

# The Determinants of Industrial Research and Development: A Study of the Chemical, Drug, and Petroleum Industries

---

Henry G. Grabowski\*

*Yale University*

Economists have recently grown interested in doing research on research—or R & D, as it is called in industrial circles. Several studies have tested Schumpeter's hoary hypothesis that large firms are responsible for most industrial inventive activity.<sup>1</sup> Few of these studies, however, suggest why this hypothesis is apparently valid for some industries and not for others. And statistical studies going beyond this question, to try to relate R & D expenditures to firm profit expectations and the availability of funds as in other investment decisions, are rare (Mansfield, 1964; Mueller, 1967).

This paper reports the results of an empirical investigation into the determinants of research expenditures in three industries—drugs, chemicals, and petroleum refining. These industries have three advantages for such a study: (1) they are among the leaders in total R & D expenditures; (2) most activity is concentrated in an appreciable number of large or moderately large firms; and (3) government support of research work is relatively small, so that decisions are more closely analogous to ordinary

\* Some of the results of this paper were first presented in my doctoral dissertation submitted to Princeton University in 1966. I am indebted to B. G. Malkiel, O. Morgenstern, F. M. Scherer, and R. E. Quandt for helpful suggestions and criticisms. For making unpublished data available to me on R & D expenditures, I am especially grateful to Dennis C. Mueller. In addition, data collected by McGraw-Hill, Inc., and compiled under the direction of Professor Robert Eisner of Northwestern were very helpful in the earlier stages of the work.

<sup>1</sup> Data collected by the National Science Foundation (hereafter "NSF") definitely show that firms below certain threshold sizes perform very little organized R & D. For example, firms with 1,000 employees or less in 1958 accounted for 60 per cent of total industrial employment but performed only 5 per cent of company-financed R & D (NSF, 1964, p. 6). However, the question which has been extensively investigated in recent studies is whether research intensity increases with size among firms clearly larger than the threshold levels. For a survey of these studies and an attempt to reconcile some conflicting results, see Markham (1965).

investment decisions.<sup>2</sup> Data samples were constructed for the three industries by requesting total R & D expenditure figures, as defined by the NSF standards,<sup>3</sup> from firms in the 1960 *Fortune 500* listing. The response rate in each industry was almost exactly 70 per cent, with R & D data over the period 1959–62 provided by sixteen firms in chemicals, fifteen in petroleum refining, and ten in drugs.

Since the main variation in the data occurs over the cross-section of firms, rather than for each firm over time, the regression models will be primarily explaining interfirm differences in R & D. However, pooled cross-section time-series samples will be used in order to increase the number of available observations in each industry sample. For each equation estimated with these pooled samples, the hypothesis of time-invariant parameters was tested. Because this hypothesis cannot be rejected at the normal confidence intervals, data pooling may be advantageously employed in the present analysis of R & D determinants.

### I. The Characteristics of the Model

When analyzing the R & D behavior of a group of firms in which substantially different scales of operation exist, research intensity, rather than the absolute level of a firm's expenditures, is a more appropriate dependent variable. When absolute figures are used, heteroscedasticity is invariably present, and scale effects tend to dominate the regression equations. In order to avoid these problems, the procedure adopted in this paper is to measure R & D expenditures as well as size-correlated independent variables, relative to the total sales of the firm.<sup>4</sup> Although there are some good a priori reasons for choosing sales as the particular size deflator,<sup>5</sup> the

<sup>2</sup> Of the six NSF industry classifications performing approximately \$300 million or more of company-financed R & D in 1963, government supported R & D was over 25 per cent of total expenditures except in chemicals and allied products and in petroleum refining. In these two industries it amounted to 21 per cent and 6 per cent, respectively.

<sup>3</sup> The NSF definitions of R & D are given in the technical notes of all its published data reviews. These data have definite advantages over company published figures in that (1) they are based on a common externally devised definition; and (2) firms are assured of confidential treatment, thereby removing any incentives to inflate the figures artificially. While the NSF data are undoubtedly the best available statistics, it must be emphasized there are still many difficult conceptual problems in measuring inventive efforts by dollar outlays. See the discussions in Kuznets (1962) and Sanders (1962).

<sup>4</sup> An alternative approach for dealing with heteroscedasticity is to estimate the equations using logarithmic transformations when the specifications of the models permit one to do so. Both procedures have been extensively used in dealing with samples that are cross-sectional in nature and span a large size spectrum. For a discussion of the problem associated with using the ratio approach, see Kuh and Meyer (1955).

<sup>5</sup> This question has been analyzed by Scherer (1965*b*, pp. 256–61) in a commentary on some results of Hamberg (1964). Among the reasons for choosing sales are: (1) sales is a more neutral size measure, and (2) firms have emphasized the use of sales as a landmark for budget decisions in various interview studies.

results are substantially unaltered if either total assets or the number of employees is used instead.

The determinants of research intensity to be considered first relate to the returns from R & D activity. Since there is a considerable lag in the payoff to R & D, a firm's estimated set of returns from current projects will likely depend on its past results. While expectations based on past performance may not vary markedly from one budget period to the next for an individual firm, substantial variations should exist across firms due to the cumulated effects of past differences in firm capacities and attitudes toward R & D. The first explanatory variable of research intensity is, therefore, an index of the research productivity of each firm over a prior period.

In order to construct such a proxy variable of research productivity, three possible measures of a firm's research output were considered: new-product sales, the number of patents granted to the firm, and the number of significant inventions made by the firm. Given the strong product orientation of industrial R & D, perhaps the best measure of the three from a conceptual standpoint is new-product sales. Unfortunately, new-product sales are not generally available, and furthermore, substantial differences in definitions and product classification among firms make the use of these numbers in a cross-section analysis quite hazardous. Of the two other measures of research output, patents is the more attractive, since all patented inventions must pass certain uniform criteria of the U.S. Patent Office, and patent statistics are readily available.<sup>6</sup>

The measure of research outputs used in the model is, therefore, the number of patents granted to a firm in a specified prior period. To form the productivity variable, this output measure is divided by a research-input measure—specifically, the number of scientists and engineers employed by the firm over the approximate period when the patented inventions were conceived.<sup>7</sup> The first hypothesis put forth here is thus that firms with higher patented output per scientific worker in the past will, *ceteris paribus*, be more research-intensive than their rivals.<sup>8</sup> Of course, this

<sup>6</sup> Attempts at constructing a series of significant inventions made by each firm were beset by serious methodological problems and therefore discarded. For some interesting attempts to measure inventive output in this manner, see Mansfield (1964, pp. 334–37).

<sup>7</sup> Scherer (1965a, p. 1097) has estimated that over a period somewhat coincident with our sample the patent office took an average of three and one-half years to process a patent. Adding several more months as time necessary to draw up a patent application, the "patent lag" is taken somewhat arbitrarily here to be four years in length.

<sup>8</sup> This relationship is likely to operate in a forward as well as backward manner. That is, more research-intensive firms now should realize a higher level of patented outputs in the future. There is ample evidence that firm behavior often follows such recursive patterns. While firms may certainly radically change past modes of operation, R & D in particular is not an activity that lends itself to frequent and marked changes in emphasis.

variable at best measures only one aspect of the firm's returns from R & D—the technical quality of its research outputs. Another aspect, relating to the applicability of R & D to the firm's operations, will be considered below.

Before discussing the other explanatory variables, however, it should be mentioned that patents have been used elsewhere as a measure of inventive output, and their limitations in this regard have been extensively discussed (Kuznets, 1962; Sanders, 1962). One of the more serious problems incurred in using patent statistics is the possibility that the propensity to patent might be systematically correlated in a positive way with a firm's degree of research intensity. If this were so, a spurious positive relation between research intensity and patent output per worker would result. However, there are no strong a priori arguments why this will be the case, and some data are available to analyze this question for the firms under study. This problem will be taken up in the next section.

The second explanatory variable of research intensity in the present model is an index of each firm's output diversification. It has been postulated in the literature (see Nelson, 1959) that a firm's degree of diversification will positively influence its profit expectations from R & D. This hypothesis follows from the belief that a more diversified firm will be better able to exploit unexpected research outputs than one with a narrower base of operations. While the original formulation was meant to apply to scientific work at the research end of the spectrum rather than to development expenditures, this hypothesis also implies greater expenditures on total R & D unless these two component activities are complete substitutes, which is highly unlikely.

The index of diversification used in the present model is based on the number of separate five-digit SIC product classifications in which a firm produced during a middle year of the sample period.<sup>9</sup> This variable, like the index of past research productivity, is designed to capture interfirm differences in expected returns from R & D. It therefore varies only over the cross-section of firms and not over time. Over a four-year interval, the size of the firm's product mix does not change significantly, and the

<sup>9</sup> All the data on diversification come from the 1961 edition of the *Fortune Plant and Product Directory*. For the chemical and petroleum industries, an examination of the various product classes revealed that they are virtually all of sufficient technical character to have at least some potential relevance to the R & D activity. For these two industries, the index will therefore be the total number of separate SIC classifications encompassed by the firm's product mix. For the drug industry, however, product mixes of highly diversified drug firms indicate a tendency for them to expand into products only tangentially related to the manufacture of drugs proper (that is, adhesives, brushes, glass bottles, toilet preparations, plastic products, and so forth). These product lines offer little opportunity for the applicability of research outputs. It was therefore decided that a more appropriate index of diversification for the drug industry would be obtained by counting only the number of SIC classifications that are concerned directly with the manufacture of drugs proper.

construction of the index as "timeless" during this period is a satisfactory approximation.

Aside from expected returns, financial factors form another basic set of considerations relevant to industrial R & D expenditures. The relationships between financial variables and investment have been extensively explored in the literature on the determinants of fixed capital expenditures (for a bibliography, see Eisner and Strotz, 1963). There is considerable evidence from these studies that retained earnings and other internally generated funds have an especially significant effect on investment expenditures. This has been attributed to the general reluctance of firms to raise funds externally because of the added risks and transaction costs entailed in this type of financing. If this is so, it would seem applicable with some qualifications to investment in R & D.

In the short run, R & D expenditures can be expected to be much less sensitive to changes in cash flow, especially in the downturn. This is true because of the higher fixed cost component in R & D activity. Research workers, whose salaries constitute a sizable percentage of total expenditures, are not perfectly elastic in supply and cannot be alternately fired and rehired in accordance with temporary changes in business conditions.<sup>10</sup> In the long run, however, a significant positive relationship between R & D and cash flow should be evident if firms behave as many previous investment studies suggest. Since the data samples investigated here are basically cross-sectional and span four years of reasonably stable growth for the firms involved, a long-run effect should be observed in the present case. The third explanatory variable in our model is, therefore, a measure of the firm's internally generated funds deflated by sales. The specific internal funds variable that will be used is the sum of the firm's after-tax profits plus depreciation and depletion charges, lagged one period.

The above three variables reflect some of the main technological, marketing, and financial factors that one would expect to influence the R & D expenditure decision. Some previous interview studies of R & D suggest some other relevant considerations (see NSF, 1956). One strong trend of thought running through these studies is that firm decisions on R & D are strongly influenced by the behavior of competitors, and, in particular, that a great deal of imitation exists among firms with respect to R & D allocations. Since most R & D is performed by firms operating in oligopolistic market structures and it is an activity presumably involving greater uncertainty than other undertakings, firms may imitate each other

<sup>10</sup> That is, there will likely be significant downward rigidities in this relationship due to the technological necessity of maintaining a reasonably stable staff of researchers. Also, the increasing cost associated with rapid expansion will act to constrain this relationship in upturns in the short run as well. While material costs are more flexible, they account for a smaller percentage of total expenditures (for representative cost figures, see the NSF data reviews).

as a conservative strategy for minimizing risks. Because of the multitude of forms which such imitation may take, however, it is difficult to deal with this phenomenon in the framework of the present empirical analysis.

One particularly simple type of imitation discussed in these interview studies and frequently mentioned in the trade literature is the adherence by firms to a general industry R & D to sales ratio. If this kind of imitative behavior is present to any significant degree in the present industrial samples, it should be evident from the empirical analysis. Since research intensity as measured by the R & D to sales ratio is the dependent variable, imitation by firms of an "industry" ratio would imply less variability in the dependent variable and cause the intercept of the regression equation to become statistically significant relative to the explanatory variables postulated above. The sign and statistical significance of the intercept term in the regression model therefore provide a first level test of this proposition. However, more subtle and complex forms of imitation are best analyzed in a more disaggregative context than the data permit here.

The present discussion leaves us essentially with three explanatory variables of firm research intensity—a research-productivity variable consisting of the level of patented output realized by the firm relative to its input of scientific personnel over a prior period, an index of firm product diversification, and a variable dealing with the financial resources of the firm which is equal to the level of internally generated funds of the firm as a percentage of its total sales. The model may be expressed formally as

$$\frac{R_{i,t}}{S_{i,t}} = b_0 + b_1 P_i + b_2 D_i + b_3 \frac{I_{i,t-1}}{S_{i,t}}, \quad (1)$$

where  $R_{i,t}$  is the level of R & D expenditures of the  $i$ th firm in the  $t$ th period,  $S_{i,t}$  is the level of sales of the  $i$ th firm in the  $t$ th period,  $I_{i,t-1}$  is the sum of after-tax profits plus depreciation and depletion expenses of the  $i$ th firm in the  $t - 1$  period,  $P_i$  is the number of patents received per scientist and engineer employed by the  $i$ th firm in a prior four-year period (1955–59), and  $D_i$  is the index of diversification of the  $i$ th firm (the number of the separate five-digit SIC product classification in which it produces).

The hypotheses discussed above suggest that coefficients  $b_1$ ,  $b_2$ , and  $b_3$  will all be positive. In addition, the intercept term,  $b_0$ , of this equation serves in a sense as a possible fourth explanatory variable since it shows the influence of sales on research expenditures and in particular provides a test of the proposition that firms adhere to an industry-wide R & D to sales ratio. The results of including some additional variables in equation (1) are also given in the section which follows.

## II. The Empirical Results

Using the data samples described above, least-squares estimates of the coefficients of equation (1) were obtained. The results are presented in

Table 1. All of the regression coefficients are of the postulated sign and are significant at the 1 per cent level except for the diversification variable coefficient in the petroleum industry, which is positive but statistically insignificant. The over-all explanatory power of our model is quite good in the case of the chemical and drug industries, given the nature of the samples under study, ( $R^2 = .63$  and  $.86$ , respectively), but is low for the petroleum industry ( $R^2 = .29$ ). An examination of the results in Table 1 also shows that the estimates of the intercept terms of equation (1) are such as to cast considerable doubt on the proposition that firms in this

TABLE 1

ESTIMATION OF REGRESSION EQUATION  $R_{i,t}/S_{i,t} = b_0 + b_1P_i + b_2D_i + b_3(I_{i,t-1}/S_{i,t})$   
FOR THE CHEMICAL, DRUG, AND PETROLEUM INDUSTRIES FOR  
THE PERIOD 1959-62

Industry	$b_0$	$b_1$	$b_2$	$b_3$	$R^2$	$F$	$N$
Chemicals . . . .	0.006 (0.004)	0.12* (0.02)	0.019* (0.004)	0.078* (0.023)	.63	29.76	60
Drugs . . . . .	-0.03* (0.01)	0.54* (0.12)	0.41* (0.07)	0.26* (0.05)	.86	73.71	40
Petroleum . . . .	0.002 (0.002)	0.016* (0.005)	0.0049 (0.0071)	0.020* (0.006)	.29	5.46	55

\* Significant at .01 level.

NOTE.—Numbers below coefficient estimates are estimates of the standard errors; technological and diversification variables ( $P_i$  and  $D_i$ ) have been multiplied by scale factors in order to present results more conveniently.

sample adhere to an industry-wide R & D to sales ratio. The only statistically significant intercept is negative (the drug industry), and the positive coefficients for the other two industries were negligible in value.<sup>11</sup>

The above regression results indicate that interfirm differences in technology, diversification, and the availability of funds all are important in explaining differences in research intensity with no single factor having a dominant influence. Table 1 also shows that the size of the regression coefficient associated with each of these variables increases with the research orientation of the industry involved—being the lowest in the petroleum industry and the highest in the drug industry in every case. Thus, as research looms more important as a competitive strategy to the firms of an industry, each of our independent variables exerts a correspondingly greater effect on the level of research that a firm performs.

<sup>11</sup> All of the results presented in Table 1 are substantially unchanged when the profit component of the available funds variable is measured by the firm's retained earnings rather than its total after-tax profits. The results on the significance of the variables are the same, and the fit of the regression becomes slightly better for chemicals ( $R^2 = .67$ ) and slightly poorer for drugs ( $R^2 = .80$ ).

The much poorer performance of the model in explaining research intensity in petroleum refining can be traced in part to certain structural factors that distinguish it from chemicals and drugs. Among these factors are: (1) R & D is much more process oriented in petroleum refining;<sup>12</sup> (2) the degree of integration within this industry is very uniform, and the amount of outward diversification is slight; and (3) R & D is more of a peripheral activity, consuming a portion of the budget which is at least an order of magnitude smaller than in the other two industries. In an industry where research is process oriented, patents will likely be a poorer measure of technological output because firms will often wish to keep knowledge of such inventive activity concealed from their competitors. Furthermore, where R & D is a competitive strategy of lesser importance, as in petroleum refining, allocations to it tend to be more vulnerable to fluctuations in other uses of scarce funds. While there is no way to quantify the effects of the above three structural factors, together they probably explain a substantial part of the poorer performance of the model in this industry.

The positive relation observed between patented output per research input and the research intensity of a firm has been interpreted above as a measure of the effects of interfirm technology differences in research intensity. If it were true that more research-intensive firms have a greater propensity to patent than less intensive ones, this interpretation would be open to serious question. To investigate this latter possibility, data on the number of in-house patent attorneys for all the firms in our sample over the period 1955–59 were obtained. This is the most meaningful measure of patent activity available, although it is far from a complete index of it.<sup>13</sup>

Utilizing these data on patent attorneys, the correlation coefficient between a firm's research intensity and the number of in-house attorneys engaged per scientist and engineer was calculated for each of the three industry samples. Since the regression estimates have shown a significant relation between a firm's research intensity and the number of patents granted to it per research employee, one would also expect a significant relation between this variable and the number of in-house patent attorneys

<sup>12</sup> The fifteenth annual McGraw-Hill *Survey of Business* for 1962 indicated that R & D planned by petroleum firms for that year was 42 per cent process oriented and only 6 per cent so in the chemical and allied products industries. In addition, the 1960 survey showed that new-product sales amounted to 16 per cent of total sales in the latter industry, while constituting only 2 per cent of total sales in the petroleum industry.

<sup>13</sup> The crucial factor here is the relation of in-house patent attorneys to those hired from outside the firm. No statistics are available at present on the latter variable. The use of in-house attorneys by the firms in our sample was, however, quite extensive—almost all firms had at least one attorney, and most had several. In addition, the number of patent attorneys was well correlated with the number of patents ( $r = .7$ ).



per research employee if this relation were merely the result of more research-intensive firms having a greater propensity to patent. The resulting correlation coefficient between these variables, however, is  $-.2$  in chemicals,  $.1$  in drugs, and  $.3$  in petroleum. These correlations, none of which is significantly positive at the 5 per cent level, do not support the hypothesis that more research-intensive firms tend to patent a greater proportion of patentable inventions. Although this is admittedly a rather slim reed of evidence, it is all that is currently available. However, since there is no strong a priori case for any kind of correlation, it may be tentatively accepted as support for the position that no systematic relations exist between these variables.

Finally, in order to investigate other possible factors affecting R & D which are omitted from the present model, a few variables that have been used with success in explaining investment in fixed capital were added to equation (1).<sup>14</sup> In particular, the relevance of the traditional accelerator mechanism to R & D was investigated by including the first differences of sales in a given period (deflated by sales) as an additional explanatory variable of research intensity. While the normal rationale underlying the accelerator does not hold directly for R & D, a relation may still exist if expectations about future business conditions are strongly influenced by current changes in sales.<sup>15</sup> However, the results do not indicate that this effect is important for R & D expenditures in the present industrial samples. When the sales difference term is added to equation (1), the estimate of the regression coefficient is quite insignificant and alternates in sign among the three industries. Specification of the accelerator with other time lags also yielded insignificant regression estimates. The fact that R & D is essentially an activity directed to the discovery and development of new products and processes with long periods until payoff may account for this apparent insensitivity to current sales changes.

### III. Research Intensity and Firm Size

All of the independent variables used in the above model—research productivity, the degree of internal liquidity, and diversification—have been cited in the literature as firm attributes positively associated with large size. It has been argued, therefore, that the large firms in a given

<sup>14</sup> In addition to the accelerator term discussed in the text, some variables relating to external conditions in financial markets and to the financial position of the firm (that is, the interest rate, the size of the firm's external debt, and the debt-equity ratio) were also included, but these all proved to be very insignificant.

<sup>15</sup> There is some question as to the expected sign of this relationship. If firms undertake investments in R & D on the basis of optimistic expectations and vice versa, a positive sign would or could be anticipated. It has been postulated by Hall (1964, p. 9), however, that firms will turn to R & D as a principal strategy for reversing poor sales performance, and a negative sign would then be expected.

industry will be more research-intensive than their smaller competitors, and vigorous government antitrust activity may have a harmful effect on technological progress. As noted above, there have been several recent investigations of this form of the Schumpeterian hypothesis, but the results have been quite mixed in nature. A study by Mansfield (1964, pp. 333–37), for example, found a significantly positive relation between research intensity and firm size in chemicals, whereas the drug and petroleum industries exhibited significantly negative ones. Since the variables supposedly underlying the Schumpeterian hypothesis yielded very good fits of research intensity in the chemical and drug industries—and yet these two industries apparently exhibit quite different structural relationships between research intensity and firm size—it is worth investigating the potential source of these differences.

First of all, in order to investigate the Schumpeterian hypothesis for the particular industrial samples under study, the following regression is estimated:

$$\frac{R_{i,t}}{S_{i,t}} = \frac{a_0}{S_{i,t}} + a_1 + a_2 S_{i,t}, \quad (2)$$

which in non-ratio form is the quadratic:

$$R_{i,t} = a_0 + a_1 S_{i,t} + a_2 S_{i,t}^2. \quad (3)$$

A quadratic estimated in the above fashion should indicate whether there is any tendency for research intensity to increase or diminish significantly as size increases through the behavior of the  $a_2$  coefficient. The regression estimates of equation (2) for the chemical and drug industries are presented in the top of Table 2.<sup>16</sup> While  $a_1$  is positive and statistically significant as expected in both cases, the estimate of  $a_2$  is significantly positive for chemicals and significantly negative for drugs.<sup>17</sup> A plot of these estimated relations is given in Figure 1. It indicates that, for the drug industry, research intensity initially increases with firm size but is characterized by a decreasing relation for most of the relevant range. For the chemical industry, the estimated relation indicates that research intensity increases

<sup>16</sup> The petroleum industry could also be included here, but it is omitted because the independent variables of equation (1) explained only one-third of the total variance for that industry. If one repeats the procedures described in this section for it, however, the interpretation of the results are consistent with those presented for chemicals and drugs. For the details, see Grabowski (1966, pp. 75–83).

<sup>17</sup> It may be noted that while the  $R^2$ 's are quite modest for this regression equation, the standard errors of estimate indicate quite good fits to the data. This, of course, is because, in the ratio form of estimating equation (3), much of the explanatory power now comes from the intercept term  $a_1$  which affects the goodness of fit of equation (2) but not the  $R^2$ ; that is, the estimates of equation (2) explain a very large portion of the variation in absolute R & D expenditures (due to the high explanatory power of the  $a_1$  term) but only a moderate amount of the variation in the research intensity of firms.

TABLE 2<sup>a</sup>

A. ESTIMATION OF REGRESSION EQUATION  $R_i/S_i = (1/S_i)(a_0 + a_1S_i + a_2S_i^2)$   
FOR POOLED TIME-SERIES CROSS-SECTIONS OF FIRMS IN THE CHEMICAL  
AND DRUG INDUSTRIES OVER THE PERIOD 1959-62

Industry	$a_0$	$a_1$	$a_2$	$R^2$	$F$
Chemicals . . .	0.04 (0.03)	0.03* (0.02)	$0.9 \times 10^{-5}$ * ( $0.2 \times 10^{-5}$ )	.28	22.22
Drugs . . . .	-6.21* (1.29)	0.17* (0.02)	$-0.4 \times 10^{-3}$ * ( $0.1 \times 10^{-3}$ )	.40	19.35

B. ESTIMATION OF THE REGRESSION EQUATION  
 $R_{i,t}/S_{i,t} = a_0/S_{i,t} + a_1 + a_2S_{i,t} + a_3(I_{i,t-1}/S_{i,t}) + a_4D_i$  FOR SAME SAMPLES

Industry	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$R^2$	$F$
Chemicals .	0.05 (0.04)	0.009 (0.005)	$0.006 \times 10^{-7}$ ( $0.28 \times 10^{-7}$ )	0.09* (0.03)	0.020* (0.006)	.50	13.81
Drugs . . .	-2.18 (1.46)	0.01 (0.03)	$-0.14 \times 10^{-3}$ ( $0.09 \times 10^{-3}$ )	0.45* (0.13)	0.44* (0.10)	.72	22.07

<sup>a</sup> See footnotes to Table 1.

steadily throughout. Thus, the results of testing the Schumpeterian hypothesis on these samples show quite different behavior in the two industries. This is in essential agreement with Mansfield<sup>18</sup> and others investigating this question using different firm samples and time periods.<sup>19</sup>

Given these results, let us now turn to an examination of the relation of size to the three research determinant variables used in equation (1). In Table 3, the simple correlations between size as measured by sales and our technological, diversification, and availability of funds variables

<sup>18</sup> Mansfield's estimates (1964, pp. 333-34) were based on a logarithmic model (log R & D regressed on log sales) for pooled time-series cross-sectional data over the period 1945-59. For a review of other recent investigations of this question, see Markham (1965, p. 328-32).

<sup>19</sup> A qualification to the interpretation of the results obtained here and in the studies cited above arises from the fact that total firm sales is used as the independent variable rather than sales only in the particular three-digit industry for which the firm produces its principal products. Now, if it is the case that the nature of firm products in each of these industries varies significantly with size, then the above relationships may be due principally to this phenomenon rather than to any scale effects, such as those postulated above. As indicated earlier (see n. 9 above), the large firms in the drug industry do tend to expand and diversify into non-pharmaceutical areas where the opportunities for R & D applications are low. It is therefore very desirable to estimate equation (2) using only pharmaceutical sales for each firm. Until the required data are available, however, the present approach must be used in such studies and, accordingly, is quite tentative in nature.

## R and D

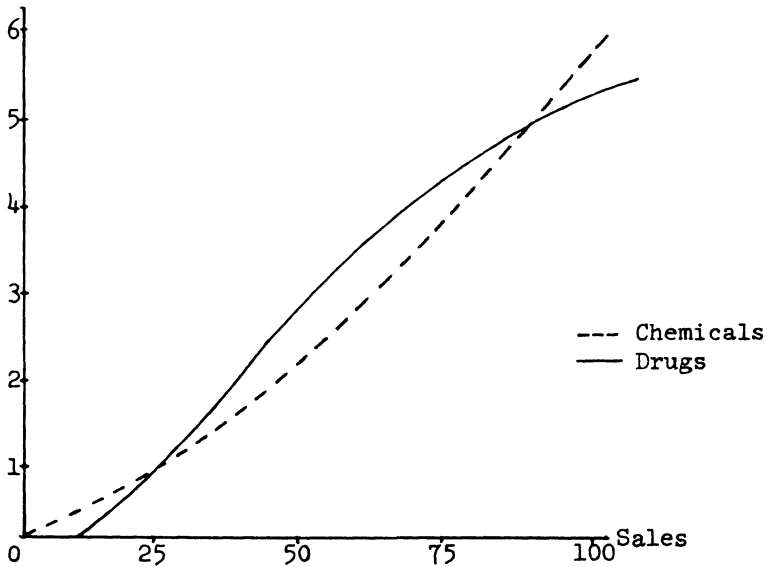


FIG. 1.—Estimated quadratic regression equation of R & D on sales for chemical and drug firms. The units of sales have been changed so that firms in the sample cannot be identified. The sales of the largest firms in each sample are represented in the new units by 100, and the smallest firms in each sample are between 1 and 10.

( $P_i$ ,  $D_i$ , [ $I_{i,t-1}/S_{i,t}$ ]) are presented. First, it may be noted that, of the three variables, the patent variable is the least size-correlated and varies oppositely to the relationship between firm size and research intensity in the two industries. The absence of a significant correlation in this variable is particularly important for the policy questions at stake here. Of all the arguments put forth by the proponents of corporate bigness, the one maintaining large firms have higher research productivity is potentially the most significant. However, the results of Table 3 clearly do not offer

TABLE 3  
SIMPLE CORRELATION OF SIZE AND VARIABLES  $P_i$ ,  $D_i$ , AND  
( $I_{i,t-1}/S_{i,t}$ ) FOR THE CHEMICAL AND DRUG INDUSTRIES

Industry	$P_i$	$D_i$	( $I_{i,t-1}/S_{i,t}$ )
Chemicals . . .	-0.2	0.8**	0.5**
Drugs . . . . .	0.3*	0.2	-0.4**

\* Significant at .05 level.

\*\* Significant at .01 level.

any real support for this position.<sup>20</sup> With regard to the other two determinants, the internally generated funds and diversification variables are significantly positive in the chemical industry, which exhibited a positive structural relation with size. However, in drugs, where the relation was negative, the internally generated funds variable is significantly negatively correlated with size, and diversification exhibits a positive but insignificant correlation coefficient. Thus, the results suggest that the difference in the observed relation between research intensity and firm size in the two industries is due primarily to the flow of funds variable and, to a lesser extent, the diversification index.

In the current analysis, it has been assumed that the significant relation between research intensity and firm size follows from the size behavior of the underlying independent variables specified in the model. An alternative interpretation is that size is the relevant determinant variable and that the significance of these explanatory variables of the model is spurious, arising from their mutual correlation with size. In order to investigate this alternative hypothesis, a further regression equation was estimated. This involved adding the two significantly size-correlated variables, diversification and the measure of cash flow, to equation (2) and observing the behavior of the previously estimated coefficients between R & D and the size of the firm. As shown in the lower half of Table 2, when these two new variables are present, the  $a_1$  and  $a_2$  coefficients lose their statistical significance in both industries. On the other hand, the estimates associated with the two determinant variables,  $a_3$  and  $a_4$ , are significant, as in Table 1. These results are thus consistent with the general interpretation of the regression estimates presented in this paper and not with the alternative hypothesis advanced above.

In summary, the analysis here indicates that there is no basis for the presumption that larger firms will necessarily possess the characteristics that promote a high degree of research intensity. Therefore, it is not surprising that tests of the Schumpeterian hypothesis have yielded such diverse results across industries. While it must be kept in mind that the results here only concern the firms in the *Fortune 500* listings, it is also true that this range of firms is most relevant from the standpoint of anti-trust policy.

#### IV. Summary and Conclusions

The main conclusion that emerges from the analysis is that interfirm differences in technology, product diversification, and availability of

<sup>20</sup> The empirical results here would also be open to question if the propensity to patent were significantly correlated with firm size. By similar procedures to those discussed in the previous section, this was not indicated to be the case. See also discussion of this issue in Scherer (1965a, p. 1110-13).

funds are all significant in explaining firm research intensity. The model presented here fits the two more research-oriented industries, chemicals and drugs, much better than petroleum refining. From a policy standpoint, these results indicate that the level of R & D expenditures will be sensitive to the broad class of government policy devices that affect the financial incentives confronting the firm. In particular, fiscal devices and other policy measures can be expected to influence the level of R & D expenditures through both profitability and flow of funds effects. Government policy actions directed toward stimulating growth should, therefore, be concerned with both of these effects if efficient programs are to be devised.

The major limitations of the present study arise from the substantial conceptual and empirical difficulties in measuring items such as R & D activity and its outputs. Considerable effort was therefore expended to obtain the best set of data that are currently available. While the results must be viewed as tentative, particularly because the total number of firms is small, they are nevertheless internally consistent and in general agreement with some of the past work in this area. They should, therefore, be of interest to economists and policy makers concerned with the economics of R & D.

## References

- Eisner, Robert, and Strotz, Robert H. "Determinants of Business Investment," in Commission on Money and Credit, *Impacts of Monetary Policy*. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1963.
- Grabowski, Henry. "The Determinants and Effects of Industrial Research and Development." Ph.D. dissertation, Princeton Univ., 1967.
- Hall, Marshall, "The Determinants of Investment Variation in Research and Development," *IEEE Transactions on Engineering Management*, EM-11 (March, 1964), pp. 8-15.
- Hamberg, Daniel. "Size of Firm, Oligopoly, and Research," *Canadian J. Econ. and Polit. Sci.*, XXX (February 1964), 62-75.
- Kuh, Edwin, and Meyer, John R. "Correlation and Regression Estimates When Data are Ratios," *Econometrica*, XXIII (October, 1955), 400-416.
- Kuznets, Simon. "Inventive Activity: Problems of Definition and Measurement," in National Bureau of Economic Research, *The Rate and Direction of Inventive Activity*. Princeton, N.J.: Princeton Univ. Press, 1962.
- Mansfield, Edwin. "Industrial Research Expenditures: Determinants, Prospects and Relation to Firm Size and Inventive Output," *J.P.E.*, LXXII (August, 1964), 319-40.
- Markham, Jesse W. "Market Structure, Business Conduct and Innovation," *A.E.R.*, LV (May, 1965), 323-32.
- Mueller, Dennis C. "The Firm Decision Process: An Econometric Investigation," *Q.J.E.*, LXXXI (February, 1967), 58-87.
- National Science Foundation. *Science and Engineering in American Industry*. Washington: Nat. Sci. Found., 1956.
- . *Research and Development in Industry, 1961*. Washington: Nat. Sci. Found., 1964.

- Nelson, Richard R. "The Simple Economics of Basic Scientific Research," *J.P.E.*, LXVII (June, 1959), 297-306.
- Sanders, Barkev S. "Some Difficulties in Measuring Inventive Activity," in National Bureau of Economic Research, *The Rate and Direction of Inventive Activity*. Princeton, N.J.: Princeton Univ. Press, 1962.
- Scherer, F. M. "Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions," *A.E.R.*, LV (December, 1965), 1097-1125. (a)
- . "Size of Firm, Oligopoly, and Research: A Comment," *Canadian J. Econ. and Polit. Sci.*, XXXI (May, 1965) 256-66. (b)