A STATISTICAL-ENGINEERING APPROACH TO ESTIMATING RAILWAY COST FUNCTIONS

Joseph S. DeSalvo

Lester B. Lave

March 1968

P-3781

A STATISTICAL-ENGINEERING APPROACH TO ESTIMATING RAILWAY COST FUNCTIONS

Joseph S. DeSalvo and Lester B. Lave*

The RAND Corporation, Santa Monica, California

I. INTRODUCTION

Both developing and developed nations share a common problem in the construction and regulation of their transportation sectors. In few nations are transportation companies left free of government control and subject only to the competition of the market place. In order to perform the tasks that governments have assumed, it is essential that they have accurate, detailed estimates of the cost functions of the various transportation modes. These cost functions are used to determine whether society's interests are served by constructing a new mode or expanding existing ones; they are also necessary for determining the proper rate structure in regulated modes.

Attempts to estimate cost functions typically follow one of two approaches: the "engineering approach" or the "statistical approach."** Each of these approaches has received considerable attention in recent years, and each has its own adherents. We view the approaches not as

---

*The RAND Corporation and Carnegie-Mellon University. Any views expressed in this paper are those of the authors. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper is based on research performed for the Army Corps of Engineers at the Econometrics Research Center, Northwestern University.

**A selection of engineering studies may be found in Vernon L. Smith, Investment and Production (Cambridge: Harvard University Press, 1961). A selection of statistical studies may be found in J. Johnston, Statistical Cost Analysis (New York: McGraw-Hill Book Co., 1960).

competitive but rather as complementary. Each has its inherent weaknesses and strengths. The statistical approach describes the present operations of firms and is inherently aggregate in character. The engineering approach can describe any proposed or existing operation (which can be described in detail) and is inherently detailed in character.

We go on to characterize the two approaches and then show how the two can be combined to form a "statistical-engineering approach." This marriage is illustrated with an application to railroad cost functions.

STATISTICAL APPROACH

One approach to determining the cost of producing transportation services involves examining the costs incurred by existing firms. While they contain some random occurrences, these data probably do characterize the current costs of operation. However, firms are not likely to be in long run equilibrium over the period observed. The costs incurred during a period depend on past investment decisions, past hiring decisions, and variations in demand during the period. If an observation is not on the long run average cost curve, it hardly seems likely that it can be used to generate an estimate of this curve. It seems fruitless to enter the controversy of just what these figures do measure; instead we regard them as a summary of current operating procedures and conditions. We do not regard the estimated cost functions as being the minimum cost locus of theory.

Apart from this basic criticism of the statistical approach, there are other problems. In order to estimate a function, its form must be specified, and estimated marginal costs are sensitive to the form specified. While econometricians usually assume that the specification is derived from underlying economic theory, the accepted theory is of little help here, for a host of alternative forms is consistent with theoretical expectations. In general, it is unlikely that "the data will decide which form is correct," i.e., it is unlikely that one form will have a much higher coefficient of determination than its competitors. Unfortunately, many of the most important properties of a cost function are sensitive to its specification.

One final problem with this approach is that the data are generally accounting data. These data are designed for purposes other than the estimation of cost functions (e.g., to take advantage of tax laws) and therefore present many problems to the analyst.

ENGINEERING APPROACH

The data on which engineering production functions (and the cost functions derived from them) are based consist either of technological information from physical or chemical theory or from empirical analyses of controlled experiments.

There are several advantages to estimating production functions from engineering data and principles. The range of applicability of the function is known in advance; it is not subject, in general, to data limitations. Unlike the information used in cross-section and time-series studies, engineering data are available over a wide range of observations. Moreover, the results of production investigations do not depend on the pattern of investment in a plant. For these reasons engineering production functions conform more closely to the production functions of economic theory. The cost functions derived from such production functions also better approximate their theoretical counterparts than do statistical cost functions.

These engineering cost functions also have disadvantages. Since they are theoretical, there is no assurance they could be realized in an actual operation. Other difficulties stem for the estimation of the production function. This function is typically represented as the sum of functions describing individual processes. Unless these processes are additive and independent, estimation is extremely difficult. However, this problem may not loom importantly in practice since engineers attempt to avoid interaction effects in specifying their processes.* Even if it were possible to estimate the various process functions, a firm production function would still not emerge. Entrepreneurial inputs are not introduced explicitly into engineering

---

*This observation is made by A. A. Walters, "Production and Cost Functions: An Econometric Survey," Econometrica XXXI (Jan.-Apr. 1963), 12.

processes (although these inputs may be reflected in "average experience" data). In addition, it may be difficult or impossible to include non-technical processes, such as selling activities, in an engineering production function.

A STATISTICAL-ENGINEERING APPROACH

The function estimated from an engineering approach is quite detailed and can characterize any operation which can be described in detail. This approach is particularly relevant where an innovation is to be evaluated or where some other "structural change" is to be incorporated in the firm's operations. In addition, the engineering approach in many cases may be the only way to estimate the marginal cost of a specific operation. The engineering approach can be used to provide detailed information on direct costs, i.e., those involved in the transportation movement itself.

The function estimated by statistical means is probably quite aggregate and is a description of current operations. Where no structural change is involved, and where there is no need for really detailed costs, it has major advantages. In addition, it provides estimates of indirect costs, those associated with the operations of the firm which are not part of the actual transportation, e.g., selling and supervisory expense. Even where detailed costs are required, the engineering approach cannot provide estimates of these indirect costs.* In these cases, statistical cost estimation is necessary.

Finally, we observe that it is possible to splice the two approaches together at any level of aggregation. The approaches do not

---

*In using the engineering approach, it is common practice to adjust the estimate of direct cost by adding on a fixed percentage for "overhead." This practice can hardly be justified; indirect expenses will vary from one context to another. Using a fixed percentage of direct costs to measure these costs may work out on average but would rarely be correct for a particular application. Note that our statistical-engineering approach could be interpreted as giving a percentage loading for indirect expenses, but this will change from one application to another.

have to be used systematically, but rather can be used when they are most convenient for the case at hand (as long as care is taken that all costs are counted once and only once).

II. ILLUSTRATION OF STATISTICAL-ENGINEERING APPROACH

Meyer, Peck, Stenason, and Zwick (MPSZ) have estimated cost functions for freight train transportation.* Their general approach was to divide total rail operating costs into several components. Each component account was then subjected to individual analysis. These analyses took the form of cross-section regressions of cost on output and size variables. The output variables were usually gross ton miles of freight and of passenger traffic, although other output variables were used where they seemed appropriate or gave better results. The size variable was usually miles of track, although, again, others were used.

The components of total operating cost were: train (linehaul) expenses, station expenses, yard expenses, traffic (selling and marketing) expenses, general (administrative and legal overhead) expenses, variable portion of maintenance expenses, variable portion of depreciation expenses, and variable portion of capital costs.

The analyses of these cost categories were synthesized to yield overall relationships between cost and output. MPSZ did this by employing certain assumptions about the relationships between ton miles and certain other operating measures such as yard hours and train miles. Based on these assumptions and the statistical analyses, Table 1 is obtained. MPSZ argue that these figures represent the long-run marginal cost of rail mevement.

The cost functions were estimated for 1947-1950 data. Since there has been a substantial increase in railroad input prices since 1950, these figures cannot be used without adjustment. To take account of this difficulty, we adjusted all costs by a price index of railroad inputs. Another difficulty is that these functions describe early postwar operations. During this period railroads were spending large amounts of money on new equipment and on upgrading existing equipment

*John R. Meyer, Merton J. Peck, John Stenason, and Charles Zwick, The Economics of Competition in the Transportation Industries (Cambridge: Harvard University Press, 1960).

Table 1

LONG-RUN MARGINAL COST OF RAIL FREIGHT TRANSPORTATION

Cost Category	Cost Per Gross Ton Mile of Freight Traffic 1947-50 Mills	1965 Mills[b]	Percentage of Total
Train[a]	0.411000	0.809670	18.90
Station	0.001799	0.003544	0.08
Yard	0.268441	0.528829	12.34
Traffic	0.070860	0.139594	3.26
General	0.126000	0.236400	5.52
Maint. of Way & Struct.	0.284000	0.559480	13.06
Freight Car Maint.	0.162000	0.319140	7.45
Yard Engine Maint.	0.004000	0.007880	0.18
Road Engine Maint.	0.151000	0.297470	6.94
Joint Equip. Repair	0.088900	0.175133	4.09
Equip. Depreciation[a]	0.165360	0.325759	7.60
Road Struct. Depreciation	0.046640	0.091881	2.14
Capital	0.401100	0.790167	18.44
Totals	2.175100	4.284947	100.00

[a]To be estimated from an engineering production function.

[b]The price indices used for correcting the 1947-50 figures were obtained from Association of American Railroads, Railway Statistics Manual (Washington, 1964), p. 10.14, "Indexes of Average Charge-Out Prices of Railway Material and Supplies and Straight Time Hourly Wage Rate Class I Railroads," for the period to 1963, and from Association of American Railroads, "Indexes of Railroad Material Prices and Wage Rates (Railroads of Class I)," Series Q-MPW-51, May 17, 1966, for the period 1963-65.

SOURCE: John R. Meyer et al., The Economies of Competition in the Transportation Industries (Cambridge, Mass., 1960), pp. 46-47, 51, 56, 60, 62, Appendix B.

and roadbed. It is heroic to assume that these costs describe 1968 operations. We use these figures merely to illustrate the statistical portion of the proposed statistical-engineering method. We do not believe that these figures represent current rail costs or that they should be used as the basis of any decision. Instead, we believe that the method we illustrate might be used to estimate the proper costs.

To use the MPSZ cost figures, one must first estimate the number of gross ton miles to be generated in a movement. The long run marginal cost of this movement would be calculated as the product of gross ton miles and 4.28 mills per ton mile (the price corrected version of the MPSZ total long run marginal cost). Gross ton miles are calculated from the distance a train is to move and the cargo and tare weights associated with the movement.

MPSZ say, in reference to the above sort of calculation, "It should be noted that these estimates are based on a sort of central tendency and are typical figures that will apply to freight movements only of a very average or ordinary kind."* The reason for this is in part due to the nature of the estimation procedure and in part due to the assumptions required to use some of the results. The variability of costs with output was investigated, as mentioned above, by means of linear regression analysis. This by its very nature makes the estimates "sort of central tendencies." In addition, certain assumptions were employed to achieve the figure of long-run marginal cost on a gross ton mile basis. For example, it was assumed that it required "approximately 24 minutes to originate, classify, and terminate the typical merchandise car, that this car goes approximately 400 miles, and that it has a load of 25 tons."** These assumptions were used to reduce yard expenses to a gross ton mile basis. Later, it was assumed "that any shipment under analysis is normal in the sense that it approximates the 1947-1950 averages of having 3,000 gross ton miles of freight moved for every road mile of diesel engine operations and requiring one diesel engine

---

*Ibid., p. 63.

**Ibid., p. 48.

yard hour for every 69,000 gross ton miles of freight traffic."* These assumptions were used to translate the yard engine and road engine maintenance figures into gross ton mile equivalents.

It is possible to avoid making such assumptions even within the context of the MPSZ work. For example, MPSZ provide equations from which yard time can be estimated.** If one does not want to relate all costs to a gross ton mile basis, the original estimating equations can be used rather than the aggregate figures of Table 1. In most cases, they will be no different because the gross ton mile is the predominant output variable used. In some cases (for example, road engine maintenance) other variables are included in the estimation, necessitating either assumptions such as those used by MPSZ or actual values for a particular situation.

While it is possible to use the MPSZ cost function without employing all of their assumptions, it is not possible to use it to get the detailed cost of a specific movement, such as a unit train haul. However, this cost function can be made more specific by employing a different technique to estimate some of the components. For example, train expense represents almost 19 percent of total cost and was estimated on an aggregate basis. Instead of using the MPSZ figure, one might estimate this cost category by using an engineering production function. This approach would allow the analysis to estimate the specific train costs associated with particular movements. These engineering train costs could then be added to the remaining MPSZ cost categories to yield a new estimate of total long run marginal cost.

There are two of MPSZ's cost categories that are important and easily susceptible to estimation via an engineering cost function.*** These are train expense (consisting of crew and fuel costs) and

---

*Ibid., p. 52.

**Ibid., pp. 308-315.

***Another good candidate for an engineering analysis would be yard operations.

locomotive and car depreciation expense. These costs estimated via engineering cost functions comprise about 25.5 percent of total cost.*

The engineering cost function to be used here was developed by DeSalvo.** It makes use of engineering relations to determine the average speed of a train as a function of (1) number of cars, (2) tare weight of each car, (3) load limit of each car (if commodity to be transported is not very dense, then cubic capacity of each car and weight per cubic foot of the commodities must be known), (4) proportion of each car filled, (5) number of axles on each car, (6) horsepower of locomotive(s), (7) number of locomotives, (8) number of axles on each locomotive, (9) gradient of terrain, (10) degree of track curvature.

Given the speed of the train from the process function, the trip time is easily calculated, assuming no delays. If an estimate of delay time is available, this may be used in conjunction with train speed to determine trip time. For any given route and any given train, the trip cost may be determined from the cost estimation procedures discussed below, using the train process function to estimate trip time and to calculate the rate of output (in ton miles per hour) produced by the train. Moreover, the behavior of cost as horsepower and number of cars vary may be observed.

In addition to the assumptions made above, it is further assumed: (1) there are no delays on route, (2) there is no grade, (3) there is no track curvature, (4) the commodities are dense enough that the cars are weight-limited rather than space-limited, (5) all cars are fully loaded, (6) trip length is 100 miles, (7) cars are standard box cars with tare weight of 30 tons each and load limit of 70 tons each, and (8) only one locomotive is used.

---

*We might also have attempted to estimate Locomotive and Car Capital Expense directly. Certainly we have detailed equipment costs and the other information necessary to calculate this cost. However, MPSZ include in this category investment in road as well as equipment capital cost. It would be difficult to disentangle the road capital from equipment capital and so we use the MPSZ calculation.

**Joseph S. DeSalvo, "A Process Function for Rail Linehaul Operations," The RAND Corporation, Paper No. P-3729, November 1967.

Combining engineering and statistical approaches, the cost of a movement is given in Eq. (1):

$$C = W + F + D_c + D_L + A, \tag{1}$$

where

W = crew wages

F = fuel cost

D_c = depreciation of cars

D_L = depreciation of locomotives

A = all other costs (estimated by MPSZ)

Crew wages might be estimated using Eq. (2) and Table 2. To use the equation, one must know trip distance, number of engine units, number of cars, and trip time.

$$C_w = \begin{cases} r_1 D & \text{if } D \le 100 \text{ and } T \le D/12.5 \\ 100r_1 + r_2(D - 100) & \text{if } D > 100 \text{ and } T \le D/12.5 \\ 12.5r_1 T & \text{if } D \le 100 \text{ and } T > D/12.5 \\ 100r_1 + r_2(12.5T - 100) & \text{if } D > 100 \text{ and } T > D/12.5 \end{cases} \tag{2}$$

where

C_w = crew member's cost to the trip

r_1 = crew member's wage rate per mile for first 100 miles

r_2 = crew member's wage rate per mile after first 100 miles

D = actual trip distance in miles

T = actual trip time in hours.

The fuel cost of a trip might be estimated from Eq. (3). To use the equation, one must know trip time, engine horsepower, and the price per gallon of diesel fuel.

Table 2

FREIGHT TRAIN CREW WAGE RATES, 1967

Crew Member	Rates/Mile for First 100 Miles	Rates/Mile after First 100 Miles
Engineer		
1 engine unit	$ .2564	$ .2389
2 engine units	.2657	.2482
3 engine units	.2775	.2575
4 engine units	.2840	.2665
Conductor		
1-81 cars	.2216	.2041
82-105 cars	.2251	.2076
106-125 cars	.2291	.2116
126-145 cars	.2316	.2141
146-165 cars	.2326	.2151
over 165 cars	a	a
Brakeman		
1-81 cars	.2104	.1860
82-105 cars	.2139	.1895
106-125 cars	.2179	.1935
126-145 cars	.2204	.1960
146-165 cars	.2214	.1970
over 165 cars	a	a

[a]$.20 for each additional block of 20 cars or portion thereof.

SOURCE: Midwestern railroad which requested not to be identified.

$$F = .073p \cdot HP \cdot T, \tag{3}$$

where

F = fuel cost per trip
p = price per gallon of diesel fuel
HP = total horsepower of locomotive units
T = actual trip time in hours.

Equipment depreciation costs might be estimated from Eq. (4) for locomotives and Eq. (5) for boxcars.

$$D_L = \$0.007HP - \$0.0295T \tag{4}$$

$$D_c = \$0.0896T \cdot N_c, \tag{5}$$

where

D_c = total car depreciation per trip
D_L = total locomotive depreciation per trip
N_c = number of cars on train
HP = total locomotive horsepower
T = actual trip time.

The only component of total cost remaining is A, all other costs. Using the MPSZ cost function, this component is simply 3.15 mills times the number of gross ton miles generated by the trip in question. (We are assuming that this trip is small enough so that marginal cost is not affected.)

The cost per ton mile of the assumed trip, as horsepower and the number of cars are varied, is presented in Fig. 1. The figures are obtained by expressing trip cost on an hourly basis and dividing this by the appropriate output (measured in ton-mile per hour) as estimated from the engineering production function. Each curve represents a given horsepower locomotive pulling from 10 to 200 cars. Each curve might be thought of as a "plant" unit cost curve, where horsepower represents the fixed input and number of cars represents the variable input. The unit cost curves for 500-2500 horsepower display the U-shape usually assumed to characterize plant cost curves. However, beyond 2500 horsepower the curves are continually downward sloping. This implies that trains longer than 200 cars would be required to make the unit cost curves rise. An envelope curve could be drawn to the individual unit cost curves. This would represent the long-run average cost of the train trip to the firm. Although such a curve is not drawn in Fig. 1, it is clear that economies of scale exist in the range of output from zero to 180,000 ton miles per hour (a horsepower range of from 500-2500). Economies of scale, in fact, seem to persist throughout the range of output shown on the graph although at a much diminished rate beyond an output of 180,000 ton miles per hour.

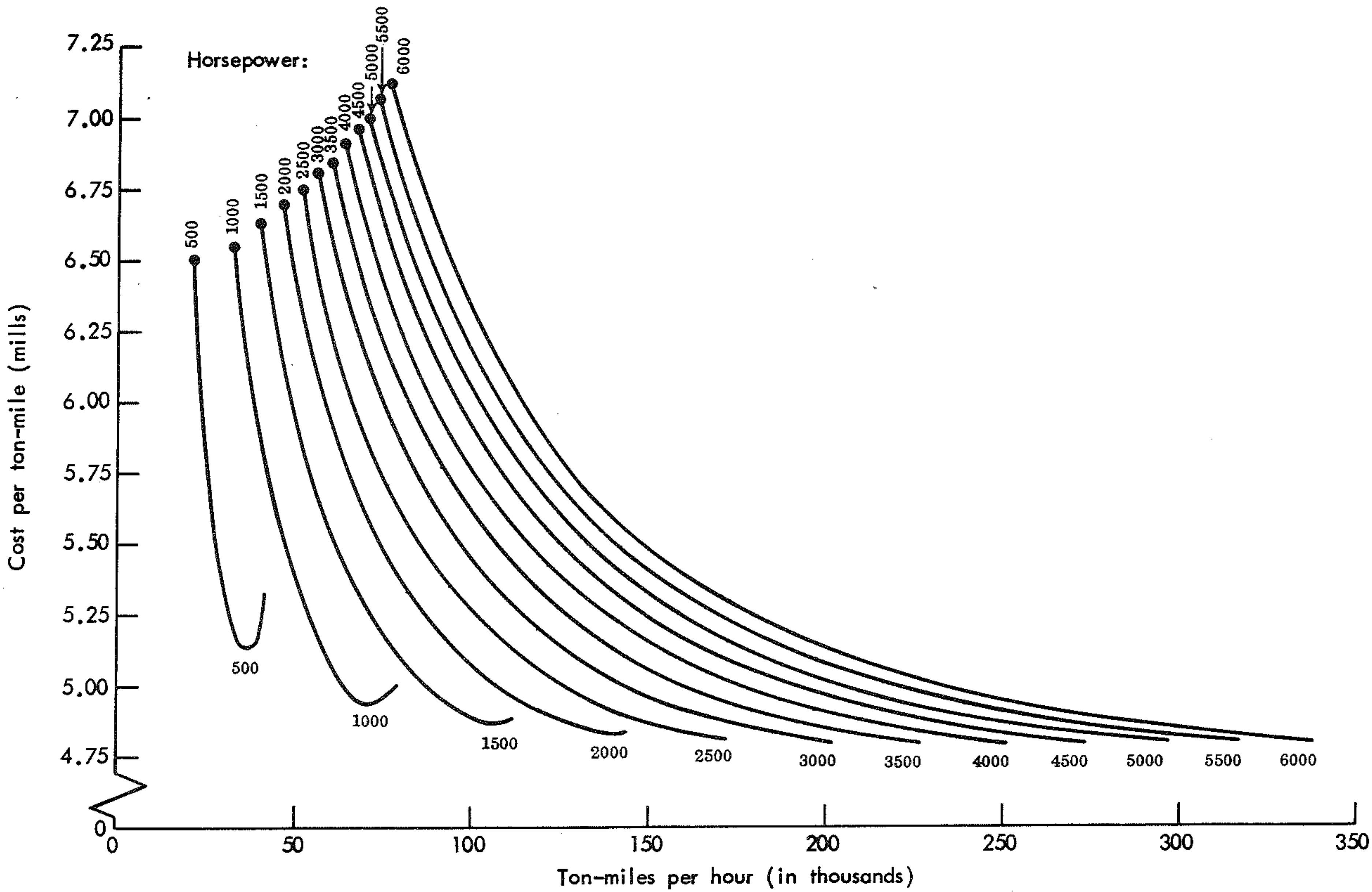

Fig. 1--Railway Firm Cost Curves

III. SUMMARY AND CONCLUSIONS

Statistical and engineering methods both possess advantages and disadvantages in the determination of cost behavior. This paper illustrates a wedding of the two approaches for the case of rail costs. The statistical approach of Meyer, Peck, Stenason, and Zwick is used for all cost components except linehaul and equipment depreciation. An engineering process function is used along with input costs to develop linehaul costs. Depreciation costs for cars and locomotives are estimated from equipment prices and estimated lifetimes. The resulting statistical-engineering cost function permits a detailed estimation of the cost of a train trip. It is much more flexible than the pure statistical approach as it permits changes in equipment, roadway, and loading conditions. In addition, it includes costs not amenable to engineering analysis, such as selling and administration.