COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 134

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

The Speed of Response of Firms to New Techniques

Edwin Mansfield

February 13, 1962
The Speed of Response of Firms to New Techniques*

by

Edwin Mansfield

Yale University and Carnegie Institute of Technology

1. Introduction

Economists have long been interested in the process whereby new techniques spread from firm to firm; but until recently our knowledge of its workings did not extend far beyond Schumpeter's simple assertion that once a firm introduces a successful innovation, a host of imitators appear on the scene. In the past few years, considerable progress has been made in investigating some important aspects of this imitation process. It is unfortunate that other important aspects are still largely unexplored.

Specifically, although it has often been observed that some firms begin using a new technique long before other firms do, there have been very few studies of the factors responsible for these differences. Can we construct models to help predict whether one firm will be quicker than another? How accurate will such predictions be? In addition, we lack information regarding the extent to which technical leadership of this sort is concentrated in the hands of a few firms. Do the same members of an industry tend to be relatively quick to introduce various new techniques, or are the leaders in one case likely to be the followers in another?
The purpose of this paper is to help answer these questions. First, hypotheses are presented regarding the effects of various factors on the length of time a firm waits before using a particular technique. Second, using data describing the diffusion of fourteen major innovations, we estimate the probability that each hypothesis will be correct and the probability that a prediction based on each hypothesis will hold. Third, using the same data, we estimate the extent to which leadership of this sort in four important industries has been concentrated in the hands of a relatively few firms.

The plan of the paper is as follows. Sections 2 - 5 take up the effects of a firm's size and the profitability of its investment in the innovation on its speed of response. Sections 6 - 7 deal with the effects of a firm's growth rate, past profits, liquidity position, and other such factors. Section 8 measures the extent to which leadership of this kind was concentrated among a relatively few firms in the bituminous coal, railroad, brewing, and iron and steel industries. Section 9 summarizes the results and provides some concluding remarks.

2. Size of Firm and the Profitability of its Investment in the Innovation

This section presents three propositions regarding the effects of a firm's size and the extent of the returns it can obtain from the innovation on how long it waits before introducing the innovation. If they hold for a particular innovation, these propositions should provide a basis for predicting whether one firm will be quicker than another to use the innovation. Section 3
converts these general propositions into testable hypotheses, and Sections 4 - 5 estimate the probability that these hypotheses will be correct and that predictions based on them will hold for a randomly chosen innovation.²

The first proposition states that, other things held equal, the length of time a firm waits before using a new technique tends to be inversely related to its size.³ There are at least three reasons for believing that this is so. First, the costs and risks involved in being among the first to use a new technique are likely to loom much larger for small firms than for big ones. Because of their larger financial resources, bigger engineering departments, better facilities for experimentation, and closer ties with equipment manufacturers, bigger firms can play the role of the pioneer more cheaply and with less risk than smaller ones can.

Second, large firms, because they encompass a wider range of operating conditions, have a better chance of containing those conditions for which the innovation is applicable at first. This is important because, when an innovation first appears, its application is often restricted to certain operating conditions, and improvements occur later that extend its usefulness. Third, because they have more units of any particular type of equipment, large firms are more likely at any point in time to have some units that will soon have to be replaced. Thus, if an innovation occurs that is designed to replace this type of equipment, they probably can begin using it more quickly than smaller firms.⁴

The second proposition elaborates on the first. It states that, as a firm's size increases, the length of time it waits tends to decrease at an
increasing rate. This proposition stems from the following model. Suppose that a new type of equipment is put on the market and that the \( j \)\(^{th} \) firm will eventually own \( \alpha_j \) units of this equipment. Suppose that \( x_{ij} \), the length of time that elapses (from the date when the innovation is first put on the market) before the \( j \)\(^{th} \) firm's \( i \)\(^{th} \) unit \((i = 1, \ldots, \alpha_j)\) is installed is a random variable with cumulative distribution function, \( F(x) \), and that the time elapsing before one of its units is installed is independent of that for another unit.\(^5\)

Under these highly simplified circumstances, the expected length of time a firm with an eventual complement of \( \alpha \) units will wait before beginning to use the innovation is

\[
E(\alpha) = \alpha \int_0^M x [1 - F(x)]^{\alpha-1} F'(x) \, dx ,
\]

where \( M \) is the maximum value of \( x_{ij} \). Integrating by parts, we have

\[
E(\alpha) = \int_0^M [1 - F(x)]^\alpha \, dx ,
\]

and it is easy to show that

\[
E(\alpha) - E(\alpha + 1) = \int_0^M [1 - F(x)]^\alpha F(x) \, dx > 0 ,
\]

\[
[E(\alpha) - E(\alpha+1)] - [E(\alpha+1) - E(\alpha+2)] = \int_0^M [1 - F(x)]^\alpha F^2(x) \, dx > 0 .
\]
Thus, if $a_j$ is proportional to the $j^{th}$ firm's size (measured in terms of sales or production), the expected length of time a firm waits should decrease at an increasing rate with increases in its size. 6

The third proposition states that, other factors held equal, the length of time a firm waits will tend to be inversely related to the extent of the returns it obtains from the innovation. If these returns are very high, the expected returns are likely to be high enough to make the gamble involved in introducing the innovation seem worthwhile at the outset. If they are not so high, the firm will wait until the risks are reduced to the point where the investment seems warranted. (Because surrogates must be used below to represent the profitability of a firm's investment in the innovation, there is little use in speculating further regarding the shape of the relationship between the profitability of a firm's investment in the innovation and its speed of response.7)

3. Hypotheses and Predictions

Since these propositions are not spelled out in sufficient detail to permit testing, we translate them into the following, more specific model:

\[(5) \quad d_j = a_1 + a_2 \ln S_j + a_3 P_j + \epsilon_j,\]

where $d_j$ is the number of years the $j^{th}$ firm waits before beginning to use this innovation, $S_j$ is its size, $P_j$ is a surrogate for the profitability
of its investment in the innovation, $e_j$ is a random error term, $a_2$ is negative, and $a_3$ is positive or negative, depending on whether $P_j$ is inversely or directly related to the profitability of the investment.

We somewhat arbitrarily use $\ln S_j$ in equation (5) because the second proposition states that, as $S_j$ increases, $C(d_j)$ decreases at an increasing rate. The logarithmic form has this property; it is convenient to work with; and some experiments suggest that the results below would not be altered in any event if somewhat different forms had been used instead. Similarly, the only reason for using a linear form in the case of $P_j$ is its simplicity.

If we were sure that this model would hold for a given innovation and if we knew each firm's value of $S_j$ and $P_j$, equation (5) could be used to help predict whether one firm would be quicker than another to begin using the innovation. Without information regarding the values of $a_2$ and $a_3$, we could make predictions only for firms with similar values of either $S_j$ or $P_j$. But such predictions, to the extent that they are accurate, are obviously useful. Taking firms where $P_j$ is approximately the same, we would predict that large firms would be quicker than small ones; and taking firms where $S_j$ is approximately the same, we would predict that firms where the innovation is more profitable would be quicker than firms where it is less profitable.
The likelihood that these predictions are correct obviously depends on the probability that the hypotheses in equation (5) will hold. How large is this probability? Given that a regression of the form of equation (5) were fitted to data for a randomly chosen innovation, how large is the probability that \( a_2 \) and \( a_3 \) would have the hypothesized signs?\(^{10}\)

To find out, we estimate \( a_2 \) and \( a_3 \) for a number of innovations in several quite different industries and determine the proportion of cases where they have the hypothesized signs. To the extent that these innovations are representative, the result is an estimate of the probability that these hypotheses hold. Then, on the basis of this result, we estimate the probability that predictions of this sort are correct. Section 4 describes the basic data that are used, and Section 5 presents the results.

4. Innovations and Data

As a first step, data were collected regarding the diffusion of fourteen innovations in the bituminous coal, iron and steel, brewing, and railroad industries: the shuttle car, trackless mobile loader, and continuous mining machine (in bituminous coal); the by-product coke oven, continuous wide strip mill, and continuous annealing (in iron and steel); the pallet-loading machine, tin container, and high-speed bottle filler (in brewing); and the diesel locomotive, centralized traffic control, mikado locomotive, trailing-truck locomotive, and car retarders (in railroads).\(^{11}\)
Three kinds of data were collected in each case. First, we obtained the date when each major firm in the industry began to use the innovation, and we subtracted it from the date when the first firm began to use it to obtain \( d_j \). Second, an estimate was made of each firm's size. Physical output was used in the coal and brewing industries, ingot capacity was used in steel, and freight ton-miles were used in the railroad industry to measure \( S_j \).

Third, because it was impossible to get a direct estimate of the innovation's profitability to each firm, surrogates were obtained where possible. For example, since the profitability of a firm's investment in a continuous mining machine was likely to vary directly with the percent of its output derived from "high seams," this percentage was used as a surrogate. Other surrogates were the percent of a railroad's mileage that was double track (centralized traffic control), the percent of a firm's revenues derived from hauling coal (diesel locomotive), the ratio of a firm's rolling capacity to its ingot capacity (continuous wide strip mill), and a firm's tinplate capacity as a percent of its ingot capacity (continuous annealing).

Two points should be noted regarding these surrogates. First, despite considerable effort, no suitable surrogates could be found for nine of the fourteen innovations, and the surrogates used in the remaining five cases are obviously very rough. Second, for centralized traffic control and the diesel locomotive, \( P_j \) is inversely related to the profitability of the innovation, and consequently \( a_j \) would be expected to be positive if the hypotheses hold. In the case of the other innovations, since \( P_j \) is directly related to the profitability of the innovations, \( a_j \) would be expected to be negative.
5. Empirical Results

Using these rather crude data, we obtained least-squares estimates of \( a_1, a_2, \) and (where possible) \( a_3 \). The results are shown in Table 1. Because practically all of the firms in the populations described in note 12 are included, sampling errors in the ordinary sense are small, if present at all. In examining the results, more weight must be given to the innovations where data for both \( S_j \) and \( P_j \) could be obtained. The results in the other cases only provide evidence regarding some, not all, of the hypotheses in Section 2 - 3.

In what proportion of these cases do the estimates of \( a_2 \) and \( a_3 \) have the expected signs? Taking the five more significant cases where data for both \( S_j \) and \( P_j \) are available, the estimates have the expected signs in every case. Taking the nine cases where only data on \( S_j \) are available, the estimates of \( a_2 \) have the expected sign in all but one case.

Thus, taken as a whole, the results suggest that there is a high probability that these hypotheses will hold for a randomly chosen innovation. Treating these fourteen innovations as a random sample, one can be reasonably sure (fiducial probability = .80) that the chances are at least 4 out of 5 that \( a_2 \) will be negative for a randomly chosen innovation. Similarly, one can be reasonably sure that the chances are at least 3 out of 4 that \( a_3 \) will have the expected sign for a randomly chosen innovation. Of course, we have data
Table 1 - Estimates of $a_1$, $a_2$, and $a_3$, Fourteen Innovations, Coal, Steel, Brewing, and Railroad Industries.

<table>
<thead>
<tr>
<th>Innovation</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>Number of Firms</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel locomotive</td>
<td>15.0</td>
<td>-2.67</td>
<td>-.08</td>
<td>71</td>
<td>.70</td>
</tr>
<tr>
<td>Continuous mining machine</td>
<td>11.5</td>
<td>-2.71</td>
<td>-.03</td>
<td>12</td>
<td>.53</td>
</tr>
<tr>
<td>Continuous wide strip mill</td>
<td>34.1</td>
<td>-3.04</td>
<td>-16.51</td>
<td>12</td>
<td>.88</td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>42.6</td>
<td>-1.33</td>
<td>-.20</td>
<td>9</td>
<td>.56</td>
</tr>
<tr>
<td>Centralized traffic control</td>
<td>30.2</td>
<td>-8.27</td>
<td>.24</td>
<td>23</td>
<td>.59</td>
</tr>
</tbody>
</table>

Data on $S_j$ and $P_j$:

<table>
<thead>
<tr>
<th>Innovation</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>Number of Firms</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shuttle car</td>
<td>63.4</td>
<td>-3.91</td>
<td>-</td>
<td>11</td>
<td>.75</td>
</tr>
<tr>
<td>Trackless mobile loader</td>
<td>18.6</td>
<td>-.83</td>
<td>-</td>
<td>11</td>
<td>.17</td>
</tr>
<tr>
<td>Tin container</td>
<td>13.8</td>
<td>-.16</td>
<td>-</td>
<td>19</td>
<td>.02</td>
</tr>
<tr>
<td>Pallet loading machine</td>
<td>19.5</td>
<td>-.97</td>
<td>-</td>
<td>18</td>
<td>.30</td>
</tr>
<tr>
<td>High speed bottle filler</td>
<td>23.6</td>
<td>-1.12</td>
<td>-</td>
<td>15</td>
<td>.23</td>
</tr>
<tr>
<td>By-product coke oven</td>
<td>31.9</td>
<td>-2.59</td>
<td>-</td>
<td>12</td>
<td>.29</td>
</tr>
<tr>
<td>Trailing-truck locomotive</td>
<td>8.9</td>
<td>.16</td>
<td>-</td>
<td>28</td>
<td>.02</td>
</tr>
<tr>
<td>Car retarders</td>
<td>24.0</td>
<td>-4.28</td>
<td>-</td>
<td>22</td>
<td>.22</td>
</tr>
<tr>
<td>Mikado locomotive</td>
<td>14.2</td>
<td>-2.06</td>
<td>-</td>
<td>32</td>
<td>.44</td>
</tr>
</tbody>
</table>

Data on $S_j$ only:

Source: See notes 12-17.

1Dashes are shown in cases where no estimate of $a_3$ could be made.

The units in which a firm's size is measured here are billions of freight ton-miles (railroad innovations), thousands of tons of capacity (strip mill and coke oven), dollars of sales (annealing), production in barrels (brewing innovations), production in tons (shuttle car and mobile loader), and millions of tons produced (continuous mining machine). Naperian logarithms are used throughout. For further details, see note 12. Surrogates for $P_j$ could be obtained for only the first five innovations. For a description of these surrogates see notes 14 and 15.
for only a small number of innovations, and the assumption that we can act as if they were randomly chosen is rather bold, but nonetheless the results in Table 1 certainly suggest that the probability is high that these hypotheses will hold.

In addition, there is evidence of another sort that seems to support these findings. When about thirty executives in these industries were interviewed, their impression seemed to be that these hypotheses usually held in their industries. However, there were some interindustry differences in their impression of the effect of a firm's size. In the coal, brewing, and railroad industries, their almost unanimous impression was that the smaller firms were slower than the large ones to install important new techniques, but in the steel industry many believed the opposite to be true.

Of course, all of this pertains to the probability that these hypotheses will hold. Although this probability may be quite high, they may be of little use for predictive purposes (Cf. notes 21 and 22). What is the probability that a prediction based on these hypotheses will hold for a randomly chosen innovation? Suppose that one firm is \( X \) times as large as another, the value of \( P_j \) being the same for each firm, and that we predict that the larger firm will be quicker than the smaller one to introduce the innovation. What are our chances of being correct?

Of course, the answer depends on how large \( X \) is. If \( X = 4 \), our chances of being correct seem to be at least .75, whereas if \( X = 2 \), our chances seem only to be about .65. (To obtain these figures, we assume that, if these hypotheses hold, \( \epsilon_j \) is normally distributed with variance \( \sigma^2_\epsilon \).
If so, our chances of being right are at least $P_H \left[ -a_2 \ln X / \sigma_e \sqrt{2} \right]$, where $P_H$ is the probability that these hypotheses will hold. Estimates from Table 1 were inserted in this expression to obtain these figures.\(^{21}\)

Thus, so long as the difference between firm sizes is quite large, it appears that predictions of this sort have a very good chance of being correct. Moreover, although the situation is somewhat more complicated in cases where predictions are based on differences in values of $P_j$ ($S_j$ being held constant), the results in Table 1 suggest that these predictions also have a much better than even chance of being right if the difference in $P_j$ is quite large.\(^{22}\)

In conclusion, three additional points should be made regarding the empirical results in Table 1. First, they provide the first quantitative estimates of the effect of a firm's size on the length of time it waits before introducing an innovation. In these cases, a firm that was half as large as another waited about two years longer on the average. Second, they show that a considerable amount of the variation in $d_j$ can be explained by $\ln S_j$ and $P_j$. For those cases where data for both $S_j$ and $P_j$ are available, almost half of the variation can be explained by these two variables alone. Third, the underlying scatter diagrams seem to indicate that as $S_j$ increases, $\bar{C}(d_j)$ decreases at an increasing rate. Thus the evidence seems to support the second proposition in Section 2.
6. Growth Rate and Profitability of the Firm

Besides a firm's size and the extent of the returns it obtains from the innovation, there are many other factors that may influence its speed of response to an innovation. Sections 6 - 7 contain propositions regarding the effects of five such factors and estimates of the probability that each will hold for a particular innovation. If one could be reasonably sure that these propositions would hold, they could be useful -- in the same way as the propositions in Section 2 -- in predicting whether one firm would be quicker than another to use an innovation.

The first proposition states that, other factors held equal, the length of time a firm waits before introducing a new technique tends to be inversely related to its rate of growth. If a firm is expanding at a relatively rapid rate, it might be expected to install a new technique relatively quickly because it can introduce it in its new plants, whereas a firm experiencing little or no growth must wait until it can profitably replace existing equipment. Of course, the significance of this factor will vary somewhat, but for a large class of innovations, it would seem to be important.\(^{23}\)

The second proposition states that, other factors held equal, the length of time a firm waits before introducing a new technique tends to be inversely related to the firm's profitability. One would suppose that less profitable firms -- with their smaller cash inflows and poorer credit ratings -- would have more difficulty in financing the necessary investment and that they would be in a poorer position to take whatever risks are involved in being among the first to use it.\(^{24}\)
If these propositions (as well as those in Section 2) hold, we assume that

\[ d_j = b_1 + b_2 \ln S_j + b_3 P_j + b_4 g_j + b_5 \Pi_j + e'_j, \]

where \( \Pi_j \) is the \( j^{th} \) firm's profits as a percent of its net worth during a three-year period soon after the innovation was first put on the market, \( g_j \) is the percentage increase in the \( j^{th} \) firm's production or capacity (plus 100) during the (approximate) period during which the imitation process was going on, \( b_2 \) and \( b_3 \) have the same signs as \( a_2 \) and \( a_3 \), both \( b_4 \) and \( b_5 \) are negative, and \( e'_j \) is an error term. 25

To estimate the probability that each of these hypotheses will hold, we estimate \( b_4 \) and \( b_5 \) in every case for which we have the necessary data and determine the proportion of cases where they have the appropriate signs. Table 2 shows the least-squares estimate of the \( b \)'s, the data on \( g_j \) and \( \Pi_j \) having been obtained from trade sources and Moody's. There is no evidence that the probability that these hypotheses will hold is very high. The estimates of \( b_5 \) have the expected sign in about half of the cases; the estimates of \( b_4 \) have the expected sign in only about one-third of the cases. Moreover, if either \( b_4 \) or \( b_5 \) is assumed to be zero, the results for the other parameter remain the same. 26
Table 2 - Estimates of $b_1$, $b_2$, $b_3$, $b_4$, and $b_5$, Eleven Innovations, Coal, Steel, Brewing, and Railroad Industries.

<table>
<thead>
<tr>
<th>Innovation</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>$b_4$</th>
<th>$b_5$</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel locomotive</td>
<td>14.13</td>
<td>-2.66</td>
<td>.08</td>
<td>-.0002</td>
<td>.15</td>
<td>65</td>
</tr>
<tr>
<td>Continuous mining machine</td>
<td>5.76</td>
<td>-1.50</td>
<td>-.06</td>
<td>.003</td>
<td>.46</td>
<td>8</td>
</tr>
<tr>
<td>Continuous wide strip mill</td>
<td>62.57</td>
<td>-5.03</td>
<td>-.18</td>
<td>-.004</td>
<td>-1.14</td>
<td>10</td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>44.66</td>
<td>-1.40</td>
<td>-.21</td>
<td>.005</td>
<td>-.49</td>
<td>9</td>
</tr>
<tr>
<td>Centralized traffic control</td>
<td>28.40</td>
<td>-3.39</td>
<td>.23</td>
<td>.001</td>
<td>.30</td>
<td>23</td>
</tr>
</tbody>
</table>

Data on $S_j$ only:

<table>
<thead>
<tr>
<th>Innovation</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>$b_4$</th>
<th>$b_5$</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shuttle car</td>
<td>98.18</td>
<td>-6.03</td>
<td>-</td>
<td>-.002</td>
<td>-.12</td>
<td>9</td>
</tr>
<tr>
<td>Trackless mobile loader</td>
<td>22.80</td>
<td>-1.50</td>
<td>-</td>
<td>.031</td>
<td>-.68</td>
<td>9</td>
</tr>
<tr>
<td>Pallet loading machine</td>
<td>26.19</td>
<td>-1.27</td>
<td>-</td>
<td>-.004</td>
<td>-.25</td>
<td>7</td>
</tr>
<tr>
<td>High-speed bottle filler</td>
<td>27.63</td>
<td>-2.55</td>
<td>-</td>
<td>.107</td>
<td>.41</td>
<td>6</td>
</tr>
<tr>
<td>Trailing-truck locomotive</td>
<td>1.90</td>
<td>-0.47</td>
<td>-</td>
<td>.008</td>
<td>.81</td>
<td>21</td>
</tr>
<tr>
<td>Car retarders</td>
<td>10.06</td>
<td>-4.13</td>
<td>-</td>
<td>.038</td>
<td>-.28</td>
<td>22</td>
</tr>
</tbody>
</table>

Source: See section 6.

The units in which $g_j$ is measured are described in Section 6; $\Pi_j$ is measured in percentage points. See notes 12-14 for a description of the measures of $S_j$ and $P_j$. 

---

1 The units in which $g_j$ is measured are described in Section 6; $\Pi_j$ is measured in percentage points. See notes 12-14 for a description of the measures of $S_j$ and $P_j$. 

---
7. Age of President, Liquidity, and Profit Trend

Of course, if these hypotheses really hold in the great majority of the cases, the inclusion of other independent variables in equation (6) may make this more apparent. As an experiment, we included \( A_j \) -- the age of the \( j \)th firm's president -- as an additional independent variable. It is often asserted that younger managements, being less bound by traditional ways, are more likely than older ones to introduce a new technique relatively quickly. Moreover, in agriculture, there is some evidence that this is the case [3, 9]. If this proposition (as well as the previous ones) holds, we assume that

\[
d_j = c_1 + c_2 \ln S_j + c_3 P_j + c_4 g_j + c_5 \Pi_j + c_6 A_j + \varepsilon_j'',
\]

where \( c_2, \ldots, c_5 \) have the same signs as \( b_2, \ldots, b_5 \); \( c_6 \) is positive; and \( \varepsilon_j'' \) is a random error term.\(^{27} \)

To estimate the probability that each of these hypotheses will hold, we determine the proportion of cases where the estimates of \( c_4 \), \( c_5 \), and \( c_6 \) have the appropriate signs. The results, shown in Table 3, indicate that the chances are only about 50-50 that each will hold. Moreover, the results remain essentially the same when one of these parameters is set equal to zero and when each pair is set equal to zero. In all, about seventy such regressions were run, the results in each case being similar to those in Table 3.
Table 3 - Estimates of \( c_1, c_2, c_3, c_4, c_5, \) and \( c_6 \), Nine Innovations, Coal, Steel, Brewing, and Railroad Industries.

<table>
<thead>
<tr>
<th>Innovation</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
<th>( c_3 )</th>
<th>( c_4 )</th>
<th>( c_5 )</th>
<th>( c_6 )</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data on ( S_j ) and ( P_j ):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous mining machine</td>
<td>6.35</td>
<td>-1.49</td>
<td>-.067</td>
<td>.003</td>
<td>.480</td>
<td>-.009</td>
<td>8</td>
</tr>
<tr>
<td>Continuous wide strip mill</td>
<td>65.17</td>
<td>-5.54</td>
<td>-1.85</td>
<td>-.004</td>
<td>-1.32</td>
<td>.053</td>
<td>10</td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>69.16</td>
<td>-1.85</td>
<td>-.226</td>
<td>.013</td>
<td>-.019</td>
<td>-.420</td>
<td>9</td>
</tr>
<tr>
<td>Centralized traffic control</td>
<td>48.14</td>
<td>-6.82</td>
<td>.297</td>
<td>-.001</td>
<td>.328</td>
<td>-.417</td>
<td>23</td>
</tr>
<tr>
<td>Data on ( S_j ) only:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shuttle car</td>
<td>100.56</td>
<td>-6.06</td>
<td>-</td>
<td>-.001</td>
<td>-1.07</td>
<td>-.040</td>
<td>9</td>
</tr>
<tr>
<td>Trackless mobile loader</td>
<td>8.47</td>
<td>-1.18</td>
<td>-</td>
<td>.029</td>
<td>-.470</td>
<td>.190</td>
<td>9</td>
</tr>
<tr>
<td>Pallet loading machine</td>
<td>31.22</td>
<td>-1.23</td>
<td>-</td>
<td>-.003</td>
<td>-.263</td>
<td>-.092</td>
<td>7</td>
</tr>
<tr>
<td>Trailing-truck locomotive</td>
<td>-13.91</td>
<td>-1.36</td>
<td>-</td>
<td>.008</td>
<td>.728</td>
<td>.311</td>
<td>21</td>
</tr>
<tr>
<td>Car retarders</td>
<td>-9.57</td>
<td>-5.34</td>
<td>-</td>
<td>.040</td>
<td>-.351</td>
<td>.378</td>
<td>22</td>
</tr>
</tbody>
</table>

Source: See Section 7.

\(^1\) \( A_j \) is measured in years. For further explanation, see note 1, Table 2.
Finally, we take a less detailed look at the effects of two other factors -- a firm's liquidity and its profit trend. With regard to liquidity, one might expect that more liquid firms would be better able to finance the investment in the innovation and that consequently they might be quicker than less liquid firms to use it. With regard to a firm's profit trend, one might suppose that firms with decreasing profits would be stimulated to search more diligently than other firms for new alternatives [17] and that, other things equal, they might tend to be quicker than others to begin using a new technique.

If these propositions (and those in Section 2) hold for a given innovation we assume that

\[ d_j = \theta_1 + \theta_2 \ln S_j + \theta_3 P_j + \theta_4 L_j + \theta_5 t_j + \epsilon_j^{''''}, \]

where \( L_j \) is the average value of the \( j^{th} \) firm's current ratio (current assets divided by current liabilities) during the three years up to and including the year when it began using the innovation, \( t_j \) is the slope of the linear regression of the \( j^{th} \) firm's profit rate against time (measured in years) during a six-year period just before it began using the innovation, \( \theta_2 \) and \( \theta_3 \) have the same signs as \( a_2 \) and \( a_3 \), \( \theta_4 \) is negative, \( \theta_5 \) is positive, and \( \epsilon_j^{''''} \) is an error term.\[28\]

To estimate the probability that these hypotheses will hold, we estimate the \( \theta \)'s for those innovations for which data regarding \( S_j \) and \( P_j \) could be obtained and compute the proportion of cases where \( \theta_4 \) and \( \theta_5 \) have the
appropriate signs. The very small number of cases prevents us from coming to any real conclusion, but the results in Table 4 seem somewhat discouraging with regard to the liquidity hypothesis (the estimate of $\theta_4$ having the expected sign only once) and somewhat encouraging with regard to the profit-trend hypothesis (the estimate of $\theta_5$ having the expected sign in three out of four cases).\textsuperscript{29}

In conclusion, there is no real evidence that any of the hypotheses in Sections 6 - 7 are of use in predicting whether one firm will be quicker than another to introduce a new technique. Judging from Tables 2 - 4, one would have only about a 50-50 chance, if that much, of being correct if he predicted, holding $S_j$ and $P_j$ constant, that a more profitable firm would be quicker than a less profitable one, that a faster growing firm would be quicker than a slower growing one, that a more liquid firm would be quicker than a less liquid one, or that a firm with a younger president would be quicker than one with an older president. With regard to the hypothesis concerning the effects of a firm's profit trend, there is too little data to be sure one way or the other, but the available data seem fairly encouraging.
Table 4 - Estimates of $\theta_1$, $\theta_2$, $\theta_3$, $\theta_4$, and $\theta_5$, Four Innovations, Coal, Steel, and Railroad Industries.

<table>
<thead>
<tr>
<th>Innovation</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
<th>$\theta_4$</th>
<th>$\theta_5$</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diesel locomotive</td>
<td>13.59</td>
<td>-2.48</td>
<td>.088</td>
<td>.491</td>
<td>-.682</td>
<td>65</td>
</tr>
<tr>
<td>Continuous mining machine</td>
<td>10.02</td>
<td>-3.03</td>
<td>-.008</td>
<td>.189</td>
<td>.153</td>
<td>9</td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>47.70</td>
<td>-1.01</td>
<td>-.099</td>
<td>-2.73</td>
<td>.306</td>
<td>9</td>
</tr>
<tr>
<td>Centralized traffic control</td>
<td>30.03</td>
<td>-8.44</td>
<td>.011</td>
<td>.181</td>
<td>.299</td>
<td>23</td>
</tr>
</tbody>
</table>

Source: See Section 7.

1 See note 1, Table 2.
8. Concentration of Technical Leadership

To what extent do the firms that are quick -- or slow -- to introduce one innovation tend to be quick -- or slow -- to introduce others as well? How high is the correlation between how rapidly a firm introduces one innovation and how rapidly it introduces another? The answer to this question is important because it shows the extent to which technical leadership is concentrated in the hands of only a few members of an industry. In this section, we see how closely it has been concentrated among firms in the bituminous coal, railroad, brewing, and iron and steel industries.

As a first step, we note that the coefficient of correlation between how rapidly a firm introduces one innovation and how rapidly it introduces another is likely to be inversely related to the time interval separating the date when the one innovation first appeared from the date when the other innovation first appeared. Put differently, as the time interval separating the appearance of two innovations increases, there is likely to be less tendency for the same firms to be relatively quick -- or slow -- to introduce both. This seems reasonable because, as time goes on, technical leadership, if at all polarized, is likely to pass from one group of firms to another.\textsuperscript{30}

Assuming that this hypothesis is correct and that a linear function is satisfactory, we have

\[
\rho_{qr} = v + w t_{qr} + Z_{qr},
\]
where $\rho_{qr}$ is the coefficient of correlation between the length of time a firm waits before introducing the $q^{th}$ innovation and the length of time it waits before introducing the $r^{th}$ innovation, $t_{qr}$ is the interval (in years) between the dates when these two innovations were first used commercially, and $Z_{qr}$ is a random error term.$^{31}$

Estimates of $v$ and $w$ for each industry would allow us to estimate the average correlation coefficient, given that $t_{qr}$ is fixed. To obtain estimates of $v$ and $w$, we estimated $\rho_{qr}$ for each pair of innovations in a given industry (Table 5). Then, since the appropriate analysis of covariance provides no evidence to the contrary,$^{32}$ we assumed that there were no inter-industry differences in $w$; and using least-squares, we found that

\[
(10) \quad P_{qr} = \begin{cases} 
.40 \\
.30 \\
.66 \\
.28 
\end{cases} - .012 t_{qr} (.004)_{qr},
\]

where $Z_{qr}$ is omitted and the figures in brackets (reading from top to bottom) pertain to brewing, coal, steel, and railroads.

This result indicates at least four things. First, given that two innovations occur within a few decades of each other, one can expect some positive correlation between how long a firm waits before introducing one and how long it waits before introducing the other. Thus, if two innovations are reasonably close together in time, there is generally some tendency for the same firms to be relatively quick -- or slow -- to introduce both.$^{33}$
Table 5 - Values of $\rho_{qr}$ and $t_{qr}$ for Nineteen Pairs of Innovations, Coal, Steel, Brewing, and Railroad Industries.\(^1\)

<table>
<thead>
<tr>
<th>Pair of Innovations</th>
<th>$\rho_{qr}$</th>
<th>$t_{qr}$</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous mining machine: shuttle car</td>
<td>-.02</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Continuous mining machine: trackless mobile loader</td>
<td>-.17</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Shuttle car: trackless mobile loader</td>
<td>.54</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>By-product coke oven: continuous strip mill</td>
<td>.21</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>By-product coke oven: continuous annealing</td>
<td>.18</td>
<td>42</td>
<td>5</td>
</tr>
<tr>
<td>Continuous strip mill: continuous annealing</td>
<td>.53</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>High speed bottle filler: tin container(^2)</td>
<td>.42</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>High speed bottle filler: pallet loading machine(^2)</td>
<td>.24</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Tin container: pallet loading machine</td>
<td>-.06</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Centralized traffic control: trailing-truck locomotive</td>
<td>.02</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Centralized traffic control: diesel locomotive</td>
<td>.04</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Centralized traffic control: car retarders</td>
<td>.16</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Centralized traffic control: mikado locomotive</td>
<td>.19</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>Trailing-truck locomotive: diesel locomotive</td>
<td>.26</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Trailing-truck locomotive: car retarders</td>
<td>.32</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Trailing-truck locomotive: mikado locomotive</td>
<td>.13</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>Diesel locomotive: car retarders</td>
<td>.41</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Diesel locomotive: mikado locomotive</td>
<td>.12</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Car retarders: mikado locomotive</td>
<td>-.29</td>
<td>23</td>
<td>21</td>
</tr>
</tbody>
</table>

Source: See Section 8.

\(^1\)See Section 8 for definitions of $\rho_{qr}$ and $t_{qr}$. The latter is measured in years. The number of firms in the final column differs from that in Table 1. Only those firms for which we have data regarding both innovations could be included.

\(^2\)Only firms already using high-speed bottle fillers are included here. See note 16.
Second, although there is some such tendency, technical leadership does not seem to be very highly concentrated in most of these industries. Even if two innovations occur simultaneously, the average value of $\rho_{qr}$ in the bituminous coal, brewing, and railroad industries is only about .30; and if the innovations occur five years apart, it is only about .25. To see the implications of this, suppose that firms are ranked by how quickly they used a particular innovation and that one firm is one-quarter of the way down the ranking, whereas a second firm is three-quarters of the way down the ranking. If another innovation occurred five years later and $\rho_{qr} = .25$, the probability is only .59 that the first firm would introduce it before the second firm.34

Third, although leadership seems to be quite widely diffused in most of these industries, there is one exception -- iron and steel. If two innovations in steel occur close together in time, there seems to be a fairly high correlation (about .60) between how rapidly a firm introduces one and how rapidly it introduces the other. The higher correlation in steel may be due in part to a more unequal distribution among firms of research expenditures (and other such factors) than in other industries.35

Fourth, as we expected, the estimate of $w$ is negative and statistically significant. Thus, as the time interval separating two innovations increases, there is less correlation between a firm's speed of response to one and its speed of response to the other. But it is noteworthy that the correlation coefficient decreases very slowly, an increase in the time interval of one year resulting in a decrease of only about .01 in the correlation coefficient.
Finally, we should note that if the rank (not the product-moment) correlation coefficient had been used in equation (10), the results would have been almost exactly the same as those presented there. Moreover, if we had used the coefficient of correlation between \( e_j \) [the residual in equation (5)] for one innovation and \( e_j \) for another innovation, it would have made little difference either.

9. **Summary and Conclusions**

This paper is concerned with the process whereby the use of a new technique spreads from one firm to another. Put very briefly, the results indicate that the length of time a firm waits before using a new technique is generally inversely related to its size and the profitability of its investment in the innovation. If the innovations for which we could obtain data are representative, one can be quite sure that relationships of this sort will be present in at least 75-80 per cent of the cases. Moreover, if the differences in size or the profitability of the investment in the innovation are substantial, these relationships seem to be quite useful for purposes of prediction. For example, if one firm is four times as large as
another (the profitability of the investment in the innovation being the same for both), the chance that it will introduce a randomly chosen innovation more rapidly than its smaller competitor seems to be at least .75.

On the other hand, there is no evidence that several other hypotheses presented above are of use in predicting whether one firm will be quicker than another to introduce a new technique. If one holds firm size and the profitability of the investment in the innovation constant and if one predicts that the length of time a firm waits will be inversely related to its profitability, its growth rate, and its liquidity, or directly related to the age of its president, there seems to be little better than an even chance, if that much, of being correct. With regard to the effects of a firm's profit trend, the results are more encouraging but inconclusive because of the small amount of data.

In addition, the results indicate that technical leadership of this kind has not been very highly concentrated in most of the industries for which we have data. Even if one firm was considerably quicker than another to begin using one innovation, the chance that it will also be quicker to introduce another innovation occurring only five years later is not very much better than 50-50. Despite the tendency (all other things equal) for the larger firms to lead, other factors like the profitability of investment in the innovation are of considerable importance. Apparently, there is no particular group of firms that consistently exercises leadership of this kind and no particular group that consistently brings up the rear.
These findings have at least four implications. First, they support the general proposition that the speed at which a firm responds to an investment opportunity is directly related to the profitability of the opportunity. This proposition, which is akin to the psychological laws relating the speed of response to the extent of the stimulus, has played an important role in studies of the imitation process and will undoubtedly prove useful in other areas too.

Second, the results seem to contradict the view held by some economists that increases in size result in such sluggish performance that large firms tend to follow the lead of small ones. On the contrary, it appears that, $P_j$ held constant, one can predict with considerable confidence that a large firm will be quicker than a small one to begin using a new technique. Of course, this tells us nothing about the effects of market structure on the rate of imitation, but it does indicate that, holding market structure and $P_j$ constant, the larger firms can be expected to be the early users of new techniques.

Third, the results seem to indicate that a firm's financial health as measured by its profitability, liquidity, and growth rate, bears no close relationship to how long it waits before introducing a new technique. Whereas the more prosperous members of the industry tend to be quicker than the others to introduce some innovations, they are slower than the others to introduce about as many other innovations. However, it is possible that changes in a firm's financial health are important, some fragmentary evidence indicating that, if a firm's profits are declining, it may be somewhat quicker to seize on
new techniques.

Fourth, the results show quite clearly the dangers involved in the common assumption that certain firms are repeatedly the leaders, or followers, in introducing new techniques. It would be very misleading to take a few innovations and assume that the firms that are quick to use them are generally the leaders in this sense. Judging by our findings, there is a very good chance that these firms will be relatively slow to introduce the next innovation that comes along.

In conclusion, the limitations of the study should be noted. The data we could obtain are limited in quality and scope; and because they are almost impossible to measure, we had to omit many important factors -- the amount of research a firm conducted in the relevant area, the preferences of its management with respect to risk, the age of its old equipment, and the extent to which manufacturers of the new equipment could exert pressure on it. The results are only a first step toward understanding the factors determining whether one firm will be quicker than another to use an innovation, but despite their obvious limitations, they should be useful to economists concerned with the process of technical change.
REFERENCES


FOOTNOTES

* The work on which this report is based was supported financially by the Graduate School of Industrial Administration at Carnegie Institute of Technology, the Office of Special Studies of the National Science Foundation, The Ford Foundation, and the Cowles Foundation for Research in Economics at Yale University. It is part of a broader study of research, technical change, and economic growth that I am conducting. My thanks go to the many people in industry who provided information and to A. Meltzer, J. Muth, and R. Nelson for their helpful comments.

1/ In recent years, studies of the imitation process have been carried out by Bohlen and Beal [5], Coleman, Katz and Menzel [6], Griliches [7], Hildebrand and Partenheimer [9], Mansfield [13, 15], Sutherland [21], and Yance [22]. But only a few of these studies have been concerned with the factors influencing a particular firm's speed of response, and those that have been concerned with these factors have pertained almost always to agriculture. This seems to be the first study pertaining to the industrial sector.

2/ Of course, the universe from which an innovation would be chosen at random is an abstract one containing all new techniques that occurred in a given industry during a specified period. Needless to say, universe of this sort are often used in econometrics and are useful despite the fact that they are abstractions.

3/ Not all economists seem to agree with this proposition. For example, Stocking [20], p. 966, in discussing this question, has asserted that pioneering is unlikely to occur where "bureaucracy, red tape, and a quest for security... afflict private industry...[and] the likelihood of their doing so varies directly with the size of a firm." Although it is undoubtedly true that increases in size are sometimes accompanied by unwieldiness and conservatism, it seems unlikely to me that they generally can offset the factors noted above.

On the other hand, Carter and Williams [5, p. 186] conclude that, while technical progressiveness is possible for firms of all sizes, small firms often operate with certain handicaps that make technical progress particularly difficult for them to achieve. Scitovsky [18, p. 78], although he is concerned with a somewhat different problem seems to be impressed with the advantages of size in this connection. Bohlen and Beal [3] present evidence to the effect that the large firms tend to be quicker to introduce new agricultural techniques.
To illustrate the second point in the text, consider an innovation that can be used in only a certain type of coal mine when it is first commercialized at time $t_1$, but that is made applicable to all mines at time $t_2$. Assume (for simplicity) that each mine has the same chance -- $P$ -- of being able to accommodate the innovation at first and that this chance does not depend on whether other mines under the same ownership can accommodate it. If each firm begins using it as soon as it can, one can show that

$$E(d_j) = (1 - P)^{S_j/A} \left( t_2 - t_1 \right),$$

where $E(d_j)$ is the expected value of the number of years the $j^{th}$ firm will wait (dating from $t_1$) before introducing it, $S_j$ is its output rate, and $A$ is the output rate per mine. As expected, there is an inverse relationship between $E(d_j)$ and $S_j$.

To illustrate the third point in the text, suppose that an innovation occurs at time $t$ and that the number of years -- measured from time $t$ -- before a unit of old equipment can profitably be replaced by the innovation is a random variable, $I$, with probability density

$$P(i) = \Phi^{-1} \quad (0 \leq i \leq \Phi)$$

If we assume that the number of years before one unit can be replaced is statistically independent of that for another unit and if the number of units owned by a firm is proportional to its output rate, it follows that

$$E(d_j) = \Phi \left( S_j/C + 1 \right)^{-1},$$

where $C$ is the output rate per unit of the replaced equipment. As expected, there is an inverse relationship between $E(d_j)$ and $S_j$.

This model incorporates both the second and third reasons given in the preceding section for expecting that large firms will lead small ones. The results in note 4 are special cases. Note that the first reason is omitted entirely here.

The assumption that $F(x)$ is the same for each unit and that the $x$'s are independent is not very realistic, but small departures from it should make little difference. The results are used only to get a rough idea of the shape of the relationship between a firm's size and its speed of response.
6/ Given $F(x)$, the probability density function for the smallest $x$ can easily be shown to equal $\alpha[1 - F(x)]^{\alpha-1}F'(x)$. Thus equation (1) follows. Integrating equation (1) by parts, we find that $E(\alpha) = -x[1 - F(x)]^\alpha \bigg|_0^M + \int_0^M [1 - F(x)]^\alpha \, dx$. Since the first term equals zero, equation (2) follows.

Equations (3) and (4) follow simply, and since $0 \leq F(x) \leq 1$, the indicated inequalities hold.

7/ There is considerable agreement that the profitability of a new technique determines how quickly a firm begins using it. See Mansfield [13], Griliches [7], Mack [12], and Sutherland [21].

There is little use in speculating further about the shape of the relationship between the profitability of a firm's investment in the innovation and its speed of response because, with so little knowledge of the relationship between the actual profitability and the surrogate, we could deduce little about the relationship between the surrogate and the firm's speed of response (other than what is expressed in equation (5)) even if we knew a great deal about the relationship between actual profitability and the firm's speed of response.

8/ Instead of assuming that equation (5) held, we assumed that $\ln d_j$ was a linear function of $S_j$ and $P_j$. The results turned out to be similar to those in Table 1.

9/ Of course it would be better if we could predict $a_2$ and $a_3$ in advance. But, because $a_3$ depends on which surrogate is chosen and on the units in which the surrogate is measured and because we have little idea of the relationship between various surrogates and the profitability of the investment in the innovation, the prediction of $a_3$ at this point seems rather hopeless. However, if data could be obtained regarding the actual profitability figures, not surrogates, it might be possible to do something along this line. E.g., one might expect that the effect of differences in profitability on $d_j$ would be greater if the innovation is less profitable generally than if it is very profitable. Thus, $a_3$ might be inversely related to the average value of $P_j$.

The problem in estimating $a_2$ is not quite so great as in estimating $a_3$. One of the most important factors governing $a_2$ is the profitability of the innovation. If it is very profitable, one would expect that $\ln S_j$ would
have less effect on $d_j$ than if it is less profitable. Moreover, the evidence supports this. Another important factor is the size of the investment required to install the innovation. The larger the investment required, the greater the effect of $\ln S_j$ on $d_j$ is likely to be.

Finally, since the effect of $\ln S_j$ is likely to depend on the profitability of the innovation, it is likely that a multiplicative form of equation (5) would be preferable, e.g., $d_j = \phi_2 S_j^2 \phi_3 P_j$. But had such a function been used, it is unlikely that any of our major conclusions would have been altered.

10/ According to equation (5), if data regarding $d_j$, $P_j$, and $S_j$ are obtained for all relevant firms in the case of a particular innovation and if $d_j$ is regressed on $\ln S_j$ and $P_j$, the regression coefficients should have the signs stipulated above — for $a_2$ and $a_3$. If they have these signs, we regard these hypotheses as being fulfilled for this innovation. The question is: for what proportion of the innovations are these hypotheses correct?


12/ For the continuous wide strip mill, all firms having more than 140,000 tons of sheet capacity in 1926 were included; for the by-product coke oven, all firms having more than 200,000 tons of pig iron capacity in 1900 were included; and for continuous annealing, the nine major producers of tin plate in 1935 were included. For centralized traffic control and car retarders, all Class I railroads with over 5 billion freight ton-miles in 1925 were included. For the mikado and the trailing-truck locomotive, the roads included in Healy's sample [8] are included. All Class I railroads in 1925 are included for the diesel locomotive. Firms producing over 4 million tons of coal in 1956 were included for the coal innovations, and firms with more than $1 million in assets in 1934 were included for the brewing innovations.

These lower limits on size were imposed because of difficulties in
obtaining information concerning smaller firms and because in some cases they
could not use the innovation in any event. As it is, data could not be obtained
for all these firms because some went out of business or refused to cooperate.
But in most cases the results are complete -- or very nearly so. For a more
detailed discussion, see [13].

The date when each firm first introduced the innovation was obtained
from trade journals, industry directories, and correspondence with the firms.
There were a handful of firms for which no data could be obtained and they had
to be excluded. The specific sources of these data are listed in the Appendix
of Mansfield [13].

Note that these are the dates when firms first introduced the innovation --
regardless of the scale on which they did so. Of course, the possible objections
to this are largely removed by the fact that these innovations had to be intro-
duced on a fairly large scale. Moreover, the only alternative would be to use
the date when a firm first used the innovation to produce some specified
percentage of its output, and in almost every case it would have been extremely
difficult, if not impossible, to obtain such data. For a study of one case
where data of (roughly) this type were available, see Mansfield [15].

13/ For the railroads, the data on freight ton-miles come from the
Interstate Commerce Commission's Statistics of Railways and pertain to 1925.
For the coal firms, the production data come from Moody's and relate to 1939
for the trackless mobile loader and the shuttle car. For the continuous mining
For the by-product coke oven, the pig iron capacities come from the 1901
Directory of the American Iron and Steel Institute; for the continuous wide
strip mill, the ingot capacities come from the 1926 Directory of the American
Iron and Steel Institute; for continuous annealing, the size data are sales
volumes taken from Moody's for 1935. For high-speed bottle fillers and the
pallet-loading machine, production data came from Modern Brewery Age and they
pertain to 1955; for the tin container, the data are production estimates for
1940 from the Brewers Journal (July 15, 1943).

14/ According to interviews with coal executives, a firm with relatively
little coal from 4 to 9 feet high would almost surely not find continuous mining
machines as profitable as a firm with most of its coal in that range. The
greater profitability of the innovation in this range was also pointed out by
Bituminous Coal Research, Inc. in a private report. For each firm, the per cent
of its capacity in this range was computed from the 1949 Keystone Coal Buyer's
Guide and this rather crude measure was used as a surrogate for the profit-
ability of the investment.

As noted previously, the introduction of diesel locomotives seemed more
profitable for firms that hauled little coal. This factor was often cited in
the interviews, and its importance has been stressed by Yance [21]. Moreover,
Footnote 14 (Continued) - 36 -

according to railroad officials, centralized traffic control was probably less
profitable for roads with little double track. The per cent of trackage that
was double and the per cent of revenues derived from coal come from Moody's.
They pertain to 1926 and 1935.

We also noted previously that, according to the interviews, there were
considerable economies of scale for the continuous rolling mill and continuous
annealing. For given over-all size, it would therefore seem that those firms
that specialized most heavily in sheets or tin plate would have found them most
profitable. Moreover, these firms had the most to lose by delay. A firm's
sheet capacity divided by its ingot capacity was computed from the 1926
Directory of the American Iron and Steel Institute; a firm's tin plate capacity
as a per cent of its ingot capacity was computed from the 1940 Directory. Of
course, these measures are very crude.

15/ The estimates almost certainly are statistically significant, and it
would be misleading to present the ordinary standard errors that assume the sampling
fraction is small. However, in most cases the estimates would be statistically
significant even if we used the inflated standard errors that assume incorrectly
that the sampling fraction is small. Note too that, strictly speaking, our
results can pertain only to the universe of firms in note 12.

16/ These results indicate that firms for which the potential returns
were highest -- because of their physical set-ups, market situations, etc. --
tended to be early users of the innovation. Note that the data are such that
the line of causation can not be turned about. These results can not possibly
be a mere reflection of the fact that early users, by virtue of their quickness,
often enjoy a somewhat higher return. E.g., the per cent of a firm's revenues
derived from coal or the per cent of its output derived from high seams could
hardly be affected by its speed of response to these innovations.

Note that some firms had not yet begun using high-speed bottle fillers.
We included them by assuming that they would introduce them in 1963. Of course,
this makes the results for this innovation rather arbitrary, but it would also
have been misleading to exclude them altogether.

17/ However, because the uncontrolled effects of $P_j$ inflate the
residuals, the correlation coefficients are usually quite low in the latter
nine cases. Had it been possible to include smaller firms, the relationship
between $d_j$ and $\ln S_j$ would probably have been stronger. In only two
cases was it possible to extend the analysis on an exploratory basis. For the
continuous mining machine, data were collected for all firms producing over
100,000 tons in 1948 from the listings of equipment in the Keystone Coal Buyers
Guide. As size decreased, a smaller percentage had installed continuous miners
as yet, and those that did had installed them later. For centralized traffic
control, data for a somewhat larger group of firms were obtained from Healy [8],
and the results were much the same. Of course, the data used here are less
reliable than those on which Table 1 is based.
18/ That is, if one were to choose an innovation at random, one can be quite confident (fiducial probability = .80) that the probability is at least .80 that, if \( a_2 \) were computed for this innovation (using data for all relevant firms), it would turn out to be negative. Similarly, one can be quite confident that the probability is at least .75 that, if \( a_3 \) were computed, it would have the sign predicted by the model.

To obtain these results, we used the standard techniques to derive confidence intervals for a probability or percentage. First, we computed the proportion of cases where \( a_2 \) -- or \( a_3 \) -- had the proper sign.

Then, letting \( P_H \) be the true probability that \( a_2 \) -- or \( a_3 \) -- will have the proper sign, we determined that value of \( P_H \) such that the chance is .20 of our obtaining an observed proportion equal to or larger than the one we got. One can show that if this estimation procedure were used in repeated random samples, this value of \( P_H \) would be below the true probability in 80 per cent of the cases.

19/ For a description of these interviews, see Mansfield [13]. In passing, the following points might be noted, since they help to integrate our present data and findings with those presented in [13]. First, it should be noted that our finding in this paper that the larger firms tend to lead does not contradict our finding there that the rate of imitation tends to be lower in more highly concentrated industries. The results presented here for a particular innovation are based implicitly on a certain industry structure, and how long a firm of given size waits may depend on this structure (and the extent of concentration). Second, two innovations are included here but not in [13] because of lack of necessary auxiliary data. Third, the firms included here sometimes differ slightly from those included there. For the sake of greater homogeneity we excluded steel firms in dealing with the coal innovations and excluded the switching roads in the case of the car retarders. We also included all Class I roads for the diesel locomotive. Where size data could not be obtained, firms had to be omitted.

20/ Note two things here. (1) In the railroad, brewing, and coal industries, there was almost complete agreement with respect to innovations requiring a fairly large investment. But for techniques that could be installed very cheaply there was less agreement that the larger firms lead. On the basis of the reasoning in Section 2, one might anticipate that the larger firms would lead less consistently if very small amounts of capital were required. (2) The impression that the largest firms do not lead in the steel industry can be found in congressional hearings and popular business literature as well as in these interviews. But in almost all of this literature, the performance of U.S. Steel is contrasted with some of its smaller competitors. Although it may tend to lag behind some of them, there may nonetheless be an inverse relationship between size and delay in the whole industry. Moreover, differences among firms in the profitability of the innovation are not taken into account. See Fortune (March 1936), Stocking [20], and Stigler [19].
Given that the hypothesis holds (i.e., given that equation (5) holds and that \( a_2 \) is negative), it follows that the probability that a firm \( X \) times as large as another will be quicker to introduce the innovation is

\[
\Pr \left\{ d_k < d_x \right\} = \Pr \left\{ a_1 + a_2 \ln S_k + a_3 P_k + \varepsilon_k < a_1 + a_2 \ln S_x \right\} \\
+ a_3 P_x + \varepsilon_x \right\} \\
= \Pr \left\{ a_2 \ln S_k + \varepsilon_k < a_2 \ln S_x + \varepsilon_x \right\} \\
= \Pr \left\{ a_2 \left[ \ln S_x + \ln X \right] + \varepsilon_k < a_2 \ln S_x + \varepsilon_x \right\} \\
= \Pr \left\{ a_2 \ln X + \varepsilon_k < \varepsilon_x \right\} \\
= \Pr \left\{ \varepsilon_x - \varepsilon_k > a_2 \ln X \right\} \\
= 1 - U \left[ a_2 \ln x / \sigma_x \sqrt{2} \right] \\
= U \left[ - a_2 \ln X / \sigma_x \sqrt{2} \right]
\]

But this probability is conditional on the hypothesis holding. What we want is the unconditional probability (the probability given that we do not know whether or not the hypothesis holds). This probability equals

\[
P_H U \left[ - a_2 \ln X / \sigma_x \sqrt{2} \right] + \left[ 1 - P_H \right] V,
\]

where \( P_H \) is the probability that the hypothesis will hold and \( V \) is the probability that the prediction will be correct, given that it does not hold. Since the second term must be positive, the probability exceeds

\[
P_H U \left[ - a_2 \ln X / \sigma_x \sqrt{2} \right].
\]

The figures in the text are derived by inserting the proportion of cases where \( a_2 \) is negative in Table 1 for \( P_H \), the average value of \( a_2 / \sigma_x \) in Table 1 in cases where data for both \( S_i \) and \( P_j \) are available as an estimate of \( a_2 / \sigma_x \) in this case, and \( 4 \) (or \( 2 \)) for \( X \). The roughness of these figures should be obvious.
One can go through the same sort of procedure as that carried out (in connection with $S_j$) in note 21. But the added difficulty here is that $P_j$ is expressed in units of a surrogate. Consequently, the result is that, if two firms differ by $\Delta$ in terms of a particular surrogate and if $a_j$ is based on this surrogate, the probability of the prediction being correct is

\[ U \left\{ \frac{-a_j \Delta}{\sigma_b \sqrt{2}} \right\}. \]

Unfortunately, although this might be useful for a particular case, there is little that one can say in general because $a_j\Delta$ will differ in interpretation from one case to another.

See Mansfield [13], note 26 for further discussion of the effects of this factor.

Ruth Mack [12], p. 289, quotes one machinery manufacturer as saying that the early purchasers tend to be "either the 'wide-awake progressive companies' which were generally...in a strong financial position or the 'do or die' group which decided to play a turn of the wheel and sink or swim thereby." According to him, the first group was the more important.

Specifically, $g_j$ is 100 times the ratio of the sales (or freight ton-miles in the case of railroads) in the terminal year to that in the initial year:

<table>
<thead>
<tr>
<th>Innovation</th>
<th>Terminal Year</th>
<th>Initial Year</th>
<th>Period for $\Pi_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous mining machine</td>
<td>1958</td>
<td>1948</td>
<td>1946-48</td>
</tr>
<tr>
<td>Shuttle car and trackless mobile</td>
<td>1948</td>
<td>1938</td>
<td>1938-40</td>
</tr>
<tr>
<td>loader</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>1956</td>
<td>1956</td>
<td>1939-41</td>
</tr>
<tr>
<td>Continuous strip mill</td>
<td>1936</td>
<td>1926</td>
<td>1926-28</td>
</tr>
<tr>
<td>Bottle filler and pallet loader</td>
<td>1958</td>
<td>1950</td>
<td>1950-52</td>
</tr>
<tr>
<td>Railroad innovations</td>
<td>1949</td>
<td>1925</td>
<td>1925-27</td>
</tr>
</tbody>
</table>

The period to which $\Pi_j$ pertains, is provided in the final column. The data came from Moody's and the Statistics of Railways. Of course, only firms for which such data could be obtained were included in Table 2. In the case of the brewing innovations, only a fairly small sample remains and some of the results may not be statistically significant. But in the other industries the coverage remains virtually complete. (Three innovations were excluded altogether because of lack of data.)

Of course, if equation (6) -- or equations (7) or (8) -- is the true model there is a bias in the estimates in Table 1 because of specification error. But it does not appear that this bias is very important. The results in Tables 2 - 4 differ in no important way (from our viewpoint) from those in Table 1.
Note three things. (1) All possible combinations of the independent variables were used without affecting the results. (i.e., each of the
independent variables in equation (6) was omitted; each pair was omitted; etc.) In none of the results did $\Pi_j$ or $g_j$ persistently have the hypothesized
effect. (2) Taking only those cases when the innovation replaced quite durable old equipment and hence when the expected effect of $g_j$ should be most pronounced, one still finds it has no consistent effect. (3) The data are rather crude. The growth data pertain only to very long periods and the rate of growth within these periods was probably far from smooth. The data on $\Pi_j$ pertain only to a three-year period.

Although the hypothesis that $A_j$ is an important determinant of $d_i$ seems dubious on a number of counts, the results of the agricultural studies indicated that it was worthy of investigation. Data were obtained from Moody's, Standard and Poor's Register of Executives, Who's Who, etc., regarding the age of the president of each firm when the innovation first was used.

The results in Table 4 were obtained by K. E. Knight in a term paper. Profits plus bond interest divided by total sales was used as the profit rate for the railroads and profits less preferred dividends divided by net worth were used for the coal and steel firms. Somewhat different measures might have been used instead, but the results would almost certainly have been about the same. This part of the paper benefited from discussions with G. von der Linde.

The reason why other combinations of these five variables -- e.g., $\Pi_j$ and $t_j$ -- were not used was simply that we had used up as much computer
time as we could afford when we finished the complete analysis of equation (7). The results based on equation (8) are merely preliminary. They are probably less trustworthy than those in Tables 1 - 3 because only four innovations are included and because other combinations of exogenous variables were not tried.

Firms with aggressive managements often lose their taste for pioneering as those managements grow older or as others take their place; and laggard firms sometimes change their ways because of an injection of new blood and capital. With the passage of time, it becomes increasingly likely that those that were particularly receptive to change in a past era have given up this role to others.
Note that, although the coefficient of correlation between how rapidly a firm introduced one innovation and how rapidly it introduced another is a reasonable measure of the extent to which leadership is concentrated, it has certain obvious disadvantages and it is not the only measure that might be used. E.g., the coefficient of rank correlation is a possible alternative measure and it, too, is used below.

Note too that it is impossible for the linear function in equation (9) to be applicable throughout the entire range, but it may be a useful local approximation.

The results of the analysis of covariance are as follows:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall slope &lt;w&gt;</td>
<td>.0917</td>
<td>1</td>
</tr>
<tr>
<td>Slope of means versus w</td>
<td>.2122</td>
<td>1</td>
</tr>
<tr>
<td>Means about regression of means</td>
<td>.0273</td>
<td>2</td>
</tr>
<tr>
<td>Industry slopes versus w</td>
<td>.1895</td>
<td>3</td>
</tr>
<tr>
<td>Residual</td>
<td>.3577</td>
<td>11</td>
</tr>
</tbody>
</table>

For a description and interpretation of the analysis of covariance, see Kendall [10].

Of course, one important reason why there is a tendency for the same firms to be the leaders is that the large firms tend to be the leaders and the same firms tend to be large during periods of relevant length. However, this is not the whole story. If $e_j$ rather than $d_j$ is used here the results are much the same. Thus, if one corrects for a firm's size, there is still some tendency for the same firms to be the leaders.

Let $d_{1j}$ and $d_{1k}$ be the delays for the $j^{th}$ and $k^{th}$ firms in introducing the first innovation. Let $c_{2j}$ and $d_{2k}$ be their delays in introducing the second innovation. We suppose that in general for the $i^{th}$ firm,

$$d_{2i} = Q_0 + Q_1 d_{1i} + V_i,$$

where $Q_0$ and $Q_1$ are parameters and $V_i$ is a normally distributed random variable with zero expected value. What is the probability that $d_{2j} < d_{2k}$, given that $d_{1j} < d_{1k}$?
Footnote 34 (Continued)

It equals

\[
\Pr \left\{ Q_0 + Q_1 d_{ij} + V_j < Q_0 + Q_1 d_{ik} + V_k \right\}
\]

\[
= \Pr \left\{ Q_1 (d_{ij} - d_{ik}) < V_k - V_j \right\}
\]

\[
= 1 - U \left[ \frac{Q_1 (d_{ij} - d_{ik})}{\sigma_v \sqrt{2}} \right]
\]

If \( d_{1i} \) was normally distributed and \( d_{1k} \) was at the 75th percentile and \( d_{1j} \) was at the 25th percentile, \( d_{1j} - d_{1k} \approx 1.34 \sigma_d \), where \( \sigma_d \) is the standard deviation of the \( d_{1i} \). Thus, the probability equals

\[
1 - U \left[ -1.34 \frac{\sigma_d}{\sigma_v} \right]
\]

But it can easily be shown that \( \frac{\sigma_d}{\sigma_v} = \sqrt{r^2/1-r^2} \). Thus, we have

\[
1 - U \left[ -1.34 \sqrt{r^2/2(1-r^2)} \right]
\]

where \( r \) is the coefficient of correlation between \( d_{1i} \) and \( d_{2i} \).

Inserting .25 for \( r \), we get the result (.59) in the text.

Note too that, because the linear approximation in equation (9) is unlikely to hold very well for extreme values of \( t_{qr} \), the estimates of \( v \) are probably only rough estimates of the values of \( \rho_{qr} \) when \( t_{qr} \) is extremely small. But for values of \( t_{qr} \) of one or more, the results may be quite good.

Finally, although a firm's rate of response to a previous innovation is not very useful in most of these industries as a predictor of its rate of response to a current one, it is as good a predictor as a firm's size when \( P_j \) is not held constant. (Of course, this is not surprising since, as pointed out in note 33, one important reason for \( \rho_{qr} \) being greater than zero is the size effect.) But when \( P_j \) is held constant, a firm's size seems to be an appreciably better predictor than a firm's rate of response to a previous innovation.
If a few firms continually spend much more on research of practically all sorts (relative to their size) than the others, they may tend to be the leaders again and again. Note too that the extent to which the same firms tend to be leaders may be affected by the extent to which the innovations differ with regard to such characteristics as capital requirements, the sorts of firms for which they are most profitable, etc. The less the innovations differ in these respects, the more likely that the same firms will tend to be the leaders.