant no. PRA-7826526.

__________. 1983. "R&D and Declining Productivity Growth," American Economic Review 73:215-218.

Schumpeter, Joseph A. 1950. Capitalism, Socialism, and Democracy, 3rd. ed. New York: Harper and Row.

Stern, Robert M., Christopher F. Baum, and Mark N. Greene. 1979 "Evidence on Structural Change in the Demand for Aggregate U.S. Imports and Exports," Journal of Political Economy 87:179-192.

Terleckyj, Nestor. 1974. Effects of R&D on the Productivity Growth of Industries. Washington: National Planning Association.