Discussion Paper

Discussion Paper No. 94–01

R&D and Productivity in German Manufacturing Firms

Dietmar Harhoff

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Industrial Economics and International Management Series

R&D and Productivity in German Manufacturing Firms

by
Dietmar Harhoff

University of Mannheim and
Zentrum für Europäische Wirtschaftsforschung (ZEW)

January 1994
Revised December 1994

Abstract

This paper uses a new firm panel data set to explore the relationship between R&D and productivity in German manufacturing firms for the period from 1979 to 1989. The results confirm the view that R&D is an important determinant of productivity growth. In the cross-section, the elasticity of sales with respect to R&D capital is on the order of 14 per cent. Using fixed-effects estimators yields R&D elasticities of about 8 per cent. Differencing estimates improve considerably when growth rates are computed over longer time periods, suggesting that the divergence between time-series and cross-sectional estimates is driven by measurement errors. The paper also considers differences between high-technology and other firms. Cross-section and panel elasticity estimates of the R&D effect diverge considerably for the two groups, while the corresponding rate of return estimators display far less variation. There is some evidence that the R&D elasticity increased during the early 80s, and that it fell sharply back to its 1979 value during the period from 1985 to 1989.

Acknowledgements
I would like to thank seminar participants at NBER, in particular Zvi Griliches, Bronwyn Hall, and Jaques Mairesse for valuable comments on earlier versions of this paper. Gebhard Flaig and Georg Licht contributed numerous helpful suggestions. Remaining errors are, as ever, my responsibility. The author owes a particular acknowledgement to Christoph Grenzmann, Rüdiger Marquardt and Joachim Wudtke at the *Stifterverband für die Deutsche Wissenschaft* for their cooperation and support in assembling the data set. The data analysis was based on anonymized data and took place on the premises of the *Stifterverband*. Any opinions, findings, and conclusions expressed here are those of the author and do not necessarily reflect the views of the *Stifterverband*. This research has been supported by research grants from the Federal Ministery for Research and Technology (BMFT) and the Zentrum für Europäische Wirtschaftsforschung (ZEW).

1. Introduction

The relationship between R&D and productivity has attracted a great deal of attention over the last thirty years. Recently, this interest has been revived due to the emergence of theories of endogenous economic growth in which private R&D and R&D-related knowledge spillovers play a central role. Nonetheless, it is probably fair to say that our knowledge with respect to the productivity effects of R&D is still limited. Recent surveys of the main empirical results and methods employed in estimation have been presented by Mairesse and Mohnen (1994) and Mairesse and Sassenou (1991). Responding to the perception that productivity growth has been weak over the last two decades, some researchers have explicitly asked whether R&D has lost its potency during this period.

While there have been studies on this issue in virtually all of the major industrialized countries, the case of productivity and R&D in Germany has remained largely unexplored so far.[1] One of the key problems in the past has been the availability of suitable data for a detailed econometric analysis. The lack of results comparable to those surveyed by Mairesse and Sassenou has been unfortunate, since private R&D expenditures per capita in Germany are relatively high in comparison to those in the U.S. and France and should thus provide another interesting test case for exploring the relationship between R&D and changes in productivity.

This paper presents first estimates using firm data collected by the science statistics branch of the *Stifterverband für die Deutsche Wissenschaft*, a non-profit organization that has been surveying the R&D activities of German manufacturing firms for over thirty years. I report in some detail on the steps undertaken so far to construct a firm panel from seven cross-sections of data and explore the suitability of this data set by producing regressions results that can be compared to those published for the U.S., France and Japan. While the initial samples are quite large (between 1352 and 2858 firms per year), the sharpest constraint at this point is the frequent lack of investment data which are used to construct a measure of the capital stock. Thus the size of the panel constructed here (443 firms) is small in comparison to the initial data set, but the sample represents about 50 per cent of all private R&D in the manufacturing sector. Therefore, this dataset provides an interesting starting point for an investigation of the R&D-productivity relationship in German manufacturing firms.

By and large, the results obtained below are similar to those found by Griliches and Mairesse (1984) using comparable U.S. data. The fact that the *Stifterverband* collects data only every second year seems to be less of a concern than originally anticipated. Moreover, the lack of a capital stock variable in the data set does not appear to cause major problems, since gross investment data can be used to construct a reasonably reliable measure of the capital stock. On the positive side, the *Stifterverband* data contain a very detailed breakdown of a firm's R&D budget such that corrections for the double-counting of R&D employees and investment in R&D

---

1 Some researchers have studied the effect of R&D on productivity at the industry level or at the level of individual enterprises. See for example Brockhoff (1986). However, to the best of my knowledge there have been no empirical studies comparable to those surveyed by Mairesse and Sassenou (1991) and Mairesse and Mohnen (1994).

capital and labor or capital, respectively, can be performed easily. Not surprisingly and consistent with some of the earlier papers, the results demonstrate that the corrections for double-counting are quite important. Other corrections explored here, e.g. for reductions in the average number of hours worked per employee, have only minute effects on the estimates. Differences between the within and cross-sectional estimates are persistent, but within and differencing estimates are quite similar. Moreover, estimations using "long differences" yield results that are relatively close to cross-sectional regression coefficients.

The paper proceeds in four sections. In section 2, I describe the data source and the steps undertaken to construct a panel data set suitable for studying the relationship between productivity and R&D. Section 3 briefly reviews the econometric framework for the analysis. The respective regression results are presented in section 4. The final section concludes and elaborates on future research to be conducted on the basis of the data set described here.

2. Data Source and Construction of the Panel

2.1 Data Collection

The data used in this paper originate with the science statistics section of the *Stifterverband für die Deutsche Wissenschaft*. Since 1948, this organization has been collecting information on R&D expenditures of West German firms. The *Stifterverband* survey is the basis for the offical R&D statistics of the Federal Republic of Germany and for the respective figures reported to the OECD and the statistics office of the European Community. The *Stifterverband* survey is administered every two years and contains several questions regarding the composition of a R&D budgets. More specifically, the survey asks responding firms for information on

- total number of employees,
- revenues,
- gross investment,
- total R&D expenditures,
- internal R&D expenditures,
- external R&D expenditures,
- the composition of internal R&D expenditures by investment, personnel expenditures, and other current expenditures,
- the composition of R&D personnel by number of scientists and engineers, technicians, and other R&D employees,
- the percentage of R&D spending dedicated to basic research,
- the distribution of applied internal R&D over 34 product classes,
- and the distribution of the firm's R&D employees by geographic area.

Furthermore, in 1977 and in all surveys administered since 1987 the questionaire contained questions regarding the orientation of R&D efforts towards product or process R&D and a question regarding the sales contributions of new products. The data used here cover the years 1977 through 1989 and consist initially of seven large cross-sections. Prior to this study, the various cross-sectional data sets have not been analyzed econometrically nor has any attempt been made to construct a panel data set. Thus, the paper describes the first attempt to utilize this data set in an econometric study of R&D and productivity at the firm level in German manufacturing enterprises.

While the data set used here has been designed as a sequence of cross-sections for the purpose of estimating the overall R&D expenditures in German industry, the data collection procedures used by *Stifterverband* are nonetheless quite conducive to generating a panel from the data, since longitudinal plausibility checks have been made for most firms covered in the sample. There is one important exception to this rule, though. Until 1983, data collection in the chemical industry (including pharmaceuticals) was largely managed by the industry's employer association which did not report the identity of the respective enterprise and did not use identical identification numbers over time. Clearly, this practice did not conflict with the original purpose of the data collection - the estimation of total R&D expenditures by industry - but it reduces the representation of chemical and pharmaceutical firms in the panel data set constructed here. The implications of this and other problems encountered in assembling and "cleaning" the panel are described in more detail in the appendix.

Several other features of the data set impose limitations on the empirical analysis. First, there is no measure of value added which would be preferable to revenues for the purpose of exploring the relationship between productivity and R&D. Hence, the dependent variable in the regression analysis below is revenues or revenues per employee.[2] Second, since the questionaire is only administered every two years, we have at best six observations for each firm. In many cases, the number of observations is actually smaller. Third, for the construction of R&D capital stocks it would be desirable to have access to pre-sample information on R&D expenditures which is not available at this point. Finally, while the data include the firms' gross investments in physical capital, the data set does not contain the value of the firms' capital stocks, e.g. in the form of book value of capital.[3] Hence, an approximation has to be used to impute initial capital stocks from investment data. In essence, the problem is similar to that of computing R&D capital stocks. However, since data on capital stocks are available at a rather detailed industry level, the computation of an initial capital stock is considerably easier than in the case of R&D stocks. The

2 Cuneo and Mairesse (1984) compare sales and value added as dependent variables and find some evidence of a missing variables bias when materials are not included as an additional input in the sales regression. The effect is most pronounced in the within dimension. More recently, Mairesse and Hall (1994) find in their data that using sales as a dependent variable without correction for materials on the right-hand side might be preferable to a regression on sales with inclusion of materials.

3 Of course, the book value of capital is not necessarily a good measure of the *actual* capital stock, either.

following subsections address the deflating procedure, the computation of stock variables, and the correction for double-counting in the labor and capital variables.

2.2 Deflating the Data

Sales were deflated using price indexes computed from data published by the Federal Statistical Office (*Statistisches Bundesamt 1989*).[4] All price indexes and other aggregate data used below were computed at the two-digit SYPRO level which is somewhat more detailed than the U.S. two-digit SIC classification. Investment figures were deflated using the investment goods price index from the same source. Based on the results of a study by Brockhoff and Warschkow (1991), the same index was used to deflate R&D expenditures.[5]

2.3 Computing Capital Stocks

The data used in this study did not include information on a firm's book value of capital or other variables that could be used as a proxy for the stock of physical capital. However, data regarding the firm's investment in physical capital is available. Once an initial capital stock has been determined, the capital stock in subsequent years can readily be computed using the perpetual inventory method, leaving aside the problem of imputing investment for the years between surveys. The capital stock also has to be corrected for the double-counting bias in investment. Let $C_{i,1}*$ be an uncorrected measure for the initial capital stock of firm i in period 1. Assuming that firm i spends a constant share α_i of their total investment on R&D related investments, the corrected capital stock $C_{i,1}$ can be written as

$$(1) \qquad C_{i,1} = C_{i,1}* (1-\alpha_i) \quad .$$

As in Mairesse and Cuneo (1984), the firm's share of investment α_i is estimated by taking the mean share of R&D related investment in the firm's total investment over all available observations.[6]

The computation of the uncorrected initial capital stock itself can be solved in various ways. In order to explore the implications of this choice, I used two alternative estimates. The first and simplest method is to proceed as in the case of R&D capital and to assume that the capital accumulation process has been going on for a

4 I would like to thank Georg Licht for providing the deflating indices used in this paper and the time series on industry capital stocks, investment, and depreciation.

5 Brockhoff and Warschkow (1991) show that the aggregate investment index performs relatively well when compared to an alternative price index computed on the basis of a detailed breakdown of R&D expenditures.

6 Again it would be ideal to have pre-sample information, but the use of firm-specific sample averages is clearly preferable to industry-specific pre-sample information.

sufficiently long time with a fixed real growth rate g_I of investment and at constant depreciation τ such that the approximation

$$(2)\qquad C1_{i,1}^* = I_{i,0}^* + (1-\tau)I_{i,-1}^* + (1-\tau)^2 I_{i,-2}^* + \ldots \quad = \sum_{r=0}^{\infty}(1-\tau)^r I_{i,-r}^* \\ = \frac{I_{i,1}^*}{g_I + \tau}$$

is applicable. $I_{i,t}^*$ denotes firm i's uncorrected gross investment in period t. Investment growth in the decade prior to the sampling period was estimated at 2 per cent and depreciation rates were computed at the industry level from aggregate national accounts. Since investment tends to be rather volatile, it may not be advisable to base this estimation on a single year.[7] Hence, to estimate the initial capital stock according to (2), the average of the first two years' investment in physical capital was used.

The second approach followed here is to assign each firm a share of the respective industry's net capital stock in the initial year of the time series. Firm i's share in industry j's capital stock was computed as the ratio of the firm's gross investment over the industry's gross investment. Again, early experimentation showed that it was preferable to use an average ratio computed from the first two years of the time series to estimate the capital stock, i.e.

$$(3)\qquad C2_{i,1}^* = \frac{1}{2}\left(\frac{I_{i,1}^*}{I_{j,1}^*} + \frac{I_{i,2}^*}{I_{j,2}^*}\right)C_{j,1}^*$$

These two alternative estimates for the initial capital stock were then corrected for double-counting as indicated in equation (1). For the following years, the usual perpetual inventory method was applied to the two measures, using the industry- and time-specific depreciation rates[8] for physical capital and the *corrected* measure for investment where investment related to R&D activities is subtracted from investment in physical capital:

$$(4)\qquad C_{i,t+1} = \left(1-\tau_{j,t}\right)C_{i,t} + I_{i,t}$$

Using a linear approximation of the intermediate investment figures

$$(5)\qquad I_{i,t+1} = 0.5\left(I_{i,t} + I_{i,t+2}\right)$$

7 Clearly, it is also not advisable to base the estimation on the average of all sample years, since the capital stock path will be smoothened artificially for firms with growing or declining shares in the industry's capital stock. Early experiments led to the conclusion that using the first two years of each time series was a preferable choice.

8 The depreciation rates of capital τ were computed on the basis of data on industry capital stocks and estimated depreciation within industries.

yields

$$C_{i,t+2} = (1-\tau_{j,t+2})(1-\tau_{j,t+1})C_{i,t} + (1-\tau_{j,t+2})I_{i,t} + 0.5(I_{i,t} + I_{i,t+2}) \tag{6}$$

In addition to the two corrected measures *C1* and *C2*, I also computed the corresponding uncorrected variables *C1** and *C2** in order to evaluate the size of the double-counting distortion in the regression coefficients. The two variables are obtained by taking the uncorrected proxy for the initial capital stock in year 1 and applying the perpetual inventory method, using uncorrected investment data.

The two alternative ways of computing a proxy for the firm's capital stock (net of R&D components) share two major weaknesses which are due to the nature of the data. First, there is no firm-specific pre-sample information which would allow us get a more precise estimate of the capital stock in period 1. Second, the imputation of investment between survey years is not a trivial step, in particular with respect to time series-estimates of the R&D-productivity relationship which tend to be highly sensitive anyway. Imputation in this form is likely to introduce measurement problems beyond the extent usually found in R&D data at the firm level. At this point, an extension of the imputation method would be to correct the interpolation by taking known industry-specific growth rates into account. This step was not undertaken so far, since the concomitant improvements are likely to be small. Absent any information on gross or net capital for firms in the sample, a more refined procedure would be the use of an overlapping samples approach where the complementary sample would contain information on firms' capital stocks. Clearly, the ideal solution would be to have access to capital stock proxies from financial statements of the sample firms, but confidentiality constraints make this approach infeasible, at least for the time being. Note that the current computation of the capital stocks requires a complete (i.e. contiguous) time series of investment figures. As the description in the appendix points out, this requirement is currently the strongest constraint on obtaining a larger sample.

2.4 Computing R&D Capital Stocks

The knowledge stock variable K is computed using the perpetual inventory definition

$$K_{i,t} = (1-\delta)K_{i,t-1} + R_{i,t-1} \quad . \tag{7}$$

As in the case of physical capital, this specification is based on a "time to build" assumption. Since the Stifterverband survey is conducted only every second year, intermediate values of R&D expenditures are again approximated by linear interpolation

$$R_{i,t-1} = 0.5(R_{i,t-2} + R_{i,t}) \tag{8}$$

Rewriting (7) using this approximation we get

$$(9) \qquad K_{i,t} = (1-\delta)^2 K_{i,t-2} + (1-\delta) R_{i,t-2} + 0.5\left(R_{i,t-2} + R_{i,t}\right)$$

It is well-known from a number of studies that the particular choice of depreciation rates does not affect the estimation results greatly. Nonetheless, in order to check the robustness of the above approximations, I calculated the knowledge capital stocks of firms for depreciation rates of 15 and 25 per cent. Thus, there are two alternative measures for the R&D capital stock (K15 and K25).

The initial knowledge capital stock is computed using the pre-sample (real) R&D growth rate g_R of 0.059, computed over the period of 1967 to 1977. As 1967 to 1977 data for firms in this sample become available, it will be possible to test the quality of this approximation more thoroughly. For the time being, I compute the initial knowledge stock in 1977 using the definition

$$(10) \qquad K_{i,0} = \frac{R_{i,0}}{\delta + g_R}$$

For firms not being surveyed in 1977, the same procedure is used to compute the starting value for either 1979 or 1981. Note that the effect of the initial stock of knowledge capital is diminishing over time, while the choice of depreciation rate remains relevant no matter how long the times series is.

The parallels to the computation of the stock of physical capital are obvious, and so are the problems coming with it. However, the problem of producing more accurate measures of between-survey R&D expenditures may be solvable, using responses to another question on the *Stifterverband* questionaire. The questionaire asks respondents for the estimated growth rate of their R&D expenditures in the year between surveys. Since the survey is conducted roughly in the middle of a year, one can argue that factors determining the development of R&D expenditures should be known by then, with the exception of surprise events like surging energy prices etc. Thus, it seems promising to include the anticipated growth rates in the computation of the R&D capital stocks.

2.5 Corrections for the Labor Variable

In this paper, I use two labor measures and explore the implications of using them alternatively in the productivity estimation. The first follows the usual practice of computing labor input into production as the number of full-time employees *minus* the number of employees working in research and development. The correction for R&D employees have been shown to be quite important (Hall and Mairesse 1995, Cuneo and Mairesse 1984).

The second measure corrects for differences in average employee-hours per year at the industry level. While the number of actual hours worked per year and employee differ over time in every country, it is well-known that German manufacturing has experienced a particularly steep reduction in the number of hours worked. Over the sample period 1977 to 1989, the average number of hours worked dropped from

1768.9 to 1637.7 which represents a 7.4 per cent reduction. I computed the second labor variable (denoted LHRS) by taking 1985 as the base year and multiplying the number of full-time employees minus R&D employees by an index figure, defined as the ratio of hours per employee in the respective year and industry divided by hours worked per employee in 1985 in the respective industry.

2.6 A Descriptive Account of the Constructed Data Sets

Using cleaning procedures similar to those employed by Hall and Mairesse (1995), I constructed two panels. The first is a balanced panel with 190 firms and observations for each of the six sample years. The second panel - containing the first one - is slightly unbalanced with 443 firms and 2257 observations in total. In this section I will briefly point out a few characteristics of the data set, while details of the data set construction and cleaning procedures are described in the appendix.

Table 1 describes the sectoral breakdown of the initial sample of manufacturing firms and of the two samples constructed by applying the cleaning procedures outlined in the appendix. Since there are only few sectors with sufficient observations for industry-specific estimations, I use a distinction between high technology and other firms suggested by Legler et al. (1992, p. 38) in order to test for significant differences in the potency of R&D. The classification used here is based on four-digit product area (SYPRO) classifications. The group of high technology firms includes (among others) producers of pharmaceuticals, organic chemicals, electronic devices and components, optical devices, data processing equipment, automotive vehicles, and producers of particular mechanical engineering products, like machine tools and electrical machinery.

The original manufacturing sample is dominated by firms in mechanical engineering, electrical machinery and chemicals. The precise number of sample firms in the chemicals and pharmaceuticals sector could not be determined so far, since - unlike all other firms in the sample - most of the observations in this sector were not identified by a unique (i.e. time-invariant) identification number and thus could not be linked to a time-series (see the appendix for further details). Due to the relatively small number of chemicals and pharmaceuticals producers in the constructed samples, the R&D intensity of my two samples is likely to be below that of other data sets described in this literature. Furthermore, the number of firms producing data processing, electronics and computer equipment is quite small in comparison to the Hall and Mairesse (1995) and the Griliches and Mairesse (1984) samples. The last column in Table 1 indicates the number of high technology firms in the unbalanced panel. By definition, these firms are mostly found in the chemical and electrotechnical sectors as well as in mechanical engineering.

Table 2 presents an extensive set descriptive statistics (in 1985) for the manufacturing firms in the original Stifterverband dataset, for the balanced panel, the unbalanced panel, and the high technology firms in the unbalanced panel. The differences between the 1985 cross-section of the initial manufacturing sample and the two panel data sets become quite apparent in this table. Note first that the difference between initial sample and unbalanced panel with respect to R&D intensity is relatively small, while firms in the balanced panel tend to be more research-intensive. Size differences between the samples are quite pronounced: the median number of employees is 266 for the initial manufacturing sample, 869 employees for the unbalanced panel, and 1610 employees for the balanced panel. High technology firms in the unbalanced panel are characterized by relatively high R&D expenditures and respective capital stocks, a high share of scientists and engineers among R&D employees and relatively small stocks of physical capital.

The two capital stock variables C1 and C2 have very similar median values and interquartile ranges. Moreover, the correlation between the two adjusted capital stock measures is already high in the initial year (ρ=0.9952) and increases further until 1989 (ρ=0.9993) when the measures become virtually collinear. This high correlation measure is reassuring, since the measures for the intial year have been constructed in fundamentally different ways. Regression results using these variables alternatively as exogenous regressors should not be expected to differ substantially. The correction for double-counting of R&D-related investments reduces the capital stock variables (at the median) by about 3.5 per cent in the balanced panel and by about 5 per cent in the unbalanced panel. The correction of the labor variable is on the order of 4.5 per cent in the balanced panel, and 4 per cent in the unbalanced panel.

As in the US and France, the bulk of private R&D expenditures is accounted for by a relatively small number of enterprises. In 1985, total private R&D expenditures were estimated to be DM 39.55 Billion. Manufacturing firms expended DM 36.4 Billion on research and development. The 190 firms in the balanced panel account for DM 13.5 Billion of this amount (37.2 per cent). Firms in the unbalanced panel upon which most of the analysis will be based account for DM 17.9 Billion and thus for 49.3 per cent of 1985 total R&D expenditures in West German manufacturing firms. While selectivity remains a problem, it is comforting that about half of all private R&D expenditures in German manufacturing are covered in this study.

3 The Econometric Framework

3.1 Production Function Approach

The approach taken here is known as the production function framework and has been described in some detail by Griliches (1986) and Mairesse and Sassenou (1991), among others. It is assumed that a firm's production function can be approximated using a Cobb-Douglas specification

$$S_{it} = A_i \exp(\lambda t) C_{it}^{\alpha} L_{it}^{\beta} K_{it}^{\gamma} \exp(\varepsilon_{it}) \qquad (11)$$

where S is the firm's sales, the index i represents firms and t denotes one of the years 1979, 1981, 1983, 1985, 1987 and 1989. λ is the rate of exogenous technological change and A is a scaling parameter. C denotes physical capital, L represents labor input, and K is the value of the R&D (or knowledge) capital stock. Equation (11) can be estimated using standard OLS procedures by writing it in logarithmic form

(12) $$s_{it} = a_i + \lambda t + \alpha c_{it} + \beta l_{it} + \gamma k_{it} + \varepsilon_{it}$$

where $s = log\ S$, $c = log\ C$, $a = log\ A$, $l = log\ L$, and $k = log\ K$. In some of the estimations, I follow the usual practice of transforming this relationship such that deviations from constant returns to scale can be assessed directly from the estimates. Subtracting the labor measure on both sides of equation (12) yields

(13) $$s_{it} - l_{it} = a_i + \lambda t + \alpha(c_{it} - l_{it}) + \gamma(k_{it} - l_{it}) + (\mu - 1)l_{it} + \varepsilon_{it}$$

where $\mu = (\alpha + \beta + \gamma)$. The CRS assumption (constant returns to scale) can be rejected when the coefficient on labor input in (13) is significantly different from zero. Moreover, in this specification it is possible to impose CRS by dropping the labor variable from the regression altogether.

Time effects will be captured by including dummy variables for each of the survey years, except for the base year 1979. Furthermore, in regressions using pooled cross-sections, industry dummies will be introduced as regressors in order to limit the effect of sectoral sample composition on the estimation results. Since there is reason to assume that much of the heterogeneity across firms is latent and will not be captured using the above framework, I also use two alternative estimators. First, equation (13) will also be estimated using a fixed effects (within) estimator. If latent heterogeneity is not a problem, then the estimation results should be consistent with the cross-section results. However, there may be other (and more important) reasons why the estimators differ. In particular, the presence of measurement errors should affect time-series estimators to a greater degree than coefficients from a cross-sectional regression (Mairesse 1992). I will focus on this particular possibility although a number of other econometric problems could conceivably drive a wedge between the two types of estimates.[9]

A second alternative to estimate the model consistently - even in the presence of firm-specific effects that are correlated with the other regressors - is the use of differenced data. For examples, estimation in one-period differences involves the regression equation

(14) $$\Delta_1(s_{it} - l_{it}) = \lambda\Delta_1 t + \alpha\Delta_1(c_{it} - l_{it}) + \gamma\Delta_1(k_{it} - l_{it}) + (\mu - 1)\Delta_1 l_{it} + \Delta_1\varepsilon_{it}$$

where $\Delta_1 g_t = g_t - g_{t-1}$.

Taking this approach amounts to a regression of productivity growth (or output growth) rates on input growth rates. Taking "longer differences" is equivalent to "averaging out" serially uncorrelated measurement errors, since $\Delta_2 g_t = g_t - g_{t-2} = \Delta_1 g_t + \Delta_1 g_{t-1}$.

9 See for example Griliches and Mairesse (1984) and Klette and Griliches (1992).

3.2 Returns to R&D Estimation

An alternative approach used in the R&D literature is the measurement of the rate of return to R&D. Rather than assuming that the technical coefficient of R&D capital in the production function is the same for all firms, this approach postulates a homogeneous rate of return to R&D. However, as Hall and Mairesse (1995) have noted, there are a number of problems with this approach. Suppose that the simple production function formulation in logarithmic growth rates is given by

$$\Delta s_{it} = \lambda \Delta t + \alpha \Delta c_{it} + \beta \Delta l_{it} + \gamma \Delta k_{it} + \Delta \varepsilon_{it} \quad . \tag{15}$$

where $\gamma = (\partial S_{it} / \partial K_{it})(K_{it} / S_{it})$. Hence, we can simply restate (15) as

$$\Delta s_{it} = \lambda \Delta t + \alpha \Delta c_{it} + \beta \Delta l_{it} + \frac{\partial S_{it}}{\partial K_{it}} \frac{K_{it} \Delta k_{it}}{S_{it}} + \Delta \varepsilon_{it} \quad . \tag{16}$$

If the rate of return of R&D (i.e. the marginal effect of R&D capital on output $\partial S_{it} / \partial K_{it}$) is assumed to be constant across firms, then it can be estimated directly from equation (16). Some studies assume that the term $K_{it} \Delta k_{it}$ can be approximated by the flow of R&D R_{it}, the R&D expenditures of firm i in period t, hence

$$\Delta s_{it} = \lambda \Delta t + \alpha \Delta c_{it} + \beta \Delta l_{it} + \varrho \frac{R_{it}}{S_{it}} + \Delta \varepsilon_{it} \tag{17}$$

Since this approximation abstracts from depreciation of R&D capital, the respective estimate is often referred to as a "gross rate of return" to R&D capital. Typically, researchers argue that the net rate of return can be computed by subtracting the rate of depreciation from the estimated gross rate. This calculation obviously requires knowledge of the depreciation rate, as Hall and Mairesse (1995) point out. Moreover, this specification assumes a contemporaneous relationship between R&D flow and productivity growth which is not very plausible.

Goto and Suzuki argue that - given estimates of the knowledge stock - it might be preferable to estimate a variant of equation (16).

$$\Delta s_{it} = \lambda \Delta t + \alpha \Delta c_{it} + \beta \Delta l_{it} + \varrho \frac{\Delta K_{it}}{S_{it}} + \Delta \varepsilon_{it} \tag{18}$$

in which $K_{it} \Delta k_{it}$ is approximated by ΔK_{it}. Since the net stock of R&D capital is used in this specification, the estimated coeffient can be interpreted as the net rate of return to R&D. The relationship between the gross and net rate of return estimates has been controversial. If the interpretation of the above-mentioned gross rates of return is correct, then the estimated gross rate should roughly equal the net rate plus depreciation. Clearly, the difference between gross and net rate should be positive in this case. However, as Hall and Mairesse (1995) point out, one can just as well argue that the net rate of return should be equal across firms. Since the gross flow of R&D should be proportional to and about 3 to 4 times larger than the net flow (see their footnote 20), the coefficient for the net rate of return should be three to four times larger than the gross rate estimator. So far, the empirical results have not been conclusive. Goto and Suzuki (1989) find results in which the gross rate exceeds the net rate for some industries and is smaller than the net rate for others. However, their sample is quite small, hence the results should be interpreted with caution. Hall and Mairesse (1995) also find rather puzzling results in their application of the rate of returns approach. In their sample, gross and net rates are about equal which is not consistent with either of the two arguments.

To provide results that are comparable to these studies, equations (17) and (18) will be estimated for all firms and separarely for the two sub-groups (high technology and other firms). As mentioned before, measurement errors are likely to produce inconsistent estimation results if "short differences" are used. Thus, the estimation is done for 2- and 6-year differences to explore the effect of different specifications. Moreover, to avoid simultaneity biases I use R&D intensity measures computed from average R&D flows and past sales values in order to avoid the appearance of the output variable in the denominator of the flow variable.[10]

4. Estimation Results

4.1 Production Function Framework

Table 3 presents cross-sectional and fixed-effects panel estimates of equation (13), using alternative definitions for the R&D capital, physical capital, and labor variables. The upper two sets of coefficients in this table summarize the results from using the balanced panel with 1140 observations while the lower two sets describe the corresponding coefficients for the unbalanced panel (2257 observations). Since alternative choices for the capital stock and labor variables yield only very minor differences in the results, a comparison across all variables is presented only in Table 3.

Without inclusion of industry dummy variables (not shown in Table 3), the capital coefficient is about 0.28 and the R&D capital coefficient is on the order of 0.07. Inclusion of the industry controls reduces the capital coefficient to 24 per cent in the balanced panel and 23 per cent in the unbalanced panel, while the R&D capital coefficient increases to 16 per cent in the balanced and about 14 per cent in the unbalanced panel. Since the industry controls are highly significant in all of the specifications used here, they are retained. It is somewhat surprising that the inclusion of industry dummy variables strenghtens the estimated coefficients for R&D capital, but it should be kept in mind that the dependent variable (sales) may be affected considerably by industry characteristics, e.g. the extent of vertical integration.

The cross-sectional estimates of the R&D elasticity appear to be relatively large, but still plausible. Moreover, they are stable with respect to changes in variable

[10] See Hall (1993) and Hall and Mairesse (1995) for details. The effect of using contemporaneous output measures in the denominator of the flow intensity is profound: it raises the coefficients by approximately one third of the values displayed in Table 6.

definitions. The results in Table 3 confirm the expectation that differences in the construction of alternative measures for the capital stocks do not matter greatly. There is also no discernable effect from applying an hours correction to the labor input variable. This is not to say that further corrections at the firm level would not be promising. But industry data are of little help here: the decline in the number of work hours follows a very similar time path in most industries and is therefore reflected mostly in the time dummies and (in the cross-section) in industry effects. As in many other studies, the choice of the depreciation rate of knowledge capital is of little relevance for the estimates obtained from the cross-sectional regressions. As other researchers have noted, there is little hope to determine the "correct" depreciation rate in this way.[11]

The corrections for double-counting do matter, both in the cross-sectional and the fixed-effects estimates. In the cross-section estimates using the unbalanced panel, the correction increases the R&D capital coefficient by about 3 per cent, while the within coefficient increases by about 1.5 to 2.2 per cent. This result is broadly consistent with the findings in Hall and Mairesse (1995), but the effect of the correction is smaller in my data than in their sample. Note that the within-estimates for the R&D capital coeffients become insignificant in the case of the balanced panel if no double-counting correction is used. The assumption of constant returns to scale is consistent with the estimation results for the balanced panel while there are very small increasing returns to scale in the larger panel.

The within estimates are characterized by large decreasing returns to scale which is consistent with virtually all of the results summarized by Mairesse and Sassenou (1991) and Mairesse and Mohnen (1994). While there are a number of potential explanations for this result, the most promising candidate appears to be measurement error. The various computations used in the construction of the variables required a number of interpolations (thus causing measurement error), but have been designed such as to minimize simultaneity problems. In the within-estimates both the capital and the R&D coefficients become smaller, but the effect is much stronger for the R&D variable. However, the R&D effects are still significant in all specifications employing the double-counting corrections. Taking the within-estimates at face value would lead us to place the R&D coefficient between 6 and 8 per cent.

In order to facilitate the discussion and to avoid the presentation of very similar estimation coefficients, the following results are based on the variables *C1* (initial capital stock based on the constant growth assumption), *K15* (knowledge depreciation rate of 15 per cent) and *L* (simple labor variable without hours correction). The correction for R&D employees and R&D investment is used in each case.[12] Moreover, I only present results for the more comprehensive unbalanced panel of 443 firms.

The differences between the totals and the within estimates are considerable and require further analysis. Interesting results bearing on these differences can be ob-

[11] See Klette (1994) for an alternative approach to the estimation of depreciation rates.

[12] More comprehensive results for various variable definitions and the balanced and unbalanced data set are included in a previous version of this paper which is available from the author upon request.

tained from comparing high technology firms to other firms. The results appear by and large consistent with those obtained by Griliches and Mairesse (1984) and Cuneo and Mairesse (1984) who have reported relatively high R&D effects on productivity for a subset of "scientific" firms in their respective datasets. In the data used here, the effect of R&D capital on output is considerably higher for high technology firms than for the residual group of enterprises. Table 4 also displays within-estimates with constant returns to scale imposed by dropping the labor variable from the regression. Note that the R&D results obtained with CRS imposed are virtually identical to the cross-sectional results.

From the results summarized in Table 4, we would again conclude that R&D has a significant and large effect on labor productivity. For high technology firms, the elasticity ranges between 11.2 (within estimates, CRS not imposed) and 16.5 per cent (cross-section). For the residual group of firms, the within estimates of the R&D coefficient become insignificant. If CRS is imposed or if the cross-sectional results are considered, the elasticity is on the order of 9.2 per cent and significantly different from zero.

If measurement errors are at least in part serially uncorrelated, then differencing estimates (or estimates in logarithmic growth rates) should become more precise once one uses longer time periods to compute growth rates. Table 5 reveals that there is indeed some evidence for this type of measurement problem. Using one-period differences yields somewhat astonishing results, since the labor variable becomes insignificant while capital and R&D coefficients are significant and of plausible size. Using longer differences, the labor coefficient becomes positive and reaches about 0.6 for the "long difference" estimates. The estimated coefficients for R&D and capital are stable throughout and remarkably similar to the within estimates. Using the distinction between high technology and other firms, the effect of taking longer differences is the same, and again the coefficients are very close to those obtained from the fixed-effects estimation without imposed constant returns to scale.

Deviations from constant returns are much smaller for most of the differencing estimates than in the case of the within estimates. Using six-year differences, the deviation from constant returns to scale is on the order of -.141 for all firms and -.115 for high-technology firms (-.274 and -.234 respectively, in the within regression). The hypothesis of constant returns is rejected for the overall sample at the 5 per cent level. For long differences, the deviation from CRS is -.105 for the complete sample of firms, and virtually zero for the high technology firms. Constant returns cannot be rejected at the 5 per cent level for high-technology firms, while there are still statistically significant decreasing returns to scale in the group of non-high-technology firms. These results would be consistent with the hypothesis that estimation of the initial capital stock and interpolations cause larger measurement errors for firms with relatively small R&D expenditures. Thus, the tentative conclusion is that errors-in-variables seem to have some effect on the estimates, and that taking longer differences helps to control this problem, in particular for the high-technology firms where CRS can no longer be rejected in the case of long-difference estimations.[13]

4.2 Rate of Return Estimates

The results using the returns to R&D approach are presented in Table 6, using 2-year and 6-year differences. As should be clear from the above, the 2-year-differencing results are not very reliable, in particular with respect to the estimate of the labor coefficient. The respective parameter estimates are not significant and very close to zero. However, the 2-year and 6-year growth rate specifications yield very similar estimates for the effect of physical capital and the rate of return to R&D. The estimated gross rates of return to R&D are remarkably stable across specifications and display considerably less variation across the two subgroups of firms than the elasticity estimates in Table 5. These differences are even less pronounced if equation (17) is used to estimate the net rate of return. Moreover, in most regressions the net rate specification performs slightly better (in terms of the R-squared value) than the gross rate estimation based on equation (16). It should be noted that these regressions do not control for materials and that the estimated rates have to be interpreted carefully. Assuming a materials share of .6, a coefficient of .8 in these results should be roughly consistent with a true net rate of return of .32. Note also that the net rate is roughly proportional to the gross rate with a proportionality factor of 3 to 4 in the overall sample. Thus, these estimates conform quite closely to the Hall and Mairesse (1995) expectations.

One should note that the rate of return estimates display less variation across the two subgroups of firms than the respective elasticity estimates. Nonetheless, in the case of the net rate estimations R&D appears to have a somewhat higher payoff for high-technology firms. The net rates estimated here also conform relatively well with those implied by the elasticity estimates. Assuming an elasticity estimate of about 9 per cent for all firms and evaluating the ratio of sales over R&D capital stock (K15) at the sample median yields an implied net rate of return of 0.66. The respective estimates are 0.77 for the high-technology firms, and 0.38 for the less technology-oriented firms. The deviation between directly estimated and implied net rates of return to R&D is strongest for the latter group of firms.

4.3 Time Trends in the R&D-Productivity Relationship

The results from the simple econometric frameworks described above clearly establish that R&D has had a positive effect on the productivity of West German manufacturing firms. But irrespective of the average contribution of R&D, a question

[13] An alternative explanation for the large deviations between cross-sectional and time-series estimates is a simultaneous relationship between output and labor. One can investigate the effects of simultaneity by using the semi-reduced form approach described by Griliches and Mairesse (1984). It turns out that the deviations from CRS for the within estimates are not reduced substantially by following this approach. These results are available upon request.

of some concern to researchers and policy-makers alike is the development of R&D and productivity nexus over time. As Griliches (1986) has pointed out, there is little evidence of a slackening in the potency of R&D for the period from 1960 to 1980. However, the story may be quite different for the 80s. Hall (1993) has provided evidence that both the elasticity of output with respect to R&D capital and the gross rate of return have been decreasing dramatically in the U.S. Moreover, the stock market valuation of R&D in the U.S. has decreased to a similar extent during this period.

Little is known about the German case, since this study is the first one to cover a time-period of more than ten years using a relatively large sample. Nonetheless, there have been interesting developments in the aggregate data. Since 1989, real private R&D expenditures in Germany have been declining for the first time since official R&D statistics were published in 1965. R&D personnel figures have been on the decline as well, thus establishing a pattern that resembles the recent decline in real R&D expenditures in the U.S. (see Stifterverband 1994 and NSF 1994 for details).

While it may be hard to disentangle the effects of German reunification, the ensuing recession in 1992 and more general trends in corporate reorganization, it may nevertheless be interesting to look at the time patterns up to 1989. It should be noted that the actual decline in R&D expenditures and personnel did not yet occur in this period, and that the data available at this point do not allow me to draw strong conclusions. Nonetheless, some suggestive results can be provided. To exploit the time dimension of this rather short panel, I use within-regressions in which I let the capital and the R&D capital coefficient vary for each of the survey years. Constant returns to scale are imposed, but qualitatively very similar results emerge from regressions without this restriction. The results for the overall sample and for the two sub-samples are displayed in Table 7. The capital coefficient for the overall sample ranges between .280 and .347, but there is no apparent time pattern in the data. Looking at the two subsamples, the results are similar. The elasticity estimate is increasing towards the end of the sampling period, but the capital elasticities for the high technology firms are significantly higher than those for other firms. The estimates for the R&D capital elasticities do suggest a distinctive pattern over time, however. The estimates increase until 1985 and fall rather dramatically during the period from 1985 to 1989. The pattern emerges from estimates based on the full sample as well as from separate regressions with the two subsamples. While there may be a number of reasons for the initial rise and the subsequent decline of the R&D elasticity[14], it is at least suggestive that the decline since 1985 precedes a notable cutback in private R&D expenditures starting in 1989. Unfortunately, the currently available panel is too short to draw strong conclusions or to sharpen the empirical tests. However, the need for a more detailed analysis and search for potentially important changes in the underlying private R&D incentives in Germany is emphasized by these results.

14 The result is not driven by sample composition effects in the unbalanced panel. The time trend in the R&D elasticity and the lack of any apparent time trend in the capital coefficient are also clearly visible in regression results with the balanced panel.

5. Conclusions and Further Research

Taken together, the results provide strong confirmation for the view that R&D is an important determinant of productivity at the firm level. Despite some innate weaknesses of the data used in this paper, the overall picture is encouraging. The cross-sectional estimates of the effect of R&D on productivity are consistent with evidence from previous studies. In the cross-section, the elasticity of sales with respect to knowledge capital is on the order of 14 per cent. This estimate is weakened considerably in time-series estimates, but remains positive and significant. Using within (fixed effects) or differencing estimators yields R&D elasticities of about 8 per cent in combination with implausible decreasing returns to scale. The latter result is probably caused by measurement errors which affect panel estimates to a greater degree than cross-sectional results. This hypothesis is supported by differencing estimates which yield smaller deviations from decreasing returns to scale when longer periods for the computation of growth rates. Corrections for double-counting of R&D personnel and R&D-related investments have a significant effect on both cross-sectional and panel estimates, reducing the estimated elasticities by up to 3.5 per cent.

Contrary to previous studies, the R&D elasticity estimates differ considerably between high-technology and other firms, both in cross-sectional and in time-series estimations. For high technology firms, the estimated elasticities are about 16 per cent in cross-sections and about 12.5 per cent in time series estimates. For the residual group of firms, the respective elasticities are 9 per cent and 3.9 per cent, with time series estimators no longer being significantly different from zero. The differences between rates of return to R&D between high-technology and other firms appear to be relatively small, however. While there may be considerable problems with the interpretation of rate of return coefficients (see Hall and Mairesse 1995), the assumption of a homogeneous rate of return seems more reasonable in the case of this sample than the assumption of a production function that holds for all firms.

An analysis of time trends in the R&D-productivity relationship cannot yield strong evidence at this point, since the sample ends at the beginning of a decline in real private R&D expenditures in manufacturing. Nonetheless, the data suggest that the R&D elasticity increased during the early 80s, and that it fell sharply back to its 1979 value during the period from 1985 to 1989. If this decline can be substantiated, then further investigations, preferably comparing the development at the end of the 80s in a number of industrialized countries, should prove particularly instructive.

Table 1
Sectoral Breakdown of the Sample
(Number of Firms by Industry)

Industry	Industry Code[3]	Product Group Code	Manufacturing Sample[1]	Balanced Panel	Unbalanced Panel (All Firms)	Unbalanced Panel (High-Tech Firms)
Chemicals & Pharmaceuticals[2]	200	40	1263	9	10	9
Plastics & Rubber Products	210, 213	58, 59	127	6	12	0
Stones & Clay, Abrasive Materials, Ceramic Products	221-223, 224, 226	25, 51	100	10	23	0
Glass & Glass Products	227	52	23	3	5	0
Iron and Steel, Non-ferrous Metals, Foundries	230-232, 233, 234, 236	27, 28, 29	144	17	37	0
Metal Drawing, Steel Construction etc.	237-239, 240, 241	30, 31	184	7	21	0
Mechanical Engineering	242	32	1305	71	169	146
Data Processing and Office Equipment	243	50	63	1	6	6
Road & Railway Vehicles	244-249	33	117	11	21	8
Electr. Machinery	250, 259	36	673	29	71	70
Precision and Optical Instruments	252,254	37	160	12	19	16
Structural Metal Products, Musical Instruments, Toys	256, 257, 258	38, 39	175	5	22	0
Wood Processing and Products	260, 261	53, 54	59	2	6	0
Paper Processing & Products, Printing & Duplication	264, 265, 268	55, 56, 57	74	2	5	0
Textiles, Apparel, Leather Products	271, 272, 275, 276	62-64	120	1	8	0
Food & Tobacco	280-289	68, 69	101	4	8	0
Total			4688	190	443	255

Notes
1. Definition of the Manufacturing Sample as in the appendix (without 10 firms in the oil refining industry).
2. The number of chemical and pharmaceutical firms in the manufacturing sample is inflated, since cross-sectional data could not always be matched for panel construction. See the appendix for a more detailed explanation.
3. The correspondence between industry and product classification is only approximative.

Table 2

Descriptive Statistics
(After Cleaning and Deflation)

Median (Interquartile Range)

Variable	1985 Manufacturing Sample[1]	Balanced Panel 1985	Unbalanced Panel 1985 (All Firms)	Unbalanced Panel 1985 (High Technology Firms)
R&D Expenditures (R)	0.8 (0.3, 2.5)	5.9 (1.5, 28.0)	2.4 (0.7, 10.5)	2.6 (0.9, 12.3)
R&D/Sales (RS)	2.20% (0.9%, 4,5%)	2.6% (1.2%, 4.4%)	2.3% (1.0%, 4.1%)	3.1% (1.8%, 4.9%)
Sales (S)	42.7 (16.2, 125.0)	295.9 (69.2, 1036.0)	117.7 (44.0, 490.0)	93.1 (37.9, 416.6)
Employees (L*)	279 (109, 750)	1612.5 (449, 4786)	756 (305, 2593)	712 (278, 2481)
Non-R&D Employees (L)	266 (101, 721)	1544 (432, 4593.4)	735 (294.5, 2486)	654 (271, 2322)
Total Investment (I*)	1.7 (0.5, 7.6)	11.4 (3.6, 50.8)	5.3 (1.4, 21.0)	3.9 (1.2, 18.0)
Non-R&D Investment (I)	1.5 (0.4, 7.1)	11.1 (3.4, 47.7)	5.1 (1.4, 20.8)	3.8 (1.0, 17.9)
Share of Scientists among R&D Employees	31.8% (20.0%, 47.1%)	32.9% (23.0%, 45.7%)	32.1% (21.1%, 45.5%)	35.8% (24.7%, 48.9%)
R&D Capital Stock (K15)	-	27.8 (7.0, 125.9)	12.9 (3.4, 56.4)	13.5 (4.1, 67.1)
R&D Capital Stock (K25)	-	18.9 (4.54, 87.9)	8.6 (2.4, 38.0)	9.1 (2.8, 48,1)
Adj. Capital Stock (C1)	-	112.2 (25.2, 376.3)	41.2 (13.9, 176.7)	30.9 (11.4, 165.1)
Unadj. Capital Stock (C1*)	-	113.9 (26.3, 390.6)	43.9 (14.7, 190.1)	32.4 (11.3, 169.2)
Adj. Capital Stock (C2)	-	113.5 (25.3, 408.6)	43.3 (13.7, 185.2)	31.2 (11.2, 163.9)
Unadj. Capital Stock (C2*)	-	116.8 (26.6, 426.3)	44.3 (14.8, 197.5)	33.0 (11.6, 173.6)
N	2465 (1343)	190	443	255

Notes

Definition of the Manufacturing Sample as in the appendix. In 1985, there were 2465 observations for R&D, sales, R&D/sales, employees and non-R&D employees; and 1343 observations for investment and non-R&D investment. The difference in the number of observations is due to missing values in the investment variable. Capital stocks, R&D capital stocks and investment in 10^6 DM (1985).

Table 3

Production Function Results - Total Estimates vs. Within Estimates

Dependent Variable: log(Sales/Employee)

(Heteroskedasticity-Robust Standard Errors in Parentheses)

Variable Definitions	C1, K15, L	C1, K25, L	C2, K15, L	C2,K15,LHRS	C2*, K15, L*
Totals, Balanced Panel (1140 Obs.)					
log(C/L)	.247 (.023)	.249 (.023)	.242 (.022)	.242 (.022)	.260 (.022)
log(K/L)	.160 (.014)	.161 (.015)	.161 (.014)	.160 (.014)	.126 (.015)
log(L)	-.012 (.008)	-.013 (.008)	-.012 (.008)	-.002 (.008)	-.010 (.008)
R^2 (S.E.)	.466 (.346)	.467 (.346)	.473 (.346)	.473 (.346)	.454 (.344)
Within, Balanced Panel (1140 Obs.)					
log(C/L)	.264 (.036)	.260 (.035)	.264 (.034)	.264 (.034)	.261 (.034)
log(K/L)	.072 (.024)	.073 (.021)	.058 (.024)	.059 (.025)	.042 (.025)
log(L)	-.262 (.047)	-.269 (.046)	-.279 (.045)	-.272 (.045)	-.280 (.046)
R^2 (S.E.)	.878 (.165)	.879 (.165)	.882 (.164)	.881 (.164)	.875 (.164)
Totals, Unbalanced Panel (2257 Obs.)					
log(C/L)	.243 (.015)	.244 (.015)	.240 (.015)	.240 (.015)	.247 (.015)
log(K/L)	.136 (.009)	.137 (.009)	.136 (.009)	.136 (.009)	.106 (.009)
log(L)	.013 (.006)	.012 (.006)	.013 (.006)	.013 (.006)	.017 (.006)
R^2 (S.E.)	.457 (.342)	.458 (.341)	.464 (.341)	.464 (.341)	.440 (.341)
Within, Unbalanced Panel (2257 Obs.)					
log(C/L)	.227 (.030)	.226 (.030)	.226 (.028)	.227 (.029)	.232 (.028)
log(K/L)	.090 (.019)	.082 (.017)	.082 (.020)	.083 (.020)	.068 (.020)
log(L)	-.266 (.039)	-.279 (.038)	-.275 (.037)	-.271 (.037)	-.283 (.038)
R^2 (S.E.)	.846 (.182)	.847 (.182)	.850 (.181)	.849 (.181)	.841 (.182)

Notes

The dependent variable is defined as: log(S/L) in columns 1, 2 and 3, log(S/LHRS) in columns 4 and 5, and log (S/L*) in column 6. All variable definitions as in section 2 of the paper. All regressions include 5 dummy variables to control for time effects. All *totals* regressions include 21 dummy variables for two-digit product areas. R^2 values for within estimates refer to the least squares dummy variables (LSDV) version of the model.

Table 4
Production Function Results - "High Technology" vs. Other Firms
Dependent Variable: log(Sales/Employee)
(Heteroskedasticity-Robust Standard Errors in Parentheses)

Totals Regressions

	log(C1/L)	log(K15/L)	log(L)	R^2 (S.E.)	N
All Firms (443 firms)	.243 (.015)	.136 (.009)	.013 (.006)	.457 (.342)	2257
High Technology Firms (255 firms)	.184 (.020)	.163 (.013)	.014 (.007)	.391 (.345)	1312
Other Firms (188 firms)	.334 (.023)	.090 (.014)	.007 (.009)	.530 (.330)	945

Within Regressions

	log(C1/L)	log(K15/L)	log(L)	R^2 (S.E.)	N
All Firms (443 firms)	.227 (.030)	.090 (.019)	-.266 (.039)	.846 (.182)	2257
High Technology Firms (255 firms)	.250 (.039)	.125 (.025)	-.234 (.051)	.809 (.193)	1312
Other Firms (188 firms)	.192 (.046)	.039 (.029)	-.297 (.058)	.884 (.164)	945

Within Regressions - CRS Imposed

	log(C1/L)	log(K15/L)	log(L)	R^2 (S.E.)	N
All Firms (443 firms)	.319 (.030)	.146 (.020)	-	.842 (.185)	2257
High Technology Firms (255 firms)	.341 (.040)	.176 (.024)	-	.805 (.195)	1312
Other Firms (188 firms)	.276 (.045)	.096 (.031)	-	.879 (.167)	945

Notes
All variable definitions as in section 2 of the paper. All regressions include 5 dummy variables to control for time effects. All *totals* regressions include dummy variables for two-digit product areas. R^2 values for within estimates refer to the least squares dummy variables (LSDV) version of the model.

Table 5

Production Function Results - Differencing Estimates

Dependent Variable: log (Sales)

(Heteroskedasticity-Robust Standard Errors in Parentheses)

Independent Variables	log L	log C1	log K15	R^2 (S.E.)	N
Difference	All Firms (443 Firms)				
2 years	-.008 (.046)	.233 (.043)	.072 (.032)	.077 (.209)	1814
4 years	.376 (.051)	.218 (.034)	.076 (.027)	.182 (.258)	1371
6 years	.573 (.047)	.208 (.040)	.086 (.026)	.320 (.281)	928
Long Diff.	.608 (.060)	.199 (.053)	.103 (.030)	.394 (.313)	443
Difference	High Technology Firms (255 Firms)				
2 years	-.013 (.057)	.256 (.060)	.089 (.045)	.072 (.226)	1057
4 years	.353 (.068)	.243 (.046)	.112 (.037)	.172 (.276)	802
6 years	.548 (.064)	.229 (.053)	.128 (.035)	.314 (.295)	547
Long Diff.	.540 (.079)	.246 (.074)	.155 (.041)	.390 (.330)	255
Difference	Other Firms (188 Firms)				
2 years	.019 (.074)	.198 (.060)	.055 (.043)	.098 (.181)	757
4 years	.421 (.066)	.183 (.050)	.026 (.036)	.203 (.232)	569
6 years	.625 (.062)	.171 (.061)	.024 (.037)	.340 (.258)	381
Long Diff.	.712 (.088)	.133 (.075)	.040 (.042)	.431 (.280)	188

Notes

All variable definitions as in section 2 of the paper. All regressions include dummy variables to control for time effects. Growth rates for the "long difference" estimates are computed for the longest possible period of observation for each firm.

Table 6

Gross and Net Rate of Return Estimates

Dependent Variable: Δlog(Sales)

(Heteroskedasticity-Robust Standard Errors in Parentheses)

2-Year Differences

	Independent Variables					
	log L	log C1	R/S	ΔK/S	R^2 (S.E.)	N
All Firms (443 Firms)	-.013 (.046)	.226 (.043)	.220 (.071)		.079 (.208)	1814
	-.030 (.046)	.201 (.043)		.856 (.172)	.087 (.207)	1814
High Technology Firms (255 Firms)	-.017 (.058)	.252 (.060)	.189 (.092)		.0713 (.226)	1057
	-.042 (.058)	.213 (.061)		.906 (.206)	.0829 (.224)	1057
Other Firms (188 Firms)	.013 (.073)	.192 (.058)	.297 (.117)		.103 (.181)	757
	.010 (.073)	.186 (.059)		.719 (.322)	.104 (.180)	757

6-Year Differences

	Independent Variables					
	log L	log C1	R/S	ΔK/S	R^2 (S.E.)	N
All Firms (443 Firms)	.547 (.048)	.195 (.038)	.217 (.043)		.332 (.278)	928
	.511 (.047)	.172 (.037)		.740 (.108)	.350 (.274)	928
High Technology Firms (255 Firms)	.519 (.065)	.225 (.049)	.222 (.054)		.318 (.294)	547
	.464 (.063)	.185 (.048)		.835 (.132)	.348 (.287)	547
Other Firms (188 Firms)	.592 (.063)	.153 (.059)	.266 (.077)		.363 (.253)	381
	.589 (.063)	.150 (.059)		.626 (.193)	.360 (.254)	381

Notes

All variable definitions as in section 2 of the paper. All variable definitions as in section 2 of the paper. All regressions include dummy variables to control for time effects. The intensity variable R/S for the first panel is defined as $2^*R_{t-1}/S_{t-2}$. In the second panel R/S is computed as $6^*(R_{t-1}+R_{t-2}+R_{t-3})/S_{t-4}$. The intensity variable ΔK/S is computed as $(K_t-K_{t-1})/S_{t-2}$ in the upper panel and as $(K_t-K_{t-3})/S_{t-4}$ in the lower panel.

Table 7

Elasticity Estimates from Within Regressions with Time-Varying Slope Parameters (CRS Imposed)

Dependent Variable: log(Sales/Employee)

(Heteroskedasticity-Robust Standard Errors in Parentheses)

	Physical Capital Elasticity			R&D Capital Elasticity		
Year	All Firms	High Technology Firms	Other Firms	All Firms	High Technology Firms	Other Firms
1979	.307 (.032)	.315 (.042)	.272 (.049)	.139 (.020)	.179 (.026)	.084 (0.31)
1981	.280 (.032)	.320 (.042)	.245 (.047)	.145 (.020)	.161 (.025)	.095 (.033)
1983	.319 (.032)	.345 (.043)	.274 (.046)	.175 (.021)	.213 (.026)	.116 (.034)
1985	.304 (.033)	.317 (.042)	.259 (.050)	.203 (.021)	.258 (.027)	.134 (.035)
1987	.341 (.034)	.370 (.045)	.292 (.052)	.184 (.023)	.226 (.029)	.116 (.036)
1989	.347 (.038)	.363 (.047)	.293 (.061)	.141 (.023)	.191 (.030)	.086 (.034)
F-Test of Parameter Equality	19.17 (6, 1797)	13.36 (6, 1040)	6.24 (6, 740)	17.84 (6, 1797)	17.43 (6, 1040)	3.13 (6, 740)
N	2257	1312	945	2257	1312	945

Notes

All regressions include 5 dummy variables to control for level time effects. All F test statistics are significant at p<.005. The capital stock variables are C1 and K15 (see section 2 of the paper for definitions).

References

Brockhoff, K. (1986). "Die Produktivität der Forschung und Entwicklung eines Industrieunternehmens," Zeitschrift für Betriebswirtschaft, Vol. 56, pp. 525-537.

Brockhoff, K. and K. Warschkow (1991). Deflationierung industrieller Forschungs- und Entwicklungsaufwendungen. Unpublished Final Report to the German Research Council (Deutsche Forschungsgesellschaft), DFG Project Br311/22-1. Kiel.

Cuneo, P. and J. Mairesse (1984). "Productivity and R&D at the Firm Level in French Manufacturing," in: Z. Griliches (Ed.). R&D, Patents, and Productivity, pp. 375-392. Chicago: University of Chicago Press.

Goto, A. and K. Suzuki (1989). "R&D Capital, Rate of Return on R&D Investment and Spillover of R&D in Japanese Manufacturing Industries," Review of Economics and Statistics, Vol. 71, No. 4, pp. 555-564.

Griliches, Z. (1979). "Issues in Assessing the Contribution of Research and Development to Productivity Growth," Bell Journal of Economics. Vol. 10, pp. 92-116.

Griliches, Z. (1986). "Productivity, R&D, and Basic Research at the Firm Level in the 1970's," American Economic Review. Vol. 76, No. 1, pp. 141-154.

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

Hall, B. (1993). "Industrial R&D During the 1980s: Did the Rate of Return Fall?," Brookings Papers on Economic Activity (Microeconomics), No. 2, pp. 289-343.

Hall, B. and J. Mairesse (1995). "Exploring the relationship between R&D and productivity in French manufacturing firms," Journal of Econometrics, Vol. 65, pp. 263-293.

Klette, T. and Z. Griliches (1992). "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous," NBER Working Paper No. 4026, National Bureau of Economic Research, Cambridge, Mass.

Klette, T. J. (1994). "R&D, scope economies and company structure: a 'not so fixed effect' model of plant performance," mimeo.

Legler, H., H. Grupp, B. Gehrke, and U. Schasse (1992). Innovationspotential und Hochtechnologie. Heidelberg: Physica.

Mairesse, J. (1992). "Time-series and cross-sectional estimates on panel data: why are they different and why should they be equal?," in: J. Hartog, G. Ridder and J. Theeuwes (eds.): Panel Data and Labor Market Studies. Amsterdam: North-Holland.

Mairesse, J. and B. Hall (1994). "Comparing the Productivity of Research and Development in French and United States Manufacturing Firms," mimeo.

Mairesse, J. and M. Sassenou (1991). "R&D and Productivity: A Survey of Econometric Studies at the Firm Level," Science-Technology-Industry Review, No. 8, pp. 317-348.

Mairesse, J. and P. Mohnen (1994). "R&D and Productivity Growth: What have we learned from Econometric Studies?" Paper Presented at the Conference on Evolutionary Economics of Technological Change, Strasbourg, Oct. 6-8, 1994.

Meyer-Krahmer, F. (1989). Der Einfluß staatlicher Technologiepolitik auf industrielle Innovationen. Baden-Baden: Nomos.

National Science Foundation (1994). National Patterns of R&D Resources: 1994. Washington, D.C.: National Science Foundation.

Statistisches Bundesamt (1989). Fachserie 18. Stuttgart: Kohlhammer.

Variable and Index Definitions

Variables

S	output - total sales
I*	gross investment including R&D-related investment
I	gross investment in physical capital net of R&D-related investments
C1*	unadjusted capital stock, initial value based on constant growth approximation
C1	adjusted capital stock, initial value based on constant growth approximation
C2*	unadjusted capital stock, initial value based on industry capital stock
C2	adjusted capital stock, initial value based on industry capital stock
R	total R&D expenditures
RS	R&D intensity, R&D expenditures divided by total sales
L*	total number of employees (incl. R&D employees)
L	labor input (number of non-R&D employees)
LHRS	labor input with work hours correction
LR	full-time employees in R&D
K15	R&D capital stock, depreciation rate 15%
K25	R&D capital stock, depreciation rate 25%
τ	depreciation rate of physical capital
δ	depreciation rate of knowledge capital
$g_R = 0.059$	average growth rate of real R&D expenditures, 1967-1977
$g_I = 0.02$	average growth rate of real investment, 1967-1977

Indices

t Time (years)

i Firm Index

j Industry Index

Appendix - Data Source and Construction of the Sample

The raw data consisted of seven cross-sections for the years 1977, 1979, 1981, 1983, 1985, 1987 and 1989. In 1977, the raw data in the sample was gathered exclusively through written questionaires administered by the Stifterverband. The following cross-sections also contained data from an additional survey that was conducted in connection with a research and development subsidization programme. Starting in 1979, the Federal Ministery of Research and Technology administered a long-term programme (*PKZ* for *FuE-Personalkostenzuschußprogramm*, i.e. R&D Personnel Expenditures Subsidy Programme) which allowed small and medium-sized firms to obtain financial support for R&D personnel expenditures.[15] As a part of the application process, the applicant firms had to indicate their level of R&D expenditures, sales and number of employees. These data were later used by *SV-WiStat* to complement their own data. Once the R&D promotion expired, the recipient firms were included in the *Stifterverband* sample and may have responded to the more extensive written survey administered by *SV-WiStat*. The composition of the initial set of cross-sections is described in Table A.1.

Table A.1

Composition of the Initial Set of Cross-Sections

Year	"*Stifterverband* Observations"	Mixed Type	"*PKZ* Observations"	Σ
1977	1352	0	0	1352
1979	2006	1932	3188	7126
1981	2404	2191	3653	8248
1983	2708	2248	5787	10743
1985	2858	2342	6765	11965
1987	2635	2096	6407	11138
1989	2759	1691	5667	10117
Σ(observations)	16722	12500	31467	60689
Σ(firms)	5362	2860	14375	22597

The first column in Table A.1 ("*Stifterverband* Observations") contains the number of firms that responded in the respective year to the Stifterverband questionaire and were never part of data collection under the *PKZ* programme. The third column ("*PKZ* Observations") contains the number of firms answering the questionaire associated with the *PKZ* subsidization programme, but never entering the data collection process used by Stifterverband. Finally, the second column ("Mixed Type") contains the number of firms which appeared initially as *PKZ* observations, but did respond to at least one of the follow-up survey administered by the Stifterverband.

Since the questionaire answered by recipients of government funding under the *PKZ* programme did not include a question on investment in physical capital, no use is made of the observations in this paper. Hence, the raw data used in this analysis correspond to the first column of Table A.1. The initial sample was further reduced from 5362 firms to 4698 firms with 15334 observations by excluding non-manufacturing enterprises and firms in the nuclear fuels, air- and spacecraft, ship-

[15] For a more detailed description and evaluation of this program, see Meyer-Krahmer (1989).

building, and leather production industries. For the latter industries, either no suitable price deflators were available or the respective sectors were strongly affected by government procurement.

In the chemical and pharmaceutical industry, the questionaire was actually administered by *VCI* (*Verband der Chemischen Industrie*), the respective industry association. Prior to 1985, *VCI* did - by and large - not use the same identification numbers for firms that had been observed in the past already. Therefore, it is rather difficult to construct a meaningful time series for firms in the chemical industry and no further attempts have been made to enhance the representation of chemical firms in the sample. However, it was possible to construct the R&D and investment time series for ten firms without any problems. Observations that could not be combined were treated as separate firms (see Table 1).

The deflating and data correction procedures described in section 2 of the paper were applied to the above-mentioned sample of 4698 firms. After deflating and computation of the corrected investment and labor figures, the following additional cleaning and checking procedures were applied to the sample:

1) Observations were dropped if the corrected labor variable assumed a negative or zero value. In these cases, the number of R&D employees exceeded the number of total employees. 54 observations were deleted for this reason.

2) Similarly, 117 observations were deleted because the corrected investment variable (total investment minus R&D-related investment) assumed negative values.

3) 225 observations were deleted due to missing revenue data.

4) Observations were deleted when the logarithm of R&D expenditures per employee (including R&D employees) was outside the median plus or minus three times the interquartile range. This step led to the exclusion of 17 observations from the sample.

5) 166 observations were dropped from the sample because the logarithm of revenue per employee (including R&D employees) was outside the median plus or minus three times the interquartile range. This step effectively removed all ten petroleum refining firms in the initial sample.

6) Observations were dropped if R&D investment exceeded a share of 60 per cent of total investment or if R&D employment exceeded a share of 60 per cent of total employment. Closer inspection revealed that the 384 observations deleted due to this criterion were largely concentrated in about 80 companies, most of them in electrical and non-electrical machinery industries. While these companies are also in the manufacturing sector, they appeared to earn a large part of their revenues with technical services and related research and development. Thus, for the purpose of this study they were excluded from the sample.

Thus, these six steps led to the deletion of 960 observations (6.26 per cent of the initial 15334 observations), leaving a sample consisting of 14374 observations. Since many of these observations did not contain a response with respect to gross investment, the sample shrank to 4023 firms (10012 observations) with at least 1 observation with complete R&D, investment, revenue and labor data. Table B.2 demonstrates that this intermediate panel is highly unbalanced. Due to the lack of a capital stock variable (e.g. gross capital), it was necessary to construct a capital stock proxy based on the investment data time series. Similarly, the need to construct an R&D capital stock requires a restriction to firms that were observed "often enough." Therefore, only firms with more than four observations were taken to represent next intermediate sample, indicated by the shaded cells in Table A.2.

Table A.2 - Composition of the Manufacturing Sample

(Only Observations with Complete Data*)

Observations per Firm with Complete Data*	Number of Firms	Number of Firms in Final Sample
1	1677	-
2	792	-
3	625	-
4	328	-
5	232	148
6	179	105
7	190	190
Σ(firms)	4023	443
Σ(observations)	10012	2700

* Observations with valid data on investment, number of employees, revenues, and total R&D expenditures.

The intermediate sample was subjected to the following additional cleaning procedures:

7) Observations were deleted from the sample whenever revenues grew by more than 400 per cent or dropped by more than 150 per cent between observations. Application of this criterion removed no further observations from the sample.

8) Observations were deleted from the sample whenever employment grew by more than 400 per cent or dropped by more than 150 per cent between observations. This condition led to the exclusion of 3 observations (1 firm).

Application of these two steps yielded a sample with 600 firms (190 with 7 observations, 178 with six observations, 232 with five observations). Finally, in order to avoid imputation of missing data in the investment and R&D time series, I excluded firms from the sample which had less than five *contiguous* observations of these variables. This exercise demonstrates again clearly that the lack of capital stock or investment data is the strongest constraint for the sample construction. It would have been possible to construct complete R&D time series for 298 firms while a complete investment time series is available only for 190.

Following the steps outlined in section 2 of the paper, capital stocks and R&D capital stocks were computed on the basis of this sample. The resulting capital stocks and R&D capital stocks were again subjected to some consistency tests. Two-year growth rates for the capital stock ranged between -22.3 and 87.2 per cent. The corresponding two-year growth rates for the R&D capital stock ranged between 336.6 per cent and -42.9 per cent. All observations were retained in the sample. The resulting sample is described in the third column of Table A.2. The final result of the panel construction is a balanced panel with 190 firms and 7 observations per firm (1330 observations in total) and an unbalanced panel with 443 firms and 2700 observations in total (which contains the balanced one). Note that the larger share of the initial 15334 observations has been dropped for one or several of the above-mentioned reasons.

The selection of firms according to the criteria listed above is not a random process. As should be expected, larger firms report more reliably on their R&D activities than smaller ones and tend to respond with more complete questionaires, especially with respect to the investment variable. So far, corrections for sample selection bias have not been employed.

For all regressions in this paper the first observation of the computed capital stock time series will be dropped in order to avoid excessive measurement error. Note that the capital stock construction in this paper involves a "time to build" assumption, i.e. investment and R&D expenditures of period t will not be active components of the capital stocks until period t+1. This assumption should help to avoid simultaneity problems, since the capital stocks represent the values at the beginning of the respective time period. However, the questionaire used by Stifterverband explicitly asks firms to enter their end-of-year number of employees. Thus, in order to be consistent with the construction of the stock variables, lagged labor variables are used in all regressions. Thus, for a regression on 1989 output, the end of 1987 (beginning of 1988) labor measure is used. Dropping the initial observation for each time series necessarily reduces the unbalanced panel from 2700 to 2257 observations and the balanced panel from 1300 to 1170 observations.