A STUDY IN THE DIFFUSION OF INNOVATION: STEEL RAILS IN AMERICAN RAILROADS

Jan K. Brueckner

#424
A Study in the Diffusion of Innovation:
Steel Rails in American Railroads

by
Jan. K. Brueckner
Department of Economics
University of Illinois, Urbana - Champaign
Urbana, IL 61801

Abstract
This paper proposes and tests a model of the diffusion of innovation in the railroad industry. Specifically, the diffusion of steel rails, which supplanted iron rails in American railroads in the late 19th century, is examined. The model is supported by empirical tests.

May, 1977
A Study in the Diffusion of Innovation:  
Steel Rails in American Railroads

By Jan K. Brueckner*

A number of articles dealing with the speed of adoption of new technology by firms has appeared since 1960. In what has become a standard reference, Mansfield [4] formulated and tested a model of the spread of new technology within an industry. His goal was to relate the rate of adoption of diesel locomotives in American railroads to firm characteristics. The goal of the present study is to propose and test another model of the diffusion of a specific innovation in the railroad industry: steel rails in American railroads between 1867 and 1880.

Mansfield's analysis is somewhat ad hoc in that his choice of explanatory variables is justified only by casual theoretical arguments. One virtue of this paper is that it contains a formal model of the track replacement process which clearly indicates the importance of certain explanatory variables. The success of the empirical test of the model suggests that a more formal approach to modelling diffusion processes can be fruitful and that such an approach could be successfully incorporated in future work in this area.

In 1867 the first Bessemer steel rails were produced in the United States, and steel rapidly supplanted iron in American railroad track. By 1890, 80% of the nation's total track mileage was steel.\(^1\) This study will attempt to relate in cross section the proportion of a railroad's track that

*Assistant Professor of Economics, University of Illinois, Urbana – Champaign.
had been converted to steel by the year 1880 to a number of variables suggested by a theoretical model of the replacement process. The year 1880 was chosen because the U.S. Census of Transportation of 1880 contains excellent data on variables of interest. In addition, the variation in the proportion of track that had been converted by 1880 is large across railroads, which heightens the empirical challenge any theoretical model must face.

Section I reviews some historical facts and develops the theoretical model. Section II contains a discussion of econometric problems. Section III presents the empirical results, and Section IV offers concluding remarks.

I.

The superior durability of steels rails was the reason for their rapid diffusion in American railroads. Contemporary sources contain a wealth of information about the lifespans of steel and iron rails, although the numbers vary considerably depending on the source. The 1900 Railroad Gazette cites 15 million tons of traffic as the lifetime capacity of iron rails, while claiming that steel rails could handle 95 million tons before replacement. One source [1] claims the capacity of steel was 100 to 250 million tons depending on conditions such as track grade and curvature, while yet another source [9] gives 125 million tons as the typical capacity of steel rails. Another source [7] states that the best iron rails would wear, on the average, one-third to one-half as long as steel rails. But he also cites the experience of one railroad where steel rails outlasted 13 sets of iron rails with no perceptible signs of wear, and claims that iron rails lasted only four months on a line between New York and Philadelphia which experienced heavy traffic. The 1900 Railroad Gazette states that a consensus estimate of the life of steel rails is 17 to 24.5 years, while another source [9] states that steel lasted 30 to 40 times as long as average iron rails. While no single verdict emerges from the sources, it is clear that steel rails outlasted iron rails by a wide margin, with the lifespans of
both types of rails depending on the traffic volume per period on the line. These facts are crucial in the theoretical model which is developed below.

It is assumed that the $T$ miles of iron track owned by a given railroad at the beginning of 1867, when steel first became available, had a uniform age distribution. This assumption plays an important role in the model, but there is no way of easily checking it from available sources. If the life of iron rails on the line is $n$ years, where $n$ is an integer, then $T/n$ miles come up for replacement in each of the years numbered zero through $n - 1$, where year zero corresponds to 1867. This is guaranteed by the uniform age distribution assumption. We assume that $n - 1 < 13 = r$, which says that all original track fails before the last year of the period (year $r$ corresponds to 1880). This assumption seems justified in the light of the contemporary accounts cited above. It is also assumed that $n$ is constant across the period.

The next important assumption is that when track fails, the railroad replaces it with steel with probability $p$ and with iron with probability $1 - p$. The determinants of this probability are discussed below. Once the track is replaced with steel we assume it will not fail again until after 1880, an assumption which seems correct given contemporary information. Ignoring new construction for the moment, it will be true that the number of steel miles owned by the railroad in 1880 is the sum of $n$ random variables: $S_0, S_1, \ldots, S_{n-1}$, which represent the number of steel miles in 1880 embodied in track failing initially at time $0, 1, \ldots, n - 1$. We now examine the properties of these random variables.

Consider the ways in which track failing at time zero could become steel by 1880. The track could be switched to steel at time zero
with probability $p$ or it could be replaced by iron at time zero and be switched to steel at its next failure at time $n$ with probability $(1 - p)p$, or it could be replaced with steel at its second failure with probability $(1 - p)^2p$, and so on. The number of opportunities for failure before the end of the period obviously depends on the relationship between $n$ and $r$. For simplicity, suppose $r = 6$ instead of 13 and $n = 2$. Track failing at $t = 0$ which is always replaced by iron fails again at $t = 2$, again at $t = 4$, and again at $t = 6$. It experiences $3 = r/n$ failures before the end of the period. If $r = 5$, only two secondary failures would have occurred before the end of the period. A little reflection shows that the number of secondary failures of track initially failing at $t = 0$ and always replaced by iron is the largest integer less than or equal to $r/n$, or $\text{int}(r/n)$. Hence the probability that track failing at $t = 0$ becomes steel by $t = r$ is

$$p + p(1 - p) + ... + p(1 - p)^{\text{int}(r/n)}. \tag{1}$$

Similarly, track failing initially at $t = 1$ becomes steel by $t = r$ with probability $p + p(1 - p) + ... + p(1 - p)^{\text{int}((r-1)/n)}$.

In general, track failing at $t = i \leq n - 1$ becomes steel by $t = r$ with probability

$$p \sum_{j=0}^{k^r_n(i)} (1 - p)^j = P(i, n), \tag{2}$$

where $k^r_n(i) = \text{int}((r - 1)/n)$. The expression in (2) equals

$$1 - (1 - p)^{k^r_n(i) + 1} \tag{3}.$$
This is as it should be, since the probability of track becoming steel is one minus the probability that iron is always used for replacement, which is 

\[ k^r(i) + 1 \] 

(1 - p) \[ (1 - P) \]

since there are \( k^r(i) + 1 \) failures if iron is always used.

The expected value of the random variable \( S_i \) is then \( P(i, n)(T/n) + (1 - P(i, n)) \cdot 0 = P(i, n)(T/n) \), and the expected number of steel miles from replacement of all original iron track is

\[
\frac{T}{n} \sum_{i=0}^{n-1} P(i, n).
\]  

One might argue that this model is unrealistic because it allows a railroad to replace, for example, the track failing at \( t = 0 \) with steel and the track failing at \( t = 1 \) with iron; it postulates independence of different replacement decisions. While this argument carries some force, it is not implausible that replacement decisions for different track segments were independent; steel may have been used on certain sections of the line while replacement continued with iron on other segments.

It could also be argued that a model with a constant switchover probability \( p \) is unrealistic. The spread of information about steel rails and the decline in the price of steel relative to that of iron might have resulted in a \( p \) which increased over the period 1867-1880. Unfortunately, it is not possible to develop empirical implications from a simple model where \( p \) increases over time. However, the constant-\( p \) model is defensible as an approximation, especially given the shortness of the period 1867-1880.

Many railroads constructed new track over the period 1867-1880, and the previous analysis also applies to this new construction. The expected number of steel miles at \( t = r \) in the \( C(i) \) miles of line constructed at period \( i \) is \( C(i)P(i, n) \), assuming the probability \( p \) was the same for track replacement and new construction. The expected number of steel miles from all new construction is
Letting \( \sum_{i=0}^{r} C(i) = N \), the expected proportion \( Q \) of steel miles in all track at \( t = r \) is derived by dividing the sum of (4) and (5) by \( T + N = A \):

\[
Q = \frac{T}{nA} \sum_{i=0}^{n-1} P(i, n) + \frac{1}{A} \sum_{i=0}^{r} C(i)P(i, n).
\]  

(6)

Substituting (3) for \( P(i, n) \) we have

\[
Q = 1 - \frac{(1-p)T}{nA} \sum_{i=0}^{n-1} (1-p)^r k_r^n(i) - \frac{(1-p)}{A} \sum_{i=0}^{r} C(i)(1-p)k_r^n(i). 
\]  

(7)

Quantities that will vary across firms are \( p \) and \( n \), and we proceed to calculate the effects of changes in these variables on \( Q \). Since \( k_r^n(i) > 0 \), it is immediately apparent that \( \frac{\partial Q}{\partial p} > 0 \); the expected proportion of steel increases as the probability of switch-over from iron to steel increases. Since \( n \) is an integer, we examine the effect of unit increases in \( n \). From (7)

\[
\Delta Q = \frac{(1-p)T}{A} \left[ \frac{1}{n} \sum_{i=0}^{n-1} (1-p)^r k_r^n(i) - \frac{1}{n+1} \sum_{i=0}^{n} (1-p)^r k_r^{n+1}(i) \right] 
\]  

+ \( \frac{(1-p)}{A} \sum_{i=0}^{r} C(i) [(1-p)^r k_r^n(i) - (1-p)^r k_r^{n+1}(i)] \).  

(8)

The multiplicand of \( C(i) \) is \( (1-p)^r [1 - (1-p)^r] \), which is zero or negative since \( k_r^{n+1}(i) \leq k_r^n(i) \) by the definition of \( k_r^n \) and since \( 0 < (1-p) < 1 \). This is true because \( (1-p) \) raised to a non-positive power is greater than or equal to one. Hence the second term in (8) is zero or negative. The first term in (8) has the same sign as
The first term is zero or negative by the above argument. In addition, \( k_n^{n+1}(n) = \text{int}\left(\frac{(r - n)}{(n + 1)}\right) \leq \text{int}\left(\frac{(r - i)}{n}\right) = k_r^n(i) \) for \( i = 0, 1, \ldots, n - 1 \) since \( r - i > r - n \) and \( n < n + 1 \). Hence the second term is zero or negative as well, and we have established \( \Delta Q \leq 0 \); an increase in \( n \), the life of the iron rails, decreases or leaves unchanged the expected proportion of the track that is steel at \( t = r \).

If we let \( T/A \), the proportion of 1880 track built by 1867, equal \( q \), and define \( \bar{C} = N/(r + 1) \) and let \( C(i) = \bar{C} + \Delta(i) \), then (7) may be written

\[
Q = 1 - \frac{(1 - p)q}{n} \sum_{i=0}^{n-1} (1 - p)^r k_r^n(i) - \frac{(1 - p)(1 - q)}{(r + 1)} \sum_{i=0}^{r} (1 - p) k_r^n(i)
\]

\[
- (1 - p) \sum_{i=0}^{r} \frac{\Delta(i)}{A} (1 - p)^r k_r^n(i).
\]

(10)

It may be shown that \( \Delta Q/\Delta q \) is ambiguous in sign. If \( C(i) = \bar{C} = N/(r + 1) \), then \( \Delta(i) = 0 \) for all \( i \) and the last term in (10) is zero. The pattern of
construction, however, tended to be skewed toward the early part of the
1867-1880 period for many railroads. Suppose that C(i) assumes a high constant
value in the early part of the period and drops to a lower constant value in
the latter part of the period. In other words, suppose \( \Delta(i) = \Delta_1 > 0, \)
i = 0, 1, ..., t, and \( \Delta(i) = \Delta_2 < 0, \) i = t + 1, ..., r. Since
\[ \sum_r \Delta(i) = 0, \]
we must have \( (t + 1)\Delta_1 + (r - t)\Delta_2 = 0, \) or \( \Delta_2 = (t + 1)\Delta_1/(t - r) \equiv \)
a\Delta_1. The last term in (10) is then written

\[ \frac{-(1 - p)\Delta_1}{A} \left[ \sum_{i=0}^t \frac{k^r(i)}{(1 - p)^r} + \sum_{i=t+1}^r \frac{k^r(i)}{(1 - p)^r} \right]. \]  
(11)

Suppose \( t + 1 = r - t. \) Then \( a = -1 \) and the term in brackets in (11) is zero or
negative since each term in the second summation is not less than any term in
the first \( (k^r(i) \) is nonincreasing in \( i \) and \( 1 - p < 1 \) and there are equal
numbers of terms in each summation. Letting \( \Delta_1/A = D \) we have \( \partial Q/\partial D \geq 0; \)
holding \( A \) fixed, the more skewed the construction stream is toward the first
half of the period, the greater is \( Q. \) In addition, increasing \( A \) while holding
\( \Delta_1 \) fixed (and holding \( q \) fixed as well) decreases \( Q. \) If \( t + 1 < r - t, \) then
\( a > -1, \) and since the summations have different numbers of terms, we are
unable to say anything about the sign of (11) and hence about \( \partial Q/\partial D. \) In
this case, increasing \( \Delta_1 \) requires decreasing \( \Delta_2 \) by an amount smaller than the
increase in \( \Delta_1, \) but the decrease in \( \Delta_2 \) is felt over a larger number of years
than is the increase in \( \Delta_1. \) When \( a = -1, \) \( \Delta_2 \) falls by an amount equal to
the increase in \( \Delta_1, \) and the changes are spread over equal numbers of years.

So far we have seen that \( Q \) depends on \( n, p, q, \) and \( D. \) We have been
able to evaluate the direction of influence of \( n, p, \) and \( D \) when the construction
flow has the appropriate features. What remains is to indicate how \( p \) depends on firm parameters.

Clearly, firms desired to switch to the new technology because of its superiority; \( p \) is an inverse measure of the inertia restraining firms from adopting the innovation. We assume that for the most part, firms drew on internal funds for replacement expenditures. Since steel rails were more expensive than iron rails\(^4\), we assume that the greater the availability of internal funds for replacement expenditures, the greater the probability of the switchover from iron to steel for a fixed cost differential between steel and iron rails. Similarly, for a given availability of internal funds, the higher the cost differential between steel and iron rails, the lower the probability of switchover from iron to steel. Since mill prices for rails will not vary much if at all in cross-section, we use a measure of differential shipping distance for the two types of rails as a measure of cost difference. To cover the possibility that some railroads borrowed to finance track replacement, a measure of borrowing costs is included as a determinant of \( p \). Since the difference between the financing cost of steel and iron would increase with the cost of borrowing, we assume borrowing cost negatively influences \( p \). Finally, it is assumed that the larger the railroad, the more likely it is to switch to steel from iron. It is likely that managers were better informed in larger railroads and hence more willing to embrace new technology.

In addition, if railroads at first perceived the switch from iron to steel as involving risk, the likelihood that large firms are less risk averse than small ones also would lead to a positive influence of size on \( p \). We measure size by revenue. The theoretical model thus predicts that\(^5\)

\[
Q = Q(n, q, D, F, C, G, R), \tag{12}
\]
where \( F \) represents the availability of internal funds, \( C \) measures the steel-iron cost differential, \( G \) is borrowing cost and \( R \) is revenue. From above, \( Q \) decreases or is constant in \( n \). We have shown \( \partial Q / \partial q \leq 0 \), \( \partial Q / \partial D \geq 0 \) if \( \Delta \) is properly defined, \( \partial Q / \partial F > 0 \), \( \partial Q / \partial C < 0 \), \( \partial Q / \partial G < 0 \), and \( \partial Q / \partial R > 0 \), where the latter results use \( \partial Q / \partial p > 0 \).

II.

The actual proportion \( z \) of 1880 track that had been switched to steel differs, of course, from the expected proportion. Representing the vector of explanatory variables in (12) by \( x \), we have

\[
z = Q(x) + u_x,
\]

where \( u_x \) is a random error term, the difference between \( z \) and \( Q(x) \). The \( x \) subscript indicates that the distribution of \( u \) depends on \( x \), the values of the explanatory variables. To see that this is the case, consider two railroads with \( C(i) \) equal to zero for all \( i \) and \( n = 2 \). The random variable \( z \) can assume the values \( 0, \frac{1}{2}, \) or \( 1 \) for each railroad. Suppose other elements of \( x \) differ between the railroads, so that \( z \) equals \( 0, \frac{1}{2}, \) or \( 1 \) with probability \( w_1, y_1, \) or \( 1 - w_1 - y_1 \), for railroad one and equals \( 0, \frac{1}{2}, \) or \( 1 \) with probability \( w_2, y_2, \) or \( 1 - w_2 - y_2 \) for railroad two. The expected values of \( z \) are \( 1 - w_1 + \frac{1}{2} y_1 \) and \( 1 - w_2 + \frac{1}{2} y_2 \) for railroads one and two respectively.

For railroad one, the error term \( u \) assumes the value \(- (1 - w_1 - \frac{1}{2} y_1)\) with probability \( w_1 \), the value \( \frac{1}{2} - (1 - w_1 - \frac{1}{2} y_1) = -\frac{1}{2} + w_1 + \frac{1}{2} y_1 \) with probability \( y_1 \) and the value \( 1 - (1 - w_1 - \frac{1}{2} y_1) = w_1 + \frac{1}{2} y_1 \) with probability \( 1 - w_1 - y_1 \).

For railroad two, the error term assumes the values \(- (1 - w_2 - \frac{1}{2} y_2)\), \(-\frac{1}{2} + w_2 + \frac{1}{2} y_2 \), and \( w_2 + \frac{1}{2} y_2 \) with probabilities \( w_2, y_2 \) and \( 1 - w_2 - y_2 \).
respectively. This example shows that the distribution of u does indeed depend on the values of the x variables. If n and the C(i) not been the same for both railroads, the set of possible values for z would have been different for each railroad, further reason for the dependence of the distribution of u on the values of the x variables.

The next step is to choose a functional form to approximate Q(x), which is highly non-linear. We represent Q(x) by the logistic function 
\( (1 + e^{a(x)})^{-1} \). Since \( Q(x) = (1 + e^{a(x)})^{-1} + \varepsilon_x \), where \( \varepsilon_x \) is the approximation error which depends on x, (13) becomes

\[
z = (1 + e^{a(x)})^{-1} + s_x, \tag{14}
\]

where \( s_x = u_x + \varepsilon_x \).

Since we do not observe n, the length of life of the rails, we use the fact that \( n = n(B) \), where B is a measure of the intensity of use of the track, and replace n by B in the vector x. B negatively influences n, increases in B tend to increase Q.

Transforming (14), we have

\[
e^{a(x)} = \frac{1 - z}{z} + \frac{(1 + e^{a(x)})s_x}{z} \tag{15}
\]

Taking logs on both sides of (15) and expanding the log of the LHS around \( (1 - z)/z \) in a first-order Taylor series, we get the approximation

\[
\log\left(\frac{1 - z}{z}\right) = a(x) + \frac{(1 + e^{a(x)})s_x}{s_x - e^{a(x)}(1 + e^{a(x)})^{-1}} \tag{16}
\]

\[\equiv a(x) + v_x\]
Since the distribution of the error term $v_x$ depends on $x$ in some complicated way, ordinary least squares estimates of a linear or log-linear $\alpha(x)$ will be biased and inefficient. However, under the circumstances, it appears that the best we can do is to assume that the error term in (16) is identically distributed across the sample and uncorrelated with $x$ and compute ordinary least squares estimates. This procedure is, of course, formally unjustifiable, but the intractable nature of the estimation problem leaves no alternative.

III.

Data are available on the proportion steel of track mileage owned and operated by the railroads in 1880. Many railroads, however, leased track from other lines for their own use. Our measure of the intensity of use of the owned track, which is discussed below, will be faulty if a significant fraction of the operated track is leased from other railroads. Thus, we excluded railroads from the sample for which leased track in 1880 comprised more than 20% of total mileage operated. The great majority of railroads in the sample leased no track at all, while the average proportion leased among those that did lease track was about 10%. Other criteria for inclusion in the sample were that the railroad had to have some track in operation in 1867, that it had to own some steel track in 1880, so that the dependent variable is defined, and that it had to lease none of its owned track to other railroads in 1880. Forty-four railroads qualified for inclusion in the sample under these criteria. A broad range of firm size is represented in the sample.

The variables used in the regressions are discussed next. A contemporary source [1] claims that the best measure of the intensity of use of railroad track is the number of trains it handles per period. The source states that the product of train tonnage and train speed is a good proxy
for the track wear caused by a train. Since freight trains are heavier and slower than passenger trains, the source claims that the above product is roughly constant across trains, meaning that the number of trains handled per period is an appropriate measure of intensity of use. Consider the ratio of train miles per year to miles of track in the system. If each train passes over the entire track system, this ratio equals the number of trains per year passing over each mile of the system, the correct intensity of use measure. The regressions use two variants of a measure based on this ratio. B1 is the average of the ratio in 1874 and 1880: \( \frac{TM74}{M74} + \frac{TM80}{M80} \), where \( M \) is miles of track in the system and \( TM \) is train miles. B2 is computed using MR, miles of road, or main track, excluding sidings and auxiliary track: \( B2 = \frac{TM74}{MR74} + \frac{TM80}{MR80} \). Since road miles are more heavily used than sidings and auxiliary track, it was felt this was a better measure of use intensity. The averaging is meant to capture variations in use intensity over the period. While the regression results using B1 and B2 are very similar, B2 performs slightly better than B1. All reported results are from regressions using B2.

The variable \( q \) is computed as \( \frac{M80 - BLT}{M80} \), where BLT is total construction over the period. Unfortunately, construction data are for miles of road only, excluding sidings and auxiliary track. Thus the computed \( q \) value overstates the true \( q \) value somewhat.

Three versions of the \( \Delta_1 \) variable were computed: they are the amounts by which the average number of miles of construction per year in the subperiods 1867-73, 1867-71, and 1867-70 exceed the 1867-1880 yearly average. The first variant corresponds to the \( a = -1 \) case above, whereas the latter two correspond to the ambiguous cases where \( a > -1 \). It should
be noted that construction expenditures were typically uneven instead of con-
in subperiods as the model requires. Nevertheless, we assume construction expendi-
tures were spread evenly over subperiods of interest in computing the $\Delta_1$ measures
of the skewness of the construction flow. Below, we present results using
$D = \Delta_1/M80$ values based on the first and third versions of $\Delta_1$ (the results with
the second version have insignificant estimates of the $D$ coefficient).

Our measure of the adequacy of internal funds for replacement
expenditures is based on net revenue per mile of track in the system. A
given amount of net revenue is less adequate for replacement expenditure in
a large than in a small system, so we deflate net revenue by miles of track.
The variable is $NR_{M} = (NR74/M74 + NR80/M80)/2$, where $NR$ is net revenue.
The revenue measure of firm size is $R = (R74 + R80)/2$, where $R74$ and $R80$
are gross operating revenues.

The measure of the cost differential between steel and iron rails
is essentially a measure of differential distance between the iron and steel
mills and the railroad's line. The variable was set equal to unity plus
the smallest number of states separating a state in which the railroad
operated from a state producing steel rails minus the smallest number of
states separating a state in which the railroad operated from a state
producing iron rails. For example, if steel rails were produced two states
distant while iron rails were produced in a state adjacent to a state of
operations, the variable assumed the value $1 + (1 - 0) = 2$. The variable
was computed for 1870 and 1880 and an average was taken and denoted LOC.

As a measure of the cost of borrowing, we used the highest interest
rate on those railroad bonds outstanding in either 1867, 1874, or 1880,
denoted BY. The rationale was that raising money in the bond market for
replacement purposes would have required high yield bonds due to collateral problems. This is admittedly a crude measure of the borrowing cost for replacement.\footnote{7}

The model was estimated in linear and log-linear form. The log-linear regressions produced higher t-ratios than the linear regressions, so only the former results are reported. The \textit{a priori} signs for the coefficients are as follows (they are just opposite to those at the end of Section I, given the logistic formulation): $B_2$, negative; $D_1$ (first variant), negative; $D_3$ (third variant), uncertain; $q$, uncertain; LOC, positive; $R$, negative; NRM, negative; BY, positive.

The estimates from fitting equations using $D_1$ are reported in the first three rows of Table 1. In the equation with all the variables included, the coefficients of $B_2$ and $R$ are significantly negative, as expected, at the one-tailed 5\% level, while the coefficient of LOC is significantly positive, as expected, at the one-tailed 10\% level. The coefficient of $q$ is not significantly different from zero, which is consistent with its uncertain \textit{a priori} sign. However, the coefficients of $D_1$, NRM, and BY are also not significantly different from zero, contradicting \textit{a priori} expectations. The equation's $R^2$ seems high given the nature of the dependent variable.

Kennedy and Bancroft [3] have developed a sequential variable exclusion technique for model selection. Optional variables, those which may or may not belong in a model, are sequentially excluded in reverse order of \textit{a priori} importance when the null hypothesis that their coefficients are zero cannot be rejected at a given confidence level, optimally chosen to be .25 by numerical calculations in Kennedy and Bancroft. After each exclusion, the model is reestimated. The sequential t-test uses the variance estimate from fitting the full model. The technique is based on the assumption of normal
error terms, so its application in our model is not strictly correct. Thus, the use of the technique below must be viewed with some reservation.

The variables that definitely belong in our model are $B_2$, $q$, and $D$. The optional variables are $LOC$, $NRM$, $R$, $BY$, with $BY$ being the least important. $BY$ may be deleted according to the Kennedy-Bancroft criterion (the critical points are roughly $\pm 1.15$), and reestimation results in the estimates on the second row of Table 1. The only major change is that the coefficient of $D_1$ is now significantly negative, as it should be, at the 10% one-tailed level. It should be noted that $BY$ may also be deleted according to Sawa and Hiromatsu's minimax regret criterion [6], which sets the critical values at $\pm 1.371$.

If we believe that $NRM$ is less important than $R$, then the exclusion criterion calls for its deletion as well. Reestimation results in row 3 of Table 1; the coefficient estimates for the remaining variables are essentially unchanged from row 2. Since the argument for including $NRM$ among the determinants of the probability $p$ is more persuasive than that for including $R$, deleting $NRM$ is probably undesirable. Since we are able to reject the null hypothesis that $R$'s coefficient is zero (using the variance estimate from the full model to make the test), the exclusion procedure should stop at row 2.

The last three rows of Table 1 give estimates of the model parameters using $D_3$. The $R^2$ is somewhat higher in line 4 than in line 1, $LOC$'s significance level improves to 5%, and $D$ becomes significantly negative at the 5% level. We may not exclude $BY$ by the Kennedy-Bancroft criterion, although Sawa's criterion allows its deletion. From line 5, reestimation
without BY lowers LOC's and R's coefficients significance level to 10% and reduces the $R^2$ somewhat. As before, deleting NRM leaves the estimates of the remaining coefficients essentially unchanged.

IV.

The empirical results in Table I show that the model developed in Section I is not inconsistent with the data. The coefficient of the important intensity of use variable always has the correct sign and is strongly significant in each equation. The $q$ variable belongs in the estimating equation according to model, but its coefficient's lack of significance conforms nicely to the model's ambiguity about its sign. The $D_1$ variable performs reasonably well in the first three equations, especially after the deletion of BY. Among the optional variables, LOC and R are always significant with the correct signs. It is heartening that the only variables which do not perform well, BY and NRM, are optional variables whose inclusion in the model was discretionary.

The fact that the coefficient of the $D_3$ variable is significantly negative was not predicted by the model. However since some measure of the skewness of the construction flow belongs in the estimating equation according to the model, the good performance of $D_3$ is not surprising.

The empirical results in this paper are quite encouraging, suggesting that a complex microeconomic model of the track replacement process generates predictions which conform well to the historical facts. The empirical success of realistic, detailed economic models is always a happy occurrence, increasing our confidence in economic theory. Of course, one could argue that a completely ad hoc model of the diffusion of steel rails could lead to estimating equations similar to those in this paper. For instance,
casual theoretical arguments could be advanced supporting the inclusion of an intensity of use measure such as B2 in any estimating equation. However, it is unlikely that an ad hoc model would lead to a specification which includes the variables q and D, which measure the fraction of 1880 track existing in 1867 and the skewness of the construction flow, respectively. Thus, the empirical model specification in this paper is not likely to be deducible from an ad hoc theoretical model, and we can be confident that the success of the empirical tests vindicates our theoretical model while lending no support to a competing naive hypothesis. In conclusion, we note that the results in this paper argue for a more rigorous approach to modeling diffusion processes than has characterized past work. Rigorous theoretical models may well be successful in explaining the spread of new technology in a variety of industries and time periods.
TABLE 1

<table>
<thead>
<tr>
<th>p</th>
<th>q</th>
<th>r</th>
<th>s</th>
<th>t</th>
<th>u</th>
<th>v</th>
<th>w</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3262</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>0.3268</td>
<td>(1.75)</td>
<td>(1.75)</td>
<td>(1.75)</td>
<td>(1.75)</td>
<td>(1.75)</td>
<td>(1.75)</td>
<td>(1.75)</td>
</tr>
<tr>
<td>0.3542</td>
<td>(1.769)</td>
<td>(1.769)</td>
<td>(1.769)</td>
<td>(1.769)</td>
<td>(1.769)</td>
<td>(1.769)</td>
<td>(1.769)</td>
</tr>
<tr>
<td>0.3080</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
<td>(1.0059)</td>
</tr>
<tr>
<td>0.3080</td>
<td>(1.009)</td>
<td>(1.009)</td>
<td>(1.009)</td>
<td>(1.009)</td>
<td>(1.009)</td>
<td>(1.009)</td>
<td>(1.009)</td>
</tr>
<tr>
<td>0.995</td>
<td>(1.928)</td>
<td>(1.928)</td>
<td>(1.928)</td>
<td>(1.928)</td>
<td>(1.928)</td>
<td>(1.928)</td>
<td>(1.928)</td>
</tr>
<tr>
<td>0.649</td>
<td>(1.984)</td>
<td>(1.984)</td>
<td>(1.984)</td>
<td>(1.984)</td>
<td>(1.984)</td>
<td>(1.984)</td>
<td>(1.984)</td>
</tr>
</tbody>
</table>

Dependent variable is Log ((1 - z)/z).

Observations = 44.

Independent variables in Log form.
Footnotes

1See Fishlow.

2It is possible to formulate the model with a continuous instead of integer-valued \( n \). Then the equivalent of (5) is
\[
\frac{T}{n} \sum_{i=0}^{n} P(i, n) \text{di.}
\]
This model is somewhat less transparent than the discrete one, however.

3Suppose the switchover probability grew at a constant annual percentage rate \( \sigma \): the probability that iron was used in period \( x \) is \( p(1 + \sigma)^x \). The probability \( P(i, n) \) that track initially failing in period \( i \) becomes steel by period \( r \) is
\[
k_r^n(i) + 1
1 - \prod_{j=0}^{r-n} (1 - p(1 + \sigma)^{j+n}).
\]
The expression equivalent to (7), with \( C(i) = 0 \) for simplicity, is
\[
Q = 1 - \frac{1}{n} \sum_{i=0}^{n-1} k_r^n(i) + 1
1 - \prod_{j=0}^{n-1} (1 - p(1 + \sigma)^{j+n}).
\]
It is not possible to show an unambiguous effect for a change in \( n \), as in (8), under this model.

4See Temin.

5It could be argued that the probability of switchover should also depend on the degree of superiority of steel over iron rails, which would be positively related to the intensity of use of the track. Below, \( n \) in (13) will be replaced by \( B \), a measure of intensity of track use. Arguing that \( B \) also positively influences \( p \) does nothing to change the specification of the model or the a priori signs of estimated coefficients.

6The discrete character of \( n \) in the model and the continuous nature of \( B \) mean that \( n(B) \) should be a step function. However, our formulation involves a continuous relationship between \( B \) and the value of the logistic function. The error term \( \varepsilon_x \) is presumed to smooth over this discrepancy.

7An observation on BY was not available for one small railroad, so the BY value for this railroad was set equal to the average of the BY values for two other small firms.
References


10. ________, U.S. Census of Transportation, 1880.

11. ________, Railroad Gazette, 1900, p. 535.