MOUNTAIN-PLAINS CONSORTIUM

MPC 18-365 | D. Tolliver, P. Lu and D. Benson

Railroad Investment in Track Infrastructure

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Railroad Investment in Track Infrastructure

Prepared By

Dr. Denver Tolliver
Director
Denver.Tolliver@ndsu.edu

Dr. Pan Lu
Associate Professor/ Associate Research Fellow
Pan.lu@ndsu.edu

Doug Benson
Associate Research Fellow
Doug.Benson@ndsu.edu
Department of Transportation and Logistics/College of Business
Mountain-Plains Consortium/Upper Great Plains Transportation Institute
North Dakota State University

September 2018

Acknowledgements

This research was made possible with funding supported by the U.S. Department of Transportation though the Mountain-Plains Consortium (MPC) Transportation Center. The authors express their deep gratitude to U.S. DOT and MPC.

Disclaimer

EXECUTIVE SUMMARY

A model of investment in basic track components is estimated from 1985-2008 data for Class I railroads. Network size is measured in miles of road (MOR), while traffic is measured in revenue gross ton-miles (RGTM). In addition to MOR and RGTM, the model includes railroad indicator and time variables. The purpose of the railroad variables is to capture fixed effects (e.g., effects other than traffic and network size) that are specific to particular railroads, but which do not change over time. The time variable, on the other hand, accounts for industry-wide trends and changes that occur during the period. The study shows that when miles of road are held constant (a realistic scenario), a 100% increase in RGTM results in a 50% increase in track investment. However, it is important to consider the interpretative context described in the paper. Several data anomalies were discovered and handled statistically. The parameter estimates vary somewhat with the index used to convert nominal dollars to constant dollars.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. OVERVIEW

Class I railroads in the United States have invested over $67 billion in basic track components:[1] rails, ties, ballast, other track materials (such as tie plates, spikes, bolts, and anchors), and grading. These investments grew by 223% between 1985 and 2008, in nominal dollars, and by 137% in real dollars.[2] As shown in Table 1.1, rails and other track materials comprise the largest investment component (43%), followed by crossties (26%), ballast (16%), and grading (15%), which includes expenditures for the initial construction (and subsequent reconstruction) of the roadbed. Collectively, these investments average $563,000 per route mile. However, the replacement cost of these assets is much greater than their nominal value.

Table 1.1 Class I Railroad Track Investment Per Route Mile (Nominal 2008 Dollars)

Track Component	Investment per Mile	Percent of Total
Rail and Other Track Material	$243,439	43%
Ties	$144,027	26%
Ballast	$90,575	16%
Grading	$84,898	15%
Total: Basic Track Components	**$562,939**	**100%**

Investments in basic track components are necessary to (1) provide safe transportation of passengers and goods, (2) maintain infrastructure in a state of good repair, (3) add capacity, (4) reduce congestion, and (5) increase the overall efficiency of operations. Track investments are important from a regulatory perspective, as railroad revenues must recoup operating expenses and allow companies to earn an adequate return on invested capital.

In many areas of regulation, the Surface Transportation Board (STB) utilizes the Uniform Railroad Costing System (URCS) to provide information about railroad costs. A return of 50% on roadway investment is reflected in the URCS variable cost.[3] This long-standing assumption (that half of road capital investments are fixed) is based on traffic patterns and practices prior to 1955.[4] Since then there have been many changes, including the following:

1. Deregulation has allowed railroads greater decision-making authority and the capability to expeditiously abandon unprofitable lines.
2. Changes in regulatory policies (e.g., the interpretation of the Public Convenience and Necessity clause) have made it easier to propose new rail lines or extensions.
3. Car weights have dramatically increased.

[1] This value is estimated from data reported to the U.S. Surface Transportation Board in Schedule 416 of the R-1 Report.
[2] These percentages are estimated from data reported to the U. S. Surface Transportation Board in Schedule 416 of the R-1 Report. The real percentage increase is computed using the Rail Cost Adjustment Factor
[3] Surface Transportation Board. "Report to Congress Regarding the Uniform Rail Costing System." May 27, 2010.
[4] See: Interstate Commerce Commission (Bureau of Accounts). *Explanation of Rail Cost Finding Procedures and Principles Relating to the Use of Costs*, Statement No. 7-63, Washington, D.C., November 1963. In developing the 50% variability estimate, the ICC used data from 1939 through 1951, including traffic and investment data for the World War II period. The analysis includes "road-to-road comparisons" for 1944, 1946, and 1951. In synthesizing the results of several studies, wartime and prewar traffic densities were adjusted to 1951 levels. Based on these studies, the ICC found "operating expenses to be between 80 and 90 percent variable and plant investment to be upwards of 50 percent variable" [page 86]. In reaching its conclusion, the ICC noted: "The use of a figure of 50 percent variable for road property and 100 percent variable for equipment is approximately equivalent to the use of an overall figure for road and equipment of 60 percent" [page 86].

4. A much greater proportion of traffic moves in unit trains.
5. Improvements in materials, metallurgy, and manufacturing techniques have resulted in improved track durability and response.

For these (and many other reasons), a current analysis of railroad investment practices is needed. The objectives of this study are to describe patterns of track investment in the United States and show how track investments vary with network size, traffic, and other factors.

It is important to note that the previously mentioned 50% variability ratio, which was developed by the Interstate Commerce Commission (ICC), applies to all road investments, not just basic track components. It is not clear if the ICC intended this ratio to apply specifically to track. This paper does not intend to assess the process by which the factor was originally developed or interpret the ICC's original intent. Instead, the variability ratio is used in a general sense as a "null hypothesis." An assessment will be made at this paper's conclusion to determine if sufficient evidence exists to conclude that it is not applicable to basic track components.

2. TRACK INVESTMENT MODEL

This study is based on R-1 reports submitted to the STB from 1985 through 2008. Elements of the R-1 database include miles of road and track (derived from Schedule 700), gross ton-miles (derived from Schedule 755), and investments in basic track components from Schedule 416. All investment data have been restated in constant dollars. The increment to investment in each year is computed by subtracting the gross investment in year $t + 1$ from the investment in year t. Each yearly increment is restated in 1985 dollars and the recomputed increments are added back to the 1985 base to compute an adjusted value for each year.

The track investment model reflects the sum of investments in density classes I and II (Table 2.1) from Column L of Schedule 416 and includes capital expenditures for rails, ties, ballast, other track materials, and grading. The latter category includes the preparation and reconstruction of roadbed. Collectively, these elements are referred to as *basic track components*. The hypothesized model is $I = f(K, Q, F, T)$, where I denotes capital expenditures for track. K represents network size or scope. Q is a measure of traffic activity. F symbolizes firm (railroad-related) effects. And T stands for time.

Table 2.2 Density Categories used in Uniform System of Accounts

Class	Description
I	Lines carrying at least 20 million gross ton-miles per mile on an annual basis and not designated as belonging to Density Class III
II	Lines carrying less than 20 million gross ton-miles per mile on an annual basis and not designated as belonging to Density Class III
III	Lines identified as potentially subject to abandonment pursuant to Section 10904 of the Interstate Commerce Act
IV	Yard and way switching tracks
V	Electronic yards

Capital expenditures for basic track components include installation costs. For example, the costs of new rails reflect their placement in the track. In addition to the cost of materials, capital expenditures reflect labor, logistics, equipment, and other costs incurred in moving and installing components. However, the cost of maintaining and preserving the track is treated as an annual expense. Capital expenditures include replacements, additions, improvements, and rebuilding activities—when those activities extend the service lives of components. Repairs are classified as maintenance.

When track components are replaced, they are considered to be “retired” and are no longer reflected in the investment base. The same track segment may experience capitalized expenditures and retirements several times over its life, as older light rails are replaced with new heavier ones; grades and/or curves are reconstructed to improve alignments; and passing tracks, side tracks, switches, and turnouts are added.

2.1 Traffic Measures

There are several potential traffic measures, including revenue ton-miles and gross ton-miles. A ton-mile represents the movement of one ton in one mile. It is a composite measure of weight and distance. The ton can be transported (i.e., hauled) or travel under its own power, as in the case of locomotives. Revenue ton-miles are computed by multiplying the cargo weight by the distance traveled. Gross ton-miles include the weights of locomotives, freight cars, containers, trailers, cargo, and other equipment, as well as the distance traveled. A subset of gross ton-miles (train or revenue gross ton-miles) excludes work-related and track equipment, but includes locomotive, car, container/trailer, and cargo ton-miles.

RGTM is the most appropriate measure for this study for the following reasons: (1) Cargo ton-miles alone do not describe the type of track structure that is needed. A track must be designed to support gross vehicle weights. (2) Revenue gross-ton-miles exclude non-revenue activities (e.g., work train miles).

2.2 Network Size

There are several potential measures of network size, including miles of road (MOR) and miles of running track (MRT). Both measures have been used in previous studies. MOR (or route miles) represent a railroad's base network. Most rail lines were originally built as single-track lines to connect points or nodes within a network. As defined by the STB, MOR reflect only the first main track. In addition to the main track, a rail line may include second, third, and fourth main tracks and/or side tracks. For example, a 10-mile segment between two junctions may consist of two main tracks and two miles of crossover or passing track. Altogether, this segment comprises 22 miles of running track, which includes 10 miles of road and 12 miles of "other running track."

As traffic grows, railroads may add capacity by adding second or third main tracks and/or passing and side tracks, i.e., other running tracks. Similarly, if traffic declines, other running tracks may be disassembled and the assets liquidated or used elsewhere in the network. However, the first main track can only be abandoned if local traffic disappears and through traffic moving over the line can be rerouted. Even then, the railroad must petition the STB for authority to abandon the line. In the short to intermediate run, miles of road are relatively fixed. Miles of other running track can be more easily adjusted.

2.3 Main Effects

A certain level of investment in the base network is necessary regardless of the level or composition of traffic. Initially, lines may be built with lighter rails and thinner ballast sections suitable for traditional (e.g., carload) traffic at lower volumes. Capacity may be provided by a single main track with periodic sidings or passing tracks. However, when unit trains and heavy axle load cars are added to a network and faster speeds are desired, the quality of the track infrastructure must be improved through investments in heavier (more durable) rails, heavier tie plates, more ballast, and, in some cases, concrete ties.

Base investment is strongly correlated with miles of road and may not change substantially with modest increases in traffic. However, incremental investments—those designed to handle unit trains and heavier railcars—are a function of traffic. As traffic grows, other running tracks (such as passing and side track) may be added to increase capacity. Eventually, some lines may be doubled-tracked. Changes in miles of other running track are a function of traffic. If MRT is used to represent network size (instead of MOR) these investments will be attributed to the network, not to traffic.

Conceptually, MOR and RGTM are correlated. However, in practice, they are independent, at least over the analysis period. This fact is illustrated in Figure 2.1, which shows distinctly different trend lines for the two variables. Miles of road have declined since 1985, but at a decreasing rate. In comparison, revenue gross ton-miles have increased. The decline in miles of road owned by Class I carriers is largely a function of line sales to local and regional railroads and line abandonments. However, MOR has remained relatively constant since 1998. The recent drop in RGTM reflects a downturn in the global economy.

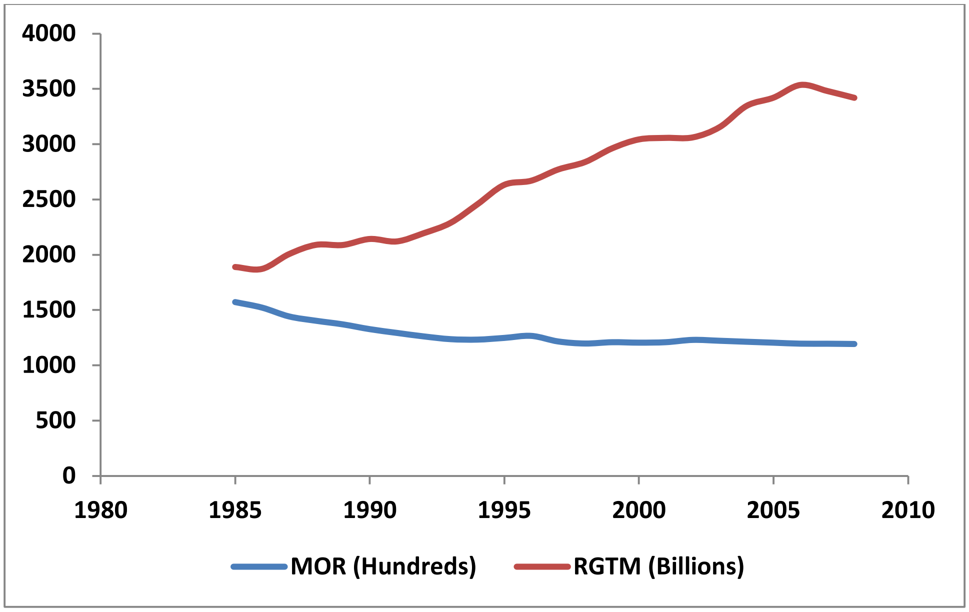

Figure 2.1 Trends in Miles of Road and Revenue Gross Ton-Miles

2.4 Treatment of Other Effects

A density variable is not included in the model because miles of road and gross ton-miles implicitly capture density effects. Increasing revenue gross ton-miles (while holding miles of road constant) results in higher traffic densities. Alternatively, increasing miles of road (while holding RGTM constant) reduces traffic density.

Most variations in basic track components result from the scope and quality of the base network and traffic. Nevertheless, investments may be made over time for other reasons. Throughout much of the analysis period, Class I railroads were making incremental track investments to effectively handle 286,000-lb. and 315,000-lb. railcars. While RGTM is the best traffic measure available, it does not explicitly account for axle weights. Two groups of traffic may generate the same RGTM, but have different effects on track because of differences in axle loads. Heavier axle loads require higher-quality track. However, the use of heavier railcars may result in fewer car-miles (thus, fewer tare ton-miles) and fewer locomotive-miles to move the same quantity. Because of these trade-offs, the effects of heavier railcars on gross ton-miles are mixed.

Axle loads are not reported in the R-1 data and cannot be computed directly from public sources. Given the mixed relationship between axle weights and RGTM, the effects of increasing axle loads may be subsumed in the time trend variable rather than being reflected in RGTM, which is expected to be positive. The time variable may reflect other changes in investment patterns over time that are not associated with traffic, network size, or specific railroads. As described later, the effects of mergers and consolidations are explicitly accounted for.

2.5 Statistical Model

The theoretical model is transformed into a statistical model in Equation 1. The subscript "i" denotes an observation for a particular railroad, while the subscript "t" indicates a particular year of the data series. Using this notation and letting epsilon (ϵ) represent the error term, the regression equation may be written as:

$$(1) \quad I_{it} = \beta_0 + \beta_1 MOR_{it} + \beta_2 RGTM_{it} + \beta_3 T_t + \beta_4 F_i + \epsilon_{it}$$

The model includes two main explanatory variables (traffic and network size), time (T), and an array of railroad indicator variables (F_i). The purpose of the railroad variables is to capture fixed effects that are specific to particular railroads but do not change over time. T, on the other hand, accounts for industry-wide trends and changes that occur over time. Even when all of these variables are considered (21 altogether, including the indicator variables), a great many factors are not accounted for in the model and are subsumed in the error term (epsilon).

F_i can assume values of 0 or 1. F_i is equal to 1 when the observation comes from a particular railroad. Once *i* is specified (i.e., the observation is determined to come from a particular railroad), the effect of β_4 is to shift the intercept (β_0) for that railroad.[5] T is an integer that measures the elapsed time in years since 1984. For example, *t* assumes a value of 1 in 1985, 5 in 1989, 10 in 1994, and so forth. Once *t* is specified (i.e., the observation is determined to belong to a particular year), the contribution of time is computed as $\beta_3 \times t$. Once computed in this manner, the contribution of time becomes a constant that shifts the intercept for a particular year. The slope of the regression is determined by MOR and RGTM.

2.5.1 Functional Form

The choice of functional form is based on data and statistical issues. A plot of track investment against miles of road is shown in Figure 2.2. In addition to revealing non-constant variance, the graph highlights the vast differences in scale between smaller Class I railroads (e.g., the Kansas City Southern and Soo Line) and the largest carriers (e.g., BNSF and UP).

While the apparent heteroscedasticity can be accounted for, the differences in scale are problematic. A linear model results in a negative intercept for MOR in a simple regression equation and a negative (counterintuitive) sign in a multiple regression model.

For comparative purposes, a plot of the natural log of track investment against the natural log of miles of road is shown in Figure 2.3. A graph of track investment and RGTM is presented in Figure 2.4, while Figure 2.5 depicts the logarithmic relationship between these two variables. Comparisons of Figures 2.2 and 2.3 and 2.4 and 2.5 suggest that the variances of the log relationships are relatively constant—more so than the linear ones.

[5]For purposes of simplification, β_4 is used in a collective sense in this description. In actuality, each railroad indicator variable has its own beta coefficient in the model (e.g., β_4–β_{21}).

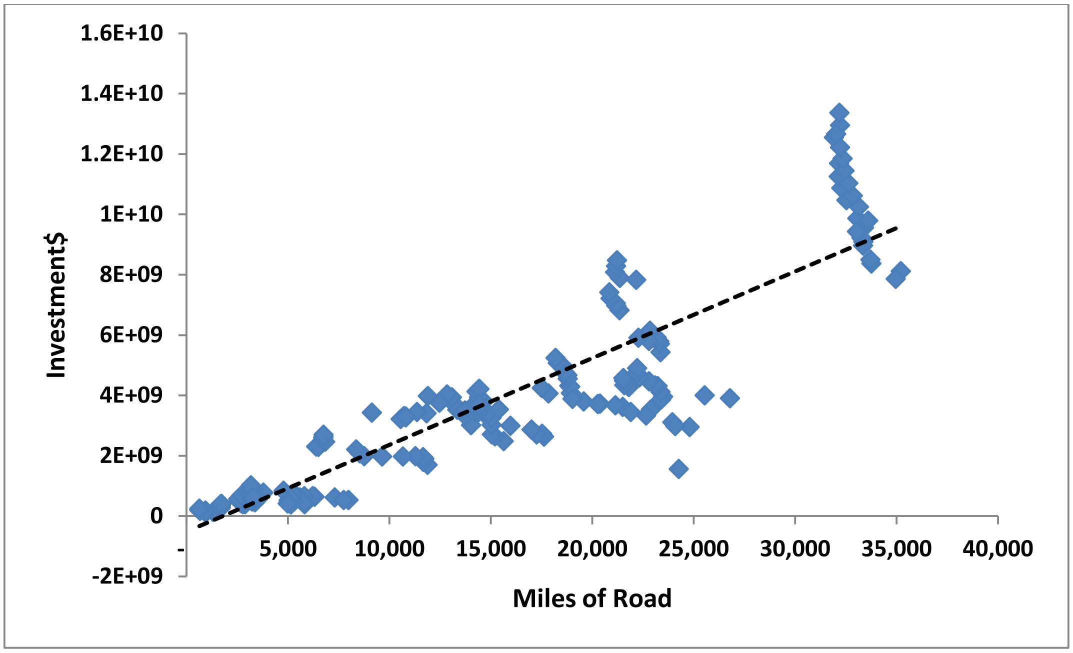

Figure 2.2 Plot of Track Investment against Miles of Road

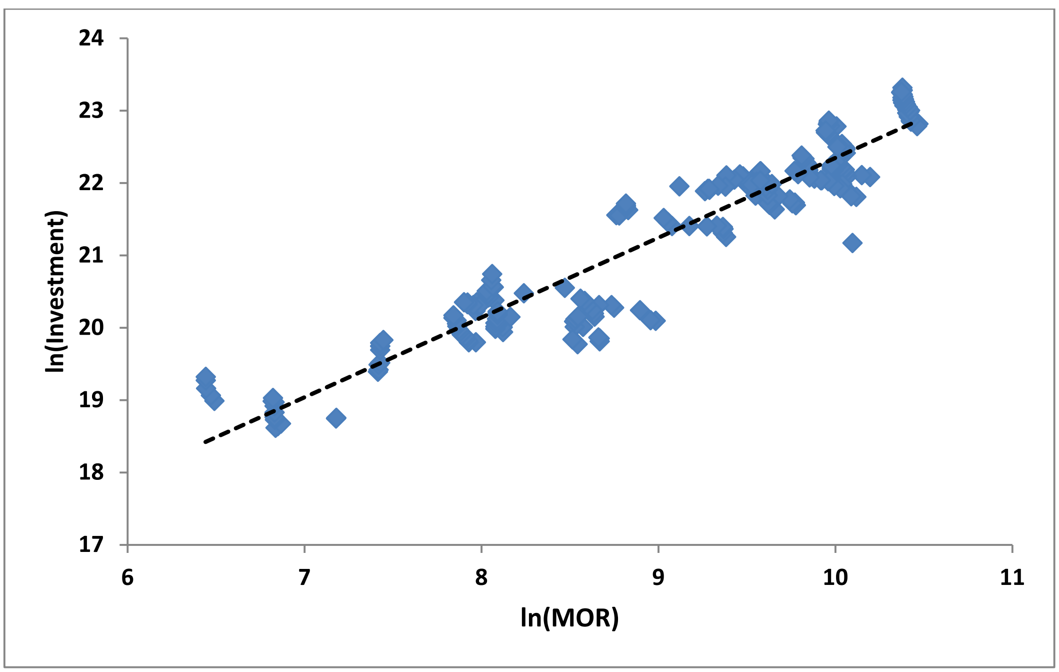

Figure 2.3 Plot of Log of Track Investment against Log of Miles of Road

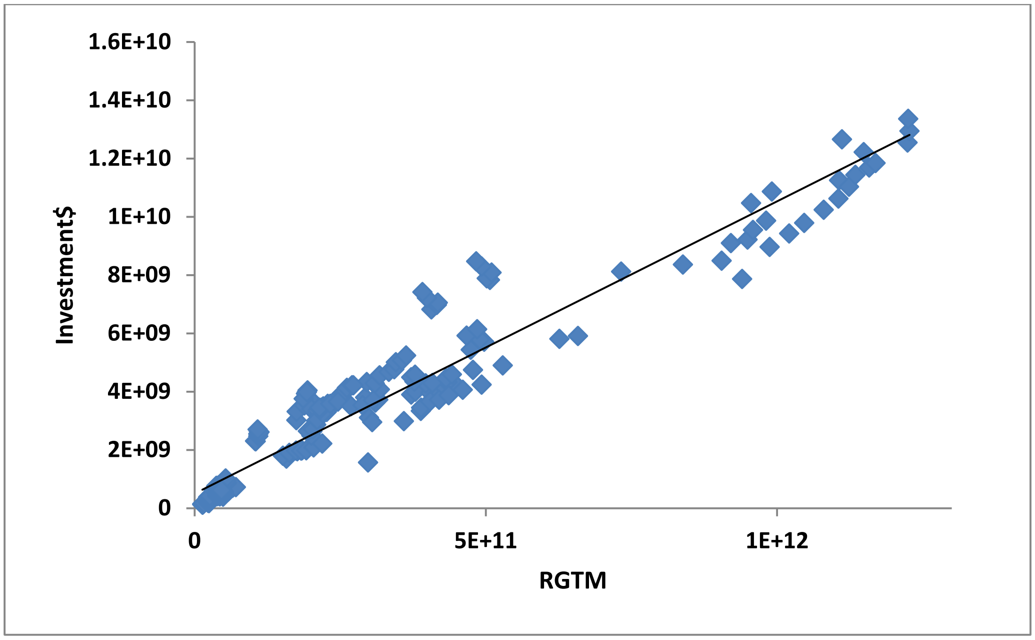

Figure 2.4 Plot of Track Investment against Revenue Gross Ton-Miles

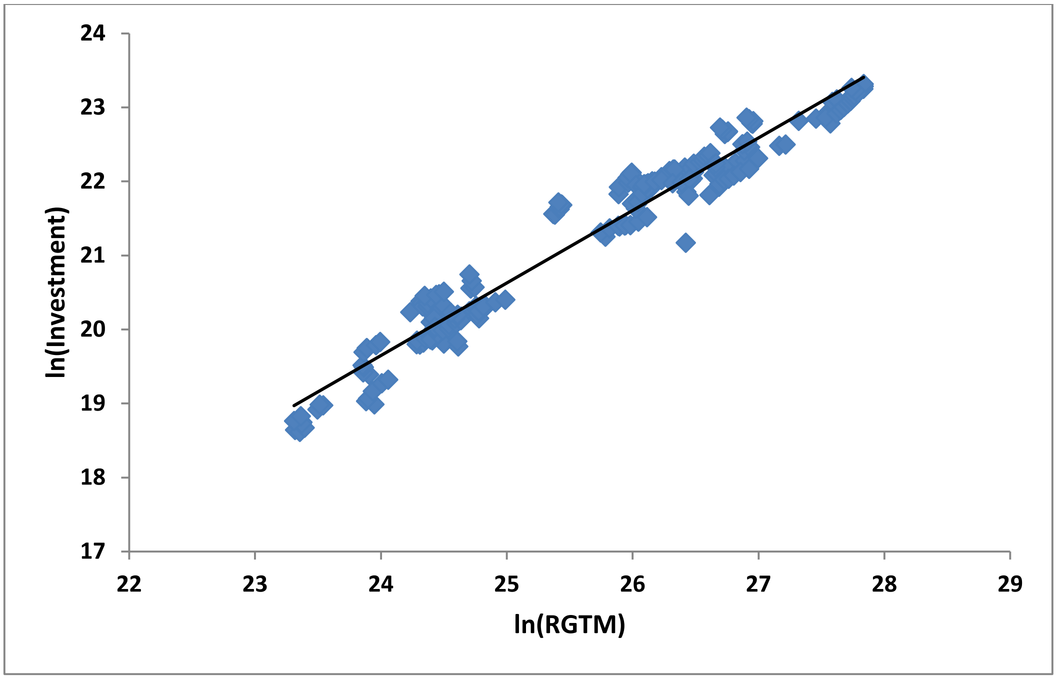

Figure 2.5 Plot of Log of Track Investment against Log of RGTM

2.5.2 Initial Model

The parameter estimates and standard errors from a logarithmic regression model are shown in Table 2.2. The dependent variable is the natural log of track investment, where investments are expressed in constant 1985 dollars using the Rail Cost Adjustment Factor (RCAF). The primary explanatory variables are the logs of MOR and RGTM. However, each Class I railroad that existed during the 1985-2008 period is represented by an indicator variable, e.g., KCS. When the observation is for the Kansas City Southern Railway, KCS equals 1. Otherwise, KCS equals zero. Additional indicator variables are defined for mergers. For example, the UP system includes three railroads that appear in the database: Union Pacific (UP), Southern Pacific (SP), and Chicago and North Western (CNW). CNW was acquired by UP in 1995. UP merged with SP in 1997. In the analysis, UP-CNW assumes a value of 1 in 1995, and each year thereafter, but is zero otherwise. Similarly, the variable UP-SP assumes a value of 1 in 1996, and each year thereafter, but is zero otherwise.

Table 2.3 Parameter Estimates from Logarithmic Model of Track Investment

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	-3.9008	1.93940	-2.01	0.0456
ln(MOR)	0.34868	0.06448	5.41	<.0001
ln(RGTM)	0.83199	0.08085	10.29	<.0001
ln(T)	0.08896	0.02067	4.30	<.0001
ATSF	0.40643	0.08134	5.00	<.0001
BNSF	-0.41320	0.13223	-3.12	0.0020
BN	0.14038	0.05686	2.47	0.0144
UPSP	-0.53416	0.13430	-3.98	<.0001
UPCNW	-0.31632	0.19125	-1.65	0.0996
SP	0.53123	0.06809	7.80	<.0001
CNW	0.38754	0.14885	2.60	0.0099
SOO	0.42883	0.17406	2.46	0.0146
ICG	0.68665	0.18027	3.81	0.0002
ICG89	-0.14949	0.13323	-1.12	0.2631
GTW	0.65526	0.25117	2.61	0.0097
GTC	-0.17770	0.32128	-0.55	0.5808
KCS	0.94761	0.19579	4.84	<.0001
CR	0.79455	0.07507	10.58	<.0001
CSX	0.42085	0.05668	7.43	<.0001
CSXCR	-0.78369	0.10486	-7.47	<.0001
NSCR	-0.93980	0.11257	-8.35	<.0001
NS	0.56670	0.06820	8.31	<.0001

Analogous variables are defined for other mergers or acquisitions. For example, Burlington Northern merged with Atchison, Topeka, and Santa Fe (ATSF) in 1996 to form the Burlington Northern-Santa Fe (BNSF). CSX and Norfolk Southern (NS) acquired parts of Conrail in 1999. In 2002, the Canadian National Railway consolidated the Illinois Central Gulf (ICG), Grand Trunk Western (GTW), and other rail lines into the Grand Trunk Corporation (GTC). In the Grand Trunk system, GTC is 1 if the year is 2002 or later; however, GTC is zero otherwise. The ICG indicator variable assumes a value of 1 when GTC is 1, or when the observation is for the old ICG prior to 2002. The GTW variable works in a similar manner. The sign and estimate of each railroad indicator variable is relative to the variable omitted from the equation, which is the unmerged UP railroad. The meaning of the variable ICG89 is discussed later.

The results of a linear model, which includes the same fixed (railroad) and time variables as the log model, are shown in Table 2.3. In the linear model, MOR has a negative sign and weaker statistical relationship than in the log model. For statistical reasons, subsequent analyses are based on the logarithmic model.

Table 2.4 Main Parameter Estimates from Linear Regression Model of Track Investment

Variable	Parameter Estimate	Standard Error	t Value	Prob. > |t|
Intercept	-688008058	994797992	-0.69	0.4899
MOR	-35083	38001	-0.92	0.3570
RGTM	0.01187	0.00086289	13.75	<.0001

As shown in Tables 2.4 and 2.5, the logarithmic model has excellent statistical properties, including an R-Square of 0.99 and a coefficient of variation of less than 1%. The model explains nearly all of the variation in the log of investment and provides a very precise fit. The low coefficient of variation (0.6%) suggests that the model could be an excellent predictor within the range of observed values. However, the Durbin-Watson test (Table 2.6) indicates autocorrelation, i.e., the errors are correlated over time. This leads to the formulation of an autoregressive model.

Table 2.5 Mean Square Error and F-Value for Log Model of Track Investment

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Prob. > F
Model	21	367.48915	17.49948	1051.34	<.0001
Error	210	3.49544	0.01664		
Corrected Total	231	370.98459			

Table 2.6 R-Square and Coefficient of Variation for Log Model of Track Investment

Root Mean Square Error	0.12902	R-Square	0.99
Coefficient of Variation (%)	0.60547	Adjusted R-Square	0.99

Table 2.7 Results of Test for Serial Correlation in Log Model of Track Investment

Durbin-Watson Statistic	0.725
Prob. < DW	<.0001
Prob. > DW	1.0000
1st Order Autocorrelation Coefficient	0.638

2.5.3 Autoregression Model

In regression analysis, each *t* is assumed to be normally and independently distributed with a mean of zero and a variance of σ^2 (i.e., $\varepsilon_t \sim IN(0, \sigma^2)$. Violation of this assumption may affect statistical tests and parameter estimates. In the revised model, the original regression equation is augmented with an autoregressive sub-model of the error term. This process is described in Appendix A.

As shown in Table2.7, the parameter estimates of the structural variables have changed. The coefficient of the log of MOR indicates that track investment increases by roughly 0.59% when miles of road increase by 1%. The estimate for the log of RGTM indicates that track investment increases by roughly 0.50%

when gross ton-miles increase by 1%. As expected, the time-related variable is positive and highly significant, indicating that track investment has been increasing over time for other reasons.[6] Many of the railroad and merger variables are highly significant, capturing differences among railroads attributable to economic, managerial, and locational factors and post-merger synthesis and rationalization.

Table 2.8 Results of Autoregression Model of Track Investment

Variable	Parameter Estimate	Standard Error	t Value	Approx. Pr. > |t|
Intercept	2.2451	2.0308	1.11	0.2703
ln(MOR)	0.5902	0.0650	9.08	<.0001
ln(RGTM)	0.5026	0.0877	5.73	<.0001
ln(T)	0.1780	0.0196	9.07	<.0001
ATSF	0.5035	0.0834	6.04	<.0001
BNSF	-0.3110	0.1222	-2.54	0.0118
BN	0.1699	0.0554	3.07	0.0025
UPSP	-0.5807	0.1124	-5.17	<.0001
UPCNW	0.0592	0.1770	0.33	0.7383
SP	0.6215	0.0743	8.37	<.0001
CNW	0.2117	0.1420	1.49	0.1375
SOO	0.1372	0.1716	0.80	0.4250
ICG	0.5070	0.1771	2.86	0.0047
ICG89	-0.2709	0.0657	-4.12	<.0001
GTW	0.5299	0.2371	2.24	0.0266
GTC	-0.0843	0.3025	-0.28	0.7809
KCS	0.6806	0.1912	3.56	0.0005
CR	0.7689	0.0709	10.84	<.0001
CSX	0.4392	0.0497	8.83	<.0001
CSXCR	-0.7520	0.0926	-8.12	<.0001
NSCR	-0.9769	0.0982	-9.95	<.0001
NS	0.5558	0.0578	9.62	<.0001

The Durbin-Watson statistic for first order autocorrelation in the revised model is essentially 2.0. The probability values shown in Table 2.8 indicate that the null hypothesis (independence of errors) should not be rejected. Because the transformed model is estimated via generalized least squares, the error variances are homoscedastic. The regression R-square is essentially unchanged. The error sum of squares is 1.4145 and the mean square error is 0.00756.

[6] In this study, the model is estimated from a population of observations, not a sample. The relationships between the parameter estimates and standard errors are important in assessing the fit and precision of the regression. Technically, the probability or p-values based on sampling theory are not applicable to the interpretation of results. Nevertheless, the population of Class I railroads may be thought of as a sample consisting of railroads that were classified as Class I carriers during a given year (based on the revenue definitions established by the Surface Transportation Board) from a larger population of railroads. In this way, the familiar interpretations of p-values can be applied. It is also instructive to note that the null hypothesis for a t-test is that the slope of a parameter estimate is zero. The t-ratios and p-values are instructive in this regard, indicating the likelihood of observing a larger value of the parameter estimate when the null hypothesis is true, its value is actually zero.

Table 2.9 Durbin-Watson Test for First Order Autocorrelation

DW	Prob. < DW	Prob. > DW
1.9981	0.2128	0.7872

2.5.4 Data Issues

Most of the data series are consistent throughout the period. However, data for the Illinois Central Gulf (ICG) stand out (Table 2.9). Line investment suddenly drops by 46% between 1988 and 1989. Distinct trends exist before and after 1989. The sudden drop is captured by the indicator variable ICG89, which assumes a value of 1 if the railroad is ICG and the year is 1989. Otherwise, ICG89 is zero.

Table 2.10 Data for Illinois Central Gulf (ICG) Railroad

Year	Miles of Road	Nominal Line Investment (millions)
1988	2,900	$691
1989	**2,887**	**$396**
1990	2,773	$398
1991	2,766	$405

As shown in Table 2.7, ICG89 is highly significant and negative, suggesting that the indicator variable is capturing the sudden drop in investment without a corresponding drop in miles of road. The actual reason for the sudden decrease in reported investments is unknown. Without a detailed inquiry, it must be assumed that the data are correct but anomalous. Irrespective of the reason for the sudden drop, the parameter estimates are largely unaffected when the indicator variable is included in the model.

Table 2.10 indicates a second anomaly in the data. Conrail was acquired by Norfolk Southern and CSX in 1999. Conrail appears in the data series for the last time in 1998. In 1999, the miles of road reported by CSX and NS collectively increased by 38%, reflecting the integration of Conrail into the two networks. Similarly, the collective RGTM of CSX and Norfolk Southern increased by 37% between 1998 and 1999. However, the reported investments in basic track components increased by only 4%.

Table 2.11 Data for Conrail, CSX, and Norfolk Southern Before and After Acquisition

Year	Railroad	Miles of Road	RGTM (millions)	Nominal Line Investment (Thousands)
1998	CR	10,797	209,069	$3,169,190
	CSX	18,181	337,311	$5,742,229
	NS	14,423	249,840	$4,633,736
1999	**CSX**	**23,357**	**440,836**	**$6,024,295**
	NS	**21,788**	**364,826**	**$4,728,444**
2000	CSX	23,320	461,935	$6,467,962
	NS	21,759	376,550	$4,751,575

There could be many reasons for this inconsistency. As shown in Table 2.7, all five indicator variables associated with these railroads (CR, CSX, CSXCR, NSCR, and NS) are highly significant. The two indicator variables associated with the post-acquisition railroads (CSXCR and NSCR) are highly significant and negative, suggesting that these variables are capturing the anomaly, where MOR and RGTM jump while line investment remains largely unchanged. The indicator variables may be capturing other effects as well.

2.6 Model Interpretations

2.6.1 Defining Predictive Equations

The model can be used to predict the log of investment for individual railroads. The intercept and all applicable indicator variables are used in these predictions. For example, the mean-value formula for BNSF (Equation 2) uses the average values of MOR and RGTM for the 1996-2008 period.

$$(2)\quad \widehat{ln(I)} = \beta_0 + \beta_1 \ln(\overline{MOR}) + \beta_2 \ln(\overline{RGTM}) + \beta_3 \ln(T^*) + \beta_4 ATSF + \beta_5 BN + \beta_6 BNSF$$

Where:

$\widehat{ln(I)}$ = Predicted log of investment for BNSF

$\ln(\overline{MOR})$ = Log of mean value of miles of road for BNSF

$\ln(\overline{RGTM})$ =Log of mean value of RGTM for BNSF

$\ln(T^*)$ = Log of T, where T* represents the midpoint of the period

Similar equations can be developed for other railroads using different indicator variables. The predictions for Norfolk Southern utilize the variables CR, NS, and NSCR. Predictions for CSX utilize CR, CSX, and CSXCR. When the appropriate indicator variables are selected, the model yields a series of predictive equations for individual railroads.

2.6.2 Implications of Constant Elasticity

The log model is a constant elasticity model, e.g., the percentage change in track investment resulting from a 1% change in RGTM is the same for all output levels. However, this does not mean that the increase in investment is the same at all levels. A 1% increase starting from an investment base of $1 billion is much greater than a 1% increase starting from a base of $500 million. The slope of the log model reflects the same (relative) rate of change in investment over the range of observations. In comparison, the slope of a linear model represents a constant (absolute) rate of change.

Even though the elasticity of the log model is constant, the effects are nonlinear. This is illustrated in Figure 2.6, which shows how the predicted values of investment for a particular railway (BNSF) derived from Equation 2 change when RGTM is varied, while holding miles of road constant at its mean value: $(\overline{MOR})$. The graph is juxtaposed against a linear trend line to illustrate economies of density.

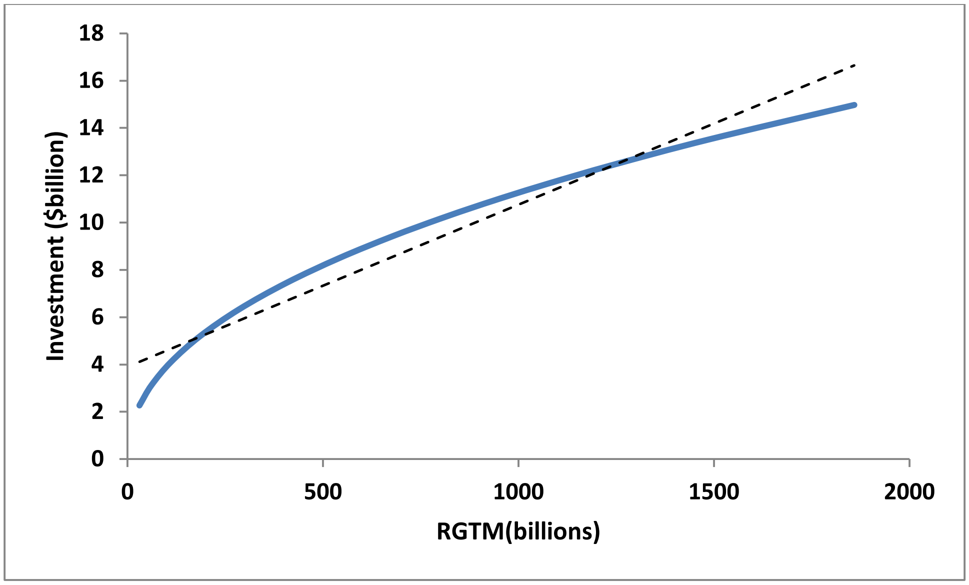

Figure 2.6 Variations in Predicted Track Investment from Log Model Holding Miles of Road Constant at Mean Value

2.6.3 Economies of Density

RGTM and miles of road are the numerator and denominator, respectively, of traffic density as measured in revenue gross ton-miles per route mile. As Equation 3 suggests, density can be increased by scaling (reducing) the size of the network in relation to traffic or increasing traffic for a given size of network.

$$(3) \quad Density = \frac{RGTM}{MOR}$$

As shown in Figure 1, MOR decreased throughout much of the period. The elasticity of MOR suggests that track investments decrease when miles of road decrease, but at a less-than-proportionate rate. If RGTM is held constant, a 1% MOR reduction results in a 0.59% decrease in track investment. Similarly, the elasticity of RGTM indicates that track investments increase with traffic, but at a less-than-proportionate rate. If miles of road are held constant while RGTM increases, capital expenditures for basic track components will rise by approximately 0.50% for each 1% increase in RGTM. Output will increase at a greater rate than input cost, implying economies of density.

2.6.4 Magnitudes of Parameter Estimates

Several obvious questions stem from the results.

- Why is the elasticity of track investment with respect to MOR substantially less than 1.0? Should not a 1% reduction in MOR result in a proportional decrease in track investment? If the investment in each mile of road was the same (e.g., a constant $500,000 per mile) the expected elasticity would be 1.0 (ceteris paribus). However, the average investment in rail lines sold or abandoned by Class I railroads (and thus disappear from the investment base) may be less than the average investment in retained lines (which tend to be mainlines). Moreover, when miles of road are decreased while revenue gross ton-miles are held constant, the same level of traffic is concentrated on fewer route miles. While this leads to economies of density, the additional traffic may require incremental investments elsewhere in the system, i.e., on those lines that now have higher traffic levels. If this occurs, the overall reduction in track investment resulting from a 1% reduction in MOR will be less than 1%.

- Why is the elasticity of track investment with respect to revenue gross ton-miles less than 1.0? (1) *Economies of Utilization*: In many cases, significant traffic volumes can be added to lines with low traffic levels before any incremental investments are needed. When investments are needed, adding passing tracks to an existing line to accommodate traffic growth costs less than the construction of the main track. (2) *Economies of Design*: In some cases, the strength of materials increases in a nonlinear manner with size or weight. For example, a rail's moment of inertia is an indication of its tendency to resist rotational and bending forces. Moment of inertia increases with both the cross-sectional area of the rail and its weight. Upgrading a track from 115-lb. to 136-lb. rail increases the weight of the rail by only 18%, but the moment of inertia increases by 45%. As described below, incremental capital investments made to existing track and roadbed realize foundational economies.

- Why is the elasticity of track investment with respect to revenue gross ton-miles less than the elasticity with respect to MOR? Economies of design and utilization are two key factors. Some base level of investment in roadbed, ties, ballast, rails, and other track materials is necessary to initially build and operate a line, regardless of the expected traffic level. In the model, base investment is a function of MOR. However, once a line is built, further improvements (which are a function of traffic) comprise incremental capital investments, such as replacing lighter rails with heavier ones. Incremental investments such as these may not require re-grading or roadbed reconstruction. Because of foundational investments, capital projects that utilize existing roadbeds and tracks may be less expensive than initial construction, which reflects extensive grading and roadbed preparation costs. While the parameter estimates of MOR and RGTM are different, they are not divergent or inconsistent.

2.7 Sensitivity of Estimates to Cost Indexes

In the results presented thus far, track investments have been restated in constant 1985 dollars using the RCAF. The Railroad Cost Recovery Index (RCRI) is an alternative series. However, neither index is perfect for this study. Both are heavily influenced by increases in fuel costs. The mix of labor, materials, fuel, and other inputs for track construction is unique. While fuel is a significant construction cost, other railroad activities, such as train and yard operations, are more fuel-intensive than construction projects. The disadvantage of using an aggregate index is that it reflects cost increases for the railroad as a whole, not for a specific category such as track investment.

For comparison purposes, the Civil Works Construction Cost (CWCC) Index for roads, railroads, and bridges, published by the U.S. Army Corps of Engineers, is shown in Figure 7.[7] This index is specific to construction, but reflects highways and bridges as well as track. As the graph shows, the RCAF and the CWCC are closely aligned until 2008. All things considered, the RCAF may be the best index.

The sensitivities of the parameter estimates to the two indexes are illustrated in Table 2.10, where elasticities for miles of road, gross ton-miles, and time are shown using the RCRI and the RCAF. These values are compared with parameter estimates derived from a model that uses nominal investments, unadjusted by either index. The elasticities based on nominal dollars reflect the true underlying mix of labor, materials, fuel, and other inputs used in track construction each year. However, the coefficients

[7] U.S. Army Corps of Engineers, Department of the Army. *Civil Works Construction Cost Index System*, March 31, 2011.

may be misleading because the dollars are not constantly valued. All things considered, the elasticities based on the RACF may be the most relevant ones.

3. INTERPRETIVE CONTEXT

In addition to the uncertainties posed by cost indexes, other factors should be considered when interpreting the results of this study.

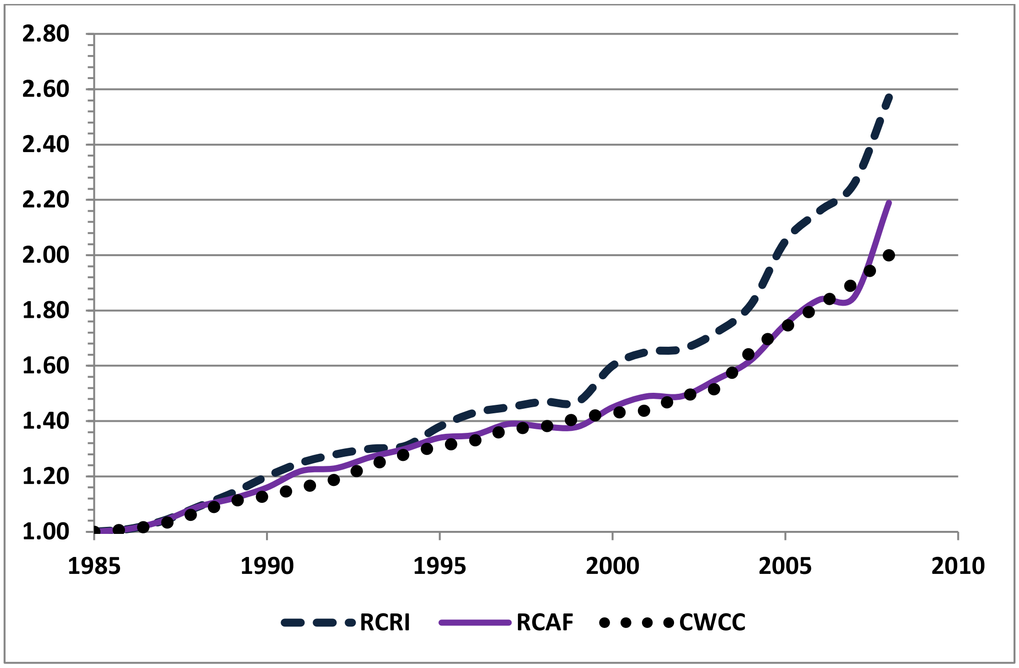

Figure 3.7 Comparison of RCRI, RCAF, and CWCC Indexes

Table 3.12 Estimated Elasticities of Track Investment with Respect to Miles of Road, Gross Ton-Miles, and Time Under Different Assumptions

		Constant Dollars Based on	
	Nominal Dollars	**Rail Cost Adjustment Factor**	**Railroad Cost Recovery Index**
Miles of Road	0.6623	0.5902	0.5831
Gross Ton-Miles	0.6088	0.5026	0.4791
Time	0.2136	0.1780	0.1721

3.2 Delayed Capital Expenditures

The full cost of owning and operating a rail line includes both capital and maintenance expenditures. If capital expenditures are delayed or deferred, maintenance costs may rise. On the other hand, timely capital investments may reduce maintenance costs.

In an earlier era, railroads may have delayed capital expenditures because of low returns on investment. However, Class I industry returns improved from 1.7% in 1970 to more than 10% in 2006 and exceeded 5% for most years since 1985. Given this trend, there is a greater likelihood that capital investments were made when needed during the analysis period.

3.3 Regulated Versus Market Investments

Some investment (and disinvestment) decisions are regulated, while others are not. For example, decisions to abandon a main track, extend a line, or construct a new rail line must be approved by the STB. In contrast, decisions to upgrade an existing line or add tracks within the existing right-of-way are often independent decisions under control of the railway. Nevertheless, investment levels must provide for safe operations, as rail lines are subject to inspection by the Federal Railroad Administration. In many respects, railroad investment decisions are mixed choices, reflecting purely private objectives as well as societal goals.

3.4 Accounting Interpretations

For the most part, the data series appear to be consistent. However, the distinction between capital and maintenance expenditures can be subjective. If expenditures are to be capitalized, the cost of a rebuilding a line should be "material" in relation to the cost of replacing it. But, what is material? Projects to improve track alignment without roadbed reconstruction pose interpretative dilemmas. Nevertheless, it is likely that these decisions are made similarly across railroads. If this is not the case, the railroad indicator variables should capture the differences.

This study utilizes gross (original) investment instead of net investment. The latter is computed by subtracting accumulated depreciation from gross investment. Depreciation is an accounting concept, based on the typical lives of assets. However, depreciation may reflect tax guidelines or incentives and include "accelerated depreciation." In some years, negative accumulated depreciation is reported in the R-1. Issues such as these would need to be addressed before net investment could be used as the dependent variable in a model. These issues do not affect gross investment, which is a reflection of the railroad's reactions to traffic and profit potential, as well as to general economic indicators.

3.5 Economies of Traffic Density

While the data and model suggest that economies of traffic density exist with respect to investments in basic track components, this conclusion cannot be generalized to track maintenance and line operating costs. Overall economies of density may be different when line operating and maintenance expenses are considered. The model is not offered as a comprehensive cost function. Rather, the study is an empirical one, in which patterns of investment are observed over time.

3.6 Forecasting with the Model

MOR have been relatively constant for the last decade. Given this stability, forecasting investments into the future based on variations in RGTM may yield valid results. Nevertheless, railway investment decisions are influenced by a variety of business and regulatory factors. Using the model for forecasting purposes assumes that these unobserved and uncontrollable factors, which were present between 1985 and 2008, will remain the same in the future. Predicting beyond the range of RGTM poses additional risks, given the nonlinear nature of the model.

3.7 Relative Contributions of Traffic and Network Size

The relative contributions of MOR and gross ton-miles to track investment are of interest from a regulatory perspective. When MOR are held constant (a very realistic scenario), the increase in track investment is roughly 50%; i.e., for a 100% increase in RGTM, track investment is expected to increase by 50%. However, this is not a completely satisfactory answer. As shown in Table 2.7, the elasticity of investment with respect to time is 18%. T could, at least in part, reflect the upgrading of tracks to handle heavier axle loads.

Investments to handle heavier cars do not represent fixed investments. It is unclear whether these effects should be attributed, wholly or in part, to "traffic." Perhaps the best conclusion that can be drawn from this study is that there is no compelling evidence to suggest that the traditional assumption (i.e., half of a railroad's investment in road varies with traffic) is no longer applicable to investments in basic track components. However, this conclusion cannot be extended to other areas of roadway investment.

Note that in the long run, track investments are primarily a function of traffic. The investment function estimated in this study is a short- to intermediate-run one. In the long run, miles of road are theoretically a function of traffic, even though MOR and RGTM are independent in the short run. A challenge for this and similar studies is that it is impossible to observe track investments measured on a consistent basis over a truly long-run period. If the same model was estimated from 75 years of consistent investment data, the parameter estimates could change.

3.8 Other Statistical Issues

3.8.1 Multicollinearity

The railroad indicator variables provide valuable information in the model and absorb data anomalies. However, the indicator variables are strongly correlated with MOR and RGTM. While multicollinearity is often a concern in multiple regression analysis, it poses no real problems for the track investment model, with the possible exception that some of the hypothesis tests for the indicator variables may be affected.

The null hypothesis for an indicator variable is that it does not significantly shift the intercept; i.e., its effect is nil. As shown in Table 2.7, only four of the indicator variables have p-values > 0.05, meaning that they are not statistically significant. It is possible that the standard errors of these variables are so inflated by multicollinearity that the hypothesis tests are misleading and that these four indicator variables are actually statistically significant. Even if this were true, it would have no real impact on the primary interpretations of the study.

However, multicollinearity has a more general effect. The parameter estimates are conditional on the indicator variables being included in the model. If the indicator variables are removed, the parameter estimates of MOR and RGTM will change. Since there are strong theoretical and practical justifications for the indicator variables being included in the model, they should not be removed. Moreover, the statistical significance of the indicator variables must be appraised collectively. Dropping the indicator variables with high p-values, while keeping the other indicator variables in the model, would not be appropriate.[8]

---

[8] The significance of the indicator variables as a group can be assessed through a partial F-test. The error sum of squares from a reduced model excluding the railroad indicator variables is 3.6592. In comparison, the error sum of squares from the full model (including the railroad indicator variables) is 1.4145. The difference in the error sum of squares attributable to the railroad indicator variables is 2.2448. This calculated value (which is reflected in the numerator of the F-statistic) has 21 minus 18, or 3 degrees of freedom. The error sum of squares from the full model (which is reflected in the denominator of the F-statistic) has 232 – 21 – 1, or 210 degrees of freedom. The computed F-value of 111 is far greater than the critical F-value of 2.65 for an alpha of .05. This test formally confirms what is apparent from Table 2.7. Collectively, the railroad indicator variables significantly improve the model. Therefore, all 18 indicator variables should stay in the model.

3.8.2 Impacts of Other Activity Variables

Introducing another highly correlated activity variable into the model will change the parameter estimates. For example, it could be argued that changes in other running track miles are linked to train-miles, more so than to revenue gross ton-miles. However, the R-square from a regression of the log of train-miles (TM) against the log of RGTM is 0.99.

As shown in Table 3.2, the log of TM is not statistically significant when it is included in a model with the log of RGTM and the log of MOR. Its p-value is 0.07. However, this is not a binding statistical conclusion, given the possible effects of multicollinearity on hypothesis tests. The primary justification for including TM would be if they contribute a unique and independent effect that RGTM does not. This is a difficult argument to make, given that the R-square from a regression of the log of TM against the log of RGTM is 0.99.

Table 3.13 Results of Track Investment Model with Train-Miles Added

Variable	Estimate	Standard Error	t Value	Approx. Prob. > |t|
ln(MOR)	0.5591	0.0664	8.42	<.0001
ln(RGTM)	0.3385	0.1354	2.50	0.0133
ln(TM)	0.2398	0.1306	1.84	0.0680
ln(T)	0.1766	0.0195	9.07	<.0001

For purposes of brevity, only the main variables from a model that also includes 18 indicator variables are shown.

The primary effect of TM (when it is included in the model) is to reduce the parameter estimates of the other variables. However, the combined partial effects of TM and RGTM are only marginally greater than the effect of RGTM in the previous model. The general conclusion regarding the elasticity of track investment with respect to “traffic” does not change substantially when TM is added to the model.

The decision in this case is not to add TM, for all of the reasons noted above. However, this is a judgmental decision as there may be differing points of view.

4. RESEARCH TO EXPAND THE ANALYSIS

The model presented in this paper includes only a subset of roadway investment costs, e.g., the basic track accounts. Therefore, it only partially addresses the gap in knowledge. No definitive statements can be made regarding the efficacy of the STB's overall assumption that roadway investment costs are 50% variable with traffic. Roadway investment includes many other cost elements. Thus, other models of roadway investment are possible.

4.1 Other Roadway Investment Models

A model of traffic control and communication infrastructure could be estimated using investments in communication systems, signals and interlockers, power transmission systems, and grade crossings. These investments may be more closely related to TM than to RGTM. Traffic control and communication investments are affected more by the number of trains per day than by train weight.

Investments in *structures*, such as tunnels, bridges, and trestles, and *miscellaneous facilities* could comprise additional clusters. Gross ton-miles may be the most logical traffic variable for a structure's sub-model, while investments in other facilities, including station and office buildings, may be more appropriately modeled as a function of revenue gross ton-miles or revenue tons. Investments in *specialized facilities*, such as COFC/TOFC terminals, could be modeled as a function of related activities (e.g., container and trailer units loaded and unloaded).

A moderate level of effort is involved in developing these models. The R-1 database developed for this project includes all of the variables. However, programs must be written to create the input datasets in proper format for the models. These data elements have not been examined for consistency or statistical issues.

4.2 Individual Component Models

It is also possible to develop models for individual track components, such as rails and other track materials, ballast, ties, and roadbed. However, a sub-modeling approach may impose restrictions on the regression functions. For example, railroads may trade off better ballast and ties against heavier rails in some cases. The track is an integrated structure. The results of individual component models must be interpreted accordingly.

4.3 Density Class Models

Using data from Schedule 720, it may be possible to develop separate regression models for density classes I and II. The consistency of Schedule 720 data has not been examined. Moreover, programming changes are needed to create a database for use with density class models. However, the time and resource costs to develop these databases are moderate. The practical applications of the models with respect to URCS are unclear.

4.4 Axle Load Effects

In theory, track impacts are a function of axle loads and speed, which determine the dynamic impacts and deflections of the track. Car axle loads cannot be effectively computed from R-1 data. In order to add this variable, a weighted average would have to be computed from the waybill sample for each railroad, for each year. The axle weights in the sample could be weighted by the car-miles of travel.

An aggregate measure of speed can be computed from R-1 data by dividing train-miles by train-hours. However, this calculated value is a broad system performance measure that includes many factors, such as train delays. It is of little use in analyzing the dynamic effects of axle loads. A more promising approach is to estimate the weighted-average speed limit from Schedule 720. This variable could serve as a proxy for the weighted dynamic effect on each railroad's system, based on the carrier's line classifications and speed limits.

The resource cost of adding these variables is moderate. However, the probability of success is unknown.

APPENDIX A: STATISTICAL MODELING PROCEDURES

In this study, the SAS REG and AUTOREG procedures are used in conjunction with the underlying data illustrated in Appendix B. Because the initial results indicate serial correlation, the regression model is transformed. To illustrate the issues associated with autocorrelation and potential solutions, the structure and assumptions of the ordinary least squares (OLS) model (the results of which are shown in Table 3) are briefly introduced.

A.1 OLS Model: The Starting Point

If the track investment model was not affected by serial correlation, the OLS procedures inherent in PROC REG could be used. Using matrix notation, the OLS model can be depicted as:

$$(\mathrm{A}.1)\quad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Y represents an ($n \times 1$) vector of observations on the dependent variable, e.g., a 232 × 1 vector of track investment data for Class I railroads over time. **X** is an augmented ($n \times (k + 1)$) matrix of observations of explanatory variables. In this case, **X** is a (232 × [21 + 1]) matrix, in which the first column is a vector of ones corresponding to the implied coefficient of the intercept term. $\boldsymbol{\beta}$ is a ($k + 1$) × 1 vector of parameters to be estimated (including the intercept). The expected value of **Y** ($\mathrm{E}[\mathbf{Y}]$) is $\mathbf{X}\boldsymbol{\beta}$. The variance of **Y** is equal to the assumed-to-be-constant variance (σ^2) times an identity matrix (i.e., $var[\mathbf{Y}] = \sigma^2\mathbf{I}$). The covariance of the errors (e) is assumed to be zero, i.e., $cov(e_t, e_{t-1}) = 0$, which is equivalent to saying that the $cov(Y_t, Y_{t-1}) = 0$. The objective is to minimize the sum of the squared errors (SSE or $\boldsymbol{ee}'$). Since $\mathbf{e} = \mathbf{Y} - \mathbf{X}$, SSE may be expressed as:

$$(\mathrm{A}.2)\quad \mathrm{SSE} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

A.2 is minimized by taking the partial derivative with respect to $\boldsymbol{\beta}$, setting the derivative equal to zero, solving for $\boldsymbol{\beta}$, and verifying that the second derivative is nonnegative.

$$(\mathrm{A}.3)\quad \frac{\partial}{\partial\boldsymbol{\beta}}\mathrm{SSE} = \frac{\partial}{\partial\boldsymbol{\beta}}(\mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta})$$ *Expand the expression*

$$(\mathrm{A}.4)\quad \frac{\partial}{\partial\boldsymbol{\beta}}\mathrm{SSE} = \frac{\partial}{\partial\boldsymbol{\beta}}(2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta})$$ *Take the first derivative*

$$(\mathrm{A}.5)\quad 2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = 0$$ *Set it to zero*

$$(\mathrm{A}.6)\quad \mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y}$$ *Rearrange the expression*

$$(\mathrm{A}.7)\quad \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$ *Solve for* $\boldsymbol{\beta}$

$$(\mathrm{A}.8)\quad \frac{\partial^2}{\partial^2\boldsymbol{\beta}}\mathrm{SSE} = \mathbf{X}'\mathbf{X}$$ *Evaluate the second derivative*

An assumption in OLS regression is that the errors (i.e., the residuals of the regression) are uncorrelated, i.e., their covariance is zero. If this is not true, the OLS parameter estimates may no longer be the minimum variance estimators, in which case an autocorrelation model offers improvements.

In the following paragraphs, a transformation and autoregression modeling process is illustrated for the simple case of first-order autocorrelation. This process—referred to as AR(1)—serves to illustrate a more complex process with higher orders of autocorrelation.

A.2 Autocorrelated Errors

An autocorrelated error term may be envisioned as consisting of two components: (1) an inertial error that is carried forward from the previous time period, and (2) an error that is specific to the current period.[9] The inertial error reflects perceptions of railroad managers about factors outside the model. These unobserved and uncontrolled influences may include perceptions of government policies, regulations, and programs; modal competition; the cost of capital and projected ROI; and a variety of risks. Such perceptions tend to change slowly.

In addition to inertial perceptions, new factors may affect decision making in any given year. Changes in tax policies, stimulus spending, new or revised loan programs and other financial changes not reflected in the model may introduce error disturbances. Changes in the competitive milieu (such as changes in highway funding and truck size and weight regulations) may have similar effects.

A.2.1 Equation of Autocorrelated Error Term

The previous theory of unobserved influences is reflected in Equation A.9, where e_{t-1} represents the inertial error carried forward from the previous time period and ε_t represents the uncorrelated disturbance in the current year.

$$(\text{A}.9) \quad e_t = \rho e_{t-1} + \varepsilon_t$$

Rho (ρ) is the autocorrelation coefficient. In a stationary process, it can assume values < |1.0|. The inertial error carried forward (ρe_{t-1}) must be less than the error in the previous period. This restriction has a practical benefit of preventing the error from increasing without bound.

A.2.2 Error Variance and Correlation

Letting σ_e^2 denote the variance of e_t (the inertial component) and σ_ε^2 represent the variance of ε_t, it can be shown that:

$$(\text{A}.10) \quad \sigma_e^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

Moreover, it can be shown that the error covariance $[cov(e_t, e_{t-1})]$ is equal to $\rho\sigma_\varepsilon^2$ and that ρ is the correlation coefficient that describes the strength of the relationship between e_t and e_{t-1}. If $\rho > 0$ successive errors are positively correlated. If $\rho < 0$ successive errors are negatively correlated. The covariance between errors more than one period apart (i.e., k periods apart) is equal to $\rho^k\sigma_\varepsilon^2$, while ρ^k is the correlation coefficient of errors separated by more than one time period. The errors in the autocorrelation model are homoscedastic because the variance of e_t is equal to $\sigma_\varepsilon^2\,/(1-\rho^2)$, which is the same for all observations.

[9] Griffiths, W., Hill, R. and Judge, G.: *Learning and Practicing Econometrics*, John Wiley and Sons, 1993

A.2.3 Transformations to Achieve Desired Error Properties

The objective of the transformation process is to derive a new equation in which the error term is ε_t instead of e_t. De-emphasizing the indicator variables that affect only the intercept, the regression equation for any observation (except the first one) can be represented as:

$$(A.11) \quad \ln(I_t) = \beta_0 + \beta_1 \ln(MOR_t) + \beta_2 \ln(RGTM_t) + \cdots + \rho e_{t-1} + \varepsilon_t$$

The equation for the previous observation (in period *t*−1) can be denoted as:

$$(A.12) \quad \ln(I_{t-1}) = \beta_0 + \beta_1 \ln(MOR_{t-1}) + \beta_2 \ln(RGTM_{t-1}) + \cdots + e_{t-1}$$

Solving Equation A.12 for e_{t-1}, multiplying both sides of the solved equation by ρ (which results in ρe_{t-1} on the left-hand side), substituting the solved equation for ρe_{t-1} into Equation A.11, and simplifying the results yields a transformed equation in which the modified terms are:

$$(A.13) \quad \ln(I_t^*) = \ln(I_t) - \rho\ln(I_{t-1})$$

$$(A.14) \quad \ln(MOR_t^*) = \ln(MOR_t) - \rho\ln(MOR_{t-1})$$

$$(A.15) \quad \ln(RGTM_t^*) = \ln(RGTM) - \rho\ln(RGTM_{t-1})$$

$$(A.16) \quad \beta_0^* = 1 - \rho$$

β_0^* is the transformed intercept. After transformation, the error term has the following properties: $\varepsilon_t \sim (0, \sigma_\varepsilon^2)$. However, the transformation results in only *n*−1 new observations, leaving the first observation unchanged. Since, the error of the first observation is not linked to previous ones, the equation for the first observation may be written as:

$$(A.17) \quad \ln(I_1) = \beta_0 + \beta_1 \ln(MOR_1) + \beta_2 \ln(RGTM_1) + \cdots + e_1$$

It can be shown that multiplying A.17 by $\sqrt{(1-\rho^2)}$ results in a variance of:

$$(A.18) \quad var(e_1^*) = (1-\rho^2)var(e_1) = (1-\rho^2)\frac{\sigma_\varepsilon^2}{1-\rho^2} = \sigma_\varepsilon^2$$

With this transformation, the errors for all observations have the same desired properties.

A.3 Autocorrelation Modeling Process

The primary steps in the process are:

1. Run the regression
2. Output the residuals (errors) to file
3. Use the outputted errors in a new regression model to estimate the autocorrelation coefficient (ρ)
4. Estimate the transformed regression equation using the estimated value of ρ from the regression in step 3 and the transformed variables shown in Equations A.13–A.16
5. Output the residuals of the regression using the transformed equation to file
6. Return to step 3 and use the outputted residuals from step 5 to estimate a revised value of rho
7. Repeat steps 4–6 until the value of rho from the previous iteration is essentially unchanged

To illustrate step 3, let $\hat{e}$ represent the residuals outputted in step 2. The new regression model can be depicted as $\hat{e}_t = \hat{\beta}_1 \hat{e}_{t-1} + v_t$, where $\hat{\beta}_1$ is an OLS estimate of the autocorrelation coefficient ($\hat{\rho}$) and v_t is an estimate of the uncorrelated component of the error term.

The process described above is broadly referred to as generalized least squares (GLS). SAS AUTOREG uses a matrix algebra procedure to simultaneously estimate a vector of autoregressive parameters that includes many lag variables, not just a single variable corresponding to the first lag period.

A.3.1 Model Specification

In many cases, the form of autocorrelation can be hypothesized from theory or observation. In addition to the AR(1) model, second- and third-order autocorrelation models are frequently hypothesized. In each case, the error process is well understood. In this case, it is not.

In the long run, investment in basic track components is a regular process. However, it can be quite irregular and periodic in the short run. Rails have long lives. When a line is rebuilt with new rail, it may be some time before significant capital investments are made in the line again. Perceptions related to ROI and risks may lag several periods. Reactions to changes in government policies may be cautious and unfold over many years. Inertial forces may extend over several periods, complicated by the scale and cyclical nature of capital investments.

First-order autocorrelation is very likely to be found in the track investment model. However, higher orders of autocorrelation may exist. Given the complex structure of the error covariances, an empirical approach is used. The R-1 database includes 24 years of observations for most railroads. Therefore, 23 lag periods are analyzed. The estimated values of rho ($\hat{\rho}$) are graphed in Figure A.1, which shows autocorrelation throughout much of the period, including significant autocorrelations in lag years 11 through 15.

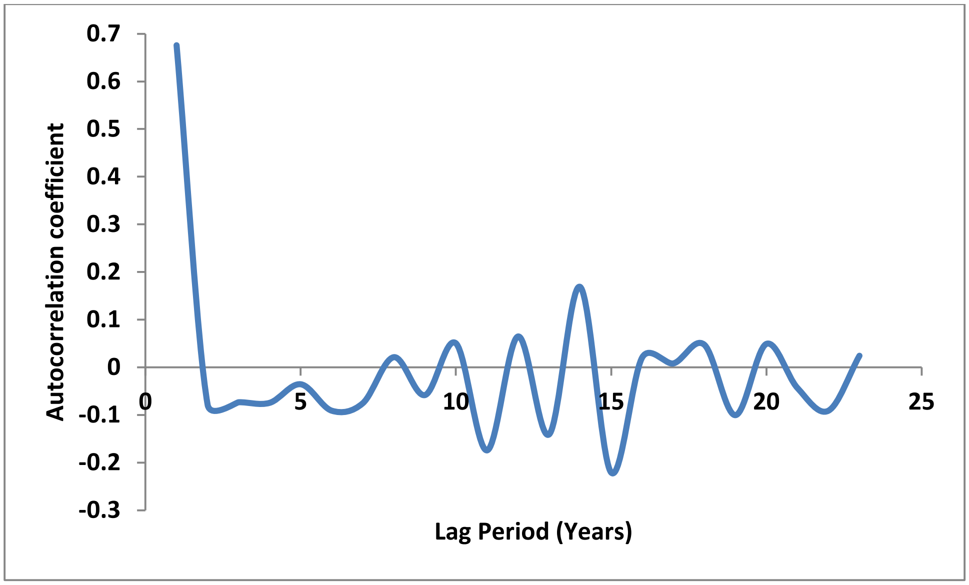

Figure A.14 Autocorrelations in Track Investment Model

A.3.2 Illustrative Manual Process

An iterative solution procedure for an AR(1) process can be derived using PROC REG. The SAS statements used in the first iteration of the process are shown below.

```
proc reg data=track ;
    model LnI=LnMOR LnRGTM LnT;
    output out=gls1(keep=ehat)
    r=ehat /* residuals */;
data gls2;
    set gls1;
    * compute lagged value of residual;
    lage=lag(ehat);
proc reg data=gls2
    outest=rho1
    (keep=lage rename=(lage=rhohat));
    model ehat=lage;
data gls4;
    if _n_=1 then set rho1;
    set track;
    * create lag variables;
    ylag = lag(LnI);
    x1lag = lag(LnMor);
    x2lag = lag(LnRGTM);
    x3lag = lag(LnT);
    * transform variables, including intercept;
    if _n_ = 1 then do;   /* first obs. */
        y = sqrt(1- rhohat**2)*LnI;
        x1 = sqrt(1- rhohat**2)*LnMOR;
        x2 = sqrt(1- rhohat**2)*LnRGTM;
        x3 = sqrt(1- rhohat**2)*LnT;
        int = sqrt(1- rhohat**2);
    end;
    else do;
        y = LnI - rhohat*ylag;
        x1 = LnMOR - rhohat*x1lag;
        x2 = LnRGTM - rhohat*x2lag;
        x3 = LnT - rhohat*x3lag;
        int= 1 - rhohat;
    end;
proc reg data=gls4;
    model y=int x1 x2 x3/noint;
```

This process could be repeated several times by outputting the residuals from the last data step (gls4) and returning to step 3 (gls2), until the estimated value of ρ (rhohat) does not change significantly from the previous iteration. While this process could be automated with an SAS macro, it is inefficient and becomes quite cumbersome when several lag periods are considered. Instead of the manual process, PROC AUTOREG is used. The essential SAS statements are shown below.

```
proc autoreg data=track;
    model LnI=LnMOR LnRGTM LnT ATSF BNSF BN UPSP UPCNW SP CNW SOO ICG ICG89
    GTW GTC KCS CR CSX CSXCR NSCR NS / nlag=23 iter;
```

Before describing how the estimation procedures in PROC AUTOREG work, the generalized least squares (GLS) process is highlighted.

A.3.3 Generalized Least Squares

When the errors of a regression model are correlated, the calculation of the variance as $\sigma^2\mathbf{I}$ is no longer valid. The covariance matrix can no longer be represented as the product of a common (scalar) variance times an identity matrix, which has ones on the diagonal and zeros elsewhere. The off-diagonal elements of the matrix, which represent the covariances among the errors from different time periods, e.g.,$[cov(e_t, e_{t-1})]$, may not be zero. Instead, the variance-covariance matrix resembles A.19, in the case of first-order autocorrelation.

$$(\text{A}.19) \quad \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix} = o^2\mathbf{V}$$

In this situation, the error covariances have a general, but not a specific, form. The variance is equal to $\sigma^2\mathbf{V}$ (as shown above) rather than $\sigma^2\mathbf{I}$. Letting $\mathbf{\Sigma} = \sigma^2\mathbf{V}$, the objective of GLS is to minimize the generalized sum of squares, as shown in A.20.

$$(\text{A}.20) \quad SSE_g = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

Equation A.21 depicts the normal GLS equation, derived in the same manner as before, which can subsequently be solved for $\boldsymbol{\beta}$.

$$(\text{A}.21) \quad (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})\boldsymbol{\beta} = \mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y}$$

With this background, the estimation procedures used in SAS AUTOREG are described.

A.3.4 Iterated Yule-Walker Method

AUTOREG uses what is called the iterated Yule-Walker method, a GLS process in which the OLS residuals are used to estimate the error covariances. Since the autocorrelation coefficient (ρ) is unknown and must be estimated from the residuals, the method may be referred to as estimated generalized least squares (EGLS), a process in which the estimators have desirable large sample or asymptotic properties only. In some publications, the Yule-Walker method has been referred to as the two-step full transform method. For an AR(1) process, Yule-Walker estimates are consistent with Prais-Winsten estimates.

In the Yule-Walker method, the initial (structural) model is augmented with a vector of autoregressive terms. By simultaneously estimating the regression coefficients and the autoregressive terms of the error model, the parameter estimates can be corrected for autocorrelation. In this process, the variance matrix $\mathbf{V}$ is formed from the autoregressive parameters (as illustrated in A.19). Afterward, $\mathbf{\Sigma}$ is computed as $\sigma^2\mathbf{V}$ and efficient parameters estimates are derived via generalized least squares. The estimation of $\boldsymbol{\beta}$ using GLS is alternated with the estimation of $\boldsymbol{\rho}$ (a vector of autocorrelation coefficients), in much the same manner as the manual process described earlier. The method starts by generating the OLS estimates of $\boldsymbol{\beta}$. Next, $\boldsymbol{\rho}$ is estimated from the OLS residuals. $\mathbf{V}$ is estimated from the estimate of $\boldsymbol{\rho}$ and $\boldsymbol{\Sigma}$ is estimated from $\mathbf{V}$ and the OLS estimate of the common variance σ^2. The estimates of the regression parameters $\boldsymbol{\beta}$ (corrected for autocorrelation) are computed via GLS, using the estimated $\boldsymbol{\Sigma}$ matrix. The only difference is that in the iterated method, the steps are repeated until the estimates of $\boldsymbol{\rho}$ are essentially stable. A convergence criterion of 0.001 is used.

Other estimation methods can be used with AUTOREG, including unconditional least squares (also referred to as nonlinear least squares) and maximum likelihood. The maximum likelihood method is recommended in cases where there are many missing values in the data series, which does not apply in this case.

The Yule-Walker estimates were shown earlier in Table 2.7. As shown in Table A.2, the maximum likelihood method yields somewhat different estimates. Nevertheless, the selection of methods does not affect the conclusions of the study.

Table A.15 Parameter Estimates from Autoregression Model Using Maximum Likelihood Method

Variable	Parameter Estimate	Standard Error	t Value	Approx. Pr > |t|
ln(MOR)	0.5886	0.0636	9.26	<.0001
ln(RGTM)	0.5415	0.0869	6.23	<.0001
ln(T)	0.1749	0.0197	8.87	<.0001

A.4 Test for Autocorrelation

The Durbin-Watson test was referred to several times in the paper. The calculation of the Durbin-Watson statistic (D) is illustrated in Equation A.22.

$$(A.22) \quad d = \frac{\sum_{i=2}^{n}(\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^{n}\hat{e}_i^2}$$

Where $\hat{e}_i = y_i - \hat{y}_i$ is the residual for observation "i" from the regression.

A more technically correct statistic for panel datasets has been proposed by Bhargava, et al. (1982).[10] In this approach, the D statistic shown in Equation A.223 is estimated within each cross-sectional class, e.g., each railroad.

$$(A.23) \quad d_{pd} = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T}\left(\hat{e}_{i,t} - \hat{e}_{i,t-1}\right)^2}{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{e}_{i,t}^2}$$

Both statistics have been calculated in this study. Because the time period is 24 years, the two approaches produce essentially the same results. Therefore, the Durbin-Watson statistic generated by the SAS software is reported in this study. Before a regression analysis is run, the data are sorted by railroad and year—a prerequisite for the calculation of either statistic and the running of PROC AUTOREG.

[10] Bhargava, A., Franzini, L., and W. Narendranathan. "Serial Correlation and the Fixed Effects Model." Review of Economic Studies (1982), XLIX, pp. 533-549.

APPENDIX B: DATA

Table B.16 Values of Variables Used in Study

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1985	ATSF	$1,699.49	$1,699.49	11,869	157,684
1986	ATSF	$1,799.32	$1,798.55	11,661	151,567
1987	ATSF	$1,906.36	$1,901.53	11,709	162,949
1988	ATSF	$1,964.07	$1,954.27	11,652	176,361
1989	ATSF	$1,999.36	$1,985.16	11,266	184,985
1990	ATSF	$1,984.65	$1,972.90	10,650	174,188
1991	ATSF	$1,990.52	$1,977.58	9,639	183,349
1992	ATSF	$1,997.78	$1,983.24	8,750	191,942
1993	ATSF	$2,129.92	$2,085.27	8,536	204,541
1994	ATSF	$2,300.21	$2,214.80	8,352	219,414
1995	ATSF	$3,971.82	$3,429.94	9,126	231,333
1985	BN	$3,903.63	$3,903.63	26,780	372,141
1986	BN	$4,000.85	$4,000.10	25,539	379,314
1987	BN	$3,955.10	$3,956.09	23,476	405,198
1988	BN	$3,849.95	$3,860.00	23,391	418,197
1989	BN	$4,140.81	$4,114.62	23,356	430,129
1990	BN	$4,379.87	$4,313.86	23,212	444,586
1991	BN	$4,404.02	$4,333.15	23,088	434,152
1992	BN	$4,572.68	$4,464.43	22,786	432,203
1993	BN	$4,740.11	$4,593.72	22,316	441,711
1994	BN	$4,937.18	$4,743.61	22,189	477,845
1995	BN	$5,155.14	$4,902.05	22,200	529,042
1996	BNSF	$9,758.43	$8,119.94	35,208	732,330
1997	BNSF	$10,118.87	$8,368.03	33,757	838,302
1998	BNSF	$11,192.65	$9,096.80	33,353	921,062
1999	BNSF	$11,371.33	$9,218.35	33,264	949,469
2000	BNSF	$11,910.07	$9,554.09	33,386	958,576
2001	BNSF	$12,429.78	$9,869.05	33,063	981,469
2002	BNSF	$13,429.25	$10,470.39	32,525	955,477
2003	BNSF	$14,117.32	$10,869.94	32,266	991,230
2004	BNSF	$14,820.72	$11,256.99	32,150	1,106,373
2005	BNSF	$15,711.84	$11,691.92	32,154	1,158,305

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
2006	BNSF	$17,558.73	$12,547.49	31,910	1,223,757
2007	BNSF	$18,472.58	$12,951.87	32,205	1,227,033
2008	BNSF	$19,542.32	$13,368.06	32,166	1,224,930
1985	CNW	$619.41	$619.41	7,301	53,455
1986	CNW	$640.60	$640.43	6,305	56,303
1987	CNW	$665.22	$664.12	6,214	56,446
1988	CNW	$664.63	$663.58	5,794	61,037
1989	CNW	$552.68	$565.58	5,650	57,719
1990	CNW	$584.67	$592.25	5,624	56,911
1991	CNW	$618.93	$619.61	5,573	57,540
1992	CNW	$673.47	$662.06	5,419	59,830
1993	CNW	$724.44	$701.42	5,337	65,651
1994	CNW	$756.69	$725.95	5,211	71,018
1985	CR	$3,014.27	$3,014.27	14,025	174,647
1986	CR	$3,322.99	$3,320.62	13,739	175,305
1987	CR	$3,539.05	$3,528.49	13,341	188,585
1988	CR	$3,794.03	$3,761.48	13,111	198,112
1989	CR	$3,991.19	$3,934.07	13,068	191,552
1990	CR	$4,115.38	$4,037.57	12,828	193,964
1991	CR	$3,765.09	$3,757.82	12,454	187,539
1992	CR	$4,055.61	$3,983.96	11,895	193,025
1993	CR	$3,310.39	$3,408.51	11,831	200,936
1994	CR	$3,369.21	$3,453.25	11,349	217,930
1995	CR	$3,171.24	$3,309.34	10,701	211,182
1996	CR	$3,034.50	$3,213.75	10,543	215,110
1997	CR	$3,114.26	$3,268.65	10,801	220,096
1998	CR	$3,169.19	$3,305.93	10,797	226,994
1985	CSX	$3,110.62	$3,110.62	23,945	299,388
1986	CSX	$3,534.61	$3,531.36	22,887	288,572
1987	CSX	$3,628.41	$3,621.60	21,494	310,651
1988	CSX	$3,743.39	$3,726.67	20,376	315,604
1989	CSX	$3,830.15	$3,802.62	19,565	293,003
1990	CSX	$4,152.94	$4,071.65	18,943	318,267
1991	CSX	$4,470.51	$4,325.27	18,854	295,766
1992	CSX	$4,418.19	$4,284.55	18,905	309,593

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1993	CSX	$4,771.31	$4,557.22	18,779	317,469
1994	CSX	$4,926.64	$4,675.36	18,759	333,507
1995	CSX	$5,049.52	$4,764.69	18,645	343,071
1996	CSX	$5,394.76	$5,006.03	18,504	345,489
1997	CSX	$5,494.33	$5,074.56	18,285	356,293
1998	CSX	$5,742.23	$5,242.81	18,181	363,024
1999	CSX	$6,024.30	$5,434.69	23,357	474,249
2000	CSX	$6,467.96	$5,711.18	23,320	497,518
2001	CSX	$6,598.98	$5,790.59	23,297	489,717
2002	CSX	$6,817.81	$5,922.24	23,160	467,258
2003	CSX	$7,194.58	$6,141.02	22,841	485,501
2004	CSX	$10,272.84	$7,834.87	22,153	507,184
2005	CSX	$10,390.04	$7,892.07	21,357	501,575
2006	CSX	$10,811.76	$8,087.43	21,114	510,137
2007	CSX	$11,273.38	$8,291.70	21,166	493,041
2008	CSX	$11,745.58	$8,475.41	21,204	483,792
2002	GTC	$3,721.43	$2,306.01	6,390	104,014
2003	GTC	$3,701.34	$2,294.34	6,493	105,363
2004	GTC	$4,013.63	$2,466.18	6,822	109,589
2005	GTC	$4,176.46	$2,545.66	6,736	109,498
2006	GTC	$4,314.97	$2,609.82	6,737	111,835
2007	GTC	$4,353.93	$2,627.06	6,738	110,833
2008	GTC	$4,538.17	$2,698.74	6,738	108,413
1985	GTW	$138.46	$138.46	1,310	14,201
1986	GTW	$139.81	$139.79	1,311	14,021
1987	GTW	$124.17	$124.76	943	13,380
1988	GTW	$120.89	$121.76	931	13,905
1989	GTW	$128.91	$128.78	959	14,436
1990	GTW	$137.82	$136.20	927	14,096
1991	GTW	$144.02	$141.15	925	13,277
1992	GTW	$156.14	$150.59	925	14,008
1993	GTW	$173.99	$164.37	925	15,998
1994	GTW	$186.51	$173.89	925	16,715
1995	GTW	$189.31	$175.93	916	16,265
1996	GTW	$201.00	$184.10	918	23,514

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1997	GTW	$189.78	$176.38	659	25,165
1998	GTW	$211.04	$190.81	646	23,877
1999	GTW	$239.52	$210.18	628	24,829
2000	GTW	$278.93	$234.74	627	26,645
2001	GTW	$298.19	$246.42	627	28,015
1985	ICG	$843.26	$843.26	4,772	54,016
1986	ICG	$780.42	$780.90	3,788	42,052
1987	ICG	$705.10	$708.44	3,205	36,262
1988	ICG	$686.17	$691.14	2,900	36,652
1989	ICG	$349.26	$396.22	2,887	36,086
1990	ICG	$351.10	$397.75	2,773	35,137
1991	ICG	$363.76	$405.46	2,766	37,037
1992	ICG	$376.43	$417.66	2,732	35,207
1993	ICG	$393.16	$430.58	2,717	37,690
1994	ICG	$407.82	$441.73	2,665	39,290
1995	ICG	$430.47	$458.20	2,642	45,337
1996	ICG	$448.65	$470.90	2,623	41,905
1997	ICG	$483.49	$494.89	2,598	42,128
1998	ICG	$511.64	$514.00	2,593	44,463
1999	ICG	$543.75	$535.84	2,591	46,451
2000	ICG	$577.39	$556.80	2,544	50,168
2001	ICG	$609.09	$576.01	2,544	50,094
1985	KCS	$263.88	$263.88	1,661	24,101
1986	KCS	$272.05	$271.99	1,666	22,999
1987	KCS	$292.09	$291.27	1,665	23,095
1988	KCS	$301.02	$299.43	1,681	22,878
1989	KCS	$366.14	$356.43	1,681	23,056
1990	KCS	$391.16	$377.28	1,681	23,669
1991	KCS	$389.98	$376.34	1,682	23,678
1992	KCS	$411.18	$392.84	1,680	25,459
1993	KCS	$433.71	$410.24	1,712	26,313
1994	KCS	$700.50	$613.16	2,880	33,412
1995	KCS	$753.51	$651.69	2,931	37,697
1996	KCS	$769.70	$663.01	2,954	36,916
1997	KCS	$769.26	$662.71	2,845	38,335

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1998	KCS	$784.21	$672.86	2,756	41,956
1999	KCS	$803.69	$686.11	2,756	42,986
2000	KCS	$812.23	$691.43	2,701	38,434
2001	KCS	$887.89	$737.28	3,102	39,271
2002	KCS	$931.52	$763.53	3,084	37,494
2003	KCS	$950.62	$774.63	3,084	41,104
2004	KCS	$1,009.75	$807.16	3,072	43,630
2005	KCS	$1,094.60	$848.57	3,197	53,943
2006	KCS	$1,118.58	$859.68	3,176	55,721
2007	KCS	$1,290.98	$935.97	3,151	54,431
2008	KCS	$1,506.99	$1,020.01	3,165	53,501
1985	NS	$2,638.60	$2,638.60	17,620	202,461
1986	NS	$2,733.98	$2,733.25	17,520	200,234
1987	NS	$2,707.71	$2,707.98	17,254	203,048
1988	NS	$2,877.33	$2,862.98	17,006	208,730
1989	NS	$3,026.94	$2,993.94	15,955	209,196
1990	NS	$3,417.24	$3,319.24	14,842	218,678
1991	NS	$3,458.20	$3,351.95	14,721	211,409
1992	NS	$3,637.36	$3,491.40	14,703	221,153
1993	NS	$3,755.77	$3,582.83	14,589	228,558
1994	NS	$4,036.88	$3,796.65	14,652	246,101
1995	NS	$4,228.76	$3,936.13	14,407	255,330
1996	NS	$4,505.18	$4,129.36	14,282	261,810
1997	NS	$4,628.21	$4,214.04	14,415	270,247
1998	NS	$4,633.74	$4,217.79	14,423	272,617
1999	NS	$4,728.44	$4,282.22	21,788	396,548
2000	NS	$4,751.58	$4,296.64	21,759	408,243
2001	NS	$4,833.75	$4,346.44	21,569	377,468
2002	NS	$5,077.69	$4,493.20	21,558	372,260
2003	NS	$5,221.11	$4,576.48	21,520	378,836
2004	NS	$9,302.68	$6,822.41	21,336	406,904
2005	NS	$9,613.29	$6,974.01	21,184	415,827
2006	NS	$9,781.08	$7,051.74	21,141	417,423
2007	NS	$10,156.48	$7,217.85	20,890	398,857
2008	NS	$10,665.95	$7,416.07	20,831	391,457

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1985	SOO	$535.27	$535.27	7,975	39,346
1986	SOO	$539.12	$539.09	7,747	42,956
1987	SOO	$396.23	$401.63	5,809	43,689
1988	SOO	$414.77	$418.57	5,807	39,856
1989	SOO	$420.99	$424.01	5,770	39,235
1990	SOO	$501.96	$491.49	5,293	43,707
1991	SOO	$502.27	$491.74	5,045	43,600
1992	SOO	$549.52	$528.52	5,033	43,556
1993	SOO	$573.62	$547.13	5,062	43,730
1994	SOO	$597.25	$565.11	5,139	40,327
1995	SOO	$351.66	$386.57	5,130	48,893
1996	SOO	$390.15	$413.48	4,980	48,547
1997	SOO	$453.06	$456.79	3,364	41,566
1998	SOO	$500.93	$489.27	3,358	40,282
1999	SOO	$520.57	$502.63	3,261	40,639
2000	SOO	$481.61	$478.35	3,225	43,329
2001	SOO	$505.98	$493.12	3,225	45,281
2002	SOO	$547.28	$517.97	3,225	45,427
2003	SOO	$562.96	$527.08	3,258	48,191
2004	SOO	$602.56	$548.86	3,251	49,945
2005	SOO	$631.23	$562.86	3,511	47,713
2006	SOO	$670.52	$581.06	3,267	48,323
2007	SOO	$697.59	$593.04	3,267	48,670
2008	SOO	$735.55	$607.81	3,267	46,122
1985	SP	$2,491.11	$2,491.11	15,624	200,706
1986	SP	$2,651.37	$2,650.14	15,194	194,792
1987	SP	$2,712.88	$2,709.32	15,046	203,470
1988	SP	$3,062.36	$3,028.67	15,023	210,530
1989	SP	$3,332.37	$3,265.03	15,023	220,390
1990	SP	$3,498.59	$3,403.57	14,846	215,851
1991	SP	$3,554.38	$3,448.12	14,389	214,183
1992	SP	$3,721.61	$3,578.29	14,389	233,049
1993	SP	$3,819.43	$3,653.82	14,099	246,077
1994	SP	$3,620.63	$3,502.62	13,715	268,935
1995	SP	$3,668.91	$3,537.72	15,388	288,759

Year	Railroad	Nominal Investment (Millions)	Real Investment Based on RCRI (Millions)	Miles of Road	Rev. Gross Ton-Miles (Millions)
1996	SP	$3,907.74	$3,704.67	14,404	306,641
1985	UP	$1,566.95	$1,566.95	24,259	297,972
1986	UP	$2,967.25	$2,956.52	24,793	304,890
1987	UP	$2,999.35	$2,987.41	24,074	359,644
1988	UP	$3,379.96	$3,335.20	22,653	388,379
1989	UP	$3,508.62	$3,447.83	21,882	389,286
1990	UP	$3,778.91	$3,673.10	21,128	404,333
1991	UP	$3,838.58	$3,720.76	20,261	419,745
1992	UP	$4,044.66	$3,881.17	19,020	436,341
1993	UP	$4,279.23	$4,062.29	17,835	460,359
1994	UP	$4,517.95	$4,243.87	17,499	492,756
1995	UP	$6,668.97	$5,807.51	22,785	626,250
1996	UP	$6,812.47	$5,907.83	22,266	658,322
1997	UP	$9,656.12	$7,865.15	34,946	939,906
1998	UP	$10,588.35	$8,497.84	33,706	905,103
1999	UP	$11,282.92	$8,970.35	33,341	987,482
2000	UP	$12,026.13	$9,433.51	33,035	1,020,951
2001	UP	$12,610.23	$9,787.49	33,586	1,046,395
2002	UP	$13,369.19	$10,244.12	33,141	1,080,195
2003	UP	$14,021.04	$10,622.64	32,831	1,105,236
2004	UP	$14,762.22	$11,030.48	32,616	1,123,480
2005	UP	$15,609.93	$11,444.22	32,426	1,134,716
2006	UP	$16,479.55	$11,847.07	32,339	1,169,215
2007	UP	$17,319.48	$12,218.75	32,205	1,148,521
2008	UP	$18,470.82	$12,666.68	32,012	1,111,650