ISSN 1046-1469
Indexed by PAIS

JOURNAL OF THE TRANSPORTATION RESEARCH FORUM

Volume 33 Number 1 **1993**

TRANSPORTATION RESEARCH FORUM

TRACK MAINTENANCE COSTS AND ECONOMIES OF DENSITY: AN ANALYSIS OF THE SPEED FACTORED GROSS TONNAGE MODEL

by Randolph R. Resor and Michael E. Smith***

ABSTRACT

It has traditionally been assumed in the railroad industry that unit costs decrease as traffic volume increases. This assumption is behind the trend to multi-car rates, unit trains, and the consolidation of more and more traffic on fewer mainlines. However, there is some reason to believe that economies of density, at least in maintenance of way, may not be as extensive as has been assumed.

One widely used MOW cost model is the Speed Factored Gross Tonnage model (SFGT). SFGT assumes that MOW costs are related to the square root of traffic density, producing a continuously declining marginal cost as traffic increases. SFGT has been used in a number of ICC cases concerning the incremental cost of new coal and grain traffic. The result has been lost revenue for the railroads.

This paper will examine the assumptions embodied in the SFGT model. From economic, engineering, strategic, and statistical viewpoints, it will be shown that continuously declining marginal costs are an impossibility, and that the assumption of continuously declining costs may have had serious consequences for the railroad industry's strategic planning.

Alternative, and more reasonable, cost functions will be proposed. One possible alternative maintenance of way cost allocation model formulation will be presented and discussed.

The conclusion of the paper is that the railroad industry in North America has likely reached a point where marginal track maintenance cost is increasing, rather than decreasing.

INTRODUCTION

For more than a hundred years, United States railroads have been regulated by the Interstate Commerce Commission (ICC). Among its many Congressionally mandated duties, the ICC has been charged with regulating railroad rates. While much of this rate regulatory authority was removed by the Staggers Act of 1980, some still remains. In particular, where railroads are found to have "market dominance" (a term whose meaning has been the subject of vigorous debate), rates may not exceed railroad "variable" cost by more than a set percentage. If the percentage is found by the ICC to have been exceeded, shippers may obtain refunds from railroads.

In order to determine prices, the ICC must have a way of defining railroad costs. As far back as the Transportation Act of 1920, this problem was addressed by Congress, which gave the ICC the power to mandate accounting rules and practices (and ultimately an entire accounting structure) for railroads. Rail Form A, as it became known, came into use in the 1920s and continued in use, in limited form, until very recently. Despite many criticisms, the ICC was slow to replace Rail Form A because it provided a consistent basis for regulatory decisionmaking. Only after the bankruptcies of the Penn Central and other Northeastern rail roads was serious consideration given to a new accounting system. Called URCS (the Uniform Railway Costing System), the new system is now being phased into use.

A full discussion of the differences between URCS and Rail Form A is beyond the scope of this paper. However, it is worth noting that a major accounting change in 1985 allowed railroads, for the first time, to depreciate capital investments in ICC reporting, rather than using the "betterment" system originally included in Rail Form A. The use of depreciation accounting is of particular significance in maintenance of way costing, since expenditures are intermittent and cyclical. MOW spending has been much more consistent since railroads were permitted to smooth out the annual variations through use of depreciation.

The ICC, in Rail Form A, used a series of regression equations to divide each category of expense into a "fixed" and a "variable" component. Since the analysis was cross-sectional across all Class I railroads, it was argued that the calculated variabilities were probably inaccurate for at least some of the railroads in the sample. URCS attempted to correct the problem by carrying out separate regressions for each railroad. However, a larger problem is that the regression-determined "variabilities" apply only at the aggregate tonnage and expenditure levels of each railroad at the time the data sample was obtained. These point estimates tell economists nothing about the change in MOW costs as traffic volume on a particular line (as opposed to the whole railroad) increases or decreases[1].

Variable track costs are very difficult to measure. Tie decay, loss of line and surface, and even (to an extent) rail degradation depend upon environment as well as traffic. Replacements are cyclical, so expenditures are "lumpy". Replacement cycles are long; rail may last 60 years or more. How can one determine, with a reasonable degree of accuracy, how much damage a single train may have inflicted on the track structure? Given that the damage can be quantified, what is the cost of the repair?

Controversy over railroad rates in market dominance cases has focused not on the overall level of railroad costs, but on the incremental costs incurred by a railroad to carry a specific traffic over a defined route. The largest number of cases has concerned the transportation of coal, in particular Western coal. Since 1977, when commercial mining began in the Powder River Basin in Wyoming, coal volume has increased enormously. Wyoming now accounts for more than 25% of all the coal mined in the United States.

In establishing rates for this new traffic, both shippers and railroads agreed that Rail Form A and URCS were inadequate tools for determining cost. Both ICC costing systems use historical data, and therefore are unable to produce meaningful results for a movement of new traffic on a new railroad. The railroads attempted to recover their costs for constructing and upgrading lines and purchasing equipment to handle this new traffic. This tactic produced very high marginal costs. Shippers, of course, argued that many of the investments would have been made in any case, and therefore were not "solely related" to the "issue traffic" (coal). Ultimately these arguments were settled in courts of law, but as the litigation continued (and still continues), a number of curious costing systems were developed by both sides to bolster their respective positions.

The focus of this paper is the determination of the relationship between traffic levels and maintenance of way costs, specifically the incremental MOW costs associated with additional traffic. To explore this relationship, the Speed Factored Gross Tonnage (SFGT) model, which has been widely used for twenty years for determination of incremental track maintenance costs, will be examined. The assumptions in SFGT regarding the behavior of costs and track component lives will be discussed in light of both economic theory and recent engineering research. The statistical basis for the SFGT model formulation will be explored. Finally, an alternative model formulation based on engineering research will be proposed.

THE SPEED FACTORED GROSS TONNAGE MODEL

The costing system most generally used in coal rate litigation was the Speed Factored Gross Tonnage (SFGT) model. SFGT was originally developed in the mid 1970s for the Interstate Commerce Commission's Rail Services Planning Office (RSPO). RSPO, created by Congress after the passage of the Amtrak act in late 1970, was intended to act as an ombudsman and consumer advocate for rail passengers. RSPO was also charged with determining a fair compensation to freight railroads for operating passenger trains. SFGT was developed to provide a methodology for allocating track maintenance costs between passenger and freight trains. As originally specified, the model considered only operating speed and amount of annual traffic (gross tons) moving over the line. Other variables that might significantly affect cost, such as axle load, were not included in the model.

SFGT was based on work conducted by an AREA committee in 1956[2]. This work, essentially a cross-sectional analysis of a number of Class I railroads, attempted to construct curves relating traffic density and MOW costs for groups of railroads sharing similar weights of rail in track and (to the extent possible) traffic types. Traffic densities varied, but most were in the 7 to 15 MGT range, very low by today's standards. Only a few data points were obtained at the 25 MGT and greater levels that characterize most of today's main lines. From this work, it was hypothesized that maintenance of way costs increased, in general, as the square root of traffic density. This assumption was included in the 1975 SFGT model.

Over the last decade, SFGT has been applied in a number of Interstate Commerce Commission cases having nothing to do with passenger service. The Staggers Act set ceilings on many bulk commodity rates, based on "direct variable" cost. Determination of direct variable cost required the use of some sort of cost allocation methodology. Rail Form A was agreed to be inadequate for the task, and SFGT was familiar and available. So, despite its having been developed specifically to identify track maintenance costs allocable to passenger trains, it was used to estimate the incremental additional track maintenance costs allocable to the operation of heavy unit coal, grain, and ore trains[3].

SFGT has been "fine tuned" several times during this period, to suit it better for estimating the incremental cost of heavy freight traffic. A separate category for heavy axle load traffic has been established, an improved tie life equation introduced, and the rail/OTM and ballast/surfacing equations modified to increase the model's sensitivity to differences in axle loads between traffics. However, the coefficients and constants in the equation remain those developed from railroad data now more than 30 years old, and the basic square-root-of-density relationship remains unchanged. Table 1 shows the SFGT equation in its most current form.

The large constants in the various SFGT equation terms have the effect -- especially in the tie model -- of relating a large part of total maintenance expense to miles of track, rather than to any traffic measure. Figures 1, 2, and 3 show the SFGT-determined cost per route mile for ties, rail and OTM, and ballast/surfacing, respectively, at tonnages varying from one million gross tons per year to 80 million. Also illustrated is the model's sensitivity to variations in percentage of curved track (for rail) and traffic speed (for ballast/surfacing). Interestingly, for a model with "speed" in its name SFGT is relatively insensitive to traffic speed.

Several observations can be made about the costs shown in Figures 1, 2, and 3. The first is that the SFGT "tie life model" (a 1977 enhancement to the original equation) is almost totally insensitive to variations in traffic density. An eighty-fold increase in traffic produces (as can be seen from Figure 1) only a 40% increase in tie renewal costs. This relationship runs counter to all recent research[4], and results from the use of a large constant in the equation (implying a large fixed cost, even if no traffic uses the track) and the square-root-of-density relationship, which insures that the constant term remains important even at very high traffic densities.

Rail and "other track material" (OTM -- spikes, plates, bolts, joint bars, and special work) costs escalate somewhat more rapidly with tonnage than do tie costs. Over the same eighty-fold traffic increase, rail and OTM costs as calculated by SFGT increase about five times. This relationship is interesting in light of much recent research which finds rail wear to be linear with increasing tonnage[5]. In other words, one additional unit of traffic causes one additional unit of rail wear. The SFGT curve plotted in Figure 2 implies that rail life (in millions of gross tons) is sixteen times greater at 80 million gross tons per year than at one million. The 1956 Association of American Railroads work often cited as the source for the square-root-of-density relationship in SFGT indicates that rail life in MGT triples as annual tonnage increases -- a very different number from a sixteen-fold increase[6]. However, even this relationship has been called into question by recent research.

The apparent relative insensitivity of rail life to tonnage variations in the SFGT model is probably due to the fact that SFGT was calibrated with data from low-tonnage lines composed entirely of bolted rail. The SFGT rail formula simply does not work on high-tonnage routes with primarily welded rail. This is a significant problem, since most recent applications of SFGT have been to determine the incremental cost of running unit train traffic -- usually coal -- over already busy lines. Again, in the rail equation as in the tie life model, a large constant term and the square-root-of- density relationship combine to reduce predicted incremental costs.

In the most recent SFGT formulation, the rail equation includes separate curve wear terms for heavy axle load traffic, general freight traffic, and passenger traffic, suggesting that

Figure 1

Tie Cost Sensitivity to Traffic -- SFGT

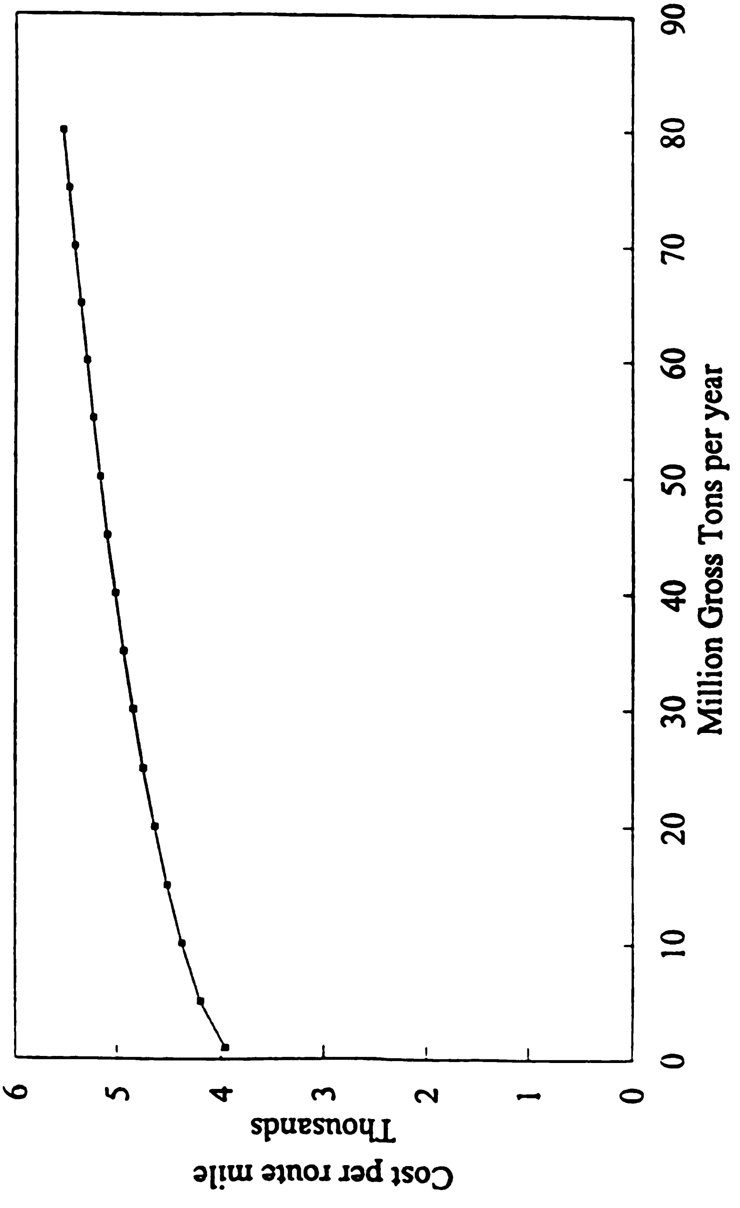


Figure 2

Rail Sensitivity to Curvature -- SFGT

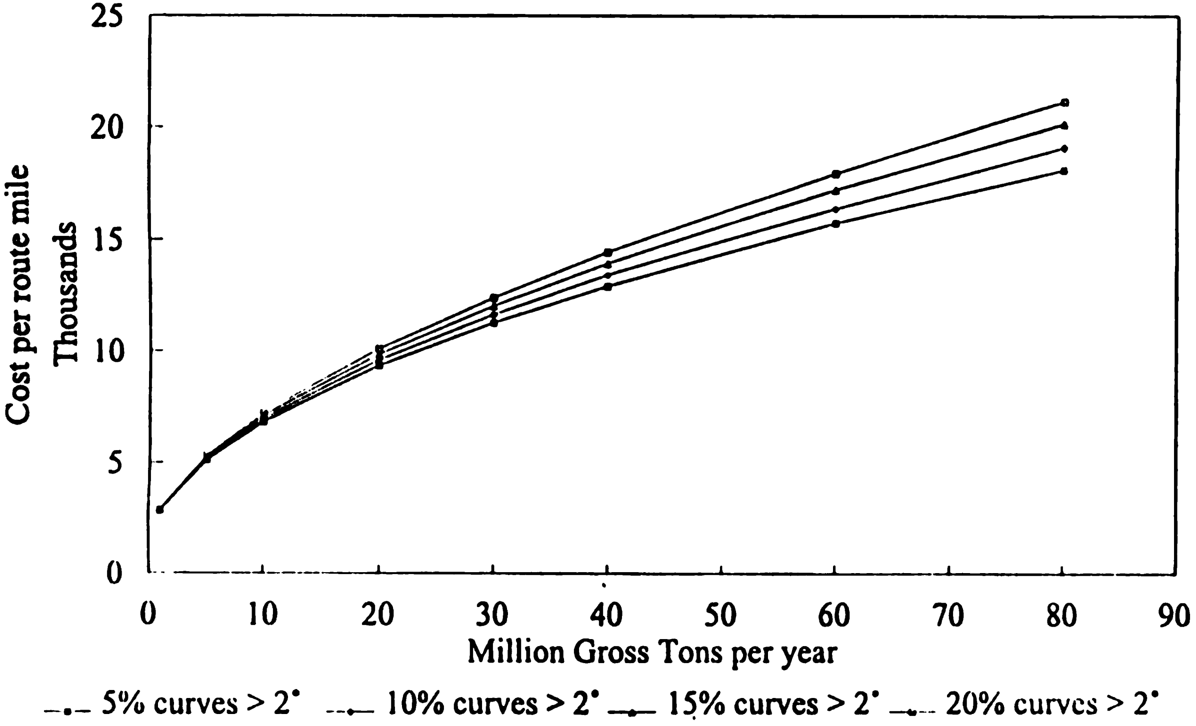


Figure 3

Ballast Sensitivity to Speed -- SFGT

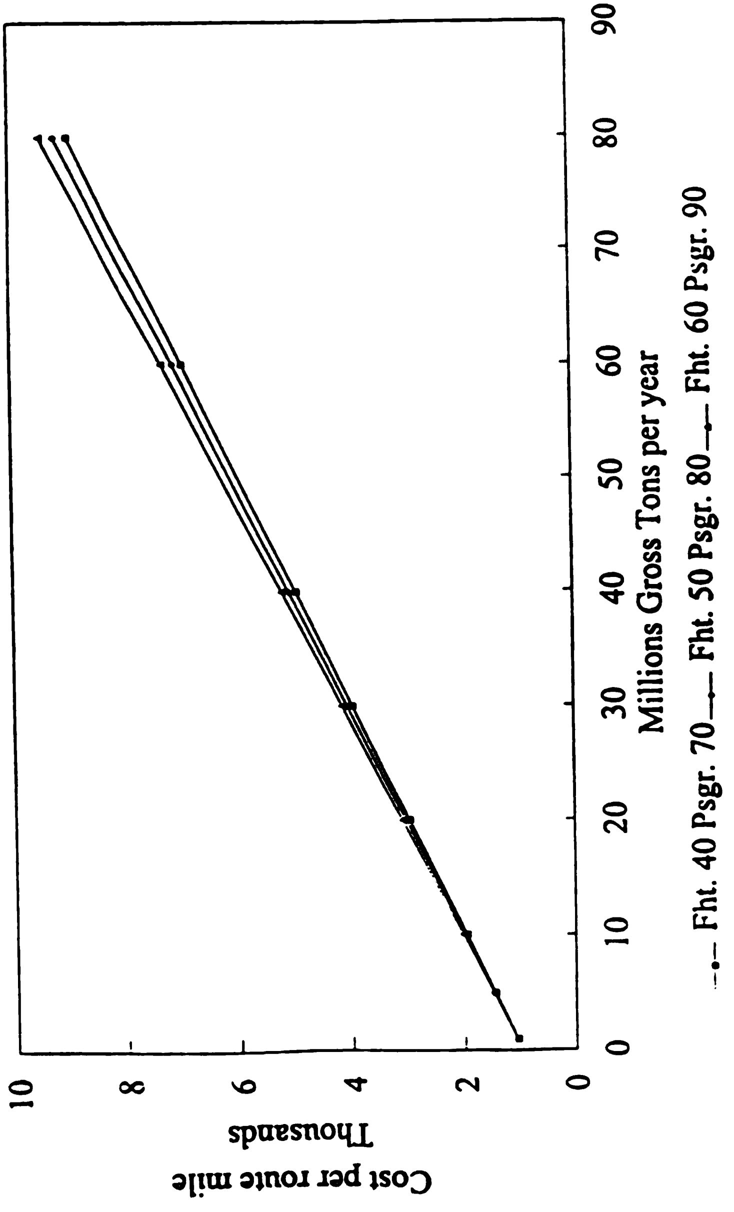


Table 1
SFGT Model used in Eastern Coal Export Case

General Form: Total Indexed Costs $E = R(A + B + C + D)$ (per route mile)

A. Roadway $A = Y(670 + 910N)$

B. Ties $B = N(1880 + 100(U + H + G + P)^{0.5}$
(for $U + H + G + P > 1$)

$B = N(940 + 100(U + H + G + P)^{0.5}$
(for $U + H + G + P < 1$)

C. Rail/OTM $C = N(360 + 900(U + H + G + P)^{0.5} + 96 + 144CV)G + (24 + 72CV)P$

D. Ballast $D = N(480 + 48J(1.26U(1 + VU/600 + VU^2/6000) + (1.26(H + G)(1 + VG/600 + VG^2/6000) + P(1 + VP/750 + VP^2/9375))$

where:
N = number of tracks per route-mile
U = unit train traffic, MGT per track mile
H = heavy wheel load traffic, MGT per track mile
G = other fht. traffic, MGT per track mile
P = passenger traffic, MGT per track mile
CV = percentage of track in curves of 2 degrees or greater
J = 1.5 - .5 (CWR miles/total track miles)
VU = operating speed (mph), unit trains
VG = operating speed (mph), other freight
VP = operating speed (mph), passenger trains
R = index to include MOW and general overheads, adjust for differences in maintenance practices, and inflate costs to reflect current price levels (underlying coefficients are at 1975 price levels and exclude overheads)
Y = adjustment to current price level (see table below)

	——— Y Factor ———	
FRA Track Class and Type of Operation	**Main Line**	**Branch**
Class 1, 2, 3: freight only < 10 MGT/yr	1.00	0.56
Class 1, 2: passenger, or Class 1, 2, 3: freight, > 10 MGT/yr	1.12	0.66
Class 3: passenger, or Class 4, 5, 6: all traffic	1.15	0.69

SFGT's authors are beginning to recognize the lack of sensitivity in the model. However, these terms are of relatively minor importance, as the sensitivity analysis in Figure 2 indicates. Further, these curve terms are the only place where axle load enters into the model in any meaningful way. In the ballast equation, the coefficients for U (unit train traffic) and H (heavy axle load traffic) are identical to those for G (general freight); the terms can simply be combined.

SFGT exhibits the greatest sensitivity to traffic increases in the costing of ballast and surfacing. Here, an increase in annual tonnage from one million to eighty million gross tons increases maintenance expenses seven and a half times. The ballast/surfacing equation includes the speed term from which SFGT takes its name. However, as with the curve term in the rail equation, the effect of the speed term is small. And, as mentioned above, the coefficients in the speed term are the same for all categories of freight traffic.

Finally, there is the matter of the "R Factor". In the first few applications of SFGT, attempts were made to index costs for inflation. This effort was soon abandoned in favor of the use of an R Factor. This factor, described as an "index to include maintenance of way and general and administrative overhead and to [adjust to] current cost level and maintenance practice" is actually a simple "make-whole" adjustment. Since the regression coefficients were derived from 1974 data, the total expenditures produced by SFGT for any other year may bear no resemblance to actual total expenditures for any railroad. Therefore, the make-whole R factor is applied to equate SFGT-produced costs with actual expenditures. Table 2 shows the R factors calculated for several representative model applications.

Table 2

R Factors for Recent Applications
of the SFGT Model

Railroad	1979	1981	1982	1983
BN	--	--	2.850	2.766
UP	--	2.868	2.026*	--
CNW	--	--	--	2,259
L&N	2.065	--	--	--

*In another model application for Union Pacific in the same year, R factor calculated as 2.082.

The variations in these factors raise some interesting questions. Is it really plausible that the Union Pacific achieved such efficiencies that, from 1981 to 1982, the R factor decreased by nearly 30%? This was, it might be recalled, at a time of high inflation. The presence of two different R factors for the Union Pacific Railroad for the same year is also worthy of note.

The explanation supplied in the SFGT documentation (that R is some kind of adjustment for overheads, material costs, maintenance practices, and inflation) is questionable in light of the above table. It is doubtful that costs, practices, overheads, or inflation change from year to year the way the R factors appear to. It is more likely that coefficients derived by regression analysis of one railroad's data for one year are inapplicable to other railroads in other years. Even the Railway Accounting Principles Board has recognized the need to calculate cost functions separately for individual railroad companies, and a recent paper by two noted statisticians has concurred[7]. SFGT claims to take these differences into account through use of the R factor; however, the use of the same coefficients and constants for every railroad will produce the same incremental costs for every railroad (holding traffic mix and track characteristics constant). Total costs will differ only because of the application of the R factor.

The most critical assumption of all in the SFGT model, however, is the square-root-of-density relationship. In practice, this relationship means that the incremental cost of all traffic on lines with average annual tonnage of about 25 MGT or more will be assigned a cost lower than the railroad's system average MOW cost for all traffic. As railroads continue to concentrate traffic on fewer and fewer lines, this frequently results in substantially more than half a railroad's traffic moving at below system average cost, a prima facie impossibility.

Equally disturbing are the economic implications of a continuously declining marginal cost. The least of these implications is that the economically optimum output level of such a firm is infinity. The economics of SFGT will be discussed in a later section.

FOUR PERSPECTIVES ON THE SFGT MODEL

There are four categories of shortcomings in the SFGT formulation, and each is discussed in turn in the following sections:

1. ECONOMIC: The SFGT relationships between cost and output violate fundamental economic principles
2. ENGINEERING: The SFGT formulation is inconsistent with the finding of research into the mechanisms producing track component deterioration.

3. STRATEGIC: The SFGT cost relationships lead to regulatory decisions that unfairly reduce railroads' ability to recover costs. Use of these cost relationships in railroad managerial costing may also result in policy decisions that may not be in the best interests of the railroads.
4. STATISTICAL: The SFGT formulation is a classic example of how not to do regression analysis. The formula was produced by postulating a relationship that happened to fit some empirical data, without regard to the underlying principles governing the causal relationships that were being modeled.

An Economic Perspective

The concept of scale economies is recognized as applying to many human endeavors. In general, average costs decline as volume increases because certain fixed costs can be spread over an ever-larger output. Note, however, that the decline in average costs does not necessarily require a decline in marginal cost (the cost of producing one additional unit of output). In fact, increasing marginal costs are generally assumed at all but the very lowest levels of production. This assumption means that average costs will decline up to a point, and then begin to increase. If a firm's level of production is at other than the minimum point on this average cost curve, changes in production will produce a reduction in the cost per unit.

In the railroad industry, costs behave no differently than in other industries. However, in railroading it is important to distinguish economies of scale from economies of density. Firm size, by itself, may or may not generate economies. Traffic density, on the other hand, has a major impact on costs[8]. On lightly used rail lines, increases in traffic can be accommodated with only small increases in the cost of track maintenance. But as traffic continues to increase, reason dictates that at some output level the economies of density disappear. First, the railroad will become congested as physical capacity is reached. Trains will operate more slowly, and track maintenance will become more difficult due to the difficulty of scheduling time on track for maintenance gangs. Second, the environmental mechanisms that limit track component life when no traffic uses a track (rust and decay) become insignificant on heavily trafficked railroads. Rust and decay are replaced by abrasive wear and crushing; rails and ties do not remain in track for long enough to permit the environmental mechanisms to work.

Several widely employed costing methodologies (notably the Speed Factored Gross Tonnage model) do not recognize the disappearance of economies of density. They assume that track maintenance costs change as the square root of density over the entire possible range of traffic densities. This assumption, if incorrect, leads to an increasing misspecification of incremental track maintenance costs as traffic density increases.

The shape of the MOW cost curve is vitally important in determining the economics of branch lines, making decisions on plant rationalizations such as removal of second main track, and in determining the incremental cost of additional traffic on an already-busy main line. A misspecification of the cost function can lead to suboptimal decisions, possibly including the abandonment of revenue-producing lines and the removal from service of needed assets.

Let us consider first the implications of the cost curve postulated in the Speed Factored Gross Tonnage model. Describing SFGT in terms of economic theory is quite straightforward. Figure 4 plots the shape of the marginal (or "incremental") cost of MOW, and total MOW cost, as a function of traffic density. In the figure, the cost of supplying track infrastructure is plotted against the quantity of demand (total traffic density). According to SFGT, costs rise with the square root of density. This is the curve labeled "total cost (TC)" in Figure 4, with the generic mathematical formula

$$a + b\sqrt{d}$$

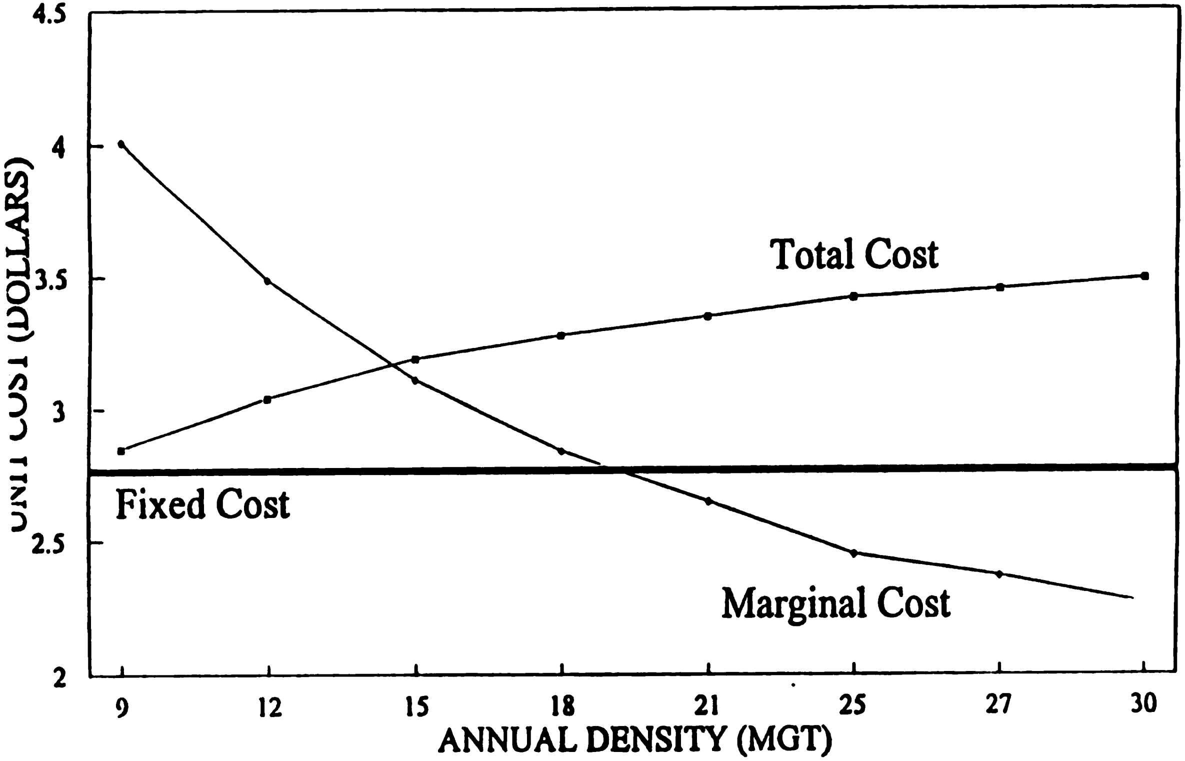
The marginal cost (MC) can be easily derived by taking the derivative of total cost (TC):

$$MC = aTC/aD = b/2\sqrt{D}$$

This curve is also plotted in Figure 4.

Figure 4

SFGT as Economic Theory



Note that the marginal cost curve slopes downward along its entire length. This leads to some non-viable conditions. To explain this, consider a case where a separate firm is established to provide track maintenance. In such a case, the firm's marginal cost would decline continuously with output. In a competitive marketplace, firms price at marginal cost. Thus this firm would behave like no other; as demand for its services increased, the price charged would decrease. This is clearly not consistent with economic reality.

It could be argued, of course, that the track maintenance activity is not a product in the strictest sense, but merely one of a number of inputs used to produce rail transportation. However, it is then necessary to believe that the input price will fall continuously as more of the input is demanded. No other input to rail transportation or any other industrial process behaves in this manner. As demand increases, eventually supplies run short and scarcity raises the price – but not, we are expected to believe, of track maintenance.

This is clearly absurd. Consider the case where the safety problems of running trains nose to tail have been solved, and the railroad is occupied continuously 24 hours per day. At this extremely high density, marginal track maintenance costs (if the SFGT model is to be believed) approach zero, though it is now impossible to get access to the track to maintain it!

As will be shown in the following section, the cost behavior predicted by SFGT is as impossible from an engineering as from an economic standpoint. To illustrate this inconsistency, let us consider the rail cost portion of the SFGT model.

Referring to Figure 4, we see that SFGT requires the marginal cost of rail to decrease with every additional ton of traffic added to the track. Stated another way, rail life (measured in million gross tons or MGT) increases with traffic. Rail that might last 500 MGT under light traffic will last, according to SFGT, much longer than that under heavy traffic. This is the explanation for the cost behavior shown in Figure 1 on page 9. At one MGT, rail cost per mile per year per MGT is about $2862. At 80 MGT, cost is $226 per MGT, or about 1/10 the cost at one MGT. What physical mechanism might be responsible for this continuously declining cost? At all but the lowest annual tonnages, environmental degradation of rail is insignificant. Wear is a linear phenomenon, and fatigue is severely non-linear (conventional wisdom relates fatigue to the square of axle load). At the very best, then, rail cost should rise in step with traffic, such that the cost per MGT should be the same at almost any density level.

An Engineering Perspective

While SFGT appears to be misspecified from an economic perspective, its outputs did, for some time, track well with observed cost behavior. This should not be surprising; after all, the original formulation was based on empirical data, and even a misspecified function may track well with reality over a limited range of observations. However, once outside that limited range, a misspecified model will fail. Indeed, the railroad industry has changed significantly in the last decade, moving far from the limited range over which SFGT produces plausible results. For example:

- Average car capacity is near 100 tons today; three decades ago it was closer to 50 tons
- Welded rail is now widespread. The elimination of joints, along with the trend to heavier cars, has completely changed the mechanisms that result in condemnation of rail and its removal from track.
- Especially in the years since deregulation, there has been a trend toward concentrating rail traffic on fewer and fewer mainlines. Partially this trend is a result of mergers and the contraction of the rail network, but it also reflects a deliberate effort by railroads to make maximum use of their investment in physical plant. The result is an increase in per-track annual tonnages to levels never before reached in the U.S., or anywhere else.

The consequences of these changes for maintenance of way costs have been great. Huge annual volumes of traffic are now being regularly and safely operated over thousands of single track route miles. Much has been learned about the effects of heavy cars on rail, ties, and roadbed, and this increased knowledge has been reflected in changed maintenance practices. Larger tie plates and better grades of ballast have extended tie life and surfacing cycles. Widespread use of welded rail, high-hardness steels, and profile grinding has extended rail life (in terms of total tonnage carried) far beyond historical levels.

These advances in maintenance technology have led some observers to conclude that the unit cost of track maintenance has in fact declined as tonnage has increased, just as had been predicted by the AREA and later by the SFGT model. However, what has in fact happened is that the ratio between maintenance spending (which is expensed) and major track reconstruction (which is capitalized) has changed. Track components are lasting longer but costing more. In short, the railroad industry is spending more per ton mile of traffic to maintain track, not less[9]. Heavier cars and more intensive use of tracks have produced economies, but not in the maintenance of way accounts. The square-root-of-density relationship does not appear to be supported by the data.

If it is assumed that MOW costs are in fact related to the square root of traffic density, then it is also necessary to assume that track component lives -- measured in MGT -- increase continuously as traffic increases. At relatively low traffic levels (below 25 MGT) there may in fact be an effect of this kind (since some component degradation may be due to environmental factors, depending on traffic volume). However, at higher tonnages (25 MGT and above) environmental degradation is supplanted by mechanical wear resulting from the passage of traffic. At these typical mainline tonnages, there is a substantial body of knowledge to suggest that the relationship between traffic and the degradation of major track components (rail, ties, and ballast) becomes linear. That is, for each additional unit of traffic there is an equivalent unit of component degradation (and therefore cost).

The development of a full understanding of the relationship between component lives and traffic is relatively recent. The SFGT square-root-of-density assumption is in fact taken from work by the American Railway Engineering Association during the 1950s, where cross-sectional analysis of a number of Class I railroads did produce an approximate square root relationship between traffic and costs. However, the analysis was confined largely to lines with moderate traffic densities (by 1991 standards), bolted rail, and relatively light axle loads (100-ton cars were still a decade away from wide use). It also did not adequately address differences in maintenance practices between railroads (e.g., heavier densities were found mostly on larger, wealthier railroads with a higher degree of mechanization in maintenance). However, the square root relationship became embedded in railway costing practice, with far-reaching consequences for both railroad rates and railroad policies.

Recent engineering research has suggested cost relationships very different from the square root of density. The shift from wear to fatigue as the primary determinant of rail life has increased the importance of the strongly non-linear effects of axle load. Increases in tonnage per track mile have moved much of the North American rail system to traffic levels where the effect of environment is negligible. Maintenance of way activities have been almost fully mechanized on all Class I railroads, meaning that there is no longer a significant difference in production rates (and therefore costs) between low-tonnage and high-tonnage lines[10].

A detailed discussion of track engineering research over the last fifteen years, at the Facility for Accelerated Service Testing (FAST) and elsewhere, is beyond the scope of this paper. However, it is sufficient to note that rail wear has been determined to be linear with tonnage, as has the degradation of track geometry and ballast under traffic. However, rail fatigue is a function of the square (or more) of axle load, resulting in a rapid deterioration as axle loads have increased. Ties are more affected by environmental factors than rail or ballast, but when traffic reaches the level where mechanical, traffic-induced factors supplant environmentally-caused decay, tie deterioration also becomes linear with traffic.

In conclusion, the accepted engineering finding is that track component deterioration is at least linearly related to traffic at typical mainline traffic densities. The economies of density so often predicted appear to apply only to activities not directly related to the passage of trains, such as track inspection, snow clearing, and the maintenance of the signal system (if any). However, these costs are small compared to the costs incurred for maintenance and replacement of rails, ties, and turnouts, and to the costs of maintaining track to proper geometric standards.

A Strategic Perspective

The regulatory costs of the SFGT cost misspecification are clear. Railroads have been the losers in a number of important rate cases, and shippers have been able to exact compensation for demonstrated "overcharges" based on SFGT cost calculations. This damage is obvious; it can be measured in terms of lost revenues.

Other effects are more subtle. The railroad industry embraced the square root of density relationship twenty years ago because it addressed two perceived railroad problems in ways the industry found favorable. First, SFGT-calculated costs for branch line operations are very high. When branch line abandonments were the focus of much railroad strategic planning (and when the process was heavily regulated), railroads were happy to demonstrate very high branchline costs. Of course, some of these unprofitable lines are now profitable shortlines and regional railroads. (Their success, however, is said to be due to "low labor costs").

The second effect of this misspecification is even more serious, however. The branch line problem has largely been solved by deregulation, abandonments, and short line spin-offs. The problem of high density on mainlines is still very real. If costing models used by railroad marketing departments produce ever-declining incremental costs as traffic density increases, incremental traffic on already-busy lines may be seriously underpriced (and this before even considering the costs of traffic congestion and the difficulty of maintenance access). SFGT-type analysis may indicate major savings from removal of second main track, where if costs are indeed linear the MOW savings may be small (from reduced track inspections and maintenance of turnouts), while the consequences in terms of service deterioration may be large.

In any industry, a complete understanding of costs is essential if production is to be profitable. A producer who is unaware of costs will not survive. Railroads have made great strides in reducing costs, but have they reduced the right ones? At least some of the gains from "rationalization" of railroad fixed plant may be illusory, leaving the railroads in the position of Milo Minderbinder (the character in Joseph Heller's novel Catch 22 who lost money on every unit but claimed a profit on the volume).

A Statistical Perspective

If SFGT appears to fail on economic, engineering, and strategic grounds it does so because it appears to be based on flawed statistical analysis. It appears that SFGT was formulated based on what might be called a "functional fishing expedition". That is, some analyst examining data on maintenance of way costs versus tonnage for a number of railroads simply selected an arbitrary mathematical function that appeared to fit the data. As far as can be determined, there was no analysis of the underlying physical mechanisms, maintenance practices, or other factors critical to the proper understanding of the relationship between maintenance of way costs and traffic. The pitfalls of this approach are well described even in basic statistical texts[11]. However, for clarity an example is provided here.

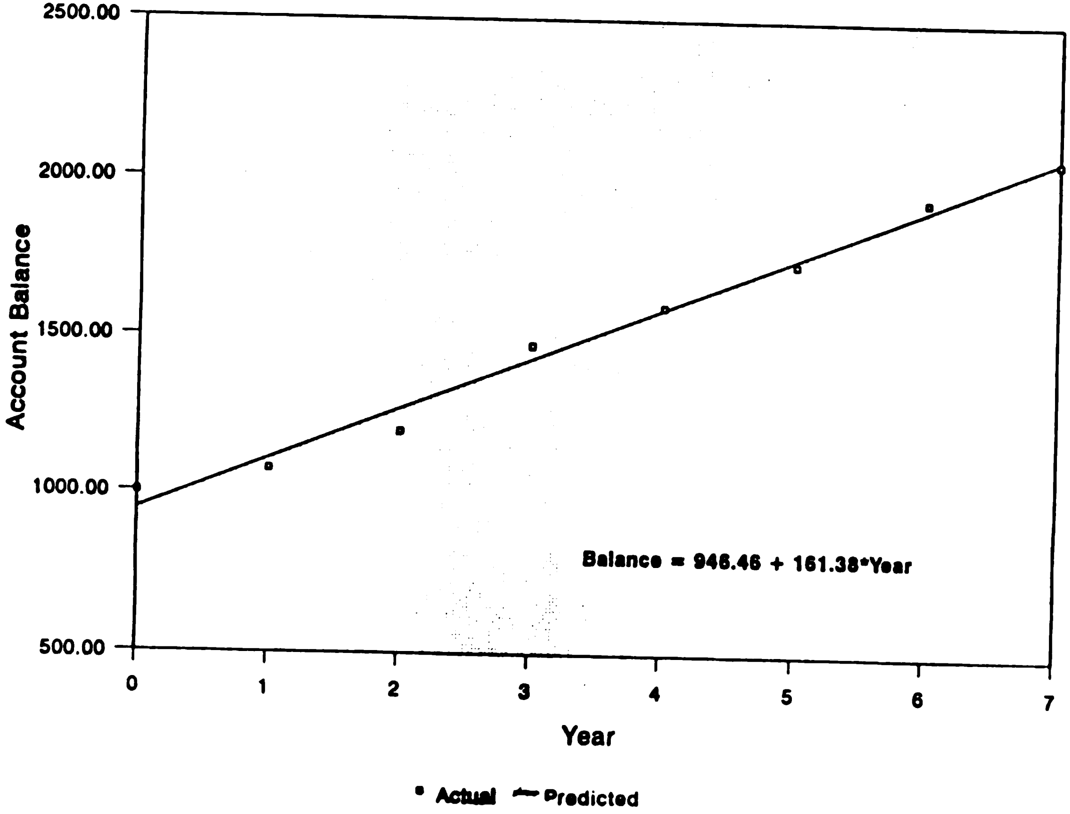
For this example, we will use a linear regression to predict the balance in a mutual fund account. Making a $1,000 investment in year 0, we have generated an annual return for a seven-year period, using a normal distribution of interest rates with an average of 10% and a standard deviation of 8%. Using a random number generator to produce interest rates for the seven years, we obtain the results shown in Table 3:

Table 3
Example of Return on Investment
Mean of 10%, Std. Deviation of 8%,
Randomized for Seven Years

Year	Return (%)	Account Balance
0	N/A	1,000
1	7.55	1,075
2	11.27	1,197
3	23.12	1,473
4	8.53	1,599
5	8.54	1,736
6	11.84	1,941
7	6.61	2.069

These results are plotted in Figure 5. A linear regression may be performed on these results, to predict the performance of the mutual fund over a longer time period. Figure 5 shows the regression results based on the data in Table 3. The regression has a R^2 of

Figure 5:Prediction of Account Balance
Using Linear Regression



98.9%, suggesting that it may indeed be a good predictor of future fund performance. Using the regression equation, we can predict the balance in the mutual fund after twenty-five additional years (32 years total) as follows: 32 x $161.38 + $946.46 = $6,111.

There is a way to check this prediction. WE can generate an additional 25 years of random numbers with the same mean and standard deviation to represent expected future annual returns. When we do this, however, we produce an account balance of $26,194 at the end of the 32-year period -- four times the value predicted by the regression!

The problem, of course, is one of specification. A much better method for predicting the expected rate of return on this investment would be to use a geometric average rather than a linear projection. This is done by taking the balance in the account after seven years, dividing by the original investment, and raising that ratio to the power of 1/7. This computation yields a rate of return of 10.948%. Applying this rate of return to the $2,069 in the account after seven years, for another 25 years, produces a final balance of $27,786 -- much closer to the actual balance than was the regression result.

In this same way, SFGT infers a square-root-of-density relationship from a limited data set gathered in 1956. While the square-root relationship may have held true for that data set, the fact that it did was coincidental. A regression is a hypothetical model. Just as in our example, a high R^2 does not necessarily indicate good predictive power, unless the underlying hypothesis is correct. In the case of the mutual fund example, the variance in interest rates was not linear but geometric, and a linear regression was therefore not a suitable form. In predicting the relationship between track maintenance costs and traffic, the hypothesis must be based on known physical relationships between traffic and track components. If it is not, its predictions will almost certainly be wrong.

The preceding sections have discussed the shortcomings of a square-root-of-density function to relate track maintenance costs to traffic density. If marginal cost does not continuously decline as tonnage increases, what is the relationship?

As indicated in the previous discussions of engineering research, it appears that rail life is linearly related to traffic density at all but the very lowest annual traffic volumes. That is, the life of rail in MGT can be expected to remain the same as traffic increases (while the life in years will of course decrease). The situation for ties and ballast is more complex. It is undeniable that environmental damage plays a part in determining the life of these components. It was almost certainly environmental decay that produced the cost relationships first plotted by the AREA in 1956 and later included in the SFGT model. However, at some traffic level, mechanical wear and crushing must become the dominant life-determining factors.

One possible alternative to SFGT is presented here. It is known as the Weighted System Average Cost (WSAC) model, and was originally developed for the Association of American Railroads. In this model, engineering equations are used to produce engineering adjustment factors which reflect the relative track damage caused by different types of rail traffic. Traffic types are defined by axle load and speed. Track characteristics such as curvature, grade, and weight of rail are also considered in the model.

The engineering adjustment factors can be applied to a unit cost, such as system aver age cost (hence the name of the model) to raise or lower it in accordance with the characteristics of each traffic type. While the original AAR formulation of the model worked only with system average cost, it is possible to extend the model to take account of both costs and track characteristics on specific line segments. While the WSAC model does require significant amounts of data on track characteristics and traffic types, the weighted costs it generates are based on known engineering relationships and therefore can be expected to reflect actual track damage done by each traffic type using a railroad line.

However, the most critical difference between WSAC and other common costing methodologies (such as the Speed Factored Gross Tonnage model discussed above) is that WSAC assumes a linear relationship between costs and traffic at typical mainline traffic densities. Therefore, when the aim is to determine incremental cost (as opposed to allocating common costs) WSAC will produce dramatically different results than SFGT and similar models. At lower tonnages, where

SFGT produces high marginal costs, WSAC indicates that the actual marginal cost is low, and at high tonnages (where cost relationships are linear) the WSAC marginal cost will be high and constant where an SFGT-determined cost declines with increasing traffic.

The reason for this behavior of the marginal cost curve in WSAC has to do with the relationship between environmentally-caused damage and traffic-caused damage to track components. As an illustration, consider a segment of track with 100 ties, 10 of which have been weakened by environmental factors and will fail in a year's time even in the absence of any rail traffic. Assume that operation of a single train over these ties will destroy one weakened tie and will also cause one sound tie to fail (due to mechanical damage). In addition, it is reasonable to assume that some number of the remaining nine weak ties will be pushed close to failure by the first train. Assume for the moment that two of the nine ties are now near failure.

Now, when a second train runs over the track, it will probably destroy the two weakened ties. In addition, the second train may also cause one of the otherwise sound ties to fail (as did the first train). So the first train destroys two ties and the second train three. This increase in marginal cost will continue until all the weakened ties fail and are replaced. At this point, the relationship between traffic and track damage will become linear.

The preceding scenario, while hypothetical, serves to show the interrelationship between environmental decay and mechanical damage to track components. This is the mechanism which accounts for increasing marginal costs at low traffic densities. WSAC is an engineering model, and does not address such factors as difficulty of maintenance access or reductions in gang productivity due to traffic on busy main lines. Costs are assumed to be linear with traffic at levels above 25 MGT per mile per year. Because of this, WSAC probably understates marginal costs at very high traffic densities. Additional research is needed to quantify the loss of productivity and the increase in costs, however. In any case, the WSAC assumption of constant marginal cost certainly will yield more accurate results than the SFGT assumption of continuously declining marginal cost.

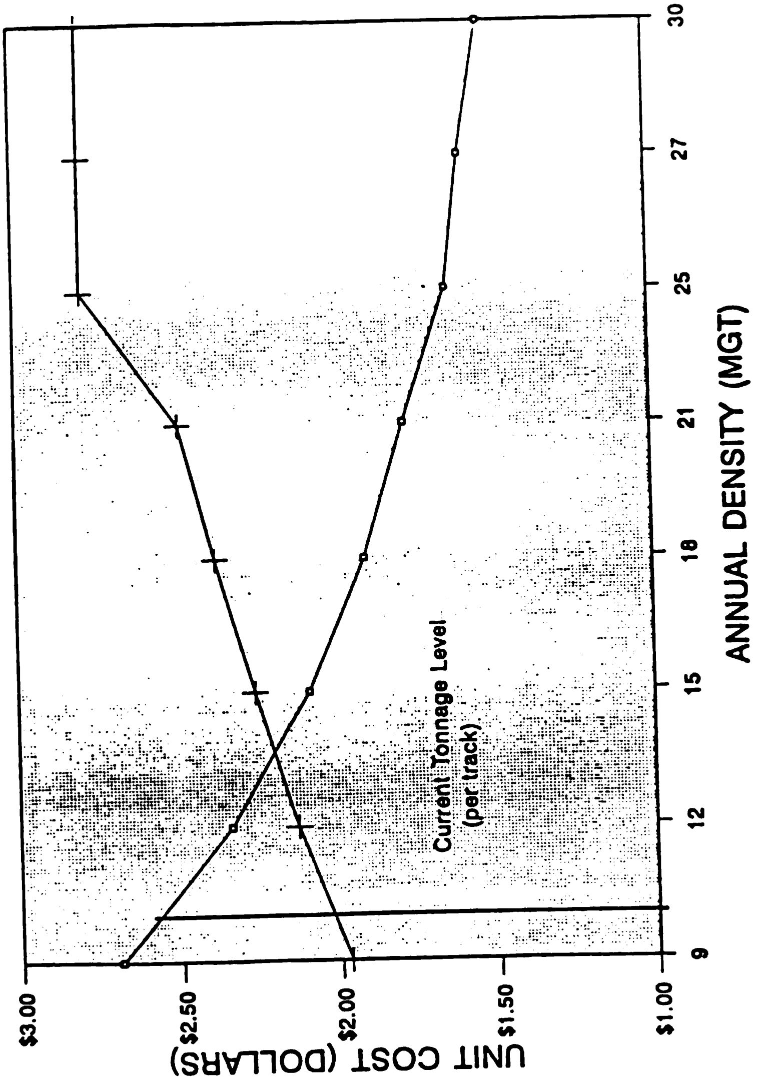
A comparative application of WSAC and SFGT to actual data from a Class I railroad is shown in Figure 6. The difference between the postulated marginal costs of the two models is immediately apparent. What is most important from a strategic standpoint is the large difference between predicted incremental costs at high annual tonnages. It approaches 50% at 30 MGT per year, and will continue to increase at higher tonnages.

The strategic implications of this difference have been discussed. There are also other possible model formulations than WSAC. However, what is most important is that incremental MOW costs be correctly specified. SFGT does not do this.

CONCLUSIONS

There seems to be no economic, engineering, strategic, or statistical reason for believing that marginal MOW costs decline continuously with traffic. Indeed, there is evidence to the contrary, and at least one costing model has been formulated using that evidence. So how has the apparent misunderstanding of MOW costs persisted in the railroad industry?

The concept of economies of scale is widely accepted, and when the railroad industry operated enormous excess capacity in a regulated environment, there may in fact have been significant economies of density. However, overcapacity has vanished on many railroads. Some are facing capacity constraints, and even investing in additional fixed plant. Nevertheless, the idea that the largest possible volume moved on the smallest track mileage is the high-profit solution remains seductive. At a time when service has become important for the first time in many years, such thinking will make the railroads' competitive burden even heavier. Traffic congestion degrades service quality, and it may be that the railroad industry has gone too far in reducing multiple track and multiple main lines. Even if the unit costs of maintenance do not increase due to the difficulty of access, the projected economies may not be realized. It may be time to look at moderate-traffic lines as potential generators of profit, rather than potential candidates for sale or abandonment. It is also time to understand that, at some point, economies of scale disappear. Costs do not decline forever. The



Comparison of Incremental MOW Costs

SFGT and WSAC, Medium Tonnage Line

railroad industry in North America has likely reached the point where marginal track maintenance cost is increasing, rather than decreasing.

ENDNOTES

* Zeta-Tech Associates, Inc.
Cherry Hill, New Jersey

** Burlington Northern Railroad
Overland Park, Kansas

1. Westbrook, M. Daniel, and George F. Rhodes, "Economic Analysis of Costing System Components in Rail Rate Regulation", *Journal of Business and Economic Statistics*, vol. 4, no. 3 (July 1988).

2. *Proceedings of the American Railway Engineering Association* (AREA), vol. 58, no. 532 (November 1956). Information also taken from letter, G.M. McGee to M.B. Hargrove, 1981, concerning the origin of the square-root-of-density hypothesis.

3. Several ICC rate cases, including San Antonio Power and Eastern Coal Export, have used SFGT to establish marginal MOW costs.

4. See Lewis, W.F., "Final Report of the Engineering Panel (Phase I) of the Association of American Railroads Cost Analysis Organization", *Bulletin of the AREA no. 694*, (January 1984). Also see Danzig, C.J. et.al., *Procedures for Analyzing the Economic Cost of Railroad Roadway for Pricing Purposes: volume I.* Federal Railroad Administration report RPD-11-CM-R (Washington, 1986).

5. AREA Proceedings, op.cit.

6. Kalousek, J., "Wear and Contact Fatigue Model for Railway Rail". National Research Council of Canada, report no. 27491 (October 1986). Many other similar citations can be found in recent engineering literature.

7. Westbrook and Rhodes, op.cit.

8. Lee, Tenpao, and C. Philip Baumel, "The Cost Structure of the U.S. Railroad Industry Under Deregulation", *Journal of the Transportation Research Forum*, vol. XXVII, no. 1.

9. See recent issues of *The Yearbook of Railroad Facts*, Association of American Railroads (Washington, published annually), for source of these calculations.

10. Hargrove, M.B. and Alvaro Auzmendi, "The Effect of Traffic Density on Maintenance of Way Productivity: An Empirical Assessment of North American Experience", presented at International Heavy Haul Workshop (Vancouver, BC: June 1991).

11. See for example Freund, John, *Elementary Modern Statistics*, (New York: 1967), as an example. All basic texts contain similar cautions regarding uses of statistical analysis.