Technical Change and the Rate of Imitation

Edwin Mansfield


Stable URL: http://links.jstor.org/sici?sici=0012-9682%28196110%2929%3A4%3C741%3ATCATRO%3E2.0.CO%3B2-W

*Econometrica* is currently published by The Econometric Society.
TECHNICAL CHANGE AND THE RATE OF IMITATION*

By Edwin Mansfield

This paper investigates the factors determining how rapidly the use of a new technique spreads from one firm to another. A simple model is presented to help explain differences among innovations in the rate of imitation. Deterministic and stochastic versions of this model are tested against data showing how rapidly firms in four industries came to use twelve important innovations. The empirical results seem quite consistent with both versions of the model.

1. INTRODUCTION

Once an innovation is introduced by one firm, how soon do others in the industry come to use it? What factors determine how rapidly they follow? The importance of these questions has long been recognized. For example, Mason, Clark, Dunlop, and others pointed out in 1941 the need for studies to investigate how rapidly "... an innovation spreads from enterprise to enterprise."¹ Eighteen years later, the need is perhaps even more obvious than it was then.

This paper summarizes some theoretical and empirical findings regarding the rate at which firms follow an innovator. A simple model is presented to help explain differences among innovations in the rate of imitation. The model is then tested against data showing how rapidly firms in four quite different industries came to use twelve innovations. The results—though by no means free of difficulties—seem quite encouraging, and should help to fill a significant gap in the literature concerning technical change.

The plan of the paper is as follows. Section 2 lists the twelve innovations and shows how rapidly firms followed the innovator in each case. Sections 3–5 present and test a deterministic model constructed to explain the observed differences among these rates of imitation. A stochastic version is discussed.

* The work on which this report is based was supported initially by research funds of the Graduate School of Industrial Administration and then by a contract with the National Science Foundation. It is part of a broader study of research and technological change that I am conducting under this contract. The paper has benefitted from discussions with a number of my colleagues at Carnegie Institute of Technology, particularly A. Meltzer, F. Modigliani, J. Muth, and H. Wein (now at Michigan State University). A version of it was presented at the December, 1959 meeting of the Econometric Society. My thanks also go to the many people in industry who provided data and granted me interviews. Without their cooperation, the work could not have been carried out.

¹ Committee on Price Determination [7, p. 169].
in Section 6. Some of the study's limitations are pointed out in Section 7 and its conclusions are summarized in Section 8.

2. RATES OF IMITATION

This section describes how rapidly the use of twelve innovations spread from enterprise to enterprise in four industries—bituminous coal, iron and steel, brewing, and railroads. The innovations are 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 line for tin plate (in iron and steel); the pallet-loading machine, tin container, and high-speed bottle filler (in brewing); and the diesel locomotive, centralized traffic control, and car retarders (in railroads).

These innovations were chosen because of their outstanding importance and because it seemed likely that adequate data could be obtained for them. Excluding the tin container, all were types of heavy equipment permitting a substantial reduction in costs. The most recent of these innovations occurred after World War II; the earliest was introduced before 1900. In practically all cases, the bulk of the development work was carried out by equipment manufacturers and patents did not impede the imitation process.\(^2\)

Figure 1 shows the percentage of major firms that had introduced each of these innovations at various points in time. To avoid misunderstanding, note three things regarding the data in Figure 1. (1) Because of difficulties in obtaining information concerning smaller firms and because in some cases they could not use the innovation in any event, only firms exceeding a certain size (given in the Appendix) are included.\(^3\) (2) The percentage of firms


\(^3\) For the innovations in the steel industry, we imposed the particular size limits cited in the Appendix because, according to interviews, it seemed very unlikely that firms smaller than this would have been able to use them. For the innovations in the coal and brewing industries, there were no adequate published data concerning the dates when particular firms first introduced them, and we had to get the information directly from the firms. It seemed likely that "non-response" would be a considerable problem among the smaller firms. This consideration, as well as the fact that the smallest firms often could not use them, led us to impose the rather arbitrary lower limits on size shown in the Appendix. For the railroad innovations, we took firms
FIGURE 1.—Growth in the Percentage of Major Firms that Introduced Twelve Innovations, Bituminous Coal, Iron and Steel, Brewing, and Railroad Industries, 1890–1958.

a. By-product coke oven (CO), diesel locomotive (DL), tin container (TC), and shuttle car (SC).

b. Car retarder (CR), trackless mobile loader (ML), continuous mining machine (CM), and pallet-loading machine (PL).

c. Continuous wide-strip mill (SM), centralized traffic control (CTC), continuous annealing (CA), and highspeed bottle filler (BF).

Source: See the Appendix.

Note: For all but the by-product coke oven and tin container, the percentages given are for every two years from the year of initial introduction. Zero is arbitrarily set at two years prior to the initial introduction in these charts (but not in the analysis). The length of the interval for the by-product coke oven is about six years and for the tin container, it is six months. The innovations are grouped into the three sets shown above to make it easier to distinguish between the various growth curves.
having introduced an innovation, regardless of the scale on which they did so, is given. The possible objections to this are largely removed by the fact that these innovations had to be introduced on a fairly large scale. By using them at all, firms made a relatively heavy financial commitment.4 (3) In a given industry, most of the firms included in the case of one innovation are also included for the others. Thus the data for each of the innovations are quite comparable in this regard.

Two conclusions regarding the rate of imitation emerge from Figure 1. First, the diffusion of a new technique is generally a rather slow process. Measuring from the date of the first successful commercial application,5 it took 20 years or more for all the major firms to install centralized traffic control, car retarders, by-product coke ovens, and continuous annealing. Only in the case of the pallet-loading machine, tin container, and continuous mining machine did it take 10 years or less for all the major firms to install them.

Second, the rate of imitation varies widely. Although it sometimes took decades for firms to install a new technique, in other cases they followed the innovator very quickly. For example, it took about 15 years for half of the major pig-iron producers to use the by-product coke oven, but only about 3 years for half of the major coal producers to use the continuous mining machine. The number of years elapsing before half the firms had introduced an innovation varied from 0.9 to 15, the average being 7.8.

3. A DETERMINISTIC MODEL

Why were these firms so slow to install some innovations and so quick to install others? What factors seem to govern the rate of imitation? In this section, we construct a simple deterministic model to explain the results in

---

4 The only alternative would be to take the date when a firm first used the innovation to produce some specified percentage of its output. In almost every case, such data were not published and it would have been extremely difficult, if not impossible, to obtain them from the firms. To install a strip mill, by-product coke ovens, continuous annealing, or car retarders, a firm had to invest many millions of dollars. Even for shuttle cars, trackless mobile loaders, canning equipment, and pallet loaders, the investment (although less than $100,000 usually) was by no means trivial in these industries.

5 Note that we measure how quickly other firms followed the one that first successfully applied the technique. Others may have tried roughly similar things before but failed. By a successful application, we mean one where the equipment was used commercially for years, not installed and quickly withdrawn.
Figure 1. In Sections 4 and 5, we test this model and see the effects of introducing some additional variables into it.

The following notation is used. Let $n_{tj}$ be the total number of firms on which the results in Figure 1 for the $j$th innovation in the $i$th industry are based ($j = 1, 2, 3; i = 1, 2, 3, 4$). Let $m_{tj}(t)$ be the number of these firms having introduced this innovation at time $t$, $\pi_{tj}$ be the profitability of installing this innovation relative to that of alternative investments, and $Su_{ij}$ be the investment required to install this innovation as a per cent of the average total assets of these firms. More precise definitions of $\pi_{tj}$ and $Su_{ij}$ are provided in Section 4. Let $\lambda_{tj}(t)$ be the proportion of "hold-outs" (firms not using this innovation) at time $t$ that introduce it by time $t + 1$, i.e.,

$$\lambda_{tj}(t) = \frac{m_{tj}(t+1) - m_{tj}(t)}{n_{tj} - m_{tj}(t)}.$$  

(1)

Our basic hypothesis can be stated quite simply. We assume that the proportion of "hold-outs" at time $t$ that introduce the innovation by time $t + 1$ is a function of (1) the proportion of firms that already introduced it by time $t$, (2) the profitability of installing it, (3) the size of the investment required to install it, and (4) other unspecified variables. Allowing the function to vary among industries, we have

$$\lambda_{tj}(t) = f_t\left(\frac{m_{tj}(t)}{n_{tj}}, \pi_{tj}, Su_{ij}, \ldots\right).$$  

(2)

In the following few paragraphs, we take up the presumed effects of variation in $m_{tj}(t)/n_{tj}$, $\pi_{tj}$, and $Su_{ij}$ on $\lambda_{tj}(t)$ and the reasons for interindustry differences in the function.

First, one would expect that increases in the proportion of firms already using an innovation would increase $\lambda_{tj}(t)$. As more information and experience accumulate, it becomes less risky to begin using it. Competitive pressures

---

6 That is, the total number of firms for which we have data. See Table I for the $n_{tj}$.

7 The profitability of installing each of these innovations was viewed at first with considerable uncertainty. For example, there was great uncertainty about maintenance costs for diesel locomotives, "down-time" for continuous mining machines, the safety of centralized traffic control, and the useful life of by-product coke ovens. The perceived risks seldom disappeared after only a few firms had introduced them; in some cases according to interviews, it took many years. This helps to account for the fact (noted above) that the imitation process generally went on rather slowly.

This also suggests the potential importance of two variables not recognized explicitly in (2): (1) the extent of the initial uncertainty concerning the profitability of an innovation (depending in part on the extent of prior field testing by manufacturers), and (2) the rate at which this uncertainty declined. For some types of innovations, a few installations and a relatively short period of use can cut the risks to little or nothing. For other types, installations under many sorts of conditions and a long period of use
mount and "bandwagon" effects occur. Where the profitability of using the innovation is very difficult to estimate, the mere fact that a large proportion of its competitors have introduced it may prompt a firm to consider it more favorably. Both interviews with executives in the four industries and the data in Figure 1 indicate that this is the case.\(^8\)

Second, the profitability of installing the innovation would also be expected to have an important influence on \(\lambda_{ij}(t)\). The more profitable this investment is relative to others that are available, the greater is the chance that a firm’s estimate of the profitability will be high enough to compensate for whatever risks are involved and that it will seem worthwhile to install the new technique rather than to wait. As the difference between the profitability of this investment and that of others widens, firms tend to note and respond to the difference more quickly. Both the interviews and the few other studies regarding the rate of imitation suggest that this is so.\(^9\)

Third, for equally profitable innovations, \(\lambda_{ij}(t)\) should tend to be smaller for those requiring relatively large investments. One would expect this on

\(\text{(to be sure of useful life, maintenance costs, etc.) are required. See Section 7 for further discussion.}

Note that all these innovations are new processes or (in the case of the tin container) a packaging innovation. This model would not be applicable to some innovations, like an entirely new product, where, as more firms produce it, it becomes less profitable for others to do so. There is no evidence of significant decreases of this sort in these cases. (E.g., as more firms introduced diesel locomotives, this did not make it less profitable for others to do so.) Moreover, the model presumes that \(n_{ij}\) is appreciably greater than unity (which is the case here). Finally, the data support the hypothesis in the text (see note 8).

\(^8\) Beginning with the date when \(m_{ij}(t) = 1\), we computed \(\lambda_{ij}(t)\) and \(m_{ij}(t)/n_{ij}\), using the intervals in the footnote to Figure 1 as time units and stopping when \(m_{ij}(t) = n_{ij}\). Then we calculated the correlation between \(\lambda_{ij}(t)\) and \(m_{ij}(t)/n_{ij}\). The correlation coefficients were .77 (continuous mining machine), .85 (by-product coke oven), .49 (tin container), .65 (centralized traffic control), .85 (continuous wide strip mill), .66 (diesel locomotive), .52 (car retarders), .55 (bottle fillers), .96 (pallet loaders), .81 (continuous annealing), .46 (trackless mobile loader), and .94 (shuttle car). Using a one-tailed test (which is appropriate here), all coefficients but those for the continuous mining machine, bottle filler, and trackless mobile loader are significant (0.05 level). All are positive.

Of course, others have stated similar hypotheses before. E.g., Schumpeter [20] noted that "accumulating experience and vanishing obstacles" smooth the way for imitators, and Coleman, et al. [6] noted a "snow-ball" effect. For what it is worth, almost all the executives we interviewed considered this effect to be present. The number of installations of the innovation or the number of firms using it might have been used rather than the proportion of firms, but the latter seems to work quite well.

\(^9\) See Griliches [9]. Another recent study of the rate of imitation is found in Yance [25]. Both papers focus on only one innovation (hybrid corn and the diesel locomotive). Of course, the interfirm variation in profitability, as well as the average, could influence \(\lambda_{ij}(t)\).
the grounds that firms tend to be more cautious before committing themselves to such projects and that they often have more difficulty in financing them. According to the interviews, this factor is often important.

Finally for equally profitable innovations requiring the same investment, $\lambda_{ij}(t)$ is likely to vary among industries. It might be higher in one industry than in another because firms in the former industry have less aversion to risk, because markets are more keenly competitive, because the attitude of the labor force toward innovation is more favorable, or because the industry is healthier financially. Casual observation suggests that such interindustry differences may have a significant effect on $\lambda_{ij}(t)$.

Returning to equation (2), we act as if the number of firms having introduced an innovation can vary continuously rather than assume only integer values, and we assume that $\lambda_{ij}(t)$ can be approximated adequately within the relevant range by a Taylor's expansion that drops third and higher order terms. Assuming that the coefficient of $(m_{ij}(t)/n_{ij})^2$ in this expansion is zero (and the data in Figure 1 generally support this),

\[
\lambda_{ij}(t) = a_{i1} + a_{i2} \frac{m_{ij}(t)}{n_{ij}} + a_{i3} \pi_{ij} + a_{i4} S_{ij} + a_{i5} \pi_{ij} \frac{m_{ij}(t)}{n_{ij}} + a_{i6} S_{ij} \frac{m_{ij}(t)}{n_{ij}} + a_{i7} \pi_{ij} S_{ij} + a_{i8} \pi_{ij}^2 + a_{i9} S_{ij}^2 + \ldots,
\]

where additional terms contain the unspecified variables in (2). Thus,

\[
m_{ij}(t+1) - m_{ij}(t) = [n_{ij} - m_{ij}(t)]\left[a_{i1} + a_{i2} \frac{m_{ij}(t)}{n_{ij}} + \ldots + a_{i9} S_{ij}^2 + \ldots\right].
\]

Assuming that time is measured in fairly small units, we can use as an approximation the corresponding differential equation\footnote{To test this assumption for these innovations, we used $(m_{ij}(t)/n_{ij})^2$ as an additional independent variable in the regression described in note 8 and used the customary analysis of variance to determine whether this resulted in a significant increase in the explained variation. For all but continuous annealing, car retarders, and the diesel locomotive, it does not (and in these cases, the increase is often barely significant). Hence, in most cases, there is no evidence that this coefficient is nonzero.}

\[
\frac{dm_{ij}(t)}{dt} = [n_{ij} - m_{ij}(t)]\left[Q_{ij} + \phi_{ij} \frac{m_{ij}(t)}{n_{ij}}\right],
\]

the solution of which is

\[
m_{ij}(t) = \frac{n_{ij} [e^{t \pi_{ij} + (Q_{ij} + \phi_{ij})t} - (Q_{ij} + \phi_{ij}) \xi]}{1 - e^{[1 + (Q_{ij} + \phi_{ij})t]}},
\]

\footnote{This is like some approximations commonly used in capital theory. E.g., one often replaces equations like $x(t+1) - x(t) = r(x(t)$ with $\dot{x}(t) = rx(t).$}
where \( l_{ij} \) is a constant of integration, \( Q_{ij} \) is the sum of all terms in (3) not containing \( m_{ij}(t)/n_{ij} \), and

\[
\phi_{ij} = a_{i2} + a_{i5} \pi_{ij} + a_{i6} S_{ij} + \ldots. \tag{7}
\]

Of course, \( \phi_{ij} \) is the coefficient of \( m_{ij}(t)/n_{ij} \) in (3).

To get any further, we must impose additional constraints on the way \( m_{ij}(t) \) can vary over time. One simple condition we can impose is that, as we go backward in time, the number of firms having introduced the innovation must tend to zero,\(^{12} \) i.e.,

\[
\lim_{t \to -\infty} m_{ij}(t) = 0. \tag{8}
\]

Using this condition, it follows that

\[
m_{ij}(t) = n_{ij}[1 + e^{-(l_{ij} + \phi_{ij} t)}]^{-1}. \tag{9}
\]

Thus, the growth over time in the number of firms having introduced an innovation should conform to a logistic function, an S-shaped growth curve frequently encountered in biology and the social sciences.\(^{13} \)

If (9) is correct, it can be shown that the rate of imitation is governed by only one parameter—\( \phi_{ij} \).\(^{14} \) Assuming that the sum of the unspecified terms in (7) is uncorrelated with \( \pi_{ij} \) and \( S_{ij} \) and that it can be treated as a random error term,

\(^{12} \) Of course, other conditions could be imposed in addition or instead. For example in Section 6, we take as given the date when a particular number of firms had installed an innovation and we force \( m_{ij}(t) \) to equal that number at that date. But (8) is all we need for present purposes. Note that it implies that \( Q_{ij} \) is zero. The data described in note 8 are consistent with this, but even if \( Q_{ij} \) were nonzero but small (and it certainly could not be large), (9) should be a reasonably good approximation. Note too that, if the model holds, \( \phi_{ij} > 0 \).

\(^{13} \) Note that things are simplified here by the fact that all firms we consider eventually introduced these innovations. (A few went out of business first, but not because of the appearance of the innovation. See the Appendix.) Had this not been the case, it would have been necessary either to provide a mechanism explaining the proportion that did not do so or to take it as given. Of course, by taking only the larger firms, we made sure that all could use these innovations. The smaller firms (that were not potential users) are omitted. For the high-speed bottle filler, we make the reasonable assumption that all the major firms will ultimately introduce it. Note that the argument here is different from that generally used in biology to arrive at the logistic function and that it explicitly includes variables affecting its shape.

\(^{14} \) It seems reasonable to take, as a measure of the rate of imitation, the time span between (1) the date when 20 per cent (say) of the firms had introduced an innovation, and (2) the date when 80 per cent (say) had done so. According to the model, this time span equals \( 2.77 \phi_{ij}^{-1} \) and is therefore independent of \( l_{ij} \). If, rather than 20 and 80, we take \( P_1 \) and \( P_2 \), it can be shown that the time span equals \( \phi_{ij}^{-1} \ln \left[ (1 - P_1)P_2/P_1(1 - P_2) \right] \).
TABLE I
PARAMETERS, ESTIMATES, AND ROOT-MEAN-SQUARE ERRORS: DETERMINISTIC AND STOCHASTIC MODELS

<table>
<thead>
<tr>
<th>Innovation</th>
<th>Parameters*</th>
<th>Estimates and Root Mean Square Errorsb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n_i$</td>
<td>$\pi_i$</td>
</tr>
<tr>
<td>Diesel locomotive</td>
<td>25</td>
<td>1.59</td>
</tr>
<tr>
<td>Centralized traffic control</td>
<td>24</td>
<td>1.48</td>
</tr>
<tr>
<td>Car retarders</td>
<td>25</td>
<td>1.25</td>
</tr>
<tr>
<td>Continuous wide strip mill</td>
<td>12</td>
<td>1.87</td>
</tr>
<tr>
<td>By-product coke oven</td>
<td>12</td>
<td>1.47</td>
</tr>
<tr>
<td>Continuous annealing</td>
<td>9</td>
<td>1.25</td>
</tr>
<tr>
<td>Shuttle car</td>
<td>15</td>
<td>1.74</td>
</tr>
<tr>
<td>Trackless mobile loader</td>
<td>15</td>
<td>1.65</td>
</tr>
<tr>
<td>Continuous mining machine</td>
<td>17</td>
<td>2.00</td>
</tr>
<tr>
<td>Tin container</td>
<td>22</td>
<td>5.07</td>
</tr>
<tr>
<td>High-speed bottle filler</td>
<td>16</td>
<td>1.20</td>
</tr>
<tr>
<td>Pallet-loading machine</td>
<td>19</td>
<td>1.67</td>
</tr>
</tbody>
</table>

* For definitions of these parameters, see Sections 3 ($n_i$), 4 ($\pi_i$ and $S_i$), and 5 ($d_i$, $\sigma_i$, $\mathbf{t}_i$, and $\delta_i$).

b For definitions of these measures, see Sections 4 ($\hat{\varphi}_i$, $\hat{\varphi}_i$, $\hat{r}_i$, and error (det.)) and 6 ($\hat{\varphi}_i$ and error (st.)).

Source: See the Appendix and notes 6, 15, 16, 17, 20, 22, 25, 29, 32, and 34.
(10) \[ \phi_{ij} = b_i + a_{i5} \pi_{ij} + a_{i6} S_{ij} + z_{ij}, \]

where \( b_i \) equals \( a_{i2} \) plus the expected value of this sum and \( z_{ij} \) is a random variable with zero expected value. Hence, the expected value of \( \phi_{ij} \) in a particular industry is a linear function of \( \pi_{ij} \) and \( S_{ij} \).

To sum up, the model leads to the following two predictions. First, the number of firms having introduced an innovation, if plotted against time, should approximate a logistic function. Second, the rate of imitation in a particular industry should be higher for more profitable innovations and innovations requiring relatively small investments. More precisely, \( \phi_{ij} \), a measure of the rate of imitation, should be linearly related to \( \pi_{ij} \) and \( S_{ij} \).

4. TESTS OF THE MODEL

We test this model in two steps: (1) by estimating \( \phi_{ij} \) and \( l_{ij} \) and determining how well (9) fits the data, and (2) by seeing whether the expected value of \( \phi_{ij} \) seems to be a linear function of \( \pi_{ij} \) and \( S_{ij} \). The results of these tests suggest that the model can explain the results in Figure 1 quite well.

To carry out the first step, note that, if the model is correct, it follows from (9) that

(11) \[ \ln \left[ \frac{m_{ij}(t)}{n_{ij} - m_{ij}(t)} \right] = l_{ij} + \phi_{ij} t. \]

Measuring time in years (from 1900), treating this as a regression equation, and using least squares (after properly weighting the observations [4]), we derive estimates of \( l_{ij} \) and \( \phi_{ij} \) (Table I).\(^{15}\) To see how well (9) can represent the data, we insert these estimates into (9) and compare the calculated increase over time in the number of firms having introduced each innovation with the actual increase.

Judging merely by a visual comparison, the calculated growth curves generally provide reasonably good approximations to the actual ones. However, the "fit" is not uniformly good. For the railroad innovations, there is evidence of serial correlation among the residuals, and the approximations are less satisfactory than for the other innovations. Table I contains two rough measures of "goodness-of-fit": the root-mean-square deviation of

\(^{15}\) When \( \ln \left[ m_{ij}(t)/n_{ij} - m_{ij}(t) \right] \) was infinite, the observation was omitted. The observations were one year apart for the continuous wide strip mill, continuous mining machine, and centralized traffic control, six years apart for the by-product coke oven, one month apart for the tin containers, and two years apart for the others. For the high-speed bottle filler, we used all the data available thus far, but the imitation process is not yet complete (cf. Figure 1).
actual from the computed number of firms having introduced the innovation and the coefficient of correlation between \( \ln [m_{ij}(t)/n_{ij} - m_{ij}(t)] \) and \( t \). They seem to bear out the general impression that (9) represents the data for most of the innovations quite well.

To carry out the second step, we assume that \( a_{i5} \) and \( a_{i6} \) do not vary among industries. Thus (10) becomes

\[
\phi_{ij} = b_i + a_{5} \pi_{ij} + a_{6} S_{ij} + z_{ij},
\]

and, if we assume that errors in the estimates of \( \phi_{ij} \) are uncorrelated with \( \pi_{ij} \) and \( S_{ij} \), we have

\[
\hat{\phi}_{ij} = b_i + a_{5} \pi_{ij} + a_{6} S_{ij} + z'_{ij},
\]

where \( \hat{\phi}_{ij} \) is the estimate of \( \phi_{ij} \) derived above. Assuming that \( z'_{ij} \) is distributed normally with constant variance, standard tests used for linear hypotheses can be applied to determine whether \( a_{5} \) and \( a_{6} \) are nonzero.

Before discussing the results of these tests, we must describe how \( \pi_{ij} \) and \( S_{ij} \) are measured. We obtained from as many of the firms as possible estimates of the pay-out period for the initial installation of the innovation and the pay-out period required from investments. They were derived primarily from correspondence, but published materials were used where possible.

16 For each innovation, the difference between the computed and actual values of \( m_{ij}(t) \) was obtained for \( t = t_{ij}^*, t_{ij}^* + 1, \ldots, t_{ij}^{**} \), where time is measured in the units described in note 15, \( t_{ij}^* \) is the first date when \( m_{ij}(t) = 1 \), and \( t_{ij}^{**} \) is the first date when \( m_{ij}(t) = n_{ij} \). Then the root-mean-square of these differences was obtained. Note that this is a measure of absolute, not relative, error, and when comparing the results for different innovations, take account of differences in \( n_{ij} \).

17 This coefficient is labeled \( \tau_{ij} \) in Table I. Note that it is based on weighted observations (the weights being those suggested by Berkson [4]).

18 Of course, the logistic function is not the only one that might represent the data fairly adequately. (E.g., the normal cumulative distribution function would probably do as well.) And, since the actual curve has to be roughly S-shaped, it is not surprising that it fits reasonably well. There seemed to be little point in attempting any formal “goodness-of-fit” tests here. The number of firms considered in each case is quite small, and the tests would not be very powerful. All that we conclude from Table I is that (9) provides a reasonably adequate description of the data and hence that \( \phi_{ij} \) is generally reliable as a measure of the rate of imitation.

19 Because of the small number of observations, we are forced to make this assumption. If data were available for more innovations, it would not be necessary. Of course, if interindustry differences in these coefficients are statistically independent of \( \pi_{ij} \) and \( S_{ij} \), \( a_{5} \) and \( a_{6} \) can be regarded as averages and there is no real trouble.

20 A letter was sent to each firm asking (1) how long it took for the initial investment in the innovation to pay for itself, and (2) what the required pay-out period was during the relevant period. (For a definition of the pay-out period, see Swalm [22]. Both the realized and required pay-out periods are before taxes.) Replies were received from 50 per cent of the rail roads and coal producers and 20 per cent of the steel companies and breweries. The estimates for the railroad innovations are probably most accurate since the firms referred us to fairly reliable published studies (see the Appendix).
Then the average pay-out period required by the firms (during the relevant period) to justify investments divided by the average pay-out period for the innovation was used as a measure of $\pi_{it}$. To measure $S_{it}$, we used the average initial investment in the innovation as a percentage of the average total assets of the firms (during the relevant period). Table I contains the results.

Using these rather crude data, we obtained least squares estimates of the parameters in (13) and tested whether they were zero. The resulting equation is

$$
\hat{\phi}_{it} = \begin{pmatrix} -.29 \\ -.57 \\ -.52 \\ -.59 \end{pmatrix} + .530\pi_{it} - .027S_{it} \\
(.015) \\
(.014) \\
(\tau = .997),
$$

for the coal innovations are probably quite good too. The estimates for the steel innovations and the tin container are probably less accurate because they occurred so long ago and because fewer firms provided data. Despite the fact that the averages are probably more accurate than the individual figures and that the results were checked with other sources in interviews (see the Appendix), the estimates of the $\pi_{it}$ are fairly rough.

For relatively long-lived investments, the reciprocal of the pay-out period is a fairly adequate approximation to the rate of return. See Gordon [8] and Swalm [22]. Hence this measure of $\pi_{it}$ is approximately equal to the average rate of return derived (ex post) from the innovation divided by the average rate of return firms required (ex ante) to justify investments. Of course, it would be more appropriate for firms to recognize that introducing it next year, the year after, etc., are the alternatives to introducing it now and to make an analysis like Terborgh's [23]. But as he points out, most firms seem to make these decisions on the basis of the "rate of return" or "pay-out period"; and thus it seems reasonable to use such measures in a study (like this) where we try to explain behavior, not prescribe it.

Even so, this measure is only an approximation. (1) The rate of return from the investment in an innovation is measured ex post, not ex ante. (2) The average rate of return that could have been realized by the "hold-outs" at a particular point in time probably differed from the average rate of return actually realized by all firms. Our estimates are based implicitly on the factor prices and age of old equipment (where replacement occurred) that prevailed when the innovation was installed, and these (and other) factors, as well as the innovation itself, do not remain unchanged. According to the interviews, the average return that could have been realized probably varied over time about an average that was highly correlated with our estimate, and hence the latter provides a fair indication of the level. But the parameters describing the temporal variation about this level are among the "other" variables in (2). See Section 7 and note 40.

For the steel, coal, and brewing industries, we obtained the total assets of as many of these firms as possible (during the relevant period) from Moody's. For the railroads, the 1936 reproduction costs in Klein [13] were used for the firms. Estimates of the approximate investment required to install each innovation were obtained primarily from the interviews (described in the Appendix), and each was divided by the average total assets of the relevant firms. Since the investment could vary considerably, the results are only approximate.
where the top figure in the brackets pertains to the brewing industry, the next to coal, the following to steel, and the bottom figure pertains to the railroads. The coefficients of \( \pi_{ij} \) and \( S_{ij} \) have the expected signs (indicating that increases in \( \pi_{ij} \) and decreases in \( S_{ij} \) increase the rate of imitation), and both differ significantly from zero. As expected, there are significant interindustry differences, the rate of imitation (for given \( \pi_{ij} \) and \( S_{ij} \)) being particularly high in brewing. These differences seem to be broadly consistent with the hypothesis often advanced that the rate of imitation is higher in more competitive industries, but there are too few data to warrant any real conclusion on this score.\(^{23}\)

![Figure 2](image.png)

**Figure 2.**—Plot of Actual \( \hat{\phi}_{ij} \) Against That Computed from Equation (14), Twelve Innovations.

*Source:* Table I and equation (14).
*Note:* The difference between an actual and computed value of \( \hat{\phi}_{ij} \) is equal to the vertical difference between the point and the 45° line.

\(^{23}\) For a more extended statement of this hypothesis, see Robinson [19]. We encounter here the usual difficulty in measuring the “extent of competition.” But most economists would undoubtedly agree that brewing and coal are more competitive than steel and railroads (and the average value of \( b_i \) is larger in the former industries). When six of my colleagues were asked to rank them by “competitiveness,” they put brewing first, coal second, iron and steel third, and railroads fourth. If these ranks are correlated with the estimates of \( b_i \) in (14), the rank correlation coefficient is positive (.80), but not statistically significant. (The .05 probability level is used throughout this paper.) There are too few industries to allow a reasonably powerful test of this hypothesis even if our procedures were refined somewhat, but for most rankings that seem sensible, the data are consistent with the hypothesis (the correlation is positive).
The scatter diagram in Figure 2 shows that (14) represents the data surprisingly well. When corrected for the relatively few degrees of freedom, the correlation coefficient is .997. Of course, one point (the tin container) strongly affects the results. But if that point is omitted, the interindustry differences remain much the same, the coefficients of $\pi_l$ and $S_l$ keep the same signs (the latter becoming non-significant), and the correlation coefficient, corrected for degrees of freedom, is still .97.$^{24}$

Hence, the model seems to fit the data quite well. In general, the growth in the number of users of an innovation can be approximated by a logistic curve. And there is definite evidence that more profitable innovations and ones requiring smaller investments had higher rates of imitation, the particular relationship being strikingly similar to the one we predicted. Though it is no more than a simple first approximation, the model can represent the empirical results in Figure 1 surprisingly well.

5. ADDITIONAL FACTORS

Of course, other factors may also have been important, and their inclusion in the model may permit a significantly better explanation of the differences among rates of imitation. Moreover, it may show that the apparent effect of the previously mentioned variables on the rate of imitation is partly due to their influence. We turn now to a discussion of the influence of four other factors on $\lambda_l(t)$.

First, one might expect $\lambda_l(t)$ to be smaller if the innovation replaces equipment that is very durable. In such cases, there is a good chance that a firm's old equipment still has a relatively long useful life, according to past estimates. Although rational economic calculation might indicate that replacement would be profitable, firms may be reluctant to scrap equipment that is not fully written off and that will continue to serve for many years.

$^{24}$ Five points should be noted. (1) The tin container is a somewhat different type of innovation from the others and the profitability data for it are probably least reliable (see note 20). Thus, besides its being an extreme point in Figure 2, there are other possible grounds for excluding it, and it is reassuring to note that its exclusion has so little effect on the results. (2) Because of the way it is constructed, the model can only be expected to work if $\pi_l$ and $S_l$ remain within certain bounds. If $\pi_l$ is close to one or $S_l$ is very large or both, it is likely to perform poorly. (3) One-tailed tests are used to determine the significance of the coefficients of $\pi_l$ and $S_l$. (4) It is noteworthy that $\hat{\delta}_l$, not a very obvious measure of the rate of imitation, should be linearly related to $\pi_l$ and $S_l$, as the model predicts. Using the result in note 14, the number of years elapsing between when 20 per cent and when 80 per cent introduced the innovation can easily be used instead as the dependent variable in (14). The resulting relationship is nonlinear. (5) This model may also help to explain the empirical results Griliches [9] obtained in his excellent study of the regional acceptance of hybrid corn, but for which he provided no formal theoretical underpinning.
If so, \( d_{ij} \)—the number of years that typically elapsed before the old equipment was replaced (before the innovation appeared)—may be one of the excluded variables in (2), and hence \( \phi_{ij} \) may be a linear function of \( \pi_{ij} \), \( S_{ij} \), and \( d_{ij} \).\(^{25}\) If \( d_{ij} \) is included in (14), we find that

\[
(15) \quad \phi_{ij} = \begin{pmatrix}
-0.28 \\
-0.56 \\
-0.51 \\
-0.57
\end{pmatrix}
+ 0.528 \pi_{ij} - 0.020 S_{ij} - 0.0017 d_{ij} \quad (r = 0.997).
\]

Though there is some apparent tendency for the rate of imitation to be lower in cases where very durable equipment had to be replaced, it is not statistically significant. (For the values of \( d_{ij} \) and the other factors discussed below, see Table I.)

Second, one might expect \( \lambda_{ij}(t) \) to be higher if firms are expanding at a rapid rate. If they are convinced of its superiority, the innovation will be introduced in the new plants built to accommodate the growth in the market. If there is little or no expansion, its introduction must often wait until the firms decide to replace existing equipment. If \( g_{ij} \)—the annual rate of growth of industry sales during the period—is included in (14),

\[
(16) \quad \phi_{ij} = \begin{pmatrix}
-0.32 \\
-0.56 \\
-0.53 \\
-0.58
\end{pmatrix}
+ 0.484 \pi_{ij} - 0.025 S_{ij} + 0.042 g_{ij} \quad (r = 0.998).
\]

Thus, there is some apparent tendency for the rate of imitation to be higher where output was expanding at a very rapid rate, but it is not statistically significant.\(^{26}\)

\(^{25}\) If we carry \( d_{ij} \) from (2) on and make the same assumptions we did with regard to \( \pi_{ij} \) and \( S_{ij} \), it follows that \( \phi_{ij} \) should be a linear function of \( \pi_{ij}, S_{ij}, \) and \( d_{ij} \). This, of course, is also true for the other variables discussed in this section. Estimates of \( d_{ij} \) were obtained primarily from the interviews. If an innovation was largely a supplement or addition to old plant (like centralized traffic control and car retarders), if it served a different purpose than the old equipment (like canning equipment vs. bottling equipment), and if it displaced only labor (like pallet loaders), we let \( d_{ij} \) equal zero. In such cases, there appeared to be no equipment whose durability could influence the decision significantly. Note that the age distribution and number of units of old equipment are also important in determining how long it takes before the first one “wears out.” It would also have been preferable to have included the profitability of replacing old equipment of various ages rather than just the average figure used. But the available data would not permit this. For further discussion, see Mansfield [15].

\(^{26}\) Of course, the effect of \( g_{ij} \) may depend on whether old equipment must be replaced, how durable it is, the difference between the profitability of replacement and of installing the innovation in new plant, the extent of excess capacity at the beginning of the period, the size of a plant relative to the size of the market, etc. (For some relevant
Third, $\lambda_{ij}(t)$ may have increased over time. This hypothesis has been advanced by economists on numerous occasions. Presumably, the reasons for such a trend would be the evolution of better communication channels, more sophisticated techniques to evaluate machine replacement, and more favorable attitudes toward technological change. If $t_{ij}$—the year (less 1900), when the innovation was first introduced—is included in (14),

\begin{equation}
\hat{\phi}_{ij} = \begin{pmatrix} -0.37 \\ -0.64 \\ -0.55 \\ -0.63 \end{pmatrix} + 0.535\pi_{ij} - 0.027S_{ij} + 0.0014t_{ij} \quad (r = 0.997)
\end{equation}

There is some apparent tendency for the rate of imitation to increase over time, but it is not statistically significant.

Finally, one might suppose that $\lambda_{ij}(t)$ would be influenced by the phase of the business cycle during which the innovation was first introduced. Let $\delta_{ij}$ equal one if the innovation is introduced in the expansion phase and zero if it is introduced in the contraction phase. Including $\delta_{ij}$ in (14), we have

\begin{equation}
\hat{\phi}_{ij} = \begin{pmatrix} -0.26 \\ -0.55 \\ -0.48 \\ -0.57 \end{pmatrix} + 0.530\pi_{ij} - 0.033S_{ij} - 0.022\delta_{ij} \quad (r = 0.997)
\end{equation}

and the effect of $\delta_{ij}$ turns out to be non-significant.

Thus, our results regarding these additional factors are largely inconclusive. Though each might be expected to have some effect, their inclusion in the analysis does not lead to a significantly better explanation of the observed differences in $\hat{\phi}_{ij}$. Their apparent effects are in the expected direction, but data for more innovations will be required before one can be reasonably sure of the persistence and magnitude of their influence. There is no evidence

discussion, see Scitovsky [21].) The effect of this factor, like $d_{ij}$, reflects the possible unwillingness of firms to scrap existing equipment. It would have been preferable to have used figures on the profitability of using the innovation in new plant as well to replace equipment of various ages, but data were not available. Note too that $g_{ij}$ may be affected by the appearance of the innovation. For a description of the data on $g_{ij}$, see the Appendix. If $S_{ij}$ is omitted from (15)–(18), the coefficients of $d_{ij}$, $g_{ij}$, etc. remain non-significant.

27 E.g., Mack [14, p. 295] and Jerome [11, p. XXV].
28 I.e., $t_{ij} = t^*_ij - 1900$. See note 32.
29 We used the National Bureau's reference dates (given in Moore [18]), to determine whether $t^*_ij$ was a year of contraction or recovery. The residuals in Figure 2 do not seem to be affected by whether or not an innovation was being accepted sometime during the depression of the nineteen-thirties.
that the effects of the previously considered variables on the rate of imitation are due to the operation of these factors. When these other factors are included, the coefficients of \( \pi_{ij} \) and \( S_{ij} \) and the interindustry differences remain relatively unchanged (though the coefficient of \( S_{ij} \) becomes non-significant.)

6. A STOCHASTIC VERSION OF THE MODEL

In this section, we present and test a somewhat more sophisticated, stochastic version of the model. For the \( j \)th innovation in the \( i \)th industry, let \( P_{ij}(t,k) \) be the probability that any one of the “hold-outs” at time \( t \) will introduce it by time \( t + k \), and assume (for small \( k \)) that

\[
P_{ij}(t,k) = \theta_{ij} \frac{m_{ij}(t)}{n_{ij}} k,
\]

\[
\theta_{ij} = b_i + a_5 \pi_{ij} + a_6 S_{ij} + z_{ij},
\]

where the coefficients in (20) are analogous to those in (12). To see how closely this resembles the deterministic model, note that the expected increase in the number of users between time \( t \) and time \( t + 1 \),

\[
[n_{ij} - m_{ij}(t)] \theta_{ij} \frac{m_{ij}(t)}{n_{ij}},
\]

is almost identical with the expression given before for the actual increase in the number of users (see (4) and (5)). The only (apparent) difference is that the terms corresponding to \( Q_{ij} \) are assumed to be zero from the start here, whereas in Section 3 this followed from (8).30

Whereas our task in the deterministic model was to determine how the actual number of users grows over time, our problem here is to see how the expected number grows. To do so, we first obtain an expression for \( P_{ij}(t) \)—the probability that at time \( t \) there are exactly \( r \) firms in the \( i \)th industry that have not yet introduced the \( j \)th innovation. From (19),

\[
P_{ij}(t + k) = P_{ij}(t) \left[ 1 - r \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r)k \right] + P_{ij}^{r+1}(t)(r + 1) \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k + o(k),
\]

where \( o(k) \) represents terms that, if divided by \( k \), tend to zero as \( k \) vanishes.31

30 To simplify things, we also assume from the beginning that the coefficients of \( \pi_{ij} \) and \( S_{ij} \) do not vary among industries (Cf. (20)). In the deterministic model we introduced this assumption later in the analysis.

31 To derive this expression, note that \( P_{ij}^x(t+k) = \sum_{x=r}^{\infty} P_{ij}^x(t) D_{ij}^{x-r}(t) \), where \( D_{ij}^{x-r}(t) \) is the probability that \( (x-r) \) firms begin using the innovation between time \( t \) and time \( t + k \), given that \( (n_{ij} - x) \) firms are using it at time \( t \). From (19),

\[
D_{ij}^0(t) = [1 - (\theta_{ij}/n_{ij}) (n_{ij} - r)k]^r.
\]
And if we subtract $P_{ij}(t)$ from both sides, divide by $k$, and let $k$ tend to zero, the following differential-difference equation results:

\[
\begin{cases}
- \frac{\theta_{ij}}{n_{ij}} r(n_{ij} - r)P_{ij}(t) + \frac{\theta_{ij}}{n_{ij}} (r + 1)(n_{ij} - r - 1)P_{ij}^{r+1}(t), & \text{if } r < n_{ij} - 1, \\
- \frac{\theta_{ij}}{n_{ij}} r P_{ij}^{r}(t), & \text{if } r = n_{ij} - 1,
\end{cases}
\]

with the initial conditions that $P_{ij}(t_0^*)$ equals unity for $r = n_{ij} - 1$ and zero otherwise, and where $\dot{P}_{ij}(t)$ is the time derivative of $P_{ij}(t)$ and $t_0^*$ is the date (taken as given) when the innovation was first introduced. From (21), it follows [3] that $M(t_0^* + V)$—the expected number of firms using the innovation at time $t_0^* + V$—equals

\[
M(t_0^* + V) = \sum_{r=1}^{n_{ij} - 2r} \frac{(n_{ij} - 1)!}{(n_{ij} - r - 1)! (r - 1)!} \left[ (n_{ij} - 2r)^2 - \frac{\theta_{ij}}{n_{ij}} V + 2 \right.
\]

\[
- (n_{ij} - 2r) \sum_{u=r}^{n_{ij} - r - 1} e^{-r(n_{ij} - r)(\theta_{ij}/n_{ij})V},
\]

where $r$ runs up to $(n_{ij} - 1)/2$ for $n_{ij}$ odd and up to $n_{ij}/2$ for $n_{ij}$ even.33

Thus, the stochastic version of the model leads to the following two propositions. First, the expected number of firms having introduced an innovation at any date subsequent to $t_0^*$ should be given by (22). Second, $\theta_{ij}$—the parameter that, for given $n_{ij}$, determines the expected rate of imita-

\[
D_{ij}(t) = (r + 1) \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k \left[ 1 - \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k \right]^{r-1};
\]

etc. Thus,

\[
\dot{P}_{ij}^{r}(t + k) = \dot{P}_{ij}^{r}(t) \left[ 1 - \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r)k \right]^r + \dot{P}_{ij}^{r+1}(t) (n + 1) \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k \left[ 1 - \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k \right]^{r-1} + o(k)
\]

\[
= \dot{P}_{ij}^{r}(t) \left[ 1 - r \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r)k \right] + \dot{P}_{ij}^{r+1}(t) \left[ (r + 1) \frac{\theta_{ij}}{n_{ij}} (n_{ij} - r - 1)k \right] + o(k).
\]

Note two things, (1) The assumption that the probability in (19) is the same for all “hold-outs” is obviously only a convenient simplification. Without it, the analysis becomes extremely difficult. (2) We assume that the decisions of the “hold-outs” between time $t$ and time $t + k$ are independent. Hence, the probability that one “hold-out” will introduce it between time $t$ and time $t + k$ is $[n_{ij} - m_{ij}(t)] \theta_{ij} m_{ij}(t) k/n_{ij} + o(k)$, and the probability that none will introduce it is $1 - [n_{ij} - m_{ij}(t)] \theta_{ij} m_{ij}(t) k/n_{ij} + o(k)$.

32 For all but the tin container, we use December 31 of the year during which the innovation was first introduced by one of these firms (or if we know the month when it was first introduced, we use that December 31 which was closest in time to it). For the tin container, we use February 1, 1935. See Table I for the year in which each $t_0^*$ falls.

33 When $n_{ij}$ is even, some adjustment has to be made to this expression. See Bailey [3] for the adjustment and for a much more extensive discussion of the argument.
RATION OF IMITATION

should be a linear function of $\pi_{ij}$ and $S_{ij}$, the intercept of the function differing among industries.

To see how well the first proposition seems to hold, we estimate $\theta_{ij}$, insert it, $n_{ij}$, and $t_{ij}$ into (22), and compare the growth over time in the computed number of users with the actual growth curve. Table I contains a rough measure of "goodness-of-fit": the root-mean-square deviation of the actual from the computed number of users.\textsuperscript{34} Although the agreement between the actual and computed growth curves is sometimes quite good, Table I shows that (22) is never so accurate as (9) in representing the data. Perhaps this is due in part to differences in the way parameters are estimated as well as differences in the model.\textsuperscript{35}

To test the second proposition, we assume that errors in the estimates of $\theta_{ij}$ are uncorrelated with $n_{ij}$ and $S_{ij}$ and hence that

\begin{equation}
\hat{\theta}_{ij} = b_t + a_5 \pi_{ij} + a_6 S_{ij} + z_{ij}^{\prime \prime},
\end{equation}

where $\hat{\theta}_{ij}$ is our estimate of $\theta_{ij}$ and $z_{ij}^{\prime \prime}$ is a random error term. Then, proceeding as we did in Sections 4 and 5, we get much the same sort of results as those obtained there. (1) We estimate the parameters in (23), find that these estimates are almost identical with those given in (14) for the analogous parameters,\textsuperscript{36} and conclude that the resulting equation represents the data very well ($r = .997$). (2) We introduce the additional variables discussed in Section 5 into (23) and find that their coefficients have the expected signs, but that (except for the coefficient of $t_{ij}$) all are non-significant.

Thus, the model, in either its deterministic or stochastic form, seems to represent the empirical results in Figure 1 quite well. Although (22) does not fit the data as well as (9) in the deterministic version, the purpose and interpretation of these equations are quite different and hence this may not be very surprising.\textsuperscript{37} With regard to the explanation of differences in the

\textsuperscript{34} To estimate $\theta_{ij}$, we used the statistic suggested by Bailey [3, p. 41] to estimate $\theta_{ij}/n_{ij}$ and multiplied the result by $n_{ij}$. This estimate of $\theta_{ij}$ is biased, but we can show that the bias is approximately equal to $\theta_{ij}/(n_{ij} - 1)$, which is usually quite small. These estimates, together with the tables in Mansfield and Hensley [16], were used to calculate the growth over time in the expected number of users. The root-mean-square errors are computed in the way described in note 16.

\textsuperscript{35} Two parameters ($b_t$ and $\phi_t$) are fitted to the data in the deterministic model whereas only ($\theta_{ij}$) is fitted here. However, this is only part of the explanation (see note 37). Note too that the differences between the computed and actual curves are by no means random for a number of these innovations. The computed curves sometimes lie below the actual ones.

\textsuperscript{36} The only appreciable difference is that the coefficient of $\pi_{ij}$ is .580 rather than .530.

\textsuperscript{37} In the deterministic model, we fit (9) to the data in such a way as to minimize differences between the actual and computed curves. In this model, we use the data to estimate the curve of averages that would result if the process were repeated indefinite-
rate of imitation, (23), like (13) in the deterministic model, provides an excellent fit. The stochastic version of the model seems somewhat more reasonable to me, but it is difficult to tell at this point whether it is a significant improvement over the simpler, deterministic version.

Before concluding, one final point should be noted regarding the dispersion about the expected value in (22). This dispersion is often relatively large and hence a prediction of the rate of imitation based on (22) would be fairly crude, even if the model were correct and \( \theta_{ij} \) were known in advance. To illustrate this, consider an innovation where \( n_{ij} = 12, t^*_i = 1924, \) and \( \theta_{ij} = .42. \) Table II shows the expected number of firms having introduced it at various points in time and the relatively large standard deviation of the distribution about this expectation. Of course, despite this variation, the model could set useful upper and lower bounds on how long the entire process would take.38

<table>
<thead>
<tr>
<th>Date (Dec. 31)</th>
<th>Expected Number</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1924</td>
<td>1.0</td>
<td>—</td>
</tr>
<tr>
<td>1925</td>
<td>1.5</td>
<td>.9</td>
</tr>
<tr>
<td>1926</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>1927</td>
<td>2.7</td>
<td>1.6</td>
</tr>
<tr>
<td>1928</td>
<td>3.6</td>
<td>2.3</td>
</tr>
<tr>
<td>1929</td>
<td>4.5</td>
<td>2.7</td>
</tr>
<tr>
<td>1930</td>
<td>5.5</td>
<td>2.9</td>
</tr>
<tr>
<td>1931</td>
<td>6.4</td>
<td>3.1</td>
</tr>
<tr>
<td>1932</td>
<td>7.2</td>
<td>3.2</td>
</tr>
<tr>
<td>1933</td>
<td>8.0</td>
<td>3.1</td>
</tr>
<tr>
<td>1934</td>
<td>8.8</td>
<td>2.9</td>
</tr>
<tr>
<td>1935</td>
<td>9.4</td>
<td>2.6</td>
</tr>
<tr>
<td>1936</td>
<td>10.0</td>
<td>2.4</td>
</tr>
<tr>
<td>1937</td>
<td>10.5</td>
<td>2.1</td>
</tr>
<tr>
<td>1938</td>
<td>10.8</td>
<td>1.8</td>
</tr>
<tr>
<td>1939</td>
<td>11.1</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Source: Tables in Mansfield and Hensley [16]. Small errors due to interpolation are present in both the expected values and standard deviations.

\[38\] For further discussion of the distribution about these expected values, see Mansfield and Hensley [16] and the literature cited there.
7. LIMITATIONS

In evaluating these results, the limitations of the data, methods, and scope of the investigation must be taken into account. For example, although its scope far exceeds that of the few studies previously conducted, it nonetheless is limited to only four industries and to only a few innovations in each one. Before concluding, it seems worthwhile to point out some of these limitations in more detail.

First, there is the matter of scope. Although the four industries included here vary with respect to market structure, size of firm, type of customer, etc., they can hardly be viewed as a cross-section of American industry. For example, none is a relatively young industry with a rapidly changing technology. In addition, the innovations included here are all important, they generally required fairly large investments, and the imitation process was not impeded by patents. To check and extend these results, similar studies should be carried out for other industries and other types of innovations. Moreover, international comparisons might be attempted.

Second, the data are not always as precise as one would like. For example, much of the data regarding the profitability of installing an innovation had to be obtained from questionnaires and interviews with firms. Although these estimates were checked against others obtained from suppliers of the equipment, trade journals, etc., they are probably fairly rough. Similarly, the estimates of the pay-out period required for investments and the durability of old equipment are only rough. Because of these errors of measurement, the regression coefficients in (14)–(18) are probably biased somewhat toward zero.39

Third, the data measure the rate of imitation among large firms only. As noted above, firms not exceeding a certain size (specified in the Appendix) were excluded. Moreover, the rate at which firms imitated an innovator, not the rate at which a new technique displaced an old one or the rate at which investment in a new technique mounted, is considered here. Although these are related topics also being studied, they were not taken up in this paper.

Fourth, various factors other than those considered here undoubtedly exerted some influence on the rate of imitation. Variation over time in the profitability of introducing an innovation (due to improvements, the business etc.)

39 It can easily be shown that measurement errors, if random, have this effect. Another limitation of the analysis is that the model can only be expected to hold for relatively profitable innovations. Certainly, it will not hold in cases where $\pi_{ij} < 1$ and it may do poorly if $\pi_{ij}$ is not appreciable greater than unity. But this limitation is not very serious because if $\pi_{ij}$ does not exceed unity, the innovation almost certainly will not become generally accepted, and if $\pi_{ij}$ is not much greater than unity, the innovation is probably not very important.
cycle, etc.) is one potentially important factor that is omitted.40 Others are the sales and promotional efforts made by the producers of the new equipment and the extent of the risks firms believed they assumed at first by introducing an innovation.41 I could find no satisfactory way to measure them, but perhaps further research will obviate this difficulty and disclose the effects of these and other such factors on the rate of imitation. I suspect that the results underestimate their influence and that data for more innovations would show that the unexplained variation (due to their effects) is greater than is indicated in Figure 2.

8. CONCLUSION

In their discussions of technological change, economists have often cited the need for more research regarding the rate of imitation. Because an innovation will not have its full economic impact until the imitation process is well under way, the rate of imitation is of considerable importance. It is unfortunate that so little attention has been devoted to it and that consequently we know so little about the mechanism governing how rapidly firms come to use a new technique.

My purpose in this paper has been to present and test a simple model designed to help explain differences among process innovations in the rate of imitation. This model is built largely around one hypothesis—that the probability that a firm will introduce a new technique is an increasing func-

---

40 One cannot tell with any accuracy how, and to what extent, the profitability varied in each of these cases. But, to illustrate the factors at work, note a few broad changes that occurred in the case of three of the innovations. (1) The appearance of improved diesel locomotives in the Thirties and Forties increased the profitability of introducing them and hastened their acceptance. In addition, first costs declined, the relative prices of coal and oil changed, and the composition of the "hold-outs" varied. (2) The demand for toluene during World War I undoubtedly accelerated the imitation process in the case of the by-product coke oven. (3) The Depression reduced the profitability of introducing centralized traffic control, whereas the wartime traffic boom increased its profitability. For further discussion, see note 21. Of course, the relative profitability may have varied less than the absolute profitability.

41 In each of these cases, there was considerable promotional effort by the producers of the new equipment and considerable help given to the firms that used it. E.g., Semet-Solvay and Koppers helped to finance by-product coke plants and provided personnel, General Motors helped to train diesel operators, kept maintenance men at hand, etc. For some discussion of the perceived risks, see note 7.

Another factor that could be very important for some innovations is the presence of patents. Still another is the development of a marked improvement in the innovation. E.g., an innovation may be applicable at first to only a few firms and a significant improvement is required before others use it. In a case like this, the model is clearly not applicable. See Mansfield [15]. In each of these cases, improvements occurred, but they were of a more gradual and less significant nature.
tion of the proportion of firms already using it and the profitability of doing so, but a decreasing function of the size of the investment required. When confronted with data for twelve innovations, this model seems to stand up surprisingly well. As expected, the rate of imitation tended to be faster for innovations that were more profitable and that required relatively small investments. An equation of the form predicted by the model can explain practically all of the variation among the rates of imitation.

As expected, there were also interindustry differences in the rate of imitation. We have too few industries really to test the hypothesis often advanced that the rate of imitation is faster in more competitive industries, but the differences seem to be generally in that direction. When several other factors are included in the model, the empirical results are largely inconclusive. There was some apparent tendency for the rate of imitation to be higher when the innovation did not replace very durable equipment, when the firms’ output was growing rapidly, and when the innovation was introduced in the more recent past. But it was almost always statistically non-significant.

In view of the relatively small number of innovations and the uneven quality of the data, these results are quite tentative. But even so, they should be useful to economists concerned with the dynamics of firm behavior and the process of economic growth. To understand how economic growth is generated, we must know more about the way innovations occur and how they become generally accepted. Further attempts should be made to check and extend the results and to pursue other approaches to the particular area considered here. We need to know much more about this and other aspects of technological change.

Carnegie Institute of Technology and Yale University

APPENDIX

Basic Data

Iron and steel industry: 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 with over 200,000 tons of pig iron capacity in 1901 were included; and for continuous annealing of tin plate, the nine major producers of tin plate in 1935 were included. In the case of the strip mill and coke oven, a few of these firms merged or went out of business before installing them, and there was no choice but to exclude them.

The date when each firm first installed a continuous wide strip mill was taken from the Association of Iron and Steel Engineers [2]. Similar data for the coke oven were obtained from various editions of the Directory of Iron and Steel Works of the American

42 This is the stochastic version of the model. In the deterministic version, we use the proportion of “hold-outs” that introduce it rather than this probability.
Iron and Steel Institute and issues of the *Iron Trade Review* and *Iron Age*. The date when each firm installed continuous annealing lines was obtained from correspondence with the firms.

The size of the investment required and the durability of replaced equipment came from interviews. (The estimates of the pay-out periods were also checked there. Cf. note 20.) The interviews (each about 2 hours long) were with major officials of three steel firms, the president and research manager of a firm that builds strip mills and continuous annealing lines, officials of a firm that builds coke ovens, and representatives of a relevant engineering association and of a trade journal. The data on growth of output were annual industry growth rates and were for sheets during 1925–37 (strip mill), pig iron during 1900–25 (coke oven), and tin plate during 1939–56 (continuous annealing). They were taken from the *Bituminous Coal Annual* of the Bituminous Coal Institute, Association of Iron and Steel Engineers [2], and *Annual Statistical Reports* of the American Iron and Steel Institute.

**Railroad industry:** For centralized traffic control, all Class I line-haul roads with over 5 billion freight ton-miles in 1925 were included. For the diesel locomotives and car retarders, essentially the same firms were included. (The Norfolk and Western, a rather special case, was replaced by the New Haven and Lehigh Valley in the case of the diesel locomotive. Some important switching roads were substituted in the case of car retarders, an innovation in switching techniques.) An entire system is treated here as one firm.

The date when a firm first installed centralized traffic control was usually derived from a questionnaire filled out by the firm. For those that did not reply, estimates by K. Healy [10] were used. The date when each firm first installed diesel locomotives was determined from various editions of the Interstate Commerce Commission's *Statistics of Railways*. The date when each firm first installed car retarders was taken from various issues of *Railway Age*.

For centralized traffic control and car retarders, some of the pay-out periods were estimates published in the *Signal Section, Proceedings of Association of American Railroads*, and the rest were obtained from questionnaires filled out by the firms. All estimates of the pay-out period for the diesel locomotive and the pay-out period required for investment were obtained from questionnaires. Information regarding the size of the investment required and the durability of old equipment was obtained primarily from interviews with eight officials of six railroads (ranging from president to chief engineer) and three officials of a signal manufacturing firm and a locomotive manufacturing firm. The data on growth of output were annual growth rates for total freight ton-miles during 1925–41 (diesel locomotive) and 1925–54 (centralized traffic control and car retarders.) They were taken from the *Statistics of Railways*.

**Bituminous coal industry:** Practically all firms producing over 4 million tons of coal in 1956 (according to McGraw-Hill’s *Keystone Coal Buyers Manual*) were included. A few firms that did strip mining predominantly were excluded in the case of the continuous mining machine, and a few had to be excluded in the case of the shuttle car and trackless mobile loader because they would not provide the necessary data. The date when each firm first introduced these types of equipment was usually obtained from questionnaires filled out by the firms, but in the case of the continuous mining machine, data for two firms that did not reply were derived from the *Keystone Coal Buyers Manual*. 
Data regarding the size of the investment required and the durability of old equipment were obtained from interviews with two vice-presidents of coal firms, several executives of firms manufacturing the equipment, employees of the Bureau of Mines, and representatives of an independent coal research organization. The data on growth of output were annual growth rates for bituminous coal production during 1934–51 (trackless mobile loader), 1937–51 (shuttle car), and 1947–56 (continuous mining machine). They were taken primarily from the *Bituminous Coal Annual*.

*Brewing industry:* We tried to include all breweries with more than $1 million in assets in 1934 (according to the *Thomas Register*), but several would not provide the necessary data and they could not be obtained elsewhere. The date when a firm first installed each type of equipment was usually taken from a questionnaire that it filled out, but in a few cases it was provided by manufacturers of the equipment or articles in the *Brewers' Journal*. The size of the necessary investment and the durability of equipment that was replaced were determined from interviews with a number of officials in two breweries and sales executives of two can companies. The data on growth of output were annual growth rates for beer production during 1935–37 (tin containers) and 1950–58 (pallet loading machine and high-speed bottle filler). They were taken from the *Brewers' Almanac* and *Business Week* (June 20, 1959). The 1950–58 figures refer only to half of these larger firms.

REFERENCES

[1] **American Mining Congress:** *Coal Mine Modernization Yearbook* (annual).


You have printed the following article:

**Technical Change and the Rate of Imitation**
Edwin Mansfield
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28196110%2929%3A4%3C741%3ATCATRO%3E2.0.CO%3B2-W

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

**Footnotes**

2 A Model for the Location of a Railroad Classification Yard
Edwin Mansfield; Harold H. Wein
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28195804%294%3A3%3C292%3AMFTLO%3E2.0.CO%3B2-O

8 The Diffusion of an Innovation Among Physicians
James Coleman; Elihu Katz; Herbert Menzel
Stable URL:
http://links.jstor.org/sici?sici=0038-0431%28195712%2920%3A4%3C253%3ATDOAIA%3E2.0.CO%3B2-N

9 Hybrid Corn: An Exploration in the Economics of Technological Change
Zvi Griliches
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28195710%2925%3A4%3C501%3AHCAEIT%3E2.0.CO%3B2-A

21 The Payoff Period and the Rate of Profit
Myron J. Gordon
Stable URL:
http://links.jstor.org/sici?sici=0021-9398%28195510%2928%3A4%3C253%3ATPPATR%3E2.0.CO%3B2-A

NOTE: The reference numbering from the original has been maintained in this citation list.
24 Hybrid Corn: An Exploration in the Economics of Technological Change
Zvi Griliches
Stable URL: http://links.jstor.org/sici?sici=0012-9682%28195710%2925%3A4%3C501%3AHCAEIT%3E2.0.CO%3B2-A

26 Economies of Scale, Competition, and European Integration
Tibor Scitovsky
Stable URL: http://links.jstor.org/sici?sici=0002-8282%28195603%2946%3A1%3C71%3AEOSCAE%3E2.0.CO%3B2-8

29 Measuring Recessions
Geoffrey H. Moore
Stable URL: http://links.jstor.org/sici?sici=0162-1459%28195806%2953%3A282%3C259%3AMR%3E2.0.CO%3B2-T

References

6 The Diffusion of an Innovation Among Physicians
James Coleman; Elihu Katz; Herbert Menzel
Stable URL: http://links.jstor.org/sici?sici=0038-0431%28195712%2920%3A4%3C253%3DATDOAJA%3E2.0.CO%3B2-N

8 The Payoff Period and the Rate of Profit
Myron J. Gordon
Stable URL: http://links.jstor.org/sici?sici=0021-9398%28195510%2928%3A4%3C253%3ATPPATR%3E2.0.CO%3B2-A

NOTE: The reference numbering from the original has been maintained in this citation list.
9 Hybrid Corn: An Exploration in the Economics of Technological Change
Zvi Griliches
Stable URL: http://links.jstor.org/sici?sici=0012-9682%28195710%2925%3A4%3C501%3AHCAEIT%3E2.0.CO%3B2-A

17 A Model for the Location of a Railroad Classification Yard
Edwin Mansfield; Harold H. Wein
Stable URL: http://links.jstor.org/sici?sici=0025-1909%28195804%294%3A3%3C292%3AAMFTLO%3E2.0.CO%3B2-O

18 Measuring Recessions
Geoffrey H. Moore
Stable URL: http://links.jstor.org/sici?sici=0162-1459%28195806%2953%3A282%3A3259%3AAMR%3E2.0.CO%3B2-T

21 Economies of Scale, Competition, and European Integration
Tibor Scitovsky
Stable URL: http://links.jstor.org/sici?sici=0002-8282%28195603%2946%3A1%3C71%3AEOSCAE%3E2.0.CO%3B2-8

NOTE: The reference numbering from the original has been maintained in this citation list.