1978

A New Approach to Railroad Cost Estimation *

Andrew F. Daughety

Graduate School of Management
and Transportation Center

Northwestern University
Evanston, Illinois 60201
(312) 492 - 3550

Mark A. Turnquist

Department of Civil Engineering
and Transportation Center

Northwestern University
Evanston, Illinois 60201
(312) 492 - 5183

Ronald R. Braeutigam

Department of Economics
and Transportation Center

Northwestern University
Evanston, Illinois 60201
(312) 492 - 5142

*This work was supported by the Office of University Research, U.S. Department of Transportation, Contract DOT-OS-70061

A New Approach to Railroad Cost Estimation

Andrew F. Daughety
Mark A. Turnquist
Ronald R. Braeutigam

ABSTRACT

Cost functions for railroad firms are an important element of management decision-making. They are also a principal basis for regulatory policy. Standard ICC cost-finding procedures which are based on a number of untenable assumptions regarding the allocation of common and joint costs, are inadequate. Furthermore, traditional econometric attempts at rail cost estimation generally have suffered from extensive aggregation of firms, insufficient disaggregation of output, and inadequate representation of the physical process of producing services. This paper presents the results of estimating a rail cost function using a new technique based on incorporation of engineering models of the production process into the econometric estimation process. The data base is a time series comprising nine years of monthly data from a single railroad, including detailed information on both operations and financial transactions.

The major contribution of this work is the explicit incorporation of engineering performance models in the analysis. This allows the inclusion of a measure of service quality, average speed of shipment, as a characteristic of the firm's output. This inclusion provides a significant first step in allowing management to better understand the cost implications of improving service quality. The basic structure developed in this paper is flexible enough to allow other measures of service quality, such as trip time reliability and loss-and-damage, to be included.

1.0 Introduction

An understanding of the nature of costs of production is important in every regulated industry, both for individual firms and their regulators. At the most basic level a firm will require cost data for corporate planning. For example, a firm may wish to know what size plant to build, whether to upgrade the quality of plant or whether, at an existing tariff, the revenues for a service cover the incremental cost of providing the service. Cost data may be used to argue for a change in tariffs. A firm may want to know how a change in the level of output of one service affects the costs of providing another service, and it may rely in part on costs data to determine whether it would be profitable to discontinue a service, introduce a new service, or attempt to merge with another firm.

Regulators and other policy makers also have many reasons to seek improved information about costs. When examined correctly, cost data can be used to determine whether there are in fact economies of scale in production, and whether regulation is a necessary tool of social control in a given industry. Regulators often ask whether a service is being subsidized by other services of a multiproduct firm, is subsidizing other services, and whether the provision of service by one mode will eliminate another mode over a given route. If regulators are interested in setting tariffs that allocate economic resources efficiently, they will require information about costs. Generally speaking, then, regulators need cost information to determine how their policies will affect market structure and economic performance. These comments apply without exception to the railroad industry.

1.1 Other Railroad Cost Estimates

A number of studies have examined costs in the railroad industry. The early work in this area attempted to characterize the output of railroads as a single product, usually ton-miles. These studies typically have examined a cross section of Class I railroads, using ICC data, to test whether there are economies of scale in rail transport. The results have generally been mixed. For example, Klein [9] used 1936 data to find economies of scale that were statistically significant, though modest. On the other hand, estimates by Borts [2] and Griliches [6] have suggested that, while there may be economies of scale for smaller railroads, scale economies are not prevalent for larger Class I railroads.

Several aspects of these studies have served to limit the inferences that can be drawn. They rely on data from the ICC acounts rather than on raw data from the firms. They typically specify a relatively simple functional form for costs, and assert that the form is appropriate without a test of that assertion. The do not adjust for quality of service, and more importantly, they do not account for the multiproduct nature of virtually every rail firm. And, they do not attempt to adjust for the fact that some railroads operate with a more complicated network than others.

More recently, Hasenkamp [7] and Keeler [8] have attempted to recognize the multiproduct aspects of railroad activities by distinguishing between freight and passenger service. Again, these studies use cross-section ICC data, and despite the limitations imposed by these constraints represent important advances in our understanding of costs.

1.2 A Time Series Estimate of a Hybrid Cost Function

Our own research on railroad transport costs represents a strikingly different approach to the problem for a number of reasons.

1) Our analysis begins at the level of an individual firm, and uses cost and production data obtained directly from the firm rather than from the ICC. This has a number of important advantages, including the avoidance of arbitrary cost allocations of the sort often found in the ICC accounts. (For a discussion of the kinds of problems arising

from the use of ICC data, see, for example, Friedlaender [4], Appendix A.) We employ a time series analysis for a single firm rather than a cross sectional analysis for a particular year.

2) The multiproduct nature of the firm is incorporated into the analysis. Output will be characterized both by the volume of freight hauled and by the average speed of a shipment through the system. We explicitly recognize that speed of service is an important determinant of rail costs, and include this in our estimates.

3) We use information about the underlying technological production process, developed through engineering process functions, to better specify the nature of technology and to improve the efficiency of our estimates.

In several respects the last point is particularly novel. Historically, most econometric estimates of cost functions have ignored valuable information generated from an analysis of engineering process functions to provide observations of service related variables. We have labeled this a "hybrid" approach for that reason, and we believe that important new insights can be gained from applications of this technique to other modes, as well as in rail transport.

2.0 Proposed Cost Function Formulation

2.1 Overview

As indicated above, we have developed and estimated a model of railroad costs that incorporates engineering models of operations and avoids the problem of relying upon data from railroads of varying size and description. Thus, our model is estimated using time-series data from a single firm. The purpose of this section is to outline how the engineering aspects of the problem are introduced into the overall model formulation, and to briefly examine the nature of the data used in the model.

2.2 The Use of Engineering Models

In general, economic theory stipulates that a long-run cost function should contain as independent variables the quantities of output(s) produced and the prices of input factors utilized. The technical basis for such a specification is discussed somewhat more fully in the Appendix. Furthermore, if one or more input factors are held fixed, so that a short-run cost function is obtained, the levels of these fixed factors will also appear as independent variables. However, while economic theory thus dictates some elements of the form of cost functions, a good deal of the success to be achieved in estimating such functions relies upon the accuracy with which the inputs and outputs of the production process are specified. It is in this respect that engineering analysis of the firm's operations can be extremely useful.

One of the crucial elements of the output of a railroad (or other transportation firm) is the quality of service provided. It is not simply the number of carloads moved, but also the associated service characteristics which are important. These service characteristics might include average speed of shipment, measures of service reliability, and loss-and-damage, for example. Since the costs of providing varying levels of service could be substantially different, it is important to capture service quality measures in the definition of the firm's output. To do so, however, requires detailed data on the nature of operations. Since such data are not normally collected on a regular basis by the ICC or other public agency, previous cost studies have tended to ignore service quality measures. The work in this study, however, demonstrates that service quality measures <u>are</u> important, and that thorough understanding of costs requires recognition of these elements of output.

One effective way of reflecting service quality measures is through the use of engineering models of network operations. This is the approach taken here, and the models used include representation of linehaul train movement and classification yard operations. The objective of these engineering models is to reflect average speed of shipment as a function of volume moved, and the motive power and physical facilities available. For the present, average speed of shipment is the only service measure utilized, but this should be viewed as a first step toward more comprehensive models incorporating other measures as well.

2.2.1 The Linehaul Model

The purpose of the linehaul model is to reflect first, the relationship between locomotive horsepower, trailing load and velocity for a train; and second, the delays enroute due to interactions among trains (meets, overtakes, etc.). This model allows us to represent the way in which several of the major input factors for the railroad (locomotives, fixed plant of varying quality, etc.) are used in a major element of the production process -- the movement of trains.

The first component of this model is an equation which reflects the relationship between power, trailing load and velocity for a train, given characteristics of the fixed plant over which it operates. This equation can be solved to find the theoretically attainable velocity of the train. However, since in many cases this velocity will exceed speed limits imposed by track quality or other factors, the actual velocity will be limited by the speed limit in effect. Additional detail on the derivation of this model is provided in [14].

This first component model describes the linehaul velocity which could be attained if there were no interactions among trains. However, trains are often delayed enroute due to passing or being passed by other trains going in the same direction, or on single track line, meeting trains going in the opposite direction. Detailed simulation models are often used by railroads to evaluate train congestion (see, for example, Lach and Skelton [10]). However, for the purposes of this study, it is desirable to have a simpler, analytic model which can be incorporated more readily into the specification of a production function for cost estimation. The model used here draws heavily on work done by Petersen[11].

In order to utilize this model, the expected number of encounters between trains must be expressed in terms of quantities available. These include traffic density of trains of different clases, their speeds and dispatching policies through time. Examples of the derivation of the expected number of encounters under different sets of conditions are provided in [14]. These results represent significant extensions of Petersen's earlier work. Together with the model of train movement discussed previously, the delay model allows determination of overall linehaul velocity for a shipment.

2.2.2 Classification Yard Activities

According to data gathered by Reebie Associates [12] the average rail car spends only 16% of its time actually moving in trains. An additional 56% is spent in classification yards. This underscores the importance of representing classification yard activities if we are to reflect railroad operations with any reasonable degree of accuracy.

While in a railyard, a car undergoes four basic operations: 1) inbound inspection; 2) classification; 3) assembly into outbound train; and 4) outbound inspection. It is quite natural to think of these activities as a series of queues through which the rail car passes, and a model developed on this premise predicts average time in the yard as the sum of the average times for the four operations. Inbound and outbound inspections consume a relatively small amount of time for each car, and the amounts of time required are not highly variable. For these reasons, they are represented by constant standard times. Explicit queuing models have been constructed for the remaining elements, classification and assembly.

The classificaiton process is one in which railcars arrive in batches (trains). Thus a batch arrival queuing model is appropriate. The nature of the service process depends upon the type of yard involved. Hump yards can generally be represented by deterministic service times, but flat yards often require more complex service time models based on distributions of cars per inbound cut, relative likelihood of various outbound blocks, and the time required per switch. Given the mean and variance of inbound batch sizes, and a mean and variance for the service time distribution, average classification delay can be computed using a formula developed by Gaver [5].

Delay due to assembly into outbound trains includes two major components: connection delay waiting for the next scheduled outbound dispatch, and the time required for the actual assembly process itself. Formulated as a queuing model, the assembly process is one in which cars arrive individually (from classification) and wait until the "server" (an outbound train) is ready to accept a batch of cars (a train) to leave the yard. In theory, the delay time to cars depends upon the interarrival time distribution of cars from classification, the distribution of times between successive dispatches of outbound trains, and the maximum train length that can be accommodated on outbound trains. From a practical perspective, the most important factor is the distribution of times between successive dispatches of outbound trains.

Combining the classification queuing model and the assembly queuing model, we can predict total average time in yard for cars. The principle advantage of using this queuing approach is that it provides an explicit link between volumes moved, input resources provided and the quality of service resulting in the yard. In this way, we have an analogous component model which can be linked with the linehaul model to predict overall quality of service in terms of speed of shipment.

2.3 Specification of the "Economic" Variables

The engineering analysis provides a procedure for developing output measures of system performance: in the case at hand we have concentrated on the overall speed of a shipment through the system. The model also uses the following economic variables:

1. Volume: loaded car-miles;
2. Inputs: a) prices of cars, locomotives, fuel, crews, non-crew labor, rail;
   b) fixed factor: quality of plant;
3. Cost: operating costs plus charges for car and locomotive use.

They will be briefly described in turn.

2.3.1 Volume

Our model considers two types of output: speed and volume. The previous subsection has described the speed variable; here we will concentrate on volume.

Shippers are viewed as buying loaded car-miles of product transported. Thus, empty-car miles are not an output so much as an intermediate product necessary to produce loaded car-miles. We have concentrated on using car-miles rather than ton-miles for two reasons. First, it is the car capacity that is generally purchased; shippers do not simply ship a ton of a good. Secondly, this allows us to tie this unit of output to the engineering process models of the yard.

Detailed records of the firm allow us to calculate carloads of various commodities (by STCC code if desired) on a monthly basis. The model to be presented has aggregated all commodities into one volume variable (Y) so as to focus our discussion on the hybrid nature of the model (i.e. the introduction of speed (V) into the model). In general, however, the degree of disaggregation of Y into various commodity categories is only limited by problems of statistical significance.

2.3.2 Inputs

As discussed in the Appendix, the model includes both variable and fixed inputs. Variable inputs are represented by their prices. There are six prices in the model, constructed as follows:

1. PCAR: monthly rental price of cars found by taking purchase

prices of cars and constructing equivilent rental prices using the firm's interest rate for borrowing funds; prices for the various car types (e.g. hopper, refrigerator cars) are combined in proportions reflecting the firm's car use at each point in time;

2. PLOCO: similar to PCAR construction;
3. PFUEL: monthly purchases of fuel divided by gallons used;
4. PCR (price for a crew): wage bill for train crews each month, divided by the number of hours actually worked, plus supplemental pay;
5. PNCR (price for non-crew labor): wage bill for management, clerical staff, maintenance labor, etc. for each month divided by the number of hours actually worked, plus supplemental pay;
6. PRAIL: cost per ton of rail amortized over expected life into monthly rental payments via the firm's interest rate on borrowing.

A price for ties was also computed, but is highly correlated (.989) with the price of rail, so only the rail price is used.

In addition, because we are estimating a short-run function, the level(s) of fixed factors must be included. We have chosen to represent the fixed factor as the quality of mainline track, measured by the proportion of the total track miles which are of sufficient quality to be considered in FRA Class 4. Changes in the fixed plant of railroads occur very slowly because the elements of the plant have very long lifetimes. Since the data available reflect few major changes in the nature of the plant, we have estimated short-run functions with plant as the fixed factor. Because the FRA classifications have associated speed limits, and these speed limits affect the speed of shipment over the system, an index of plant quality based on this classification structure is a very effective measure of the fixed facility for this study.

2.3.3 Cost

Observations on cost have been obtained from operating costs, plus estimated opportunity costs for the cars and locomotives used. Because the operating costs only include repair costs on the capital equipment, it is necessary to supplement them to account for ownership costs. We chose not to use per diem payments on cars because such payments are often complicated by special arrangements with other roads on the use of cars, and because the per diem rate does not accurately reflect opportunity costs since it is based on historical data and is distorted by being set as part of the regulatory process.

In general, locomotive leases were used when possible. However, equipment obligations entered into before the time period of the analysis do not completely reflect opportunity costs at points of time within the analysis period.

2.4 Summary

The model specification contains two measures of output: loaded car-miles and average speed; prices of six variable input factors: cars, fuel, rail, crews, non-crew labor and locomotives; and the level of one fixed factor: the quality of main line track.

The cost model is estimated simultaneously with five factor share equations as discussed in the Appendix. Thus our system of equations can be generally represented as:

$$\text{Cost} = C(Y, V, PCAR, PFUEL, PRAIL, PCR, PLOCO, PNCR; QK)$$

$$S_{FUEL} = S(Y, \ldots, PNCR; QK)$$

$$\vdots$$

$$S_{NCR} = S(Y, \ldots, PNCR; QK)$$

where QK is the quality of track variable and S is the appropriate share function foundby differentiating the cost function with respect to the appropriate factor price. Factor share equations for fuel, rail, crews, locomotives and non-crew labor have been used.

3.0 Estimation Results

The model to be estimated* (each variable is divided by its mean before taking logs so as to protect the proprietary nature of the data) is as follows:

$$\begin{aligned}\log(C/\overline{C}) = \alpha &+ \beta_1 \log(Y/\overline{Y}) + \beta_2 \log(V/\overline{V}) \\ &+ \gamma_1 \log(PCAR/\overline{PCAR}) + \gamma_2 \log(PFUEL/\overline{PFUEL}) \\ &+ \gamma_3 \log(PRAIL/\overline{PRAIL}) + \gamma_4 \log(PCR/\overline{PCR}) \\ &+ \gamma_5 \log(PLOCO/\overline{PLOCO}) + \gamma_6 \log(PNCR/\overline{PNCR}) \\ &+ \delta \log(QK/\overline{QK})\end{aligned}$$

We have used 108 monthly observations on C, Y, V, PCAR, PFUEL, PRAIL, PCR, PLOCO, PNCR and QK to provide estimates of α, β_1, β_2, $\gamma_1, \ldots, \gamma_6$ and δ. As is explained in the Appendix, the above model should be estimated jointly with factor demand equations so as to improve the efficiency of the estimation. In this case, while the factor demand equations turn out to be non-linear, the factor shares s_i (i.e. $p_i x_i/c$) of the cost are linear, of the form:

$$s_i = \gamma_i$$

Factor share equations for fuel, rail, crew labor, locomotives and non-crew labor were estimated. The sixth factor share equation for cars is unnecessary since it is linearly dependent. The results of the estimation are indicated in Table 1.

The left hand column of Table 1 lists the coefficient and the associated variable. Because the model is in logarithmic form, the coefficients are elasticities of cost with respect to the indicated variable. As can be seen, all variables are significant at the 5% level with the exception of PRAIL. Furthermore, all variables are of teh correct sign: prices and flow are positive, quality

*See the Appendix. More general (i.e. non-Cobb-Douglas)forms will be examined in the final report of our study. This model was chosen to highlight the hybrid nature of the analysis.

of plant and velocity negative. The appendix indicates the reason that quality of plant should be negative. We will discuss the reason for the sign on V below.

Table 1 Estimation Results

Coefficient	Estimated Value	t-Value
α	.6592	4.54
β_1 (Y)	.1863	4.57
β_2 (V)	-.038	-2.28
γ_1 (PCAR)	.1328	5.21
γ_2 (PFUEL)	.0457	85.2
γ_3 (PRAIL)	.00007	0.66
γ_4 (PCR)	.1521	125.7
γ_5 (PLOCO)	.0827	51.7
γ_6 (PNCR)	.3998	123.8
δ (QK)	-.5208	-8.56

Before that, however, let us consider what the coefficients mean. Because they are elasticities, the larger the absolute value, the greater the effect (all else held constant) the variable has on cost. This is especially clear since all variables are divided by their means and thus scaling issues are irrelevant.

The input factors that appear to have the most direct affect are labor and cars. The elasticity of cars probably reflects the manner of adding in projected car costs discussed in section 2.3.3, and also reflects the fact that the railroad under study has been significantly improving its car fleet over time.

The elasticities of cost with respect to crew labor and non-crew labor are the greatest among the input factors, with the non-crew labor elasticity being almost three times the crew-labor elasticity. Non-crew labor includes top management, clerical, sales and maintenance labor, and it is the last category that probably accounts for the large value of the coefficient. Historically the firm has significantly expanded its maintenance labor during the summer months so as to improve track condition. A further segmentation of labor into the three categories of crew, maintenance and all other labor should be illustrative.

Finally, we observe that the sign on the speed variable is negative and significant. Though this might seem to be peculiar, (since one would usually expect outputs to have positive coefficients), in fact the result is quite reasonable. A negative coefficient on V means that if V could be shifted, then changes in the variable factors would result in lower short-run costs. Since V reflects both line both linehaul and yard facilities, an improvement in V would require long-run investments which are not directly reflected in a short-run variable cost function. Thus, one would expect V to have a negative sign in a short-run variable cost function and a positive sign in the long-run function.

4.0 Summary and Directions for Further Research

The research has produced a number of products. First, we have been able to translate theoretical data needs specified by the model into data requirements than can be fulfilled by a firm using available information. Thus, the model allows us to specify ways of combining available firm data correctly to produce measures of cost that should be used in regulatory proceedings (e.g., "incremental" costs).

Second, the costs functions produced provide for multiple element outputs: we need not lump all flows together as ton-miles, but can allow for different commodity types and types of moves. The level of disaggregation is limited only by the number of available observations and the number of parameters to be estimated. Furthermore, our output vector includes characteristics of service as well as commodities carried. Thus, marginal costs for particular commodities and services are computable.

A third product is a complete estimated cost function for a moderate-to-small railroad, which has been our test case. Since there are many such firms in the U.S., this is a useful product in and of itself.

Finally, the approach we have developed is robust with respect to adding more service characteristics and network complexity. The procedure outlined in this paper can start with a very general model of production which makes a minimum of economic assumptions (e.g. it makes no assumption as to whether or not there are returns-to-scale). This is important since we would like to examine (i.e. test) such economic attributes rather than assume them. Engineering models are then added so as to increasingly restrict (and thereby further <u>reveal</u>) the model of production. Again, it should be noted that the restrictions will reflect physical realities and not economic assumptions that need to be tested. As more engineering relationships are added (reflecting network considerations or service characteristics) the <u>economic</u> attributes of the model become more and more refined, for the general production model becomes increasingly restricted by the engineering relationships and this, in turn, reveals more of the economic relationships.

Our future work will concentrate on three areas of extension: 1) incorporation of the volume variable, 2) incorporation of more service characteristics (such as schedule unreliability and equipment availability) and 3) application to significantly more complex network structures.

5.0 Acknowledgements

The authors thank Robert Baesemann, William Delaney, Ken Hurdle, and several students at Northwestern University for help and suggestions.

Appendix

A.1 Production and Cost

Let x be an n-vector of input levels, $x = (x_1,\ldots,x_n)$. For example, elements of x would include fuel used, amounts of various labor services, car hours, locomotive hours, amounts of rail, ties and ballast, etc. Each input factor (x_i) has associated with it a price per unit, p_i. Thus total costs are

$$\sum_{i=1}^{n} p_i x_i$$

The reason for purchasing inputs is to provide output; the relation between inputs and outputs is called a production function. In the single output case we would have a single output Y and a production function f(x) so that:

$$Y = f(x)$$

The result of minimizing the total input costs subject to the production requirements is called the firm's cost function* C(Y,p):

$$\left.\begin{array}{l} \min\limits_{x} \ \Sigma\, p_i x_i \\ \quad \text{S.T. } f(x) = Y \end{array}\right\} \Rightarrow \qquad C(Y,p)$$

The above function is called the <u>long-run</u> cost function since all factors are assumed to be variable. A general representation of it is shown in Figure A-1. From it we could theoretically find the long-run <u>average</u> cost function AC(Y,p) by dividing C(Y,p) by output Y, and we could obtain the long-run <u>marginal</u> cost function MC(Y,p) by differentiating C(Y,p) with respect to Y. Furthermore, we can get the <u>factor-demand</u> functions $x_i^*(Y,p)$ which indicate how much of factor i to use if we face input price vector p and are required to produce Y by the following result ([1], [13]):

$*p = (p_1,\ldots,p_n)$

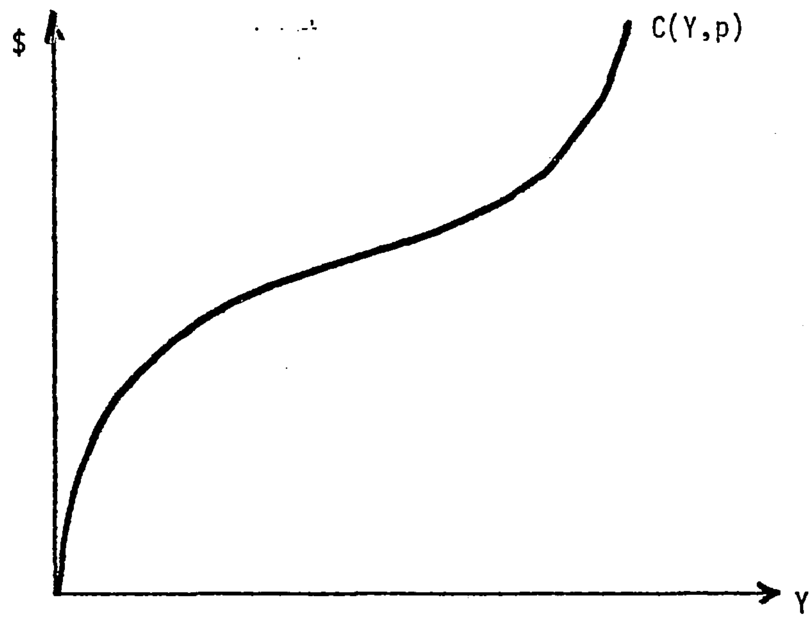

Figure A-1
The Long-Run Cost Function C(Y,p)

$$x_i^* (Y,p) = \frac{\partial C(Y,p)}{\partial p_i}$$

i.e. the i^{th} factor demand function is simply the derivative of the cost function with respect to the i^{th} factor price. We will also be interested in the i^{th} factor share equation $s_i(Y,p)$:

$$s_i(Y,p) = \frac{p_i x_i^* (Y,p)}{C(Y,p)} = \frac{p_i}{C(Y,p)} \frac{\partial C(Y,p)}{\partial p_i}$$

which is the share of cost associated with factor i; furthermore it is the elasticity of cost with respect to the factor price as seen on the right above.

A.2 Short and Long-Run Costs

As has often been observed (see, for example, [3], [8]) estimating the long-run cost function is problematical since observations should only lie on or above it, otherwise it would not represent minimal costs. Though this observation is generally associated with analyses using a cross-section of firms, this is also true of an individual firm. Here again, observations may occur on the long-run function or above it (due to adjustments in progress in various firm factor levels). Thus placing a line through such observations in order to estimate the long-run function is bound to be biased. The approach (in the cross-section railroad case, again see [8]) to be used is to estimate a family of short-run functions that together describe the long-run function.

The short-run is defined as that period of time wherein one or more factors is fixed. For convenience we take this to be the n-th factor, x_n, and partition our vectors into two parts, with the first part of dimension n-1:

$$x = (\tilde{x}, x_n)$$

$$p = (\tilde{p}, p_n)$$

Therefore our cost minimization problem

$$\min \sum_{i=1}^{n} p_i x_i$$

$$\text{S.T.} \quad f(x) = Y$$

is now posed as:

$$\left. \begin{array}{l} \min\limits_{\tilde{x}} \sum_{i=1}^{n-1} p_i x_i \\ \\ \text{S.T.} \quad f(\tilde{x}, x_n) = Y \end{array} \right\} \Rightarrow \qquad C(Y, \tilde{p}; x_n)$$

since it reflects optimization over the variable factors. Since the fixed factor provides short-run fixed costs, $p_n x_n$, then the total short-run cost function is $C(Y; \tilde{p}; x_n) + p_n x_n$. The use of the semicolon in the short-run variable cost function is to signify that the short-run costs vary as we fix x_n at different levels. This is seen in figure A-2. The figure shows the long-run cost function C(Y,p) and two short-run total cost functions, one with the fixed input at x_n^1 and the other with the fixed input at x_n^2. Notice that the short-run functions lie above or just touch the long-run function. Thus, since the observations are on the long-run function or above it (due to short-run adjustments) we could estimate the family of short-run cost functions $C(Y,\underset{\sim}{p};\, x_n) + p_n x_n$ and find the long-run cost function by minimizing the resulting function on x_n i.e.

$$\min_{x_n} \; [C(Y,\underset{\sim}{p};\, x_n) + p_n x_n] = C(Y,p)$$

In fact, for our purposes, $C(Y,\, \underset{\sim}{p};\, x_n)$ is sufficient, since our main interest is marginal costs and the technique of getting hybrid cost functions.

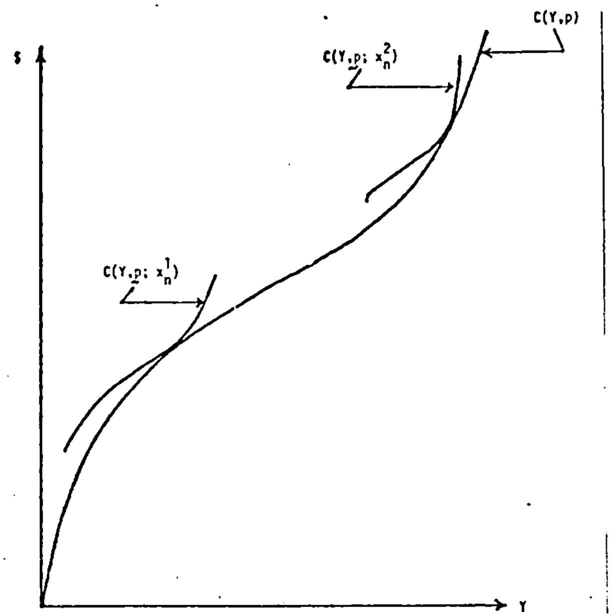

Figure A-2 Short-Run and Long-Run Cost Functions

The last point above raises the issue that output in our model should be composed of flow and speed. To allow for this we will in fact examine cost functions of the form $C(Y,\, V,\, \underset{\sim}{p};\, x_n)$ where V is speed. This can be taken as coming from an underlying transformation function (instead of production function):

$$H(x,\, Y,\, V) = 0$$

The mathematical symbolism above simply states that inputs x produces outputs flow (Y) and speed (V).

A.3 An Example of a Cost Function and Development of the Hybrid Model to be Estimated.

To better see the above, we consider the following example. Assume a firm uses capital (K), labor (L) and fuel (E) to produce output (Y) subject to the following Cobb-Douglas function:

$$Y = AK^{\alpha}L^{\beta}E^{\gamma}$$

where A, α, β, and γ are coefficients that we would like to estimate.

The long-run cost function $C(Y, P_K, P_L, P_E)$ is found by solving the following problem

$$\min \quad P_K K + P_L L + P_E E$$

$$\text{S.T.} \quad AK^{\alpha}L^{\beta}E^{\gamma} = Y$$

which yields

$$C(Y, P_K, P_L, P_E) = d\, Y^{1/f} P_K^{\alpha/f} P_L^{\beta/f} P_E^{\gamma/f}$$

where

$$f = \alpha + \beta + \gamma$$

and

$$d = f\left(\frac{1}{A\alpha^{\alpha}\beta^{\beta}\gamma^{\gamma}}\right)^{1/f}$$

Since the short-run function arises when one of the factors is fixed, then if we fix capital (K) at, say $\overline{K}$, we have a short-run variable cost function:

$$C_V(Y, P_L, P_E; \overline{K}) = g\, Y^{1/h} P_L^{\beta/h}\; P_E^{\gamma/h} \overline{K}^{-\alpha/h}$$

and the short-run fixed cost function is

$$C_F(P_K; \overline{K}) = P_K \overline{K}$$

where

$$h = \beta + \gamma$$

$$g = h\left(\frac{1}{A\beta^{\beta}\gamma^{\gamma}}\right)^{1/h}$$

The function to be estimated is C_V; it is most easily estimated by using least squares if we take logarithms of all terms giving:

$$\log C_V = \log g + \frac{1}{h} \log Y + \frac{\beta}{h} \log P_L + \frac{\gamma}{h} \log P_E + \left(\frac{-\alpha}{h}\right) \log \overline{K}$$

Notice that since α, β, γ, all should be positive numbers, the coefficient of $\log \overline{K}$, i.e. $-\alpha/h$, should be a negative number. In other words, if we estimate the following linear model:

$$Z = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4$$

where:

$$Z = \log C_V \qquad x_2 = \log P_L \qquad x_4 = \log \overline{K}$$

$$x_1 = \log Y \qquad x_3 = \log P_E$$

then the estimate of a_4 (i.e. $\hat{a}_4$) should be negative. Note that using $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$ and $\hat{a}_4$ we can recover all information of interest (i.e. estimate A, α, β, γ) by simply relating the proper coefficients (for example, an estimate of β is $\hat{a}_2/\hat{a}_1$).

Bibliography

1. Blackorby, C., D. Primont and R. Russel. Duality, Separability and Functional Structure: Theory and Economic Applications, North Holland, New York, 1978.

2. Borts, G. "The Estimation of Rail Cost Functions," Econometrica, January, 1960.

3. Eads, G., M. Nerlove and W. Raduchel. "Cost Functions for the Local Service Airline Industry," Review of Economics and Statistics, August, 1969.

4. Friedlaender, A.F. The Dilemma of Freight Transport Regulation, Brookings, 1969.

5. Gaver, D.P. "Imbedded Markov Chain Analysis of A Waiting Line Process in Continuous Time," Annals of Mathematical Statistics, 30:3, pp. 698-720, September, 1959.

6. Griliches, Z. "Cost Allocation in Railroad Regulation," The Bell Journal of Economics and Management Science. Spring, 1972.

7. Hasenkamp, G. "Specification and Estimation of Multiple-Output Production Functions," Unpublished Ph.D. Thesis, University of Wisconsin, 1973.

8. Keeler, T. "Railroad Costs, Returns to Scale, and Excess Capacity," Review of Economics and Statistics, May, 1974.

9. Klein, L., A Textbook of Econometrics, Row Peterson, Evanston, 1953.

10. Lach and Skelton, "SIMTRAC-A Simulation of Train Dispatching on a Single-Track Railway," presented at the CORS National Conference, Toronto, 1968.

11. Petersen, E.R. "Over-the-Road Transit Time for a Single Track Railway," Transportation Science, 8:1, (1974), pp. 65-74.

12. Rebbie Associates, Toward an Effective Demurrage System, prepared for U.S. Department of Transportation, Federal Railroad Administration, July, 1972.

13. Shephard, R. Theory of Cost and Production Functions, Princeton University Press, New Jersey, 1970.

14. Turnquist, M.A. and J.T. Gray. "Process Models for Railroad Operations," working paper 426-01, The Transportation Center, Northwestern University, March, 1978.