COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 136

The Expenditures of the Firm on Research and Development

Edwin Mansfield

March 22, 1962

The Expenditures of the Firm on Research and Development*

by

Edwin Mansfield

Yale University and Carnegie Institute of Technology

1. Introduction

The past few years have witnessed a much more intense and widespread interest in the economics of research and development. Several recent studies have suggested that the increase in per capita output in this country during the last fifty years has resulted largely from improvements in technology. Our rate of economic growth is now the subject of considerable national concern, and these results, despite their crudeness, have stimulated interest in the research activities that help to produce technical improvements.[1]

Despite the heightened interest in research and development, there have been few attempts, if any, to construct econometric models to help explain differences among firms in the amount spent for such purposes. Moreover, there seems to have been no attempt to determine how such expenditures are related to the number of important inventions or innovations that a firm produces. My purpose in this paper is to investigate both of these subjects. The results, though they are rough, should shed new light on these important aspects of the economics of research and development.[2]

The plan of the paper is as follows. Section 2 presents a simple model to explain the level of a firm's expenditures on research and development. Section 3 tests this model and Section 4 discusses the economic implications of the empirical results. Section 5 estimates an important parameter in this model for 35 firms and Section 6 discusses the usefulness of these estimates as forecasting devices. Section 7 examines the relationship between a firm's R and D expenditures and the number of significant inventions or innovations it produced. Section 8 concludes the paper.

2. A Simple Model

This section presents a model to help explain the level of a firm's expenditure on research and development. According to this model, a firm's managers, when they determine how much to spend on R and D during the coming year, have in mind some desired, or target, amount that they would spend if they could instantaneously make all of the desired adjustments in personnel and plant and if they could avoid the inefficiencies involved in too rapid an increase in R and D expenditures. Taking this year's expenditures as a base from which to figure, they set next year's R and D budget so as to move a certain fraction of the way toward this desired amount.[3]

If $r_i(t)$ is the i^{th} firm's R and D budget for year t, $\tilde{R}_i(t)$ is its desired expenditure for year t, and $R_i(t-1)$ is its actual expenditure in year $(t-1)$, the model implies that

$$(1) \qquad r_i(t) = R_i(t-1) + \theta_i(t) [\tilde{R}_i(t) - R_i(t-1)] ,$$

where $\theta_i(t)$ is the fraction of the way the i^{th} firm moves toward the desired expenditure. Assuming that the actual expenditure in year t equals the budgeted amount plus a random error term,

$$(2) \qquad R_i(t) = R_i(t-1) + \theta_i(t) [\tilde{R}_i(t) - R_i(t-1)] + z_i(t) ,$$

where $z_i(t)$ is a random variable with zero expected value.[4]

Before proceeding any further, note two things regarding the model. First, it is meant to apply only to industries like chemicals, drugs, petroleum, etc., where government research contracts are of little importance. Although it is important to understand the determinants of military and other government spending on R and D , this is not an area where data are easily available or where economic analysis is likely to be very useful.[5] Second, the model is a more apt representation of decision making regarding development and applied research than basic research. However, this is not too important for present purposes because the former types of expenditures dominate the total.

Returning to equation (2), there are two variables, $\tilde{R}_i(t)$ and $\theta_i(t)$, that must be explained. What determines $\tilde{R}_i(t)$? Suppose that the i^{th} firm, when planning the extent of its research activities in year t , lists various R and D projects that could be carried out and makes a rough judgment as to the technical and commercial promise of each one. Drawing on these data and its past experience, it estimates the amount it could spend in year t on projects with estimated rates of return exceeding 20 percent, 15 percent, etc. Suppose that the results show that

$$(3) \qquad I_{it}(\rho) = M_i(t)e^{-\rho/\bar{\rho}_i(t)} ,$$

where $I_{it}(\rho)$ is the amount that could be spent on projects with anticipated rates of return exceeding ρ, $M_i(t)$ is the amount that could be spent with positive expected returns, and $\bar{\rho}_i(t)$ is the average anticipated rate of return of projects with positive expected returns.[6]

In addition, suppose that the firm has a certain minimum expected rate of return, $\rho_i^*(t)$, and that, if it could make the necessary adjustments instantaneously and without inefficiencies stemming from too rapid an expansion, it would accept all projects exceeding this minimum expected rate of return. If so,

$$(4) \qquad \tilde{R}_i(t) = M_i(t)e^{-\rho_i^*(t)/\bar{\rho}_i(t)}$$

Moreover, since one very important determinant of $M_i(t)$ is the i^{th} firm's size in year t, we assume that

$$(5) \qquad M_i(t) = \nu_{1t}\, S_i(t)^{\nu_{2t}}\, k_i(t) ,$$

where v_{1t} varies from industry to industry, v_{2t} is positive, $S_i(t)$ is the i^{th} firm's sales in year t, and $k_i(t)$ is a random error term.[7] Combining equations (4) and (5),

$$(6) \qquad \ln \tilde{R}_i(t) = \ln v_{1t} + v_{2t} \ln S_i(t) - \rho_i^*(t)/\bar{\rho}_i(t) + \ln k_i(t) .$$

Of course, $\ln v_{1t}$ depends on the profitability of R and D in this industry, v_{2t} depends on the sensitivity of $M_i(t)$ to differences in size of firm, $\rho_i^*(t)$ reflects the profitability of alternative uses of the i^{th} firm's money, and $\bar{\rho}_i(t)$ depends on the actual profitability of the i^{th} firm's prospective R and D projects and the amount its competitors have been spending on R and D.[8]

What determines $\theta_i(t)$? First, one would expect $\theta_i(t)$ to be inversely related to $[\tilde{R}_i(t) - R_i(t-1)]/R_i(t-1)$. If attaining its desired level of expenditures implies that a firm's R and D activities must be increased by a very large percentage, the firm probably will move a smaller proportion of the way toward this desired level than if only a small percentage increase in its R and D activities were implied. There are considerable costs involved in a very rapid expansion of a firm's R and D department, the importance of which was stressed in interviews with various executives in the industries described below. Of course, we assume that $\tilde{R}_i(t) > R_i(t-1)$, but this almost always seems to have been the case during the postwar period, and it will probably continue to be so in the near future.[9]

Second, one would expect $\theta_i(t)$ to be directly related to the ratio of the i^{th} firm's profits to its R and D expenditures in year $(t-1)$. Holding $[\tilde{R}_i(t) - R_i(t-1)]/R_i(t-1)$ constant, this factor is directly related to the percent of a firm's profits in year $(t-1)$ that would have been absorbed by an increase in R and D expenditures to the desired level. If a large percent of a firm's profits would have been absorbed by such an increase, the firm will tend to be more cautious in moving toward the desired level. Conversely, if a firm spent less of its profits on R and D than others in its industry during year $(t-1)$, it may move more rapidly toward its desired level in an effort to catch up with its competitors. This factor was also stressed in the interviews, and we assume once again that $R_i(t) > R_i(t-1)$.[10]

Letting $P_i(t-1)$ be the ratio of the i^{th} firm's profits to its R and D expenditures in year $(t-1)$ and $\Pi_i(t-1)$ be the ratio of $P_i(t-1)$ to the average value of $P_i(t-1)$ in the industry, we assume that

$$(7) \qquad \theta_i(t) = \nu_3 + \nu_4 \frac{R_i(t-1)}{[\tilde{R}_i(t) - R_i(t-1)]} + \nu_5 \Pi_i(t-1) + k_i'(t) ,$$

where ν_4 and ν_5 are positive and $k_i'(t)$ is a random error term. Of course, equation (7) is assumed to hold only within a certain range -- presumably the relevant one. We would expect ν_3 to differ among industries, allowing for such factors as the extent to which engineering and scientific personnel are in relatively short supply. But there is no evidence of such differences between petroleum and chemicals -- the industries treated in the following section.[11]

Taken together, equations (2), (6), and (7) constitute a simple model to explain interfirm differences in the level of R and D expenditures during year t, the exogenous variables being $R_i(t-1)$, $S_i(t)$, $\rho_i^*(t)/\bar{\rho}_i(t)$, and $\Pi_i(t-1)$. Of course, this model is over-simplified in many respects. To keep it manageable and operational, we restricted ourselves to only a few exogenous variables that could be measured at least roughly. Moreover, although it could readily be generalized to explain changes over time in a firm's R and D expenditures, the model as it stands is designed only to explain interfirm differences at a given point in time.[12]

3. Tests of the Model

To test the model, I estimate the parameters in equations (6) and (7) for the chemical and petroleum industries in 1958 and test whether they have the hypothesized signs. To obtain these results, I use estimates of $\theta_i(t)$ and $\tilde{R}_i(t)$ derived from interviews and correspondence with eight major firms in these industries. The Appendix describes in detail how these data were obtained. Although the number of firms is quite small and the data are rough, the resulting estimates should be of considerable use in testing the model.

First, consider the parameters in equation (6). To estimate ν_{1t} and ν_{2t} and to see how well equation (6) can explain differences in $\tilde{R}_i(t)$, we need data regarding $\rho_i^*(t)/\bar{\rho}_i(t)$ and $S_i(t)$ as well as $\tilde{R}_i(t)$. Using interview data regarding $\rho_i^*(t)$ and $\bar{\rho}_i(t)$ for these eight firms in 1958 and data from <u>Moody's</u> regarding $S_i(t)$, we obtained the results in Table 1.[13] Using these results, we obtained least-squares estimates of V_0, V_1, and V_2, assuming that

$$(8) \qquad \ln [\tilde{R}_i(t)/S_i(t)] = V_0 + V_1 \ln S_i(t) - V_2\, \rho_i^*(t)/\bar{\rho}_i(t) + k_i''(t) ,$$

where V_0 differs between the two industries and $k_i''(t)$ is a random error term. Since R and D expenditures are customarily expressed as a percent of sales, we use $\tilde{R}_i(t)/S_i(t)$, rather than $\tilde{R}_i(t)$, as the dependent variable. If the model holds, V_2 equals one, V_1 equals $(\nu_{2t}-1)$, and V_0 equals the 1958 value of $\ln \nu_{1t}$ in each industry.[14]

The results are quite consistent with the model. After omitting the second term on the right hand side of equation (8) -- since $\hat{V}_1$ turns out to be statistically non-significant -- the regression equation is

$$(9) \qquad \ln [\tilde{R}_i(t)/S_i(t)] = \begin{Bmatrix} -3.66 \\ -1.29 \end{Bmatrix} - \underset{(.27)}{.92}\; \rho_i^*(t)/\bar{\rho}_i(t) ,$$

where the top figure in brackets pertains to the petroleum industry and the bottom figure pertains to chemicals. The estimate of V_2 is statistically significant, and it does not depart significantly from the theoretical value of one. Figure 1 shows that equation (9) fits the data very well, the coefficient of correlation (adjusted for degrees of freedom) being .95.[15]

Table 1 -- Values of $\tilde{R}_i(t)/S_i(t)$, $\rho_i^*(t)/\bar{p}_i(t)$, $\theta_i(t)$, $R_i(t-1)/[\tilde{R}_i(t)-R_i(t-1)]$, and $\Pi_i(t-1)$, Eight Firms, Chemical and Petroleum Industries, 1958[a]

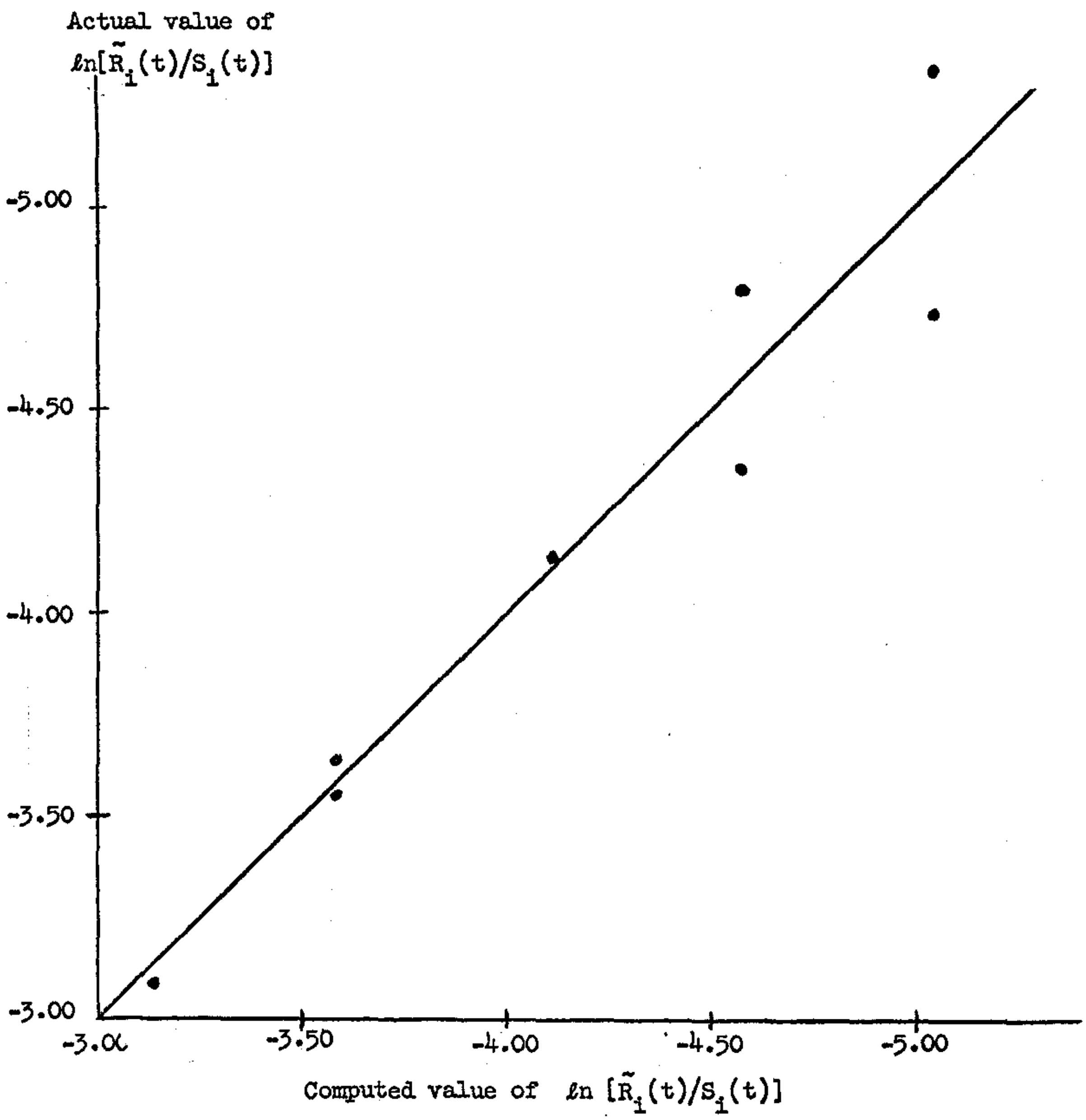
Firm	$\tilde{R}_i(t)/S_i(t)$	$\rho_i^*(t)/\bar{p}_i(t)$	$\theta_i(t)$	$R_i(t-1)/[\tilde{R}_i(t)-R_i(t-1)]$	$\Pi_i(t-1)$
Chemicals:					
1	.0262	2.5	1.00	5.6	1.2
2	.0455	2.0	.75	18.9	0.9
3	.0283	2.5	.23	3.6	0.8
Petroleum:					
1	.0048	1.5	1.00	7.6	1.3
2	.0088	1.5	.20	4.2	0.7
3	.0082	1.0	1.00	8.5	1.1
4	.0126	1.0	.10	1.1	0.9
5	.0165	.5	--	--	--

Source: see the Appendix.

[a] Symbols: $\tilde{R}_i(t)/S_i(t)$ is the ratio of the desired R and D expenditures of the i^{th} firm during year t to its sales during year t ; $\rho_i^*(t)/\bar{p}_i(t)$ is the ratio of the minimum rate of return expected from R and D projects by the i^{th} firm during year t to the average rate of return it believed it could obtain from R and D projects with non-negative expected returns; $\theta_i(t)$ equals $[r_i(t) - R_i(t-1)]/[\tilde{R}_i(t) - R_i(t-1)]$; $R_i(t)$ is the actual R and D expenditures of the i^{th} firm during year t ; and $\Pi_i(t)$ equals $P_i(t)/P(t)$, where $P_i(t)$ is the ratio of the i^{th} firm's profits to its R and D expenditures in year t and $P(t)$ is the average value of $P_i(t)$ in the industry during year t.

Most of the data were obtained with the understanding that they would remain confidential. Thus the names of the companies are not given. For company 3 in the chemical industry, the data refer to 1959, not 1958. For company 5 in the petroleum industry, $\tilde{R}_i(t) < R_i(t)$; thus, the model regarding $\theta_i(t)$ does not apply, and the last three columns of the table are left blank.

Figure 1 -- Plot of Actual Values of $\ell n[\tilde{R}_i(t)/S_i(t)]$ Against Those Computed from Equation (9). Eight Petroleum and Chemical Firms, 1958.[a]



Source: See the Appendix. The line is a 45° line through the origin.

[a] Symbols: $\tilde{R}_i(t)/S_i(t)$ is the ratio of the i^{th} firm's desired R and D expenditures to its sales in year t.

Next, consider the parameters in equation (7). Using the 1958 data regarding $\tilde{R}_i(t)$ and $R_i(t-1)$ obtained from interviews and correspondence with the eight firms and corresponding data regarding $\Pi_i(t-1)$ from Moody's, we obtained least-squares estimates of ν_3, ν_4, and ν_5. Since the estimates of ν_3 do not differ significantly between industries, we assume that they were the same and find that

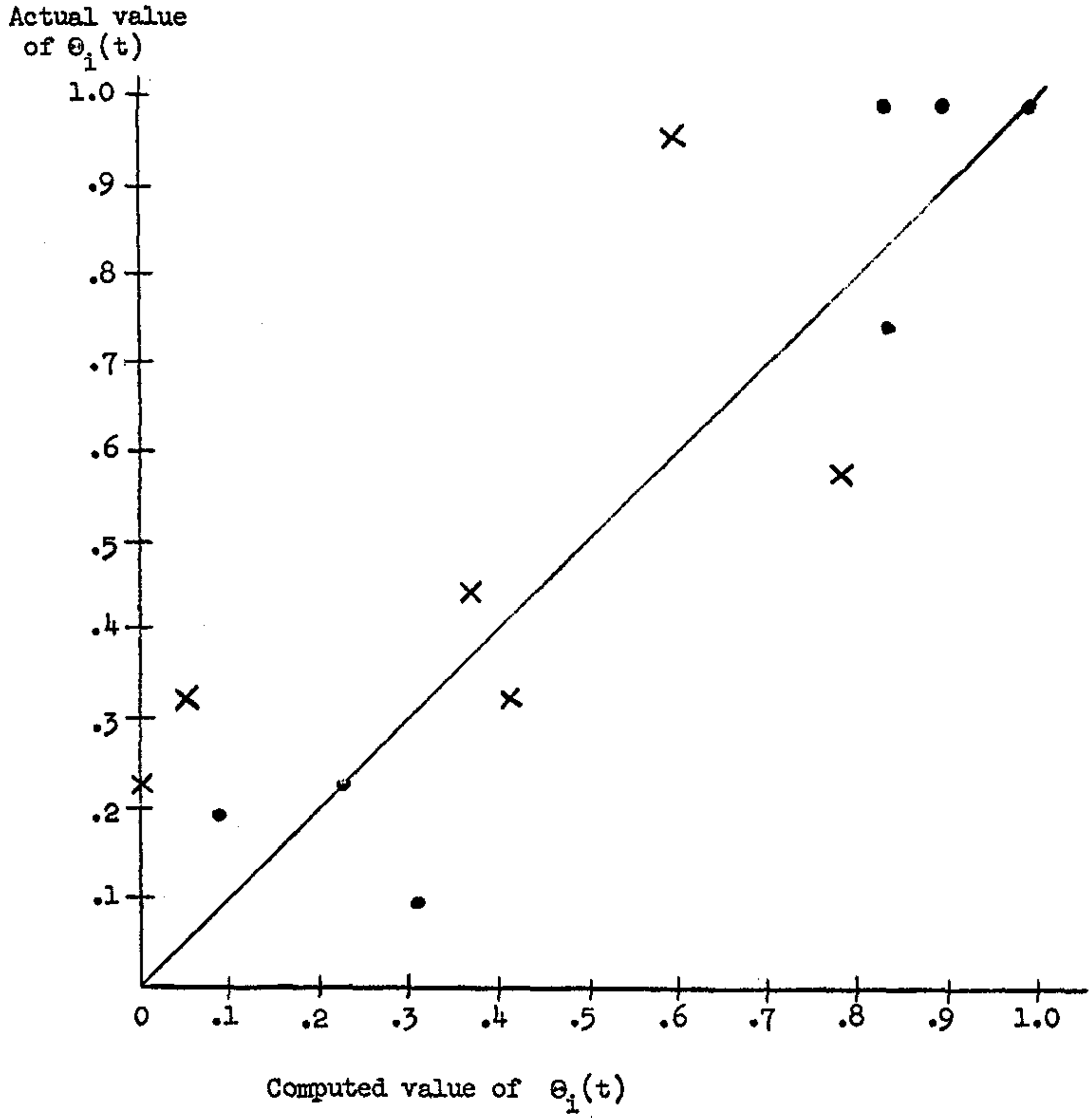
$$(10) \qquad \theta_i(t) = -1.10 + \underset{(.012)}{.029} \frac{R_i(t-1)}{[\tilde{R}_i(t)-R_i(t-1)]} + \underset{(0.31)}{1.53} \Pi_i(t-1) .$$

Again, the results are quite consistent with the model. Figure 2 shows that equation (10) fits the data quite well, the correlation coefficient being .91. The estimates of ν_4 and ν_5 have the expected signs and are statistically significant. Moreover, equation (10) seems to be reasonably effective (considering the sampling errors in ν_3, ν_4, and ν_5) in "predicting" the values of $\theta_i(t)$ for these firms in a few additional years for which we could obtain data. The average error of these "predictions" (also shown in Figure 2) is about .20.[16]

4. Implications of the Estimates

Because of the small number of observations and the roughness of the basic data, the empirical results in equations (9) and (10) are obviously tentative. But if reasonably trustworthy, they have four significant implications. First, they allow us to make rough estimates of the effect of certain kinds of

Figure 2 -- Plot of Actual Values of $\theta_i(t)$ Against Those Computed from Equation (10), Seven Petroleum and Chemical Firms, Selected Years.[a]



Source: see the Appendix. The line is a 45° line through the origin.

[a] Symbols: $\theta_i(t)$ equals $[r_i(t)-R_i(t-1)]/[\tilde{R}_i(t)-R_i(t-1)]$. The dots represent the data in Table 1 on which equation (10) was based. The X's represent "predictions" for other years for(firms where the necessary data could be obtained). Note that a computed value of $\theta_i(t)$ exceeding one was set equal to one. Similarly, a negative value was set equal to zero. These changes were made because $0 \leq \theta_i(t) \leq 1$. See note 11.

government policies on the amount a firm spends on research and development. For example, what would be the effect on a firm's expenditures of a change in policy, e.g., the tax laws, that increased the prospective profitability of each of its R and D projects by one percent? Assuming that the model holds and that the firm's actual and desired expenditures would be approximately equal, the firm's spending would increase by $\rho_i^*(t)/\bar{\rho}_i(t)$ percent -- which (for the firms for which we have data) would have been about one percent in 1958 for petroleum firms and two percent for chemical firms.[17]

Second, the fact that equation (9) fits so well seems to imply that the process by which a firm's R and D expenditures are determined is not so divorced from profit considerations as some observers have claimed. If firms "establish research laboratories without any clearly defined idea of what the laboratories could perform" [22, p. 29], and blindly devote some arbitrarily determined percentage of sales to R and D , it is difficult to see why equation (9) fits so well.[18]

Third, the fact that our estimate of v_{2t} is less than, but not significantly different from, one means that these data for large petroleum and chemical firms are consistent with the hypothesis that increases in size in this range result in no more than proportional increases in the amount of money a firm believes it can spend on R and D with positive returns. If this hypothesis holds and if $\rho_i^*(t)/\bar{\rho}_i(t)$ is independent of a firm's size (in this range), one would expect larger firms to spend no greater proportion of their sales on R and D than smaller firms in this range. This expectation is borne out by the facts.[19]

Fourth, the results allow us to make rough estimates of the effect of the amount of interfirm variation in $\bar{\rho}_i(t)$ on the total amount spent by the industry on research and development. Assuming for simplicity that all firms are of equal size, that they have the same value of $\rho_i^*(t)$, and that their actual and desired expenditures are equal, one can show that an increase in the amount of interfirm variation in $\bar{\rho}_i(t)$ results in higher industry spending if the average value of $\bar{\rho}_i(t)$ is relatively low, but that it results in lower industry spending if the average value of $\bar{\rho}_i(t)$ is relatively high.[20]

5. Estimates of $\tilde{R}_i(t)/S_i(t)$ in Five Industries

In view of the significance of $\tilde{R}_i(t)/S_i(t)$, it is important to understand its recent behavior. My purpose in this section is to estimate the value of this parameter during 1945-58 for thirty-five firms in five industries. Since it was generally impossible to obtain direct estimates like those in Table 1 and since data on $\rho_i^*(t)/\bar{\rho}_i(t)$ are generally lacking, two rather bold assumptions are made in order to obtain indirect estimates of $\tilde{R}_i(t)/S_i(t)$. Although these assumptions are not unreasonable and although they stand up to various statistical tests, their roughness should be noted.

Combining equations (2) and (7), we have

(11)

$$R_i(t) = R_i(t-1) + [\nu_3 + \nu_4 \frac{R_i(t-1)}{[\tilde{R}_i(t)-R_i(t-1)]} + \nu_5 \Pi_i(t-1)$$

$$+ k_i'(t)] [\tilde{R}_i(t) - R_i(t-1)] + z_i(t) ,$$

$$= (1 + \nu_4) R_i(t-1) + [\nu_3 + \nu_5 \bar{\Pi}_i] [\tilde{R}_i(t) - R_i(t-1)] + \left\{\nu_5[\Pi_i(t-1) - \bar{\Pi}_i]\right.$$

$$\left. + k_i'(t)\right\} [\tilde{R}_i(t) - R_i(t-1)] + z_i(t) ,$$

where $\bar{\Pi}_i$ is the average value of $\Pi_i(t-1)$ for the i^{th} firm during 1945-58.

The two assumptions are as follows: First, we assume that a firm's desired R and D expenditures, as a percent of its sales, was a linear function of time with random disturbances during 1945-58. That is,

$$\tilde{R}_i(t)/S_i(t) = \alpha_{i1} + \alpha_{i2}t + u_i(t) , \tag{12}$$

where $u_i(t)$ is a random error term, time is measured in years from 1945, and this equation is assumed to hold only during 1945-58. Second, we assume that the last term on the right hand side of equation (11) plus $u_i(t)\ S_i(t)\ [\nu_3 + \nu_5 \bar{\Pi}_i]$ can be treated as a random error term with zero expected value and that it is independent of $R_i(t-1)$, $S_i(t)$, and t .

Judging from interviews with various firms and published descriptions of post-war developments, the first assumption seems to be a reasonable first approximation. Moreover, as we shall see, some statistical tests, when applied to a sample of the firms described below, seem to support it. To test the second assumption, we estimated the last term on the right hand side of equation (11) plus $u_i(t)S_i(t)[\nu_3 + \nu_5 \bar{\Pi}_i]$ and tested whether its expected value was zero and whether it was independent of $R_i(t-1)$, $S_i(t)$, and t. Only a small amount of data for two firms were available, but these data were quite consistent with the second assumption.[21]

Combining equations (11) and (12) and bearing in mind the second assumption,

$$(13) \qquad R_i(t) = \beta_{i1} R_i(t-1) + \beta_{i2} S_i(t) + \beta_{i3}\, t\, S_i(t) + z_i''(t),$$

where $\beta_{i1} = (1 + \nu_4 - \nu_3 - \nu_5 \bar{\Pi}_i)$, $\beta_{i2} = \alpha_{i1}(\nu_3 + \nu_5 \bar{\Pi}_i)$, $\beta_{i3} = \alpha_{i2}(\nu_3 + \nu_5 \bar{\Pi}_i)$, and $z_i''(t)$ is a random error term. Moreover, if the model is correct, one would suppose that[22]

$$(14) \qquad 0 \leq \beta_{i1}, \beta_{i2}, \beta_{i3} \leq 1 .$$

Having made these assumptions, data were gathered regarding the annual R and D expenditures and sales of 35 firms in the chemical, drug, petroleum, steel, and glass industries. They pertain to 1945-58 and were obtained mainly from Langenhagen [13], Moody's, and correspondence with the firms. Before estimating $\tilde{R}_i(t)/S_i(t)$, we tested the assumption in equation (12). Taking a sample of these firms, we tested whether a quadratic function of t would fit the data better than the linear function in equation (12). In general, there was no evidence that this was the case.[23]

Next, we used equation (13) to obtain least squares estimates of β_{i1}, B_{i2}, and β_{i3} for each firm. The results -- shown in Tables 2 - 3 -- indicate that these estimates almost always have the signs predicted by the model and that they almost always are of the expected orders of magnitude. And despite relatively few degrees of freedom, they are generally statistically significant. Moreover, equation (13) seems to represent the data for each firm quite well, the correlation coefficients (adjusted for degrees of freedom) generally exceeding .95.[24]

Finally, we estimated $\tilde{R}_i(t)/S_i(t)$ for each firm. Letting $\hat{\beta}_{i1}$, $\hat{\beta}_{i2}$, and $\hat{\beta}_{i3}$ be our estimates of β_{i1}, β_{i2}, and β_{i3}, we used $\hat{\alpha}_{i1} = \hat{\beta}_{i2}/(1-\hat{\beta}_{i1}+\hat{v}_4)$ and $\hat{\alpha}_{i2} = \hat{\beta}_{i3}/(1 - \hat{\beta}_{i1} + \hat{v}_4)$ as estimates of α_{i1} and α_{i2}. These estimates are consistent but biased. To estimate $\tilde{R}_i(t)/S_i(t)$, we inserted $\hat{\alpha}_{i1}$ and $\hat{\alpha}_{i2}$ into equation (12) and ignored $u_i(t)$. The omission of $u_i(t)$ should have little effect on our conclusions.[25]

Table 4 shows the mean and dispersion of the resulting estimates of $\tilde{R}_i(t)$ $S_i(t)$ for the drug, chemical, and petroleum firms for each year from 1945 to 1958. The steel and glass firms were excluded because of the small number of cases. As one would expect, there was a considerable increase during 1945-58 in the average value of $\tilde{R}_i(t)/S_i(t)$ in each industry, the percentage increase being much greater in drugs and petroleum than in chemicals. It increased by about 110 percent in drugs, 80 percent in petroleum and 20 percent in chemicals.[26]

Table 2 -- Estimates of β_{i1}, β_{i2}, and β_{i3}, Correlation Coefficient, and Percentage Error in Forecasting 1959 R and D Expenditures, 19 Firms, Chemical and Petroleum Industries.[a]

Company	---Estimates------			--Standard Errors--			Correlation Coefficient	Forecasting Error
	β_{i1}	β_{i2}	β_{i3}	β_{i1}	β_{i2}	β_{i3}		
				---------Chemicals---------				
Hercules	.51*	.0131*	.00079	.22	.0051	.00045	.97	-5.0%
Union Carbide	.64*	.0063*	.00122*	.15	.0029	.00049	.99	3.2
Diamond Alkali	.65*	.0083*	.00035	.28	.0035	.00073	.98	5.0
Hooker	.55*	.0091*	.00037*	.06	.0012	.00010	1.00	13.1
Allied	.82*	.0079*	-.00024	.25	.0038	.00028	.98	-3.9
General Aniline	.60*	.0286*	-.00092*	.20	.0114	.00040	.86	-9.6
MMM	.19	.0168*	.00151*	.21	.0040	.00047	.99	20.5
Atlas	1.04*	-.0014	.00059	.18	.0027	.00036	.98	8.8
American Cyanamid	.70*	.0212*	-.00048	.17	.0072	.00029	.98	-6.9
Monsanto	.50*	.0104*	.00095*	.11	.0024	.00029	.99	-2.4
Total[b]	--	--	--	--	--	--		1.5
				-------- Petroleum---------				
Firm 1	.05	.00003	.000511*	.36	.00056	.000205	.99	8.0%
Firm 2	.75*	.00094*	.000047*	.17	.00047	.000021	1.00	2.0
Firm 3	.53*	.00261*	.000063	.17	.00060	.000047	.99	-2.8
Firm 4	.91*	.00080	.000063	.14	.00088	.000041	.99	4.0
Firm 5	1.09*	.00036	-.000025	.13	.00036	.000070	1.00	2.0
Firm 6	.04	.00438*	.000903*	.23	.00181	.000270	.97	47.9
Firm 7	.73*	.00240	.000019	.23	.00151	.000052	.98	2.4
Firm 8	.55*	.00247*	.000107*	.15	.00065	.000041	.99	-2.3
Firm 9	.86*	.00107	-.000004	.19	.00095	.000032	.97	6.9
Total[b]	--	--	--	--	--	--		4.4

Source: See the Appendix.

a/ The data for the petroleum firms were obtained with the understanding that it would remain confidential. Thus only numbers appear in the first column.

b/ This is the error in forecasting the sum of these firms' 1959 expenditures, using the sum of the individual firms' forecasts to forecast this total.

* Significantly different from zero (.05 probability level and one-tailed test).

Table 3 -- Estimates of β_{i1}, β_{i2}, and β_{i3}, Correlation Coefficient, and Percentage Error in Forecasting 1959 R and D Expenditures, 16 Firms, Drug, Steel, and Glass Industries.[a]

Company	Estimates β_{i1}	Estimates β_{i2}	Estimates β_{i3}	Standard Errors β_{i1}	Standard Errors β_{i2}	Standard Errors β_{i3}	Correlation Coefficient	Forecasting Error
Drugs								
Firm 1	1.07*	-.00022	.00038	.21	.00880	.00034	.99	-3.8%
Firm 2	.45*	.02644*	.00035	.17	.00785	.00044	.95	0.5
Firm 3	.63*	.00711*	.00112*	.16	.00343	.00036	.99	1.7
Firm 4	.49*	-.00566	.00444*	.17	.00738	.00110	.99	3.6
Firm 5	.66*	.00990*	.00071	.21	.00487	.00052	.97	-5.4
Firm 6	.56*	.03086*	.00117	.22	.00953	.00079	1.00	-3.1
Firm 7	.77*	.01984*	.00087	.11	.01038	.00074	.99	-6.5
Firm 8	.54	.01843	.00168*	.36	.01379	.00086	.97	-1.4
Total[b]	--	--	--	--	--	--		-1.8
Steel								
Firm 1	.28	.001609*	.000035	.20	.000410	.000023	.98	12.5%
Firm 2	.95*	-.000017	.000172*	.09	.000788	.000087	1.00	-17.7
Firm 3	.75*	-.000264	.000368*	.12	.000803	.000139	1.00	-16.0
Firm 4	-.45	-.000357	.000142*	.39	.000215	.000043	.95	4.5
Total[b]	--	--	--	--	--	--		-6.1
Glass								
Firm 1	.88*	.00284	.00018	.25	.00235	.00033	.98	-5.4%
Firm 2	.26	.01597*	.00017	.22	.00389	.00019	.98	17.9
Firm 3	.65*	.00528*	.00049	.36	.00289	.00071	.95	14.0
Firm 4	-.03	.02297*	.00117*	.29	.00668	.00049	.95	-3.4
Total[b]	--	--	--	--	--	--		3.3

Source: See the Appendix.

a/ The data for the firms were obtained with the understanding that it would remain confidential. Thus only numbers appear in the first column.

b/ This is the error in forecasting the sum of these firms' 1959 expenditures, using the sum of the individual firms' forecasts to forecast this total.

* Significantly different from zero (.05 probability level and one-tailed test).

Table 4 -- Mean and Coefficient of Variation of $\tilde{R}_i(t)/S_i(t)$, Chemical, Petroleum, and Drug Firms, 1945-58.[a]

Year	Chemical Firms Mean	Chemical Firms Coefficient of Variation	Petroleum Firms Mean	Petroleum Firms Coefficient of Variation	Drug Firms Mean	Drug Firms Coefficient of Variation
1945	.0341	.64	.0049	.46	.0371	.73
1946	.0346	.59	.0052	.42	.0404	.64
1947	.0352	.54	.0055	.38	.0436	.57
1948	.0358	.50	.0059	.36	.0469	.51
1949	.0364	.46	.0062	.35	.0502	.46
1950	.0370	.42	.0065	.34	.0534	.42
1951	.0376	.38	.0068	.34	.0567	.39
1952	.0382	.35	.0071	.35	.0599	.36
1953	.0388	.32	.0074	.35	.0632	.34
1954	.0393	.30	.0078	.36	.0664	.33
1955	.0399	.29	.0081	.37	.0697	.32
1956	.0405	.27	.0084	.38	.0730	.32
1957	.0411	.27	.0087	.40	.0762	.31
1958	.0417	.27	.0090	.41	.0795	.31

Source: Tables 2 - 3.

a/ The coefficient of variation is the standard deviation divided by the mean.

The coefficients of variation indicate that interfirm differences in $\tilde{R}_i(t)/S_i(t)$ were quite large immediately after the war, the relative variation being higher in chemicals and drugs than in petroleum. But with the passage of time these interfirm differences have narrowed appreciably. In drugs and chemicals the 1958 coefficient of variation was about half of what it had been in 1945; in petroleum it was about 10 percent below its 1945 level. Of course, this would be expected if, as we asserted in Section 2, firms tend to be influenced by their competitors' behavior and if they adjust toward the industry average.[27]

In conclusion, the estimates of $\hat{\beta}_{i1}$ in Tables 2 - 3 can also be used to test the model in the following way. The model asserts that $\beta_{i1} = 1 + \nu_4 - \nu_3 - \nu_5 \bar{\Pi}_i$. If so, it follows that the average value of β_{i1} in each industry should equal $1 + \nu_4 - \nu_3 - \nu_5$, which is .60 according to the estimates in equation (10) . Are the average values of $\hat{\beta}_{i1}$ close to .60 in each industry? It turns out that they are in all industries where the sample is reasonably large. In chemicals, petroleum, and drugs, the average value of $\hat{\beta}_{i1}$ is .61 in each case.[28]

6. Forecasts of R and D Expenditures

If one is willing to make certain assumptions about the future behavior of $\tilde{R}_i(t)$ and $\theta_i(t)$, equation (13) can be used to forecast a firm's expenditures on research and development. Although we are concerned principally with interfirm differences in R and D expenditures, we digress for a moment and discuss some forecasts of this sort. First, we see the extent to which one could have forecasted a firm's 1959 expenditures on R and D by simply assuming that $\tilde{R}_i(t)/S_i(t)$ would continue to change at the 1945-58 rate (ν_3, ν_4, ν_5, and $\overline{\Pi}_i$ remaining constant). Second, we see what would happen to the level of R and D expenditures in each industry during the Sixties if $\tilde{R}_i(t)/S_i(t)$ either remained constant at its 1960 level or if it continued to change at its 1945-58 rate (ν_3, ν_4, ν_5 and $\overline{\Pi}_i$ remaining constant).

To find out how well a firm's 1959 R and D expenditures could be forecasted in this way, the firm's 1959 sales, its 1958 R and D expenditures, and the estimates of β_{i1}, β_{i2}, and β_{i3} were inserted into equation (13), and the resulting forecast was compared with the firm's actual expenditures in 1959. The results in Tables 2 - 3 indicate that the forecasting errors for individual firms were almost always less than 10 percent and that they were only about 3 percent for the industry totals. Moreover, these forecasts were considerably better than those resulting from three standard naive models. Of course, this comparison is somewhat unfair because we assume that we have perfect information regarding the firm's 1959 sales, but when forecasts were based on 1959 sales data containing 10 percent errors, the results were still superior to the naive models. Thus, these forecasts seemed in this instance to be relatively useful.[29]

To help form some very rough judgments concerning the levels of R and D expenditures during the Sixties, it should be useful to see how such expenditures would behave in these industries under two sets of naive assumptions. First, we assume that α_{11}, α_{12}, v_3, v_4, v_5, and $\bar{\Pi}_i$ will remain at their 1945-58 levels during the Sixties -- which means in effect that $\tilde{R}_i(t)/S_i(t)$ will continue to increase at its 1945-58 annual rate. Second, we assume that v_3, v_4, v_5, and $\bar{\Pi}_i$ will remain at this 1945-58 level but that $\tilde{R}_i(t)/S_i(t)$ will stay at its 1960 level throughout the Sixties.

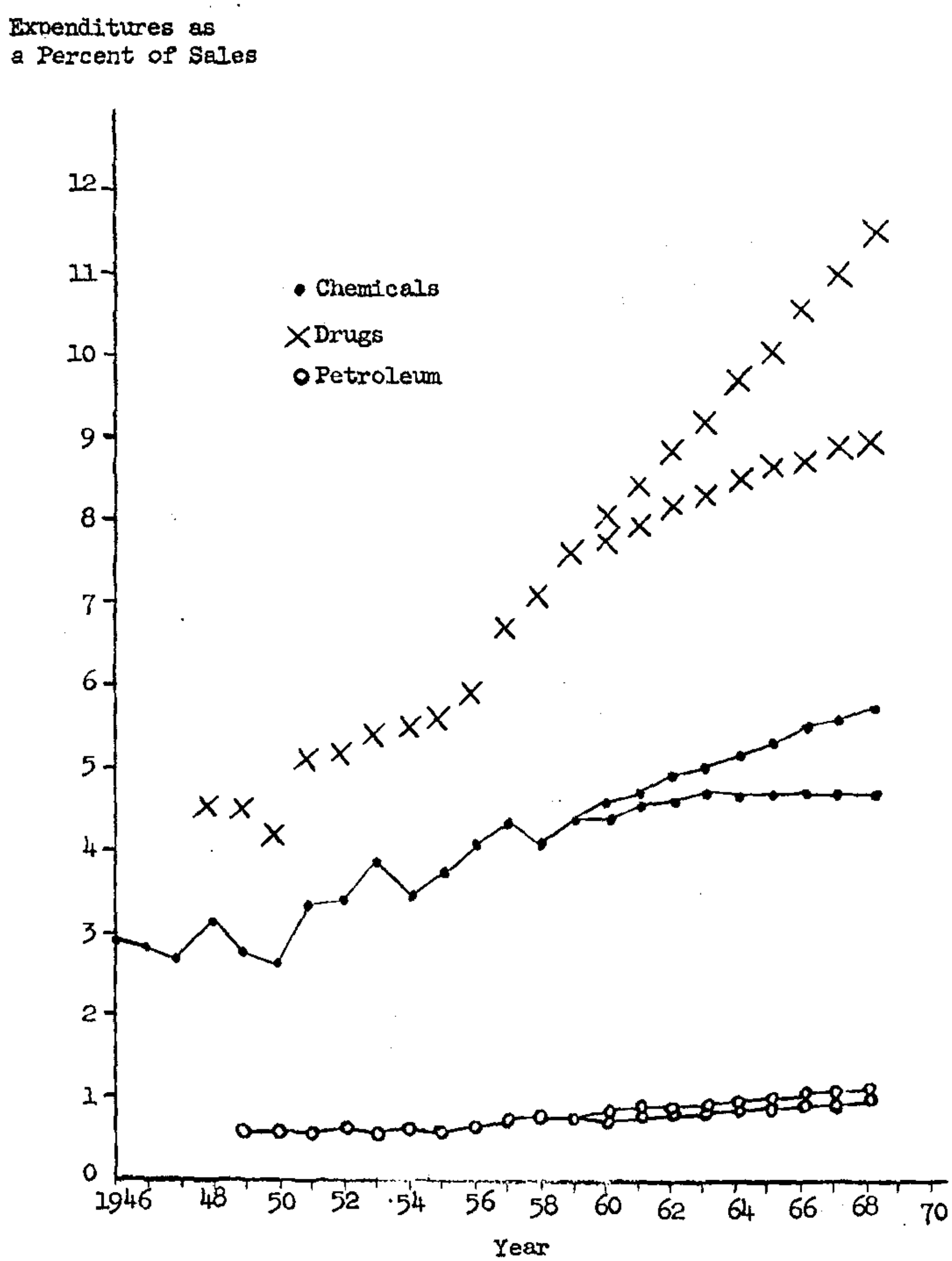
Under both sets of assumptions, we suppose that for 1960-69,

$$(15) \qquad S_i(t) = S_i(1 + r)^{t'}$$

where S_i is the i^{th} firm's sales in 1959, t' is measured in years from 1959, and $r = .04$ for petroleum firms, .05 for chemical firms, and .06 for drug firms. Using equation (15) we generate 1960-69 sales figures for each firm. To obtain the first set of forecasts, we insert these figures and $\tilde{\beta}_{11}$, $\tilde{\beta}_{12}$, and $\tilde{\beta}_{13}$ into equation (13). To obtain the second set, we do the same thing except that t is fixed at 15.[30]

Figure 3 shows the values of $\sum_i R_i(t)/\sum_i S_i(t)$ under each set of assumptions. These results suggest two interesting things. First, further increases in the ratio of R and D expenditures to sales can be expected during the Sixties in each of these industries even if there is no further increase in $\tilde{R}_i(t)/S_i(t)$. Unless the values of $\tilde{R}_i(t)/S_i(t)$ decline, which seems unlikely, R and D expenditures will continue to increase as a percent of sales. Second, unless

Figure 3 -- Expenditures on Research and Development as a Percent of Sales, Chemical, Petroleum, and Drug Firms, Actual (1946-59) and Projected (1960-69) Under Two Sets of Assumptions. a/



Source: See the text.

a/ Beyond 1959, the points are forecasts based on two naive assumptions. The higher of the points results from the first set of assumptions given in the text. The lower results from the second set of assumptions.

the values of $\tilde{R}_i(t)/S_i(t)$ increase more rapidly during the Sixties than they did during 1945-58, the increase in the ratio of R and D expenditures to sales seems likely to be smaller in most industries than during the Fifties.[31]

7. Expenditures on R and D and Inventive Output

So far, we have been concerned strictly with the determinants of a firm's expenditures on research and development. What sort of relationship exists between the level of a firm's expenditures and the number of significant inventions it produced? Even crude results bearing on this important question should be of interest because they help to indicate the extent to which there are economies of scale in R and D . Such economies of scale might be present because of the lumpiness of research equipment, the advantages of using highly specialized personnel, the greater chance that gains and losses will cancel when a large number of projects are undertaken, etc.[32] This section presents some very tentative findings regarding the chemical, petroleum, and steel industries.

For the chemical industry, we use Langenhagen's data [13] on the number of significant inventions (weighted roughly by a measure of their importance) carried out by each firm between 1940 and 1957. Letting this number be n_i for the i^{th} firm, we find that

$$(16) \qquad n_i = .18 + \underset{(.80)}{1.31}\, R_i + \underset{(.03)}{.08}\, R_i^2 , \qquad (r = .97)$$

where R_i is the average of the i^{th} firm's expenditures on R and D in 1940 and 1950. Data for fourteen firms could be included in this regression.[33]

For the petroleum industry, we use Schmookler's list [25] of important inventions in petroleum refining and my list [16] of important petrochemical innovations. Letting n_i be the combined number of refining inventions and petrochemical innovations carried out by the i^{th} firm between 1946 and 1956, we find that

$$(17) \qquad n_i = .81 + \underset{(.37)}{.59}\, R_i - \underset{(.06)}{.01}\, R_i^2, \qquad (r = .95)$$

where R_i is the average of the i^{th} firm's expenditures on R and D in 1945 and 1950. Data for eight firms could be included in this regression.

For the steel industry, my list [16] of important innovations is used. If n_i is the number of these innovations carried out by the i^{th} firm between 1946 and 1958, we find that

$$(18) \qquad n_i = .09 + \underset{(.40)}{1.06}\, R_i - \underset{(.08)}{.14}\, R_i^2 \qquad (r = .82)$$

where R_i is the average of the i^{th} firm's expenditures on R and D in 1946 and 1950. Data for eleven firms could be included in this regression.

These results suggest at least three things. First, the number of significant inventions carried out by a firm seems to be highly correlated with the size of its R and D expenditures. Although the payout from an individual R and D project is obviously very uncertain, it seems that there is a close relationship over the long run between the amount a firm spends on R and D and the total number of important inventions it produces.

Second, when a firm's size is introduced as an additional independent variable in equations (16) - (18), its effects turn out to be negative, but not quite statistically significant, in each case. Thus, the observed relationship between the number of inventions and the amount spent on R and D is not due merely to a relationship between a firm's size and its inventive output. On the contrary, when a firm's expenditures on R and D are held constant, increases in size seem to be associated with decreases in inventive output.[34]

Third, the evidence seems to suggest that increases in R and D expenditures result in more than proportional increases in inventive output in chemicals, but less than proportional increases in steel. In petroleum, there is no real indication one way or the other. Thus, except for chemicals, the results do not indicate any marked advantages of large scale research activities over medium-sized ones.[35]

Finally, the crudeness of these results should be noted. The measures of "inventive output" are obviously only the roughest approximations, since they include only inventions that are considered significant in some sense by informed observers and these inventions are weighted arbitrarily. Moreover, because of differences in accounting procedures and errors of measurement, the data regarding R and D expenditures may not be entirely comparable from one firm to another.[36]

8. Summary and Conclusions

The need for systematic research into the factors governing a firm's expenditures on R and D and the relationship of such factors to the extent of its inventive achievements seems clear. To help satisfy this need, I formulated in this paper a simple model to help explain the level of a firm's expenditures on R and D . According to this model, a firm sets its expenditures so as to move part way from last year's spending toward a desired level that depends on the firm's expectations regarding the profitability of the R and D projects at hand, the profitability of alternative uses of its funds, and its size. The firm's speed of adjustment toward the desired level depends on the extent to which the desired level differs from last year's spending and the percent of its profits spent last year on R and D.

This model seemed to fit historical data quite well and, when supplemented with additional assumptions, it seemed useful as a tool for short range forecasting. For eight firms where the necessary data could be obtained, estimates of the desired expenditures seemed in fact to be related in the way predicted by the model to the exogenous variables; and the interfirm differences in the speed of adjustment seemed to be consistent with the model. Moreover, assuming that each firm's desired expenditures (as a percent of its sales) was a linear function of time during 1945-58, the model could fit the 1945-58 data and forecast the 1959 data for 35 firms in five industries quite well.

Besides constructing and testing this model, the paper presented for the first time some rough estimates of the relationship between the level of a firm's

R and D expenditures and the number of significant inventions it produced during the relevant period. The relationship in the industries for which we could obtain data, chemicals, petroleum, and steel, seemed to be quite close, but, except for the chemical industry, there was no real indication of economies of scale within the range covered by the data. Because of the importance of such relationships, these results should be of interest, but considering the roughness of the measures of inventive output on which they are based, they should be treated with considerable caution.

Despite its obvious limitations, the paper may contribute in at least two ways to the formation of public policy in this area. First, the results provide a rough technique for estimating the effects of changes in tax laws, changes in patent laws (such as the recent Kefauver proposals), and changes in other public policies on the amount spent on R and D . To the extent that one can estimate the effect of such changes on $\rho_i^*(t)/\overline{\rho}_i(t)$, the model in Sections 2 - 3 can be used for this purpose. For example, if the effect of a tax change had been to raise the prospective profitability of each of a firm's R and D projects in 1953 by one percent, expenditures on R and D of petroleum firms would apparently have increased by about one percent and those of chemical firms would have increased by about two percent.

Second, the results provide new, but obviously tentative, evidence regarding the effects of industrial organization on the amount spent by an industry on R and D and the effectiveness of such expenditures in producing significant inventions. Among the large firms in three major industries (chemicals, petroleum, and steel), there was no tendency for larger firms to spend a greater percent of sales on R and D than smaller firms; and there was no real evidence, except in chemicals, of economies of scale in R and D in this range. Moreover, holding R and D expenditures constant, there was no evidence that the productivity of such expenditures was higher in the larger firms; if anything, the reverse seemed to be the case.

Finally, the findings may also be of interest to economists concerned with inventory behavior, investment in plant and equipment, and other areas where the acceleration principle has found extensive use. The model presented here employs a modified version of the accelerator in which the speed of adjustment depends on the discrepancy between the previous and desired level of the endogenous variable. This new version, for which there is considerable evidence in the case of research and development, is likely to be useful in other areas where the conventional accelerator has played an important role.[37]

REFERENCES

[1] Abramowitz, M., "Resource and Output Trends in the U. S. Since 1870," American Economic Review (1956).

[2] Arrow, K. and W. Capron, "Dynamic Shortages and Price Rises: The Engineer-Scientist Case," Quarterly Journal of Economics (1959).

[3] Blank, D. and G. Stigler, The Demand and Supply of Scientific Personnel (National Bureau of Economic Research, 1957).

[4] Brozen, Y., "The Future of Industrial Research," Journal of Business (1961).

[5] Gainsbrugh, M., "Allocation of Resources to Research and Development," Conference on Research and Development (National Science Foundation, 1958).

[6] Greenwald, D., "The Annual McGraw-Hill Research and Development Survey," Methodology of Statistics on Research and Development (National Science Foundation, 1959).

[7] Hitch, C., and R. Mckean, The Economics of Defense in the Nuclear Age (Harvard, 1960).

[8] Horowitz, I., "Regression Models for Company Expenditures on and Returns from Research and Development," IRE Transactions on Engineering Management (1960).

[9] Hurwicz, L., "Least-Squares Bias in Time Series," Statistical Inference in Dynamic Economic Models, ed. by T. Koopmans (Wiley, 1950.)

[10] Jewkes, J., D. Sawers, and R. Stillerman, The Sources of Invention (St. Martin's Press, 1958).

[11] Keezer, D., D. Greenwald, and R. Ulin, "The Outlook for Expenditures on Research and Development During the Next Decade," American Economic Review (1960).

[12] Kuznets, S., "Inventive Activity: Problems of Definition and Measurement," Universities-National Bureau of Economic Research Conference on the Rate and Direction of Economic Activity (in press).

[13] Langenhagen, C., An Evaluation of Research and Development in the Chemical Industry (Unpublished M.S. thesis, M.I.T., 1958).

[14] Machlup, F., "The Supply of Inventors and Inventions," Universities-National Bureau of Economic Research Conference on the Rate and Direction of Inventive Activity (in press).

[15] Mansfield, E., "Comment," Universities-National Bureau of Economic Research Conference on Rate and Direction of Inventive Activity (in press).

[16] ________, "Size of Firm, Market Structure, and Innovation," Econometrica (1960) (abstract).

[17] ________, "Technical Change and the Rate of Imitation," Econometrica (1961).

[18] Massell, B., "Is Investment Really Important?" RAND Corporation P-2088 (1961).

[19] Minasian, J., "The Economics of Research and Development," Universities-National Bureau of Economic Research Conference on the Rate and Direction of Inventive Activity (in press).

[20] Nelson, R., "The Simple Economics of Basic Scientific Research," Journal of Political Economy (1959).

[21] Ninian, A., The Role of Research in the American Steel Industry, (Unpublished M.S. thesis, M.I.T., 1959).

[22] Organization for European Economic Cooperation, The Organization of Applied Research in Europe, the United States, and Canada (OEEC, 1954).

[23] Quinn, J., "Long-Range Planning of Industrial Research," Harvard Business Review (1961).

[24] Rubenstein, A., "Setting Criteria for R and D," Harvard Business Review (1957).

[25] Schmookler, J., "Changes in Industry and in the State of Knowledge as Determinants of Inventive Activity," Universities-National Bureau of Economic Research Conference on the Rate and Direction of Inventive Activity (in press).

[26] Solow, R., "Technical Change and the Aggregate Production Function," Review of Economics and Statistics (1957).

[27] ________, "Investment and Technical Progress," Mathematical Methods in the Social Sciences, ed. by Arrow, Karlin, and Suppes (Stanford, 1960).

[28] Worley, J., "Industrial Research and the New Competition," Journal of Political Economy (1961).

APPENDIX: DATA AND METHODS

Interview data: Interviews were obtained with executives of three petroleum and three chemical firms. These executives were ordinarily the president or research director of the firm. Numerous questions were asked regarding the process by which the firm's expenditures on R and D were determined, but three of these questions are of particular importance as sources of the data in Table 1. First, to determine $\tilde{R}_i(t)$ in 1958, they were asked to estimate how much the firm would have spent in 1958 if it could have acquired instantaneously all of the personnel and equipment that it may have wanted, if it could have avoided whatever inefficiencies that may have resulted from rapid changes in R and D expenditures and if the desired change in such expenditures could have been maintained for a reasonable length of time. In a few cases, it was also possible to obtain such estimates for earlier years in the Fifties and for 1959.

Second, they were asked to estimate the frequency distribution of projects (that would have been carried out in 1958 under those circumstances) by expected profitability. These estimates generally allowed for differences among projects in risk, rough estimates of the probability of success being used by almost all firms in their measures of a project's potential profitability. The actual units in which these estimates were expressed differed from firm to firm, depending on the sorts of evaluation procedures they used.

Equation (3) means (if all projects cost the same) that there is an exponential distribution of projects by expected profitability. Judging by the frequency distributions drawn up by those executives, equation (3) would fit the data for these firms quite well. For this reason and because it is convenient, we make the assumption in equation (3). Note that we assume only that it holds in the relevant range, i.e., for ρ greater than some reasonable minimum. The lower bound of the estimated frequency distribution was used as an estimate of $\rho_i^*(t)$. To estimate $\bar{\rho}_i(t)$, we note that, if equation (3) holds, $\bar{\rho}_i(t) = \bar{\bar{\rho}}_i(t) - \rho_i^*(t)$, where $\bar{\bar{\rho}}_i(t)$ is the average expected profitability of projects that the firm would accept if it could adjust instantaneously and without the costs cited above. That is, if ρ is exponentially distributed, its average, given that it exceeds $\rho_i^*(t)$, equals $\bar{\rho}_i(t)$ plus $\rho_i^*(t)$. Estimates of $\bar{\bar{\rho}}_i(t)$ were derived from the frequency distributions drawn up in the interviews and from correspondence with the firms.

Third, they were asked how much the planned, or budgeted, expenditures on R and D for 1958 differed from the actual expenditures. Using this piece of information, it was possible to estimate $r_i(t)$, which is needed to compute $\theta_i(t)$. In a few cases, it was possible to obtain such estimates for earlier years and for 1959 too.

Finally, in the case of the remaining two petroleum firms in Table 1, answers to these three questions were obtained through correspondence with the research director. These firms were too far from Pittsburgh or New Haven for interviews to be feasible. Letters were sent to other firms among those represented in Tables 2 - 3, but their replies were not usable. See note 15.

<u>Effect of tax change on R and D expenditures:</u> We assume that the firm has a certain set of factor prices in mind and that they are not altered by the tax change. In other words, we assume that the firm is confronted by a perfectly elastic supply of engineers, scientists, etc. Otherwise, depending on the elasticity of supply, the result might be a mixture of increased "real" R and D and higher wages for engineers, scientists, etc. Moreover, because of the effect of the wage changes on $\overline{\rho}_i(t)$, the firm's expenditures would be affected.

Under these circumstances, it is easy to show that the firm's spending on R and D would increase by $\rho_i^*(t)/\overline{\rho}_i(t)$ percent if the expected profitability of all projects increased by one percent. If the profitability of all projects increased by a certain proportion, $M_i(t)$ would be unaffected but $\overline{\rho}_i(t)$ would increase by this proportion. From equation (4), it follows that

$$\frac{d\,\tilde{R}_i(t)\cdot\overline{\rho}_i(t)}{d\,\overline{\rho}_i(t)\cdot\tilde{R}_i(t)} = \rho_i^*(t)/\overline{\rho}_i(t)\,.$$

Of course, this result (as well as the fourth listed in Section 4) depends on the assumption that ρ is exponentially distributed.

For the industry as a whole, it is generally more convenient to work with the conventional supply and demand functions, if the supply of factors is not infinitely elastic. If one can express $\overline{\rho}_i(t)$ and $M_i(t)$ as functions of an index of the price during year t of a "unit" of R and D and if one is willing to allow sufficient time so that $R_i(t) \doteq \tilde{R}_i(t)$, it is easy to derive a demand function for "real" R and D. E.g., if P_t is such an index and if $M_i(t)$ and

$\overline{\rho}_i(t)^{-1}$ are proportional to $\overline{P}_t$, $Q_{it} = \phi'_{it} e^{-\rho^*_i(t)\overline{P}_t/\phi''_{it}}$, where Q_{it} is the "real" R and D the firm will purchase during year t, ϕ'_{it} is the value of $M_i(t)$ when $\overline{P}_t = 1$ and ϕ''_{it} is the value of $\overline{\rho}_i(t)$ when $\overline{P}_t = 1$. Coupled with an estimate of the supply function for "real" R and D, one could use such demand functions to estimate the effects of a tax change on the entire industry's "real" R and D effort.

Effect of interfirm variation in $\overline{\rho}_i(t)$ on R and D expenditures: For simplicity, assume that $M_i(t)$ is the same for all firms and that there are n firms in the industry. If U_1 is the average value of $\overline{\rho}_i(t)$, U_2 is the coefficient of variation about this average, and U_2 is small, the total desired expenditures in the industry will be

$$\sum_{i=1}^{n} \tilde{R}_i(t) \doteq n\, M_{it}\, e^{-\rho^*_i(t)/U_1} \left[1 + \frac{\rho^*_i(t)}{2\,U_1} \left(\frac{\rho^*_i(t)}{U_1} - 2 \right) U_2 \right]$$

Thus, increases in U_2 will increase $\sum_{i=1}^{n} \tilde{R}_i(t)$ if $U_1 < \rho^*_i(t)/2$ and decrease it otherwise.

Estimates of β_{i1}, β_{i2}, β_{i3}, and $\tilde{R}_i(t)/S_i(t)$: The data on $R_i(t)$ and $R_i(t-1)$ came mainly from correspondence with the firms and from [13]. For a few firms, they came from [8]. The National Science Foundation's definition of research and development is used throughout the paper. Nonetheless, it is possible that the data are not entirely comparable over time and among firms because, despite the instructions given the firms, different definitions were used. The data on $S_i(t)$ came from Moody's. For all firms, $R_i(t)$, $R_i(t-1)$, and $S_i(t)$ are measured in units of millions of dollars, and t is measured in years from 1945. In equation (13), the intercept is constrained to be zero, since we assume that the expected value of $z_i''(t)$ is zero.

Although the least-squares estimates of β_{i1}, β_{i2}, and β_{i3} in Tables 2 - 3 are consistent, they are not unbiased because $R_i(t-1)$, a lagged endogenous variable, is used as an exogenous variable. See Hurwicz [9]. There seems to be no simple way to eliminate this bias, but fortunately it should be fairly small. According to Hurwicz's results, it would be about 10 percent. Note however that his model includes no independent variables other than the lagged endogenous variable. Note too that for a few firms we could not get data for the entire period (1945-58) and that we had to omit the first few years for this reason.

For each industry, the means in Table 4 equal $\bar{\alpha}_1 + \bar{\alpha}_2 t$, where $\bar{\alpha}_1$ is the mean value of $\hat{\alpha}_{i1}$ and $\bar{\alpha}_2$ is the mean value of $\hat{\alpha}_{i2}$. For obvious reasons, we must omit firms where $\left(1 - \hat{\beta}_{i1} + \hat{\nu}_4\right) < 0$. For each industry, the variance of $\tilde{R}_i(t)/S_i(t)$ equals $\sigma_1^2 + 2\sigma_{12}t + \sigma_2^2 t^2$, where σ_1^2 is the variance of $\hat{\alpha}_{i1}$, σ_2^2 is the variance of $\hat{\alpha}_{i2}$, and σ_{12} is the covariance between $\hat{\alpha}_{i1}$ and $\hat{\alpha}_{i2}$. From this and the mean value of $\tilde{R}_i(t)/S_i(t)$, one can easily derive the coefficients of variation.

Besides the bias in $\hat{\beta}_{i1}$, $\hat{\beta}_{i2}$, and $\hat{\beta}_{i3}$, there are other biases in Table 4. Since the expectation of the ratio of random variables does not equal the ratio of their expectations, the transformation from these estimates to $\hat{\alpha}_{i1}$ and $\hat{\alpha}_{i2}$ results in bias. Moreover, because $U_i(t)$ is omitted, the means in Table 4 would not necessarily equal the means of $\tilde{R}_i(t)/S_i(t)$, even if there were no errors in the estimates of α_{i1} and α_{i2}. But they would be unbiased estimates. Similarly, the coefficients of variation of $\tilde{R}_i(t)/S_i(t)$ are underestimated because we ignore the variance of $U_i(t)$. But if this variance was approximately constant over time, the conclusion that the coefficients of variation declined still holds.

Despite these problems, the general conclusions in Section 5 are almost certainly correct; and although the particular numbers in Table 4 must be viewed with caution, there is evidence that they are reasonably accurate. When the direct estimates of $\tilde{R}_i(t)/S_i(t)$ in Table 1 are compared with those based on Tables 2 - 3, the results are quite close.

Finally, note that the objection has been raised that the model in equation (13) may not reveal adequately the effect on a firm's R and D expenditures of a decrease in sales. Some have said that under such circumstances R and D expenditures will fall less -- or rise more -- than the model would predict. An analysis of the residuals from these regressions shows that this objection is incorrect. There is no tendency for the residuals to be positive in years when a firm's sales had dropped below the previous year's level.

Alternative ways of estimating $\tilde{R}_i(t)/S_i(t)$: First, using equations (2) and (7), it follows that

$$\tilde{R}_i(t) = R_i(t-1) + [R_i(t) - (1 + \nu_4)R_i(t-1) - z_i(t)]/[\nu_3 + \nu_5\Pi_i(t-1) + k_i'(t)]$$

We have data on $R_i(t)$, $R_i(t-1)$ and $\Pi_i(t-1)$, and we have estimates of ν_3 , ν_4 , and ν_5 from equation (10) . Thus, if we ignore $z_i(t)$ and $u_i(t)$, we can estimate $\tilde{R}_i(t)$ from this equation. A major advantage of this approach is that we need not assume that equation (12) holds. A major disadvantage is that we must assume that our estimates of ν_3, ν_4, and ν_5 in equation (10) are good enough for these purposes.

Second, using equations (2), (7), and (12), it follows that

$$R_i(t) = \nu_3\, \alpha_{11}\, S_i(t) + (1 + \nu_4 - \nu_3)\, R_i(t-1) + \nu_3\, \alpha_{12} t S_i(t) + \nu_5 \alpha_{11} \Pi_i(t-1)\, S_i(t)$$
$$+ \nu_5\, \alpha_{12}\, \Pi_i(t-1)\, t\, S_i(t) - \nu_5\, \Pi_i(t-1)\, R_i(t-1) + z_i'''(t) ,$$

where $z_i'''(t)$ is a random error term that equals $\left\{[\alpha_{11} + \alpha_{12}t + u_i(t)]\, S_i(t) - R_i(t-1)\right\} k_i'(t) + \left\{\nu_3 + \nu_5\, \Pi_i(t-1)\right\} S_i(t)\, u_i(t) + z_i(t)$. Least-squares estimates of the coefficients in this equation could be used (in much the same way as the estimates in Tables 2 - 3 are used) to estimate α_{11} and α_{12} . An advantage of this approach is that $z_i'''(t)$ may behave more nearly like a random error term than $z_i''(t)$. The primary disadvantage is that there are relatively few degrees of freedom. (But this difficulty could be avoided by constraining the values of the ν's to be equal for all firms.)

Still other methods could have been used. E.g., if we had more data on $\rho_i^*(t)/\bar{p}_i(t)$, equation (9) could be used for this purpose. Although these techniques would result in somewhat different estimates from those in Table 4, I doubt that very substantial differences would occur. As an experiment, I used the first method described above to estimate $\tilde{R}_i(t)$ in the cases where the actual value of $\tilde{R}_i(t)$ was known. The results are quite similar to those in the text (and no more accurate, one difficulty being that the second term on the right in the first equation above sometimes is negative).

Forecasts of R and D expenditures: The root-mean-square error of the forecasts based on equation (13) is $1.16 million. Using the naive forecast that expenditures in 1959 will equal those in 1958, the root-mean-square error is $2.10 million. Using the naive forecast that expenditures in 1959 will differ from those in 1958 by the same amount that expenditures in 1958 differed from those in 1957, the root-mean-square error is $1.57 million. And using the naive forecast that expenditures in 1959 will differ by the same percent from those in 1958 as 1958 expenditures differed from those in 1957, the root-mean-square error is $1.67 million.

Our results seem to be better than forecasts based on businessmen's expectations. See Greenwald [6]. Note that the root-mean-square error of the forecasts based on equation (10) is inflated considerably by one firm -- Minnesota Mining and Manufacturing. If it is excluded, the root-mean-square error is less than $1 million.

To obtain the forecasts where sales estimates contained errors, we distributed at random errors of plus or minus 10 percent of sales. A coin was tossed to determine whether or not the 10 percent error would be plus or minus.

The estimates of r in equation (15) were obtained as follows. The figure for petroleum is based on forecasts made by the president of Shell Oil in 1960 Wall Street Journal, (May 18, 1960) and by Chemical Processing in 1960. The figure for chemicals is an average of forecasts from Chemical Week (December 24, 1960) and Chemical Processing. The figure for drugs is from the president of Eli Lilly in the Wall Street Journal (May 18, 1960). The roughness of these forecasts of $S_i(t)$ need not be labored.

Finally, the increase in the ratio of R and D expenditures to sales in Figure 3, even if $\tilde{R}_i(t)/S_i(t)$ remains constant at its 1960 level, is due to the fact that firms are currently below their desired expenditures for R and D. If R_i is the i^{th} firms 1959 expenditures on R and D and ϕ_i is the 1960 value of $\tilde{R}_i(t)/S_i(t)$, its expenditures at time t' under the assumptions above would equal

$$R_i(t') = (1 - \theta_i)^{t'} R_i + \theta_i \, \phi_i \, S_i(1 + r)^{t'} \sum_{i=0}^{t'-1} \left(\frac{1 - \theta_i}{1 + r}\right)^i \quad ,$$

if $\theta_i(t)$ were constant. As time goes on, $R_i(t')/S_i(t')$ would not tend to ϕ_i, but instead to

$$\left(\frac{1 + r}{1 + r/\theta_i}\right) \phi_i \quad .$$

Data on inventive output: For the chemical firms, Langenhagen [13] provides data on the number of significant inventions and the R and D expenditures. The data on n_i come from his Table 2, Chapter 4. The inventions are weighted by the number of times they were included on the questionnaires he describes there.

For the petroleum firms, we looked up who was credited by various trade and technical sources with having invented each of the inventions on Schmookler's list [25]. Then using biographical directories and other sources, we found out the firm that employed him when the invention was made. For most of the inventions during this period, we could find this information. Note that the data for petro-chemicals pertain to innovation, not invention. The data on $R_i(t)$ are described above.

For the steel firms, the data on n_i are described in Mansfield [16]. Note that they pertain to innovations, not inventions. No data on the latter were available. Some of the data on R_i come from Ninian [21]; the rest are described above.

Estimates of the parameters in an accelerator model with variable response coefficients: According to the conventional accelerator model, $\chi(t) = \chi(t-1) + \theta_1[\theta_2 y(t) - \chi(t-1)] + z(t)$, where $z(t)$ is an error term and $\chi(t)$ is the endogenous variable which adjusts toward a "desired" or "equilibrium" level, the latter being proportional to an exogenous variable, $y(t)$. Suppose one assumes that

$\theta_1 = \phi_1 + \phi_2 \chi(t-1)/[\theta_2 y(t) - \chi(t-1)]$ when $\theta_2 y(t) > \chi(t-1)$ and that $\theta_1 = \phi_1 - \phi_2 \chi(t-1)/[\theta_2 y(t) - \chi(t-1)]$ when $\theta_2 y(t) < \chi(t-1)$. Then if one can distinguish whether $\theta_2 y(t)$ is greater or less than $\chi(t-1)$ and if one assumes that the data always lie in the range where this expression for θ_1 holds, it is not difficult to obtain consistent estimates of ϕ_1, ϕ_2 and θ_1. Since

$$\chi(t) = (1 - \phi_1 - \phi_2)\chi(t-1) + 2\phi_2 W(t) + \phi_1 \theta_2 y(t) + z(t) ,$$

where $W(t)$ is zero if $\theta_2 y(t) < \chi(t-1)$ and $\chi(t-1)$ otherwise, conventional methods can be used.

FOOTNOTES

* A preliminary version of this paper was read at the December 1961 meetings of the Econometric Society. The work on which it is based is part of a larger project on industrial research and technical change supported by a contract with the Office of Special Studies of the National Science Foundation and by a Ford Foundation Faculty Research Fellowship. The work was completed while I was Visiting Associate Professor at the Cowles Foundation for Research in Economics at Yale University, and the paper has benefitted from discussions with my colleagues there, W. Brainard, A. Heston, J. Hooper, and E. Phelps. My thanks also go to J. Zoler for his assistance, to the thirty-five companies for data and interviews, and to Professors Horowitz, Ninian, Rubenstein and Schmookler for making unpublished data available to me.

1. Two studies by Solow [26, 27] and one by Abramowitz [1] have been particularly influential in this regard. For a list of other studies reaching similar conclusions, see Massell [18]. Of course, these studies can only be suggestive, the estimate of the effect of technical change being in reality a catch-all residual.

2. I presented a preliminary version of this model at the 1960 Conference on the Rate and Direction of Inventive Activity. Although I pointed out that $\tilde{R}_i(t)/S_i(t)$ and $\Theta_i(t)$ might vary over time, I assumed for simplicity that they were constant in the preliminary tests regarding chemical firms reported there. See Mansfield [15]. For an application of these results, see Brozen [4].

   Horowitz [8] analyzed firms' expenditure on R and D, but unfortunately his results were almost entirely negative. In another paper that is of relevance here, Minasian [19] inspected the relationship between a firm's expenditure on R and D and its profitability.

3. There are two primary reasons why the firm moves only part way toward $\tilde{R}_i(t)$. First, there are often substantial costs involved in expanding $R_i(t)$ too rapidly. It is difficult to assimilate large percentage increases in R and D staff. Second, it takes time to hire people, build laboratories, etc. (Of course this assumes that $\tilde{R}_i(t) > R_i(t)$, but this seems to have been typically the case during the relevant period. The costs of changing $R_i(t)$ are not included in determining $\tilde{R}_i(t)$.

Besides these two reasons, a firm may move only part way toward $\tilde{R}_i(t)$ because it is uncertain as to whether or not expenditures of $\tilde{R}_i(t)$ can be maintained for a sufficiently long period of time so that projects that are started can be carried out without interruption. A firm does not want to start a project that it will soon have to interrupt.

There is considerable evidence that $\theta_i(t)$ is generally less than one. E.g., the interviews in the National Science Foundation's Science and Engineering in American Industry, 1953-54, show that firms place great value on the stability of their R and D expenditures. More direct evidence to this effect is presented in the following section.

4. This assumes that, on the average, the budgeted and actual expenditures are equal. Judging from the interview data described in the Appendix, this seems to be a reasonable approximation. But if for some reason the expected value of $z_i(t)$ is non-zero, the only effect it has on the model is that equation (2) will contain an intercept. And so will equation (11), etc.

5. For some analyses of military research and development, see Hitch and Mckean [7].

6. Of course, the rate of return on the investment in a particular R and D project is notoriously difficult to measure (Cf. [24]), and we do not assume that the firm can do so with any real accuracy. What we assume is that the firm makes some rough estimates of a project's prospective profitability and that it bases its decision on them. This assumption seems to be borne out by [5, 23], and by the interview data in the Appendix. Presumably crude allowances for differences in risk are made in these estimates.

Note that the model works just as well if some other index of a project's attractiveness (rather than the prospective rate of return) is used by the firm. The interviews reported in the National Science Foundation's Science and Engineering in American Industry, 1953-54 indicate that almost all of the firms in the survey made some attempt to evaluate the returns from R and D projects, although the methods were sometimes informal. Only about one-quarter of these firms reported that they established a certain minimum rate of return as a cut-off point, but it seems likely that a cut-off point of some sort was used.

The distribution of projects by prospective rate of return in equation (3) was suggested by the interview data described in the Appendix. Although equation (3) is convenient, one could use a more general function of ρ instead (e.g., a quadratic function), with the consequence that $\tilde{R}_i(t)$ would depend in part on higher moments than the mean of ρ. If data can be obtained regarding these moments, the model can easily be generalized in this way.

Finally, note two additional points. (1) We lump research expenditures and development expenditures together throughout the paper -- although it would be preferable to separate them in cases where the distinction is a meaningful one. (2) We treat R and D expenditures as investments, although they are treated as current expenses for tax purposes. So far as the analysis is concerned, this obviously makes no real difference.

7. One would expect larger firms to believe that they could spend larger amounts on R and D with positive returns than smaller firms. Many research results would be worth much more to a large firm than to a small one because the saving per unit of production would be spread over a larger volume, because the large firm would have the necessary capital to exploit the new product or process, etc. Moreover, many projects would be cheaper for a large firm than a small one and the risks would often be less for a large firm. Cf. Nelson [6]. In passing, note that the model does not assume that the firm is necessarily a profit maximizer (the meaning of which is unclear in any event in a world containing uncertainty).

8. Certainly, it seems likely that $\bar{\rho}_i(t)$ -- the average <u>expected</u> profitability of prospective R and D projects with <u>expected</u> positive returns -- is directly related to the average <u>actual</u> profitability of prospective R and D projects with <u>actual</u> positive returns. Moreover, given this average actual profitability, one would expect $\bar{\rho}_i(t)$ to be directly related to the amount the firm's competitors are spending on R and D . In the face of considerable uncertainty regarding the actual profitability of R and D , the fact that other comparable firms spend a great deal more on R and D often results in an upward revision in a firm's estimate of the profitability of R and D . (Of course, there are also large potential losses in being too far out of step with regard to the level of R and D expenditures.) For some evidence that this process was going on during 1945-58, see Section 5. For some additional discussion, see [11].

9. According to the interviews described in the Appendix, there are considerable inefficiencies and problems caused by too large a percentage increase from one year to the next in a firm's R and D expenditures. Because of the difficulties in obtaining and assimilating a large increase in staff and equipment, some firms were reported to avoid increases exceeding 10 - 15 percent. Practically all the interviews suggested that $\theta_i(t)$ would be inversely related to $[\tilde{R}_i(t) - R_i(t-1)]/R_i(t-1)$.

10. Let $P_i(t-1)$ be the ratio of the i^{th} firm's profits to its R and D expenditures in year $(t-1)$. Since the percent of the i^{th} firm's profits in year $(t-1)$ that would be absorbed by an increase in R and D expenditures to the desired level is the ratio of $[\tilde{R}_i(t) - R_i(t-1)/R_i(t-1)$ to $P_i(t-1)$, it follows that this percent is inversely related to $P_i(t-1)$ when $[\tilde{R}_i(t) - R_i(t-1)]/R_i(t-1)$ is held constant. If a firm spent a relatively large proportion of its profits on R and D in year $(t-1)$ it seems likely to move more slowly in its plans for year t because of liquidity constraints and because it may believe it is getting out of line with its competitors. The importance of $P_i(t)$ was stressed in practically all of the interviews.

In equation (10), we use as an independent variable the ratio of $P_i(t-1)$ to the average value of $P_i(t-1)$ in the industry. Since a "large" and "small" value of $P_i(t-1)$ vary from industry to industry, this normalization seems reasonable. But the value of $P_i(t-1)$ relative to the past may also be important, although its principal effect is likely to be on $\rho_i^*(t)$ rather than $\theta_i(t)$. In this connection, it should be noted that changes in earnings are very likely to affect $\rho_i^*(t)$, the impact being particularly important for smaller firms.

11. The form of equation (7) was chosen primarily for its convenient properties and its simplicity. The following section indicates that it fits the available data reasonably well. Note, however, that because $0 \leq \Theta_i(t) \leq 1$, equation (7) can hold only for values of $R_i(t-1)/[\tilde{R}_i(t) - R_i(t)]$ and $\Pi_i(t-1)$ in a certain range. Beyond that range, the values of $\Theta_i(t)$ will violate these inequalities. We assume in the following section and in Sections 5 - 6 that firms were always in this range during the period. This probably is reasonable in the next section but it is more questionable in Sections 5 - 6. See note 31.

Of course, if the labor market is tight, it will probably take longer for firms to hire extra personnel at the wages they want to pay. Thus, this factor (which is extremely difficult to measure) is probably an important determinant of $\Theta_i(t)$. For some discussion of the market for engineers and scientists, see Blank and Stigler [3] and Arrow and Capron [2].

12. To be of use in explaining differences over time in firms' expenditures on R and D, the model would have to be expanded to explain differences over time in ν_{1t}, and ν_{2t}. Over fairly short periods, one could probably assume that ν_{2t} remain unchanged and that ν_{1t} is a function of the amount of basic research conducted before year t in this industry, the level of wage rates for scientists and engineers, etc.

An additional problem in comparing R and D expenditures over time, rather than making interfirm comparisons, is that for most purposes one is interested in changes in "real" R and D, not money magnitudes. In interfirm comparisons, this problem does not arise because the factor prices paid by all firms are about the same. (This problem should be borne in mind in Sections 5 - 6 particularly.)

Finally, note that all of the exogenous variables other than $S_i(t)$ can be measured at time t-1. If it seems desirable for all such variables to be measurable at time t-1, the firm's forecast of sales in year t or its actual sales in year t-1 probably can be used instead of $S_i(t)$.

13. For the sources of the data in Table 1, see the Appendix.

14. Biases in the estimates of $\tilde{R}_i(t)/S_i(t)$, if they are about the same proportion of $\tilde{R}_i(t)/S_i(t)$ and in the same direction in each case, will make little difference in the estimate of V_1. They will result in a corresponding bias in V_o, but V_1 should still equal one. However, errors in measuring $\rho_i^*(t)/\bar{\rho}_i(t)$ would be likely to bias the estimate of V_1 toward zero. Note too that $k_i''(t)$ is assumed to be independent of $k_i'(t)$.

15. One possible bias here should be noted. Only about half of the firms that we contacted could give us estimates of $\bar{\rho}_i(t)$ and $\rho_i^*(t)$ (see the Appendix), and the fact that certain firms could give us such data may indicate that they are more likely to conform to the model. If this is the case, there would obviously be much more unexplained variation if all firms were included.

In addition, it is always possible that a firm might overstate the figure regarding its expectations of the profitability of R and D to rationalize large R and D expenditures carried out for other reasons. But this seems rather far-fetched because the firms were not told what their estimate of $\bar{\rho}_i(t)$ and $\rho_i^*(t)$ would be used for. Moreover, it seems unlikely that they would go to this much trouble to deceive, when they could more easily claim that they could not answer.

Before omitting $\ln S_i(t)$ from the regression in equation (9), the results were

$$\ln[\tilde{R}_i(t)/S_i(t)] = \begin{Bmatrix} -2.95 \\ -0.80 \end{Bmatrix} - \underset{(.41)}{.86}\ \rho_i^*(t)/\bar{\rho}_i(t) - \underset{(.09)}{.11}\ \ln S_i(t) .$$

Thus, the available evidence suggests that ν_{2t} was less than one in these industries in 1958.

Judging from Figure 2, the variance of $k_i''(t)$ may be directly related to $\rho_i^*(t)/\bar{\rho}_i(t)$. This may be due to differences between the industries in the variance of $\ln k_i(t)$.

16. Note four things. First, the sources of the data used in the "predictions" were the interviews described in the Appendix. Unfortunately, since corresponding values of $\rho_i^*(t)/\bar{\rho}_i(t)$ could not be obtained, they could not be used in connection with equation (9). They generally pertain to 1952, 1955, or 1959.

Second, one reason why ν_3 does not differ significantly among industries may be that the labor market was about as tight in one industry as in the other. Third, there may be some tendency for the expected value of $k_i'(t)$ to vary, depending on i and t. With so few observations, this is impossible to check. Fourth, somewhat better estimates might have been obtained in equation (10) if the estimate of $\tilde{R}_i(t)$ from equation (9) had been used rather than the actual values of $\tilde{R}_i(t)$. But the differences would undoubtedly have been slight.

17. For the assumptions involved here and for more general results, see the Appendix.

18. There has been a tendency in some quarters to emphasize the importance of non-economic motives and "fashion," rather than expectation of profit as determinants of the firm's spending on R and D . See [10], p. 151 and [22], p. 29. Such factors do operate but they may not be so important as has been claimed.

Of course, in the face of considerable uncertainty, a firm's expectations of the profitability of R and D may be formulated in a way that is in some sense "irrational." For example, if it is spending much less than its competitors on R and D , it may adjust its expectations upward, even though its original expectations may have been more nearly correct. But this is quite different from saying that the firm sets its R and D budget with little or no regard to the profitability of the expenditures.

19. Note four things. First, these results regarding v_{2t} pertain only to the larger firms in these two industries. Second, since the estimate of v_{2t} in note 15 is subject to rather large error, one cannot be sure that v_{2t} is less than or equal to one. The probability is about .85 that this is the case (in the sense that an 85 percent confidence interval for v_{2t} with no lower bound has one as its upper bound). Third, for the expectation in the text to follow, $\tilde{R}_i(t)/R_i(t)$ must be independent of a firm's size in this range. Fourth, using 1957 or 1958 data for about a dozen of the largest chemical, petroleum, and steel firms, we regressed $\ln R_i(t)$ on $\ln S_i(t)$. The slopes of these regressions were .998 (chemicals), .870 (petroleum), and .932 (steel). Since all are less than one, they indicate that in this range and in this year increases in $S_i(t)$ were not accompanied by more than proportional increases in $R_i(t)$. Of course, sampling errors are present, but, if the largest fifteen or so firms in each industry are regarded as the relevant universe, these errors must be quite small because practically all of the firms are included. These results differ somewhat from

Worley's [28], but his findings are based on 1955 employment data and his cut off point with regard to firm size is smaller than ours.

20. For the assumptions involved here, see the Appendix.

21. Certainly, most observers seem to believe that $\tilde{R}_i(t)/S_i(t)$ increased for most firms in these industries during the postwar period, and the interviews indicated this as well. Moreover, with reasonable values of $\theta_i(t)$, it is hard to believe that the large postwar increases in $R_i(t)/S_i(t)$ could have occurred without such increases in $\tilde{R}_i(t)/S_i(t)$. See Figure 3 for the past values of $R_i(t)/S_i(t)$ in these industries.

Of course, if both equation (6) and equation (12) hold, it follows that $\bar{\rho}_i(t)/\rho_i^*(t) = \left\{\ell n\left[\nu_{1t}S_i(t)^{\nu_2}t^{-1}k_i(t)/\left(\alpha_{i1}+\alpha_{i2}t+U_i(t)\right)\right]\right\}^{-1}$. Whether or not this is reasonable is by no means obvious. But if it holds, $\bar{\rho}_i(t)/\rho_i^*(t)$ is likely to have increased at a decreasing rate, which seems reasonable. Of course, the 1954 changes in the tax treatment of R and D expenditures undoubtedly increased their profitability.

22. Inserting the estimates of ν_3, ν_4, and ν_5 in equation (10) in the expressions for β_{i1}, β_{i2}, and β_{i3} and noting that $\bar{\Pi}_i$ should be fairly close to one and α_{i1} and α_{i2} should be small, the inequalities in (14) seem very likely to hold.

23. To see whether a linear function was satisfactory, we tested whether the coefficient of $t^2 S_i(t)$ -- when added to equation (13) -- was significantly different from zero for five randomly selected firms. If $\tilde{R}_i(t)/S_i(t)$ were a quadratic (rather than a linear) function of t, this term would have a non-zero coefficient. The results are quite consistent with equation (12). It turns out that its coefficient is never statistically significant.

Of course, the relatively small number of years cuts down on the power of the tests. Moreover, when these additional variables were added, the coefficients of $tS_i(t)$ and $S_i(t)$ sometimes became non-significant too -- although they were closer to being significant than the additional variables. The data used here are described in the Appendix.

24. Note two points. First, the correlation coefficients in Tables 2 - 3 use $R_i(t)$ as the independent variable. If $R_i(t)-R_i(t-1)$ had been used instead, the coefficients would probably have been lower. Second, we could have tested whether equation (13) has an intercept, but the tests would have been quite weak.

25. For a discussion of the biases in these estimates, see the Appendix. Of course, this is only one of many ways in which $\tilde{R}_i(t)/S_i(t)$ might have been estimated. See the Appendix for a discussion of some alternative ways.

26. In comparing values of $\tilde{R}_i(t)/S_i(t)$ at various points in time, note that they are unadjusted for differences in the purchasing power of the dollar. Unfortunately, no appropriate deflators exist.

27. Of course, one would expect that the coefficient of variation would eventually become stabilized. Probably this decrease in the coefficient of variation is largely a postwar phenomenon. In each industry, there seems to have been some tendency for the coefficient to become more stable.

28. As a further test, we looked at individual firms in the chemical industry to see whether $\hat{\beta}_{i1} = 2.13 - 1.53\,\overline{\Pi}_i$, which would be the case if the model held and if the estimates in equation (10) were correct. For those firms where the values of $\overline{\Pi}_i$ were within the relevant range (see note 31), there seemd to be a considerable amount of agreement between $\hat{\beta}_{i1}$ and $(2.13 - 1.53\,\overline{\Pi}_i)$.

Note that the average value of $\hat{\beta}_{i1}$ in glass and steel is not very close to the theoretical value of .60. This may be due to the fact that the sample is so small that the average value of $\overline{\Pi}_i$ is not close to one or because the ν's are different in these industries due to differences in the relevant labor markets, etc.

29. The accuracy of the forecasts based on equation (13) and those based on various naive models is described in the Appendix.

30. The sources of these estimates of r are given in the Appendix.

31. Note that the results in this section and in the previous one assume that firms continually operated and will operate in the range where equation (7) holds. On the basis of the estimates in equation (10), the value of $\Pi_i(t-1)$ observed during 1945-58 in the chemical industry, and what little we can guess about $\left(\tilde{R}_i(t) - R_i(t-1)\right)/R_i(t-1)$, it seems likely that this was the case for most firms.

But for the remaining minority, it is possible that this equation did not hold and that $\theta_i(t)$ was usually one. For such firms, the estimates of $\tilde{R}_i(t)/S_i(t)$ will be underestimated slightly but the results in this section will not be affected so long as $\theta_i(t)$ continues to be one.

32. For some discussion of economies of scale in research, see the references in [16].

33. The data underlying equations (16) - (18) are described in the Appendix. A quadratic function is used because of its simplicity and because alternative functions often do not permit zero values of n_i to be used. R_i is measured in units of millions of dollars.

34. For some relevant discussion in this connection, see [10], Chapters 6-7. As measure of a firm's size we used its sales in 1940 (chemicals) or 1946 (petroleum and steel). Note that, if one regards the relevant universe here to be the largest fifteen or twenty firms in each industry, the standard errors of the regression coefficients are over-stated, and seemingly non-significant coefficients are in fact significant.

35. Note that most of the firms that are included here spent a reasonably large amount on R and D . The minimum expenditure was $270,000 (chemicals), $600,000 (petroleum), and $50,000 (steel). There may be considerable "economies of scale" in the lower ranges of spending. (These figures are averages for the years indicated in the test.)

36. For some discussion of the problems involved in measuring inventive input and output, see Kuznets [12] and Machlup [14].

37. For ways in which the parameters of such a model can be estimated, see the Appendix. Of course, one very interesting point is that an equation of the same form will result from this model, if the endogenous variable is constantly increasing or constantly decreasing, as from the conventional model with a constant response coefficient. Thus, the fact that the conventional equation fits well does not necessarily mean that the conventional model is correct. Of course, if this model rather than the conventional one is correct, the estimates of the desired or equilibrium level of the endogenous variable and the response coefficient will be biased. The direction of the bias can often be estimated.