Journal of Transport Economics and Policy, Volume 33, Part 1, pp. 9-42

A Survey of Productivity and Efficiency Measurement in Rail Transport

Tae Hoon Oum, W. G. Waters II and Chunyan Yu*

Abstract

This paper surveys alternative methodologies for measuring and comparing the productivity and efficiency of railways, and the empirical findings of applied studies. Empirical studies reveal trends and differences among railways and time periods. Almost all studies reviewed conclude that increased competition via regulatory liberalisation and deregulation has improved efficiency. Many European studies find that managerial autonomy increases efficiency. It is important that the effects of differential operating environments such as traffic density and the characteristics of a rail network should be removed in order to make a proper comparison of efficiency.

1. Introduction

Since 1945, railways in most countries have experienced a declining market share, rising input prices, and increasing competition from other modes of transport. The success of individual railways, as well as the industry as a whole, depends on improving their productive performance. In order for governments to design proper public and regulatory policies concerning the rail industry, and for railroad companies to set appropriate strategies to improve productivity, it is important to understand the determinants of productivity.

Railways are multiproduct enterprises. Their outputs have a spatial dimension as well as quality attributes. These outputs are produced via complex production processes involving numerous primary and intermediate inputs. It is difficult to compare the outputs and inputs of one firm with those of other firms, or over time within the firm. A number of methodologies have been used to assess the productivity performance of railways. Different methodologies, along with data and computational differences, lead to different empirical results and interpretations. A long history of government involvement in the industry further influences output decisions, resource use, and productivity.

There have been numerous productivity or efficiency studies of railways, but few surveys. Dodgson (1985) reviews studies on total factor productivity measurement for

* Corresponding author: W. G. Waters II, Faculty of Commerce and Business Administration, The University of British Columbia, Vancouver BC, V6T 1Z2, Canada. Tae Hoon Oum and W. G. Waters are UPS Foundation Professor of Management and Professor, respectively, at the University of British Columbia. Chunyan Yu is a post-doctoral Research Associate at the same university. The authors acknowledge with thanks the research grant support of the Social Science and Humanities Research Council of Canada, and thank the Editor and two anonymous referees for helpful comments.

railways and issues related to decomposing productivity variations into sources. He focuses on both the economic foundation of productivity measurement and the policy implications of empirical results. Hooper (1987) provides an overview of theoretical concepts in productivity measurement, but focuses mainly on empirical results; he cites nine empirical studies of rail productivity prior to 1982. Oum *et al.* (1992) examine concepts, methods, and purposes of productivity measurement in transport industries, but do not provide a detailed survey of empirical studies.

This paper reviews alternative methodologies for measuring and comparing productivity and productive efficiency in the rail industry. Section 2 sets out efficiency measurement within the framework of production theory. Section 3 describes index number procedures for measuring productivity, including partial factor productivities, total factor productivity (TFP), and the data envelopment analysis method (DEA). Data issues are discussed, along with empirical studies that use these methods. Also discussed are the methods of decomposing "gross" TFP or DEA changes into sources, such as the effect of output scale, effects of exogenous factors, and productive efficiency. Section 4 focuses on measuring productive efficiency by estimating conventional production and cost functions. Section 5 discusses frontier estimation methods and their applications. Concluding remarks are made in Section 6.

2. Productivity and Productive Efficiency

A "productivity gain" refers to increased output relative to inputs. This could be a partial measure comparing an increase in one of many output categories with that in one or more inputs, or a more comprehensive measure such as an index of total output compared to an index of total inputs (a total factor productivity index). Productivity can be compared between firms, and/or over time within a firm. One of the main objectives of productivity measurement is to make inferences about the efficiency of a firm, an organisation, or an industry. However, productivity variation can arise from different sources: differences in efficiency, economies of scale, differences in network characteristics, and other exogenous factors, that affect performance (for example, average length of haul, composition of traffic, market size, quality of service, weather, or terrain conditions) and/or technological changes. Therefore, to make inferences about productive efficiency, one must remove the effect on productivity caused by the differences in operating environment and other exogenous factors.

A production function specifies the maximal output obtainable from an input vector given the production technology, that is, a frontier. Figure 1 shows the frontier for a one-input one-output (y) production function denoted by $f(x)$, where x denotes the input level. All points on or below it, such as B, C, or D, are achievable, and hence can be observed, whereas points beyond it, such as E, are neither realisable nor observable.[1] The "distance" from an observed point to the frontier provides a measure of inefficiency of the firm.

[1] In this section we assume that there is no statistical noise, measurement error, or difference in operating environment, since we focus on theoretical aspects of productive efficiency. For a useful exposition on efficiency measurement, see Coelli *et al.* (1998).

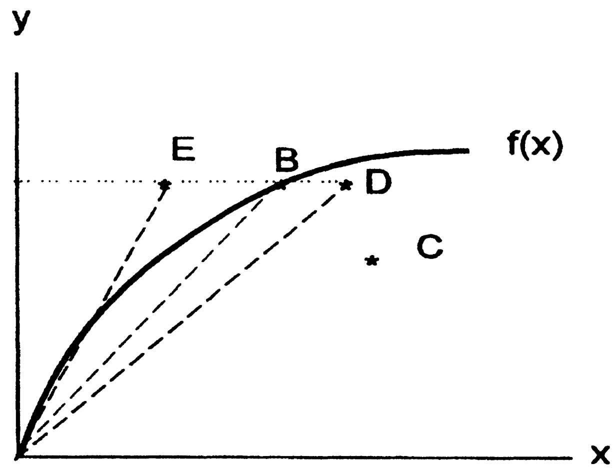

Figure 1
A Production Function as a Frontier

A measure of productivity is the ratio of outputs to inputs, indicated by a ray from the origin to the various points in Figure 1. Points D to B to E show increasing productivity. The change from D to B reflects increased efficiency of input use in existing technology, whereas achieving E requires a shift in production technology.

When multiple outputs or inputs exist, productive efficiency consists of two components: technical efficiency and allocative efficiency (Farrell, 1957). This is illustrated in Figure 2 for the case of a single output, y, using two inputs, x_1 and x_2. The production frontier is $y = f(x_1, x_2)$. To simplify further, assume constant returns to scale: the production frontier can be expressed as $1 = f(x_1/y, x_2/y)$, a unit isoquant such as ACA'A" shown in Figure 2. A firm producing at any point above ACA'A" uses more of at least one input than is necessary. Suppose the available budget is represented by the isocost line PP', which is tangent to the isoquant ACA'A" at point A'.

A firm is technically efficient if it chooses an input mix on the unit isoquant, and is allocatively efficient (pricewise efficient) if the marginal rate of substitution between two inputs is equal to the corresponding input price ratio. Technical inefficiency results from excessive use of inputs (given the level of output), while allocative inefficiency results from employing inputs in wrong proportions. Full productive efficiency requires joint satisfaction of technical and allocative efficiency conditions. This is obtained at point A' in Figure 2. Firm C is technically efficient, but is allocatively inefficient, while firm E is allocatively efficient but technically inefficient.

Consider the inefficient point D in Figure 2. Farrell (1957) defines the technical efficiency at point D as: $TE = OC/OD$; TE measures the proportion of inputs (x_1^0, x_2^0) ac-

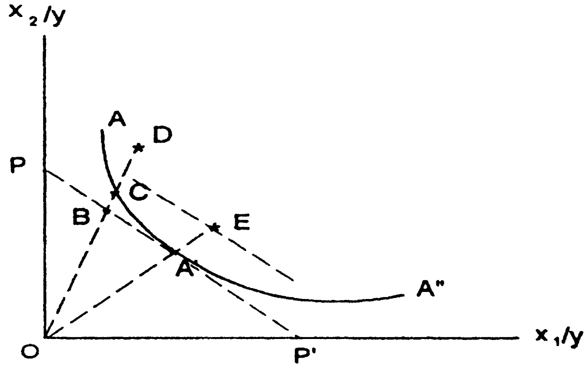

Figure 2
Technical and Allocative Efficiency

tually necessary to produce y^0. *TE* equals 1 if the firm is on the frontier ACA'A". As the observed performance of D worsens, the technical efficiency ratio falls towards zero. Thus in general: $0 \leq TE \leq 1$. Point C is technically efficient, but it costs more than necessary because the same output could be produced at the cost of *OB* by substituting x_1 for x_2; that is, point C is allocatively inefficient, caused by non-optimal input proportions. Farrell (1957) defines the allocative inefficiency (price inefficiency) at point C as: $PE = OB/OC$. As C lies closer to A', *PE* rises towards one. That is, *PE* also lies between zero and unity. Since D has the same input proportion as C, it has an allocative inefficiency of the same amount $PE = OB/OC$.

Combining technical and allocative efficiency measures gives an overall measure of the efficiency at point D. Following Farrell (1957), the productive efficiency (or economic efficiency, *EE*) of D is defined as: $EE = TE*PE = OB/OD$.[2]

While ideally we would wish to measure overall productive efficiency (OB/OD in Figure 2), economists have generally focused on technical efficiency in empirical studies.[3] There are several reasons for this. First, in the long run, changes in technical efficiency tend to dominate overall changes in productive efficiency. Second, it is not possible to develop a consistent theoretical framework for measuring and interpreting efficiency without assuming optimal behaviour of firms, such as using optimal input combination. This is especially so when firms produce multiple outputs, and thus use of the neoclassical cost function is essential to measure efficiency. Third, in many sectors of the economy, such as government services, educational and health services, informa-

[2] It is noted that Farrell's analysis does not consider the question of optimality of output level, since the optimal scale of production is indeterminate in the case of constant returns to scale. Under non-constant returns to scale, a firm is said to be "scale efficient" if it produces output where price is equal to marginal cost.
[3] Allocative efficiency has been investigated in numerous studies, including Schmidt and Lovell (1979), Kopp and Diewert (1982), Kumbhakar (1987, 1989), and Kalirajan (1990).

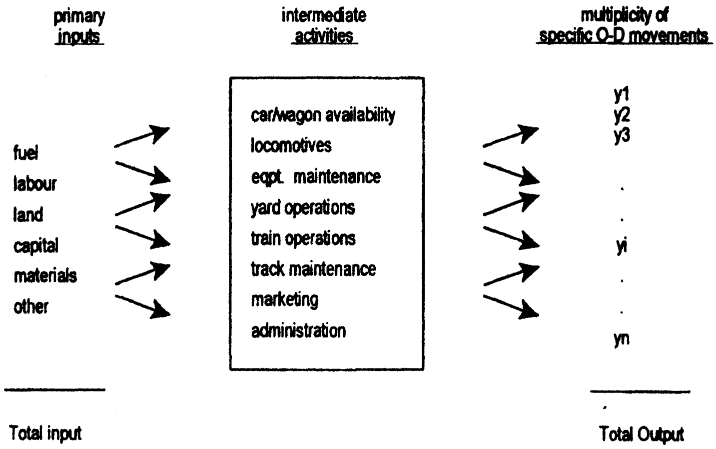

Figure 3
Schematic of Rail Production Process

tion on input prices is not available. Without input prices, it is not possible to measure allocative (input mix) efficiency. Measuring technical efficiency, on the other hand, does not require such information. In the remainder of this paper, our discussions focus on technical efficiency measurement. Other issues such as allocative efficiency, scale effects, and the influence of operating environments, are discussed to the extent that they are necessary to identify technical efficiency accurately.

Finally, Figures 1 and 2 show production from primary inputs. In fact, rail operations make use of a number of intermediate production activities, and final outputs are produced using a mix of primary inputs and intermediate outputs (for example, train operations are an intermediate output used to produce final output). This is illustrated in Figure 3. Partial productivity measures often focus on intermediate outputs, but overall efficiency is measured by linking primary inputs and final outputs as in Figures 1 and 2. Note that because of the complexity of total operations it can be extremely difficult to precisely link specific input use with specific outputs produced. Hence partial productivity measures are problematic in their accuracy/reliability.

3. Index Number Procedures for Productivity and Efficiency Measurement

Index number procedures generally construct a ratio-type productivity/efficiency measure, without the need for statistical estimation of a production or cost function. In this section, three general categories of index number procedures are reviewed and discussed: partial productivity measures, total factor productivity (TFP), and the data envelopment analysis method (DEA).

3.1 Partial factor productivity (PFP) and performance ratios

Partial (factor) productivity measures generally relate a firm's output to a single input factor. For example, revenue tonne-kilometres per employee is a labour-based partial productivity measure. A large variety of "performance ratios" are in use in practically every industry. It is possible to track one or more output or intermediate activities relative to one or more input or other intermediate activity categories. Trains despatched per hour, phone calls handled by receptionists per hour, number of trains per mile of track, loaded to empty car-miles — the list of potential measures is almost endless. These types of measures are easy to compute, require only limited data, and are intuitively easy to understand. They have thus been widely used by both academics and industry analysts. For example, the Association of American Railroads (AAR) annually publishes statistics relating miles of track, freight cars, employment, fuel, and other factors, to rail ton-miles. Similarly, British Rail reports a variety of performance measures in its annual report. Table 1 lists some examples of commonly used partial productivity measures in the rail industry. Note that some of the measures are financial rather than physical measures of performance. Barrett (1991), Nash (1981, 1985), Jackson (1991, 1992, and 1993), Schwier *et al.* (1990), Thompson *et al.* (1991), Tretheway and Waters (1990a, 1990b), Thompson and Fraser (1993), Nash and Preston (1994), and Preston (1996) are examples of studies that examine railway performance through the use of partial productivity measures, such as labour, fuel, and rolling stock productivities. However, productivity of one input depends on the level of other inputs being used; high productivity performance in one input may come at the expense of low productivity of other inputs (Brand, 1974; McGeehan, 1995a, 1995b; Barrett, 1995).

Another common problem is an inadequate output measure. A single output measure, whether it be gross ton-miles, revenue ton-miles, car miles, passenger miles, or train miles, is not a comprehensive measure of the true total economic output. Although each of these measures reflects an important dimension of railway outputs, use of a single output measure could bias performance comparisons. A better approach is to combine passenger, freight, express freight, and other service outputs, to form a total output index as described in the TFP subsection below.

Some general or "system" measures of productivity have been suggested to alleviate problems of using PFPs. These measures, although still partial in nature, employ a combination of a number of PFP measures. For example, Martland (1989, 1997) uses a number of partial rail performance measures, and attempts to link changes in them to changes in (deflated) revenues and costs. In examining US rail operations during the period 1973-83 and the update to 1995, he finds that the increased performance in some aspects of rail operations is largely (but not entirely) offset by the increased costs of other components of rail operations. Martland's approach shows promise in trying to improve understanding of net changes in productivity and the link with operations components and financial performance. However, his measurement focuses on intermediate activities rather than actual outputs, and thus there is no assurance that aggregation

Table 1

Selected Partial Indicators of Rail Productivity

General Operations
1 Average train speed (train-miles divided by train-hours)
2 Average number of cars per train (car-miles divided by train-miles)
3 Average haul (revenue ton-miles divided by revenue tons)
4 Gross and revenue tons per train (gross ton-miles divided by train-miles, revenue ton-miles divided by train miles)
5 Train-switching-hours and train-switching-miles as percentages of total train-miles and train-hours
6 Revenues and expenses per train-mile, per revenue ton-mile
7 Ratio of revenue ton-miles to gross ton-miles
8 Gross ton-miles per ton or gallon of fuel
9 Ratio of expenses to revenues
10 Ratio of loaded to total car-miles.

Locomotives
1 Locomotive unit-miles (road, train-switching, and yard-switching)
2 Locomotive switching-unit-miles as a percentage of total locomotive unit-miles
3 Locomotive unit-miles or locomotive-hours per serviceable locomotive-day
4 Gross and revenue ton-miles per serviceable locomotive
5 Average number of cars per locomotive (car-miles divided by locomotive road-unit-miles)
6 Gross ton-miles per locomotive road-unit-mile

Cars (or wagons)
1 Ratio of revenue ton-miles to carloadings or per ton of car capacity
2 Revenues and expenses per ton of car capacity
3 Average carload (revenue ton-miles divided by loaded car-miles)
4 Car-miles (or revenue ton-miles) per serviceable car-day
5 Revenue tons carried per car
6 Car cycle time or number of carloads per car per year
7 Car-hours in road movement per serviceable car-day
8 Average time in shipper's hands, in terminals, in trains, etc.
9 Revenue and expenses per car
10 Car-miles (loaded and empty)

Track
1 Number of miles of track
2 Revenue ton-miles per mile of track
3 Carloads per mile of track
4 Revenue per mile of track
5 Maintenance expenses per mile of track

Capital
1 Net investment per employee
2 Dollars of revenue per dollar of net investment
3 Revenue ton-miles per dollar of net investment
4 Investment per ton of capacity

Labour
1 Revenue and gross ton-miles per employee or per man-hour paid
2 Professional, clerical, and general: man-hour per carload
3 Maintenance of way and structures: man hours per million gross ton-miles
4 Maintenance of equipment and stores: man-hours per thousand locomotive-miles and car-miles
5 Transport (train and engine services): man-hours per thousand train-miles and train-hours

Source: Adapted from Tubb (1977), and Tretheway and Waters (1990b).

of those intermediate activities will be consistent with the true economic output of rail firms.

Despite their shortcomings, partial productivity measures can provide useful insights to causes of high or low productivity, and thus provide practical guidance for identifying productivity problems. Partial measures are useful in comparisons of performance across firms operating in similar operating environments, or over time within a firm when the operating environment and input prices remain relatively stable.

3.2 Total factor productivity (TFP)

A TFP index is the ratio of a total (aggregate) output quantity index to a total (aggregate) input quantity index. Output and input quantity indices recognise the multi-output multi-input nature of the rail industry. TFP growth is the difference between the growth of the output and input quantity indices. TFP is not an unambiguous concept either in theory or in practical measurement. Various approaches to TFP measurement lead to different interpretations and empirical results. Because of the aggregation problems inherent in multiple output production, different productivity measures lead to differences in measured results even in theory (Diewert, 1992). This is compounded by differences in data, data errors, and different assumptions in computations. We first comment on input and output measurement for rail TFP studies; this is relevant for all remaining subsections.

Measuring inputs and outputs

Some inputs are generally measured in physical quantities, for example, litres of fuel consumed or the energy equivalent. Labour inputs may be measured by the number of employees or employee-hours. The former may correct for full-time and part-time equivalent workers. Employee hours may distinguish between hours worked versus "hours paid for". It is preferable to disaggregate labour categories according to wage/skill levels. There has been a shift from low-skill to relatively high-skill workers in North American railways. Under these circumstances, a disaggregate labour index shows a lesser rate of decline in labour inputs than is calculated using undifferentiated employees or employee-hours (Tretheway and Waters, 1990a).

The most contentious input measure is capital. Capital is a stock from which a flow of services is derived. Ordinarily, capital is measured in currency units rather than physical quantities. In order to weigh capital relative to other inputs (cost share weights) it is necessary to have capital expressed in current dollars. The most common procedure is the Christensen-Jorgenson (1969) perpetual inventory method. Historical investments are accumulated for each year, converted to constant dollars by a price index for capital assets, less an assumed rate of economic depreciation. This method assumes that all capital investments were "used and useful"; that is, there is no provision for inappropriate (for example, politically motivated) investments, a dubious assumption for many railways. Obsolescence must be reflected in the assumed depreciation rates; that is, economic depreciation is used, not regulatory-mandated or tax-based depreciation rates. These are still stocks, rather than flows. Rental or leased capital, typically, is incorporated by deflating lease payments by a price index to put leased capital on an equal footing with owned capital. If we assume a constant service flow from a capital stock, then the growth of the capital stock provides the measure of the growth of capital inputs (flow) for calculating aggregate input quantity growth. This assumes that a given stock produces a flow of capital services for that year, independent of the level of actual output. This "lumpy" flow of capital services causes measured TFP to fluctuate with the business cycle; hence measured TFP may vary from year to year. TFP growth is best thought of in terms of productivity trends rather than specific year-to-year values.

Although the Christensen-Jorgenson (1969) perpetual inventory method of measuring capital is preferred methodologically, it is very data- and time-intensive. Simpler proxies for capital measurement have been used; for example, miles of track as a proxy for the size of the aggregate investment in way and structures (see Roy and Cofsky, 1985). Total rail car fleet and/or total locomotive fleet (possibly adjusted by horsepower ratings) could serve as a proxy for the equipment capital stock (see Steering Committee, 1992, chapter 5). The correspondence between these proxies and actual capital stocks is problematic; they may be reliable for equipment capital, but are less convincing for way and structures capital. It is still necessary to construct cost share weights, so it is necessary to convert whatever measure of capital into a current dollar equivalent expenditure for comparison with other input expenditures.

To construct the cost share weights, the imputed expenditure on capital is calculated by multiplying by a service price of capital. This is the imputed required return to cover the costs of using a unit of capital. This is measured as the rate of economic depreciation plus the cost of capital, and may include a capital gains component if capital assets are appreciating in value because of inflation. The cost of capital may distinguish between debt and equity capital, and adjust for taxation rates (which affect debt and equity differently), tax depreciation allowances, and so on (see Freeman *et al.*, 1987).

A sizeable proportion of total rail expenditures are the "miscellaneous" items (neither fuel, labour, nor capital). These include purchased services, materials, supplies, and so on. Typically, the quantity of these inputs is measured by a deflated expenditure approach: the total of such expenses is divided by an appropriate input price index — the GDP deflator is often used.

The aggregate input quantity index is the weighted sum of the respective indices (weighted by cost shares), with each index set at unity for some common data point. More specifically, the growth in aggregate inputs is the weighted sum of the growth rates of the individual input quantity indices.

An alternative approach to an aggregate input quantity index is to divide total expenditures (including capital) by an aggregate input price index. An example is the US Interstate Commerce Commission (ICC) (now Surface Transportation Board, 1997) approach to measuring TFP for the US Class I rail industry (this is for a productivity adjustment to limit automatic inflationary rate increases on regulated traffic). The input quantity index is calculated as total expenditure (including depreciation) divided by the Rail Cost Adjustment Factor (RCAF), a legally approved measure of the rise in rail input prices. This is calculated quarterly and divided into an aggregate output index (see Waters and Tretheway, 1991, for a discussion).

Turning to output measurement, an aggregate output quantity index is the weighted sum of the output categories. The weights are usually revenue shares for the respective outputs, although cost elasticity weights are the preferred measure (they reflect the impact of different outputs on the resources required by the firm). Revenue shares and cost elasticities are identical only if there are competitive conditions and constant returns to scale (Denny *et al.*, 1981).

Ideally, a high level of disaggregation is sought, but most productivity studies use only a few output categories. Many use just two: freight ton-miles and passenger-miles. This implicitly assumes that all ton-miles and passenger-miles are homogeneous. But different traffic types entail different input requirements. Unfortunately, disaggregate output data are relatively rare. If railways shift their traffic mix towards commodities that entail longer, heavier movements compared to smaller-volume shorter movements, this will give rise to apparent productivity gains as measured by simple ton-miles. This is because the former commodities require fewer inputs per unit than the latter categories. But a shift in traffic mix is not a real increase in productive abilities. In order to remove the effect of traffic mix in measuring TFP, Caves and Christensen (1982) introduced an output quantity index with 400 output categories and travel distances. This was superseded by an output index of 243 output categories by Reebie Associates (1988). This is used by the US ICC (now Surface Transportation Board) for regulatory purposes. The practical significance of disaggregating output is illustrated by Tretheway and Waters (1995a). Using both Canadian and US data, they show that the TFP growth during the 1980s computed from a disaggregate output measure is about a full percentage point lower than that computed from aggregate data.

An alternative way to construct an output quantity index is to deflate total revenues by an output price index. This approach is not used often in rail applications, as quantitative traffic information is usually available, but where there is more disaggregate information on prices and shares, it provides an alternative output measure. Tretheway *et al.* (1997) include a Transport Canada output price index based on disaggregate data to construct an output quantity index for combined Canadian Class I railways.

The Divisia-Tornqvist TFP index procedure

We turn from the data measurement to the indexing procedure. Theoretical analyses make use of the Divisia Index. This assumes continuous and instantaneous changes; that is, aggregate output (Y) and input (X) indices have instantaneous growth rates ($\dot{Y}$ and $\dot{X}$) (Hulten, 1973, and Diewert, 1980). Since TFP = Y/X, the TFP growth rate ($T\dot{F}P$) is defined by $T\dot{F}P = \dot{Y} - \dot{X}$, which assumes continuous and instantaneous changes.

The Tornqvist Index provides a discrete time approximation to the Divisia Index (Diewert, 1976, and Grosskopf, 1993). It replaces the continuous growth rates of outputs and inputs in the Divisia index formula with the discrete difference in logarithms (Coelli *et al.*, 1998). The change in TFP is then obtained by $\Delta TFP = \Delta \log Y - \Delta \log X$.

Studies following the Divisia-Tornqvist framework include Fishlow (1966), Deakin and Seward (1969), Hariton and Roy (1979), and Gollop and Jorgenson (1980). Deakin and Seward (1969) estimate productivity changes in the UK rail industry, Hariton and Roy (1979) examine productivity of the Canadian railways, while the other studies are based on US data. Hooper (1987) reviews these studies. Gordon (1991) uses the Tornqvist formula on industry aggregate data to examine the growth rates of the US railways over the 1948-87 period, in terms of inputs, outputs, and "multi-factor" productivity. These studies and others are listed in Table 2.

Following the same general framework, the US Bureau of Labor Statistics constructs a "multifactor productivity" measure for the US rail industry. This measure relates railway "total" output to the combined inputs of labour, capital, and intermediate purchases (Duke *et al.*, 1992). The outputs (passenger miles and freight ton-miles) are aggregated using operating expense data as weights, while the inputs (labour, capital, and intermediate purchases) are aggregated using cost shares. They find that the US rail industry experienced a 3.5 per cent productivity growth per year between 1958 and 1989, reflecting a 1.0 per cent annual growth in output and a 2.4 per cent decline in combined inputs. The US Surface Transportation Board (1997) uses a deflated expenditure measure of inputs and a disaggregated output index to calculate TFP; this is for limiting price increases on regulated rail rates.

The Industry Commission in Australia studied the TFP of various government business enterprises (Steering Committee, 1992) including Australian National (AN) and the New South Wales State Rail Authority (SRA). Data for AN included three capital categories and fuel, labour, and "other" inputs. Outputs were passenger-kms and freight tonne-kms. AN showed noticeable productivity growth through the 1980s (over 5.0 per cent per annum) but weakening after that. Data for the SRA were more difficult to obtain (physical proxies were used to measure capital), but TFP growth was positive during the reform process that was taking place between 1989 and 1992.

Brunker (1992) applies the concept of shadow prices to the Divisia-Tornqvist procedure in estimating TFP growth rates of the Australian National Railways. He points out that in the presence of surplus staff, direct use of cost shares as weights for aggregation of inputs leads to an over-estimation of labour's contribution to the aggregate input index. If labour inputs are declining more rapidly than other inputs, a lower weight results in a lower TFP calculation.

Multilateral TFP index procedure

The Divisia-Tornqvist index measures productivity changes over time. For comparisons across firms, Caves, Christensen and Diewert (1982a) (CCD) developed a multilateral index procedure. This multilateral index can be applied to cross-sectional data or panel data. The TFP formula can be written as follows:

$$\begin{aligned} \ln TFP_k - \ln TFP_j &= (\ln Y_k - \ln Y_j) - (\ln X_k - \ln X_j) \\ &= \sum_i \frac{R_{ik} + R_i}{2} \ln \frac{Y_{ik}}{\tilde{Y}_i} - \sum_i \frac{R_{ij} - R_i}{2} \ln \frac{Y_{ij}}{\tilde{Y}_i} \end{aligned} \quad (1)$$

$$-\sum_i \frac{W_{ik}+\overline{W}_i}{2}\ln\frac{X_{ik}}{\tilde{X}_i}+\sum_i \frac{W_{ij}+\overline{W}_i}{2}\ln\frac{X_{ij}}{\tilde{X}_i}$$

where Y_{ik} is the output i for observation k; R_{ik} is the revenue share of output i for observation k; $\overline{R}_i$ is the arithmetic mean of the revenue share of output i over all observations in the sample, and $\tilde{Y}_i$ is the geometric mean of output i over all observations; X_{ik} are the input quantities; and W_{ik} are the input cost shares. In this procedure, comparison of outputs, inputs or TFP between any pair of observations is accomplished by comparing each data point to geometric means of the entire data set. The multilateral index allows both absolute and growth rate comparisons so it is especially useful for performance comparisons.[4] It has the potential practical disadvantage that a new data item (for example, an additional year) requires that the index be re-computed entirely, and it is possible that values for previous year calculations will change because the mean values will change.

Freeman *et al.* (1985) use this multilateral TFP index to compare productivity growth of Canadian Pacific (CP) and Canadian National (CN), for the period 1956-81. They find that both railways achieved substantial TFP growth during the period, with CP's TFP growth (3.5 per cent) exceeding CN's (3.1 per cent). Tretheway *et al.* (1997) expand the data series to 1991, and find that although both CP and CN sustained modest TFP growth throughout the period 1956-91, their performance slipped during the 1980s, partly as a result of slower output growth. Sensitivity analysis conducted on alternative ways of computing TFP measures shows that the calculation of TFP growth rate is sensitive to a variety of underlying assumptions and calculation procedures, underscoring the importance of using a proper method for computing TFP. Even then, calculations can vary up to a full percentage point, depending on particular assumptions and computational procedures (Tretheway and Waters, 1995a, 1995b; Tretheway *et al.*, 1997).

Using the CCD multilateral index procedure, Hensher *et al.* (1995) construct two sets of gross TFP indices, demand side TFP and supply side TFP, depending on whether the final output or intermediate output measure is used, for five public rail systems in Australia over the period 1971/72 to 1991/92. They suggest the use of this TFP index as a reference for benchmarking each railway to evaluate the productivity implications of changes in the operating and managerial environment. Quite different conclusions could be drawn for intermediate or final outputs; it is possible to be very efficient at *running* trains, but not so efficient at satisfying final customers. (In terms of intermediate outputs, the most efficient way to run passenger trains is to carry no passengers because they interfere with the operation of the trains.)

[4] Both the Tornqvist index and the CCD multilateral index are "superlative" indices in that those formulae can be derived from underlying translog forms of production technology. Since the translog functional form is "flexible", and thus gives quadratic approximation to an unknown true function (Diewert, 1976, and Caves, Christensen and Diewert, 1982b), it is attractive for empirical researchers because precise forms of underlying technology are usually unknown. However, these two and other superlative indices do not necessarily produce identical TFP numbers. Diewert (1992) shows that in the multiple-output multiple-input case, different index number techniques generally produce different productivity measures.

The TFP indices discussed in this section yield a "gross" measure of productivity changes. They do not distinguish among sources of productivity growth. Furthermore, by using input cost shares for aggregation of inputs, it is assumed that input prices are "correct"; that is, there is no change in allocative inefficiency. Similarly, aggregation of outputs using revenue shares as weights assumes that relative prices of multiple outputs are proportional to their respective marginal costs. In practice, both of these input and output aggregation conditions are likely to be violated, but to some extent they can be corrected for via decomposition analysis, discussed below.

3.3 Decomposition of TFP into sources

Strictly speaking, index-number-based productivity measures can be used for making inferences about the change in overall productive efficiency only if there is no difference (change) in operating environments between the firms (over time), and no change in scale economies. In practice, operating environments and scale of outputs may be very different between firms and change over time within a firm. Therefore, in order to make inferences about productive efficiency it is necessary to separate out these influences on the "gross" measure of TFP. Two alternative procedures for accomplishing this are described below.

Formal decomposition of TFP

Denny *et al.* (1981) derive the following formula to decompose TFP growth into effects of output scale, non-marginal cost pricing of outputs, and residual productive efficiency:

$$T\dot{F}P = \dot{Y}^p - \dot{F} = (1 - \varepsilon_Y)\dot{Y}^c + [\dot{Y}^p - \dot{Y}^c] + E \qquad (2)$$

where $T\dot{F}P$ is the TFP growth rate; Y^p is the growth rate of the output aggregated by using revenue shares as the weights for aggregation; Y^c is growth rate of the output aggregated by using cost elasticities as the weights for aggregation; F is the growth rate of inputs; and

$$\varepsilon_y = \sum_i (\partial \ln C / \partial \ln Y_i)$$

is the sum of the cost elasticities with respect to outputs, which needs to be estimated via a cost function. The first term on the RHS of equation (2) is TFP growth attributable to output growth (change in scale). The second term is the effect of changes in extent of non-marginal cost pricing of outputs on TFP growth. The last term E is residual TFP growth due to productive efficiency.[5] It is important to note that this decomposition formula requires information on cost elasticities with respect to outputs, and marginal costs of all outputs, which are not normally available without estimating a neoclassical cost function. However, once a cost function is estimated, it contains all the information for computing changes in "gross" TFP and TFP changes attributable to scale of output and residual changes in TFP (productive efficiency).[6] Therefore, while this decomposition

[5] Bauer (1990a) expands this decomposition approach to distinguish the effects of productive efficiency between the effects of allocative and technical efficiencies.

formula is useful to show the accounting identity between changes in gross TFP and its components, researchers do not normally need to follow this decomposition formula to measure (residual) productive efficiency. They would compute it directly from the cost function. This point has not been recognised in the literature.

Use of regression analysis to decompose a TFP index

Some studies have adopted a different approach for TFP decomposition. Caves, Christensen and Tretheway (1981) regress the TFP index on a number of variables, such as output and network characteristics, to attribute TFP differentials to sources.[7] Essentially, a decomposition regression includes the same variables included in a cost function. Freeman *et al.* (1985) explore sources of TFP growth by regressing TFP measures on various combinations of variables including route miles, average trip length, average length of haul, firm dummy variables, and so on. They provide an estimate of productive efficiency in the form of residual or unexplained TFP levels. Their results show that some TFP growth can be explained by economies of traffic density, while economies of firm size do not appear to be an important factor. Tretheway *et al.* (1997) also decompose TFP differences into a number of sources, and compute a residual TFP growth. Their results indicate that of the average 3.4 per cent per annum TFP growth for the two Canadian carriers in the 1956-91 period, less than half is explained by TFP regressions. Average residual TFP growth is essentially the same for the two carriers, and is estimated at 1.8 ~ 1.9 per cent per annum.

Hensher *et al.* (1995) give a good example of the decomposition regression approach. They regress the gross TFP measure on variables to account for the influence of scale, density, technology, changes in management, and excess capacity, on railway performance. A residual TFP measure is derived after controlling for these sources. Results show that differences in scale, density, output composition, and excess capacity, explain a significant portion of gross TFP differentials, and a significant portion of the remaining TFP differentials can be explained by particular innovations in technology and management practices.

3.4 Data Envelopment Analysis

Data Envelopment Analysis (DEA), introduced in Charnes *et al.* (1978), is a non-parametric approach for measuring efficiency. DEA involves an application of linear programming (LP) to observed data to form a production frontier, against which to evaluate the efficiency of each firm or organisation. Specifically, DEA utilises a sequence of linear programs, one for each observation (a firm or an organisation observed at a time), to construct a piecewise linear production frontier, and to compute an efficiency index for

6 Since multiple outputs can be included in a cost function without output aggregation, differential deviations of output prices from respective marginal costs do not cause any problem for identifying productive efficiency from a cost function.

7 They have shown that the Cobb-Douglas form of TFP regression is equivalent to a Cobb-Douglas cost function.

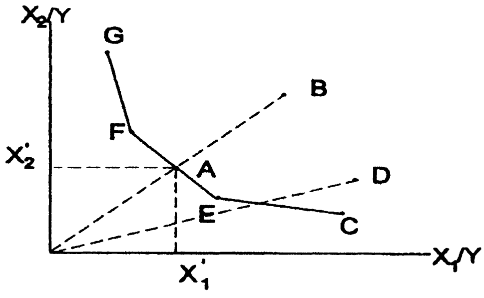

Figure 4
Diagrammatic Interpretation of DEA

each observation relative to the frontier. Observations that lie on the production frontier are deemed efficient, while those not on the frontier are regarded as being inefficient.

Figure 4 illustrates a hypothetical one-output two-input case. Firms G, F, E, and C constitute the "best practice" (frontier), since no other firms in this sample produce the same level of output using less of at least one of the inputs. Firms B and D are inefficient relative to the frontier.

DEA can accommodate multiple outputs and multiple inputs. The relative efficiency of an observation is defined as the ratio of its total weighted output (virtual output) to its total weighted input (virtual input). The weights (virtual multipliers) are determined by LP optimisation. DEA allows each observation to select the weights that maximise its own efficiency score. Generally, higher weights would be given to inputs used relatively less and outputs produced relatively more. The DEA index value of 1 (unity) implies that the observation is on the efficient or "best practice" frontier, while a value less than unity implies that its performance is poorer than that which could be achieved.

Assuming convexity of production possibility sets, Charnes, Cooper and Rhodes (1978) (CCR) define the DEA efficiency index as the maximum of a ratio of weighted outputs to weighted inputs, subject to the condition that similar ratios for every observation be less than or equal to unity. In mathematical form, efficiency of the *k*th observation can be obtained by solving the LP problem in (3):

$$h_k^* = Max \frac{\sum_{r=1}^{s} u_r Y_{rk}}{\sum_{i=1}^{m} v_i X_{ik}} \tag{3}$$

$$\text{st.}\quad \frac{\sum_{r=1}^{s} u_r Y_{rj}}{\sum_{i=1}^{m} v_i X_{ij}} \leq 1 \qquad j = 1,2\ldots n$$

$$u_r, v_i \geq \varepsilon \qquad r = 1,\ldots,s;\ i = 1,\ldots,m$$

where the Y_{rj}, X_{ij} are the known outputs and inputs of the *j*th observation; u_r, v_i are the weights (virtual multipliers) to be determined by the solution of the problem; and ε represents a small positive quantity introduced to ensure that all the observed inputs and outputs will have "some" positive value assigned to them. In this model, $h_k^* = 1$ if and only if the *k*th observation is efficient relative to other observations.

This original CCR model assumes constant returns to scale, and yields an index of overall efficiency. Extensions to this model have resulted in a variety of alternative formulations, all sharing the principle of envelopment. One widely-used formulation is the so-called BCC model (Banker, Charnes and Cooper, 1984). The BCC model distinguishes between technical and scale efficiencies by estimating pure technical efficiency at a given scale of operation.

From the above discussion, a number of features of DEA become apparent. First, the DEA efficiency ratios are greatly dependent on the observed best practices in the sample. As a result, DEA tends to be very sensitive to outliers and measurement errors. Second, since the weights for each observation are chosen so as to give the most favourable efficiency ratio possible, subject to the specified constraints, DEA evaluates an observation as efficient if it has the best ratio of any one output to any one input. As a consequence, DEA efficiency ratios are sensitive to selection of inputs and outputs included in the analysis.[8]

Bookbinder and Qu (1993) use DEA to compare the 1989 performance of .wo Canadian (CN and CP) and five US Class I railways. Three DEA models are estimated by including different numbers of railways in the sample, and the results indicate Burlington Northern (BN) as the most efficient railway, and Canadian National (CN) as the least efficient (28 per cent less efficient than BN). The study also conducts some sensitivity analysis on selection of outputs and inputs, and finds that the choice of inputs and outputs does affect efficiency ratings.

Oum and Yu (1992, 1994) use DEA to measure productive efficiency of railway systems in 19 OECD countries over the 1978-89 period. Two alternative sets of output measures are used in computing the DEA index: (1) revenue output measures (passenger-kilometres and freight tonne-kilometres); and (2) available output measures (passenger train-kilometres and freight train-kilometres). Both sets of DEA indices indicate

[8] For a detailed analysis of the effects of outliers on DEA indices and comparisons with other efficiency measurement techniques, see Yu (1995, 1998).

that DSB (Denmark) and VR (Finland) made significant improvements during the period, while CFL (Luxembourg) and NSB (Norway) experienced noticeable declines.

Similar to the TFP index, the DEA index is a "gross" efficiency indicator, which reflects both productive efficiency and the effects of operating environments. Therefore, a regression analysis can be performed on the DEA gross efficiency index to compute a residual efficiency index by removing the effects of different operating environments. Oum and Yu (1994) use a Tobit regression[9] on the DEA index to identify the effects on efficiency of public subsidies and the extent of managerial autonomy, and to compute a residual efficiency index. Residual efficiency results indicate that in 1989, British Rail (UK), NS (Netherlands), SJ (Sweden), and VR (Finland), were among the most efficient performers, while CH (Greece), and OBB (Austria), were among the least efficient ones. They find that railway systems with high dependence on public subsidies are significantly less efficient than similar railways with less dependence on subsidies, and railways with a high degree of managerial autonomy from the regulatory authority tend to achieve higher efficiency.

4. Conventional Econometric Methods

Econometric methods involve estimation of a production or cost function. The estimated production or cost function can then be used to identify changes in productivity or productive efficiency.

Solow (1957) appears to be the first to demonstrate that the rate of productivity growth can be identified with the rate of Hicks-neutral technical progress, assuming constant returns to scale and competitive input markets. Following the Solow model, the conventional econometric approach estimates the following production function:

$$y^t = f(x^t, t) + \varepsilon \tag{4}$$

for $t = 1,2,\ldots T$. The estimated parameters are then used to solve for rate of technical progress as $\partial \ln f(x^t,t)/\partial t$. There can be a substantial difference between the "gross" productivity measured via an index number procedure, and the "shift" or "technical progress" concept of productivity measured by an econometric method. These two concepts give equal empirical results only if every firm is on its production frontier and operates in an identical set of operating environments. Otherwise, the two concepts could yield very different empirical results (although a decomposition regression of TFP results might reconcile much of the difference).

Because it is difficult to estimate a production function when firms produce more than one output, cost-function approaches have been developed based on the early work on duality theory of Shephard (1953, 1970), Uzawa (1964), Diewert (1974), and McFadden (1978). A cost function, which is dual to production technology, can be easily applied to multiple-output situations. The cost function can be specified as follows:[10]

---

9 Conventional regression could not be used because the dependent variable, efficiency index, lies between zero and one.

10 The cost (production) function for railways and other transport firms generally includes other exogenous influences on production and costs, including attributes of outputs such as average load, length of haul, and other network characteristics, quality of service, and so on.

$$C^t = C(y^t, w^t, t). \tag{5}$$

Logarithmically differentiating the cost function with respect to time decomposes the rate of growth of total cost into its sources: changes in input prices, growth of output, and rate of cost reduction due to technical progress (Gollop and Roberts, 1981):

$$\frac{\partial \ln C}{\partial t} = \sum_{n=1}^{N} \frac{\partial \ln C}{\partial \ln w_n} \frac{\partial \ln w_n}{\partial t} + \frac{\partial \ln C}{\partial \ln y} \frac{\partial \ln y}{\partial t} + \frac{\partial \ln C}{\partial t}. \tag{6}$$

The rate of technical progress equals the negative of the rate of growth of total cost with respect to time, holding output and input prices constant; that is, $-\partial \ln C(w_n, y, t)/\partial t$. In a regression, this is the parameter measuring the shift in the cost function over time.

The total cost function assumes that firms adjust all inputs instantaneously as outputs change. However, in practice firms may not be able to adjust all inputs (especially capital stocks and, in some cases, labour) as outputs change. In order to account for the short-run disequilibrium adjustment in capital stock, many studies estimate variable cost functions, in which capital stock is treated as a fixed input (for example, Caves, Christensen and Swanson, 1981a, and Gillen *et al.*, 1990). Oum and Zhang (1991) observe that most of the estimated variable cost functions had incorrect (positive) signs for the capital stock variable, implying that the shadow value of capital input is negative, and show analytically that an incorrect sign is caused by a common mis-specification of the variable cost function. In order to solve this problem, they suggest that a measure of capital service flow (in place of the capital stock level) should be included in the variable cost function. Since it is difficult to measure capital service flow, they further suggest the use of the utilised capital input; that is, capital stock multiplied by utilisation rate.

Another issue in cost function specification is that there may be systematic differences between firms otherwise employing the same technology; for example, differences in terrain or market location. These exogenous influences on cost-output relationships need to be incorporated into the cost function. Network variables Z could allow for differences in terrain, weather, or exogenous characteristics of the market area served, such as more favourable directional flows of cargo and/or shorter average lengths of haul. Firm dummy variables and firm-specific trend variables are sometimes incorporated in a production or cost function to measure and compare differences in (residual) productive efficiency across firms and over time (see Friedlaender *et al.*, 1993, for an example). One danger with this approach is that there may be high collinearity between firm dummy variables (F) and output or network variables (Y or Z), especially in panel data which include both very large and very small firms (Xu *et al.*, 1994, and Oum and Zhang, 1997). Although, theoretically, high collinearity does not lead to biased parameter estimates, in a finite sample it often affects point estimates of parameters, especially for outputs (Y) and network variables Z (such as average length of haul). This situation worsens as inter-firm variations in Y and Z relative to intra-firm variations in the same variables become larger. Because the firm-dummy variables "take away" a portion of cost variation which can be legitimately explained by Y and Z, it will reduce the statistical significance and size of coefficients for Y and Z. Therefore, we question the inclusion of a firm dummy variable in a production/cost function without carefully examining its impacts on other coefficients. Identifying the variables which explain residual firm effects is a good alternative avenue to pursue. Another promising way of dealing with the problem may be stochastic frontier methods, which are described in the next section.

Many studies have estimated conventional cost or production functions to assess rail productivity and efficiency (see Table 3), although only two studies are considered in detail here.[11] Caves, Christensen and Swanson (1980) estimate cost functions based on US Class I railway data for the years 1955, 1963, and 1974, in order to determine elasticities of cost with respect to outputs. They use these cost elasticities as weights for aggregation of outputs to compute an output index, which is then used to compute the TFP growth rate of the US rail industry during the 1951-74 period. They find that productivity growth averaged 1.5 per cent per annum during the 1951-74 period, a rate substantially lower than the estimates of Kendrick (1973) and others. They also show that when they apply Kendrick's weighting scheme for output aggregation to the same data, the productivity growth estimate increases substantially, to 3.6 per cent per annum.[12] This illustrates the importance of weights and aggregation.

Subsequently, Caves, Christensen and Swanson (1981a) directly estimate productivity changes through a variable cost function and find that annual productivity growth averaged 3.5 per cent for 1955-63, 0.6 per cent for 1963-74, and an average of 1.8 per cent over the 1955-74 period. Caves, Christensen and Swanson (1981b) incorporate Canadian railways into the analysis for the same time period. Contrary to the US trend of declining productivity growth, Canadian railways had shown productivity improvement. The two large Canadian Railways achieved higher productivity growth, particularly in the later years, and this can be attributed to the partial Canadian rail deregulation. Subsequent analysis by Caves, Christensen, Swanson and Tretheway (1982), using an expanded data set, reaffirmed their earlier results.

A somewhat different method is used by De Borger (1992) to study the cost structure and productivity growth in Belgian railway operations. A generalised Box-Cox cost function for multiple-output technologies is estimated, and the two productivity growth measures are defined: (a) the common rate at which outputs can grow over time, with all inputs held constant; and (b) the common rate at which inputs can be reduced over time, with all outputs held at a fixed level. The two will differ, unless there are constant returns to scale. The results indicate that average annual productivity growth ranged be-

[11] See Oum and Waters (1996) for further discussion of transport cost functions including choice of functional form, level of output disaggregation, treatment of output attributes, and treatment of fixed or quasi-fixed inputs.

[12] Kendrick's gross TFP measure (5.1 per cent per annum) is higher than Caves, Christensen and Swanson's (3.6 per cent). The two figures are not comparable exactly. The time periods were different (1948-66 versus 1951-74, respectively), and Kendrick's analysis included labour and capital only, while Caves *et al.*'s analysis included labour, two types of capital inputs, energy, and materials. Differences in growth rates from different studies are not surprising. Tretheway *et al.* (1997) and Tretheway and Waters (1995a, 1995b) show noticeable differences in calculated growth rates from the same data depending on choice of base years and other computational procedures.

Table 2
Productivity and Efficiency Estimates with Index Number and DEA Procedures

Study	*Method*	*Sample*	*Findings*
The United States			
Martland (1989,1997)	Relate partial measures to revenues and costs	US:1973-83 US: 1977-95	net productivity savings at $65mill
Fishlow (1960)	Divisia-Tornqvist Type Index	US: 1839-1910	TFP growth at 3.4% per year
Kendrick (1961)	Divisia-Tornqvist Type Index	US: 1889-1953	TFP growth at 2.6% per year
Kendrick (1973)	Divisia-Tornqvist Type Index	US: 1948-66	TFP growth at 5.1% per year
Meyer & Morton (1975)	Laspreyre and Paasche Index	US: 1947-70	TFP growth at 1.5-2.4% per year
Gollop & Jorgenson (1980)	Divisia-Tornqvist Type Index	US: 1947-73	TFP growth at 2.2% per year
Kendrick & Grossman (1980)	Divisia-Tornqvist Type Index	US: 1948-76 US: 1966-76	TFP growth at 2.1% per year TFP growth at –1.8% per year
Gordon (1991)	Divisia-Tornqvist Type Index	US: 1948-87	MFP growth at 3.33-3.35% per year
Duke, Litz & Usher (1992)	Divisia-Tornqvist Type Index	US: 1958-89	MFP growth at 3.5% per year
STB/ICC (1997)	Divisia-Tornqvist Type Index	US: 1991-95	TFP growth at 5.0% per year
Canada			
Hariton & Roy (1979)	Divisia-Tornqvist Type Index	Canada: 1956-75	TFP growth at 3.0% per year
Freeman *et al.* (1985)	Multilateral Index TFP Regression	Canada: 1956-81	TFP growth: CP:3.5% & CN: 3.1% Residual TFP growth: 2.4% per year
Tretheway *et al.* (1997)	Multilateral Index	Canada: 1956-91 1956-81 1981-91	TFP growth 3.4% per year 2.7% per year
	TFP Regression	1981-91	Residual TFP growth: 1.8% per year
Bookbinder & Qu (1993)	DEA	Canada & US: 1989	BN most efficient; CN least efficient
Europe and OECD			
Nash (1985)	Partial Measures	Western Europe	SJ & NS: high labour productivity
Jackson (1991, 1992, 1993)	Partial Measures	European Railways	Annual performance surveys
Barrett (1991)	Partial Measures	Ireland: 1980-89	62% increase in cost per unit
Nash & Preston (1994)	Partial Measures	Europe: 1977, 1990	Market share down by 12-20% Labour productivity up by 27% Revenue/Cost ratio up by 22%
Preston (1996)	Partial Measures Multilateral TFP Index Translog Cost Function Cost-elasticity based TFP	Europe: 1977, 1990	Same as above SJ top; FS & SNCB the worst SJ top; OBB the worst SJ & CIE top; FS the worst
Deakin & Seward (1969)	Divisia-Tornqvist Type Index	UK: 1952-65	TFP growth at –0.6% per year
Oum & Yu (1994)	DEA	OECD: 1978-89	DSB & VR: most improved CFL & NSB: declined most
	DEA-Tobit		DSB & SNCB: improved CFL, CH & NSB: declined
Other Studies			
Brunker (1992)	Divisia-Tornqvist Index with shadow prices	Australia: 1979-88	TFP growth at 3.4-4.9% per year
Steering Committee (1992)	Divisia-Tornqvist Type Index	Australia - SRA: 1981-91 Australia - AN: 1979-91	TFP growth of 2.1% per year TFP growth of 4% per year
Hensher *et al.* (1995)	Multilateral Index TFP Regression	Australia: 1971-92	TFP growth: 2.4~3.2% per year Residual TFP growth: 1.56~2.4% per year
Nakajima *et al.* (1996)	Divisia-Tornqvist Type Index production divided into "3T"	Japan: 1956-86	TFP growth at –3.66% per year
Thompson *et al.* (1991)	Partial Measures	World Railways	Annual Performance Survey
Thompson & Fraser (1993)	Partial Measures	World Railways	Annual Performance Survey

Table 3

Productivity and Efficiency Estimates with Econometric Methods

Study	*Method*	*Sample*	*Findings*
The United States			
Caves *et al.* (1980)	Cost-elasticity based TFP	US: 1951-74	TFP growth at 1.5% per year
	Divisia-Tornqvist Type Index		TFP growth at 3.6% per year
Caves *et al.* (1981a)	Translog variable cost function	US: 1955-74	Productivity growth at 1.8% per year
		1955-63	Productivity growth at 3.5% per year
		1963-74	Productivity growth at 0.6% per year
Grabowski & Mehdian (1990)	Deterministic Production Frontier	US: 1950-81	Overall efficiency down by 2%
	Revenue as output		Pure efficiency down by 5%
			Scale efficiency up by 3.6%
Bereskin (1996)	Bi-level cost function	US: 1978-93	Productivity growth: 2.11% per year
Kumbhakar (1987, 1988a,b)	Stochastic Cost Frontier	US: 1951-75	Technical efficiency down
			Allocative efficiency up
			Labour the main cause of inefficiency
Wilson (1997)	Translog Cost Function	US: 1978-89	high productivity growth following deregulation
			diminish to pre-deregulation level by 1989
Canada			
Caves *et al.* (1981b)	Translog variable cost function	US & Canada: 1956-74	
		mid 1950s	Canada 30% less productive
		by 1974	Canada 10% more productive
Caves *et al.* (1982)	Cost-elasticity based TFP	Canada: 1956-79	TFP growth: 3% and 2.2% per year
	Multilateral Index	US: 1956-79	TFP growth: 1.6% per year
	TFP Regression	Canada, US	30% difference can be explained
Roy & Cofsky (1985)	Cost-elasticity based TFP	Canada: 1956-81	TFP growth: CP, 3.1%; CN, 3.25% per year
Europe and OECD			
De Borger (1991)	Hedonic Cost Function	Belgium: 1950-86	
		1950-62	Productivity growth: 2~3% per year
		1962-86	Productivity growth: 1% per year
De Borger (1992)	Box-Cox Cost Function	Belgium: 1950-86	Productivity growth: 1.3~2.4% per year
McGeehan (1993)	Translog variable cost function	Ireland: 1973-83	Productivity growth: 7~9% per year
Dodgson (1993)	Translog variable cost function	UK: 1900-12	Productivity growth: 0.3% per year
	Multilateral Index		TFP growth: –2.68% per year
Perelman & Pestieau (1988)	Deterministic Production Frontier	OECD: 1970-83	Productivity growth: 1.03% per year
			Technical progress: 0.9% per year
			Efficiency change: 0.13% per year
Deprins & Simar (1989)	Deterministic Production Frontier	OECD: 1970-83	Exogenous factors have significant effects on efficiency estimates
Compagnie *et al.* (1991)	Deterministic Production Frontier	Europe: 1962-88	NS most efficient; NSB least efficient
	Second stage regression		TCDD most efficient; CFL least efficient
Filippini & Maggi (1991)	Deterministic cost frontier	Swiss: 1985-88	Not much variation among the railways
Gathon & Perelman (1992)	Stochastic Factor Requirement Frontier (panel data models)	Europe: 1962-88	Positive correlation between technical efficiency and autonomy
Gathon & Pestieau (1995)	Stochastic Production Frontier	Europe: 1961-88	NS most efficient, DSB least efficient
	Second stage regression		VR most efficient; CFL least efficient
Coelli & Perelman (1996a)	Output-oriented. Distance Function & Multilateral Index	Europe: 1979-83	NS most efficient; BR least efficient
Coelli & Perelman (1996b)	Multi-output Distance Function (three estimations methods)	Europe: 1979-83	SNCF most efficient, BR least efficient

tween 1.3 and 2.4 per cent over the period 1950-86, and that productivity growth was driven by changes in the composition of the energy aggregate towards more electricity, rather than by improvements in labour productivity. The results are rather robust across different cost function specifications, which is contrary to the findings of Diewert and Wales (1987).

De Borger (1991) compares differences in input substitution possibilities, returns to scale, and productivity growth by estimating cost functions with two alternative output specifications: hedonic versus general models.[13] With respect to productivity growth, the general output specification implies annual increase in productivity of up to 6 per cent over the 1950-62 period, and a decline of 2 to 3 per cent per annum during the 1963-86 period. The hedonic model suggests, however, that the productivity decline in the latter period may be due to ignoring improved operating characteristics. The hedonic model shows a productivity growth of approximately 2 to 3 per cent per annum in the 1950-62 period, and slightly more than 1 per cent per annum in the 1963-86 period. This study is a vivid example of the importance of careful specification of the output measure in estimating the cost function.

Wilson (1997) estimates a translog variable cost function for American railways over the 1978-89 period; that is, overlapping the Staggers Act (1980) deregulation of the industry.[14] He uses a simple RTM output measure, but adjusts for output characteristics (average length of haul, proportion of unit train traffic, and proportion of interline traffic). He also adjusts track mileage by speed rating. His results show substantial increases in productivity following deregulation (of 6.0 to 7.5 per cent per annum), gradually diminishing to pre-Staggers rates of improvement (about 3 per cent).

5. Frontier Econometric Measures

Traditional econometric methods for estimating cost or production functions implicitly assume that all firms are successful in reaching the efficient frontier (and only deviate randomly). If, however, firms are not always on the frontier, then the conventional estimation method would not reflect the (efficient) production or cost frontier against which to measure efficiency. For this reason, many researchers now estimate frontier production or cost functions that recognise that some firms may not be on the efficient frontier. This is essentially done by specifying the following form of error terms of the production (cost) function:

$$\varepsilon = u + v \tag{7}$$

[13] The hedonic approach adjusts the output measure for variations in output quality attributes and/or differences in operational characteristics. See Oum and Tretheway (1989) for a discussion of hedonic and general specifications of translog cost functions.

[14] A very different formulation of a cost function is given in Bereskin (1996), who proposes a "bi-level" cost function incorporating four sub-components of expenditures (maintenance of way, maintenance of equipment, transport, and general plus miscellaneous freight expenditure) as micro-cost aggregates, applied to 1978-93 US Class I railways. However, the required weak separability conditions are not tested, and the use of total operating expenses as the dependent variable means that the adjustments in capital stocks are not factored into the productivity measure.

where u is the inefficiency term resulting from a firm deviating from the efficient production (cost) frontier, and v is the traditional noise term. Depending on whether u is assumed to be a deterministic or a stochastic value, the method is called a "deterministic" or a "stochastic" frontier method.

The deterministic frontier can be estimated by a variety of methods,[15] including the corrected ordinary least squares (COLS) method,[16] or by including firm dummy and/or firm-specific time trend variables in the cost (production) function.

The stochastic frontier methods postulate that some firms fail to achieve their production (cost) frontier, and the inefficiencies cannot be fully explained by measurable variables. Thus, a one-sided error term, in addition to the traditional symmetric noise term, is incorporated in the model to capture the part of inefficiency which cannot be explicitly explained. This method was first proposed by Aigner *et al.* (1977) and Meeusen and van den Broeck (1977), and has been extended by Jondrow *et al.* (1982) and Battese and Coelli (1992) among others. Specific distributional assumptions about disturbance terms must be made in order to obtain estimates of firm-specific efficiencies.[17] Statistical noise is generally assumed to be independently and identically distributed (iid) normal, while a number of distributions have been assumed for the one-sided (inefficiency) term, such as exponential, half-normal, truncated normal, or gamma distribution.

The basic stochastic frontier model is given by:

$$y = f(x,\beta)e^{v}e^{-u}, \qquad u \geq 0 \qquad (8)$$

where y represents output; $f(x,\beta)$ is the deterministic core of the frontier production function; β are the parameters to be estimated; v is a random variable that takes value over the range $(-\infty,+\infty)$ and represents the effects of non-observable explanatory variables and random shocks; and u is a random variable that takes non-negative values and captures distance from the efficient frontier (inefficiency). Specifically, $f(x,\beta)e^{v}$ is the stochastic frontier, while e^{-u} is the measure of deviation of each observation (firm) from the frontier, that is, inefficiency. The condition $u \geq 0$ ensures that all observations lie on or below the production frontier.[18]

Although stochastic frontier methods have been widely applied in many areas such as electric utility and telecommunications, only a limited number of applications have been made to the rail industry to date (see Table 3). Some of these studies are considered below.

---

[15] Aigner and Chu (1968) made the first attempt to estimate a parametric functional form for a (Cobb-Douglas) production frontier within the theoretical framework of Farrell (1957).

[16] COLS is carried out by shifting the OLS production (cost) function by the amount of the largest positive (negative) residual, thus forming the deterministic production (cost) frontier.

[17] When panel data are available, estimates of the inefficiency disturbances can be obtained without assuming a particular distribution for the efficiency terms (Schmidt and Sickles, 1984). However, one must specify how efficiency changes over time instead.

[18] Bauer (1990b) and Greene (1993) provide good reviews of recent developments in the econometric approaches to frontier estimations; see also chapters 8 and 9 in Coelli *et al.* (1998).

5.1 Deterministic frontier methods

Perelman and Pestieau (1988), Deprins and Simar (1989), Grabowski and Mehdian (1990), Compagnie *et al.* (1991), and Filippini and Maggi (1991) have applied deterministic frontier methods to measure railway efficiency. Based on the sample of 19 railways (18 European plus Japanese National Railways) over the 1970-83 period, Perelman and Pestieau (1988) construct a translog frontier production function using the corrected OLS method. A number of exogenous factors are directly incorporated in the production function to correct for their effects on railways' observed performance, and to measure net technical efficiency. NS (Netherlands) is found to be the most efficient, and CH (Greece) the least efficient. Over the 1970-83 period, average technical progress for the 19 railways was estimated at 0.9 per cent per annum, and average increase in technical efficiency (movement towards the frontier) was estimated at 0.13 per cent per annum; hence, overall productivity growth was 1.03 per cent per annum. For comparison, the study also computes TFP growth using the Tornqvist index procedure, and finds that the two sets of productivity growth estimates are quite different, not only on average over the whole period, but even more so on a yearly basis. This is not surprising, because parametric estimation of efficiency removes the effects of changes in operating environment and scale of operation, while the Tornqvist index is a "gross" TFP measure.

The Deprins and Simar (1989) study is similar to that of Perelman and Pestieau (1988), in that both studies use essentially the same data and deterministic frontier methods with corrections for exogenous factors. However, the two studies use different sets of output and input variables. Also, Deprins and Simar (1989) estimate the production frontier using three alternative estimation techniques: a corrected OLS without correction for the exogenous factors, a non-linear least square (NLIN), and maximum likelihood method. As in Perelman and Pestieau (1988), NS of The Netherlands is identified as the most efficient company. The relative efficiency rankings of other railways are also similar to those of Perelman and Pestieau (1988). The study compares efficiency rankings with and without correcting for exogenous factors, and finds that there are considerable differences between the two sets of efficiency estimates. This shows the importance of accounting for the effects of exogenous variables.

Unlike the previous two studies, Grabowski and Mehdian (1990) estimate a ray-homothetic production frontier using the COLS method to measure "revenue" efficiency of US railways over the period 1950-81. Revenue efficiency refers to maximising revenue from the production of various outputs, and is intended to measure the overall efficiency including both allocative and technical efficiency.[19] Output and inputs are measured by (deflated) revenue and input costs. Results indicate that the main source of revenue inefficiency was operation at a non-optimal scale; that is, the scale of operation was too large. In addition, it shows that revenue efficiency improved substantially during the 1950s and the first half of the 1960s, stabilised throughout the late 1960s and the first half of the 1970s, but declined consistently in the late 1970s.

[19] Revenue efficiency represents overall efficiency only if railways' outputs are priced competitively.

Filippini and Maggi (1991) estimate a cost frontier based on the 1985-88 data for 57 private Swiss railways. A cost efficiency indicator is constructed as the ratio of actual cost of a firm over the estimated frontier; thus it reflects both allocative and technical efficiency. The study finds that the efficiency estimates are relatively high (all above 90 per cent) and do not vary much among the sample railways. This finding is consistent with the fact that the sample railways are similar in size and structure, and operate in a similar political and regulatory environment. The study also finds a significant positive correlation between cost efficiency and deficit financing by the Cantons, but no significant correlation between efficiency and equity structure.

5.2 The stochastic frontier method

As the appeal of the stochastic frontier method becomes more widespread, more researchers are using this methodology. Kumbhakar (1987, 1988a,b) is among one of the first to apply the stochastic frontier method to railways. He estimates allocative and technical inefficiency, including input-specific technical inefficiency, in a cost-minimising framework, for a panel of US Class I railways, over the period 1951-75. These studies were interested primarily in methodological development. The empirical findings require further review and analysis.

Gathon and Perelman (1992) estimate a factor (labour) requirement frontier for 19 European railways using a panel data approach, in which technical efficiency is assumed to be endogenously determined. By using the factor requirement frontier, the study implicitly assumes the existence of complementarity (fixed proportions) between all the main inputs (labour, capital, and fuel) in rail production. Net measures of inefficiency are estimated after correcting for the effects of a number of explanatory variables, including an autonomy index. The results indicate a positive correlation between managerial autonomy and technical efficiency.

Gathon and Pestieau (1995) estimate a translog production frontier to compute a gross efficiency index for 19 European railways over the period 1961-88. The average gross efficiency index over the last three years (1986-88) ranges from 0.947 for NS (Netherlands) to 0.732 for DSB (Denmark). Next, in a second stage regression, they use the autonomy index constructed by Gathon and Perelman (1992) in order to correct for inefficiency caused by a lack of managerial autonomy, and to decompose the gross efficiency into managerial and regulatory efficiency. They conclude that managerial autonomy is an important determinant of the government owned railways' performance. In an earlier version of the paper, Compagnie *et al.* (1991) estimate a deterministic production frontier in the first stage. The efficiency estimates from the deterministic frontier have a larger dispersion. This illustrates that the stochastic frontier method filters out noise, resulting in smaller dispersion in efficiency estimates.

Coelli and Perelman (1996a)[20] estimate output-oriented distance functions on a panel of 17 European railways over the period 1979-83. They use two alternative estimation

[20] Coelli and Perelman (1996b) compare three alternative methods to estimate the distance function, namely, a parametric frontier by linear programming, DEA, and COLS. They find that there is a strong correlation between the parametric linear programming and COLS methods. The paper suggests the use of geometric means of the alternative measures as the final efficiency estimate.

techniques: a deterministic frontier using COLS, and a stochastic frontier using the maximum likelihood (ML) method. Comparisons lead the authors to select the COLS estimates as the preferred estimates. They also use two alternative output measures (a multilateral output index, and total revenue as aggregate output) and conclude that use of total revenue as a measure of aggregate output is fraught with danger, while the multilateral output index appears to be a suitable method of aggregating output.

6. Summary and Concluding Remarks

This paper has surveyed methodologies used in rail productivity and efficiency studies, and the empirical findings of many studies. It began by summarising the concepts of productivity as well as technical, allocative, and productive efficiency. Although partial factor productivity measures and/or performance ratios are widely used in the rail industry, they have evident shortcomings in measuring efficiency because they ignore other outputs and inputs. Hence the majority of academic and much policy-oriented research has focused on comprehensive measures of efficiency performance, either non-parametric methods of TFP indices or DEA, or econometric estimates of production or cost functions.

As observed by Dodgson (1985) and Hooper (1987), an important policy purpose of productivity studies is to investigate the relative efficiency of private versus government ownership and the relative efficiency of firms in regulated and unregulated environments. Virtually all rail productivity and efficiency studies reviewed in this paper conclude that increased competition via regulatory liberalisation and deregulation has improved productive efficiency. For example, Canadian railways have achieved higher productivity growth than their US counterparts during the 1960s and 1970s because Canada liberalised its rail pricing regulation in the 1960s. US rail productivity growth has been significantly higher following the extensive deregulation implemented by the Staggers Act of 1980.

Many productivity studies of European railways have investigated the effect of managerial autonomy of government owned or mixed ownership enterprises on efficiency. Those studies, by and large, conclude that efficiency is positively influenced by managerial autonomy. This finding is consistent with the result of Caves and Christensen (1980), who conclude that the competitive environment is a more important factor in determining railway efficiency than ownership form, after comparing the efficiencies of CN (government owned at that time) and CP (private enterprise).

Before drawing conclusions about the advantages of different methodologies, there are some basic questions about accuracy of measurement. First, productivity/efficiency comparisons by any methodology are difficult to make with precision. There are conceptual and practical issues in identifying the number and diversity of outputs and inputs and their level of aggregation. Then there are data problems including unavailability, errors of interpretation and measurement, and the need for proxies when desired data do not exist. The computation or estimation procedures can make a difference. Even a simple measure of growth rates over time can vary substantially, depending on the choice of beginning or terminating year. Although managers and policy-makers would like to measure productivity growth to a tenth of a percentage point, the real accuracy is probably an order of magnitude lower. Comparisons across studies are probably not reliable unless restricted only to broad comparisons. Certainly it is important to look closely at databases, levels of aggregation in data, computational procedures, and underlying assumptions in comparing results from different studies.

There are a few possible generalisations. As in all analyses, more data provides greater confidence. Productivity growth is best measured as a long-term trend, because the measure can fluctuate substantially from year to year. In comparing performance differences among companies, several years' observations provide a greater accuracy, and also provide the degrees of freedom to explore possible empirical explanations for observed productivity/efficiency differences. Fortunately, the development and use of panel data sets (combining cross-section and time-series data) have become more widespread. Methodologies such as the Divisia-Tornqvist index and Caves-Christensen-Diewert (CCD) multilateral procedure (which allows the comparison of absolute productivity levels and not just growth rates) have been important advances. Other non-parametric methods, such as the Data Envelopment Analysis (DEA) method, have also been applied to measure the productivity of rail and other transport firms. However, these gross productivity measures do not distinguish between sources of productivity change, which could reflect changes in the degree of technical inefficiency, changes in scale of operations, and underlying changes in traffic mix including geographical shifts, differences in service quality or operating practices, as well as actual changes in productive efficiency (technological change). Index number methods must control for — or at least explore — possible explanations for productivity differences across firms and time.

Following Denny *et al.* (1981), some studies decompose grossTFP growth into various components. To do this one needs information about cost elasticities of outputs and marginal costs, which cannot be obtained without estimating a cost function. However, once one has a cost function, it contains all the information for computing "gross" TFP and its components, such as TFP attributable to input prices, scale and mix of outputs, network characteristics and operating environment, and residual efficiency. Therefore, there is no need for separate calculation of a TFP index. This point has not been well recognised in the literature. Alternatively, many studies now use second stage or decomposition regression of the TFP or DEA efficiency indices to measure the "residual" efficiency after removing the effects of those different operating environments that are beyond managerial control. Although these TFP or DEA decomposition regressions are useful tools to approximate pure productive efficiency, there is a need for more precise clarification of the theoretical links between these regressions and econometric cost functions. But at a minimum it is important to remember that index numbers or DEA results are gross measures of productivity. It is essential that any comparisons of productive efficiency among firms or over time take into account the differences in the factors beyond managerial control. Since a cost function can be specified in such a way as to account for the effects of all variables beyond managerial control, it is recommended practice to use a cost function method for measuring efficiency, rather than non-parametric index number approaches.

Providing that data are sufficient, most economists estimate production or cost functions for efficiency measurements and comparison. Because most firms produce more than one output, it is preferable to estimate a cost function. The estimated cost function is used to compute the rate of technical progress, that is, the shift in cost function, holding output and input prices constant. However, conventional econometric methods of estimating production or cost functions implicitly assume that all firms are successful in reaching the efficient frontier. Hence it is important to recognise and try to measure sources and levels of inefficiency to improve specification of cost/production functions. The rise of frontier estimation methods is an important development for this purpose.

The translog variable cost function has emerged as the standard functional form for most researchers. While the functional form is more or less standardised, a less visible but important element is the proper specification and measurement of capital in order to obtain reliable estimates from the cost function. In order to account for the short-run disequilibrium adjustment in capital stock, it is recommended practice to estimate a variable cost function. Furthermore, as suggested by Oum and Zhang (1991) it is desirable to include a measure of capital service flow (in place of the capital stock level) in the variable cost function.

It is customary to incorporate exogenous influences on production/costs into the regression, such as average length of haul, measures of traffic mix, and ownership form. Inevitably, the choice of these variables is ad hoc, and thus the lack of theoretical consistency of the model is a constant concern. Another dilemma is the underlying assumption of efficient use of inputs. Traditional cost or production function methods implicitly assume that all firms are successful in reaching the efficient frontier. Recognition that not all firms are efficient underscores the importance of frontier estimation techniques that recognise deviations from efficient operations. Frontier estimation techniques are still relatively novel as far as railway applications are concerned, but they are the direction for future research on rail cost estimation and efficiency measurement.

It is appropriate to close on another important direction for future research. This is the need to recognise and incorporate quality changes into the analysis. All the productivity and efficiency studies reviewed implicitly assume that quality of inputs and output are constant. But an increasing emphasis on "quality" has been stressed by probably every railway manager in the world. Improving quality absorbs increased inputs, and the output produced is not homogeneous with what was produced before. There is a need to develop comprehensive quality measures that can be explicitly introduced into the analysis. Many partial performance measures are used (for example, percentage of on-time performance) but these are not sufficient for comprehensive efficiency measures. Quality must be expressed in a way that can be valued relative to traditional output and input quantity measures. Development of comprehensive quality measures both for outputs and inputs, and incorporating them into performance measurement, should be a high priority in future research.

References

Aigner, D. J. and S. F. Chu (1968): "On Estimating the Industry Production Function". *American Economic Review*, 58, 826-39.

Aigner, D., C. A. K. Lovell and P. Schmidt (1977): "Formulation and Estimation of Stochastic Frontier Production Function Models". *Journal of Econometrics*, 6, 21-37.

Banker, R. D., A. Charnes and W. W. Cooper (1984): "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis". *Management Science*, 30, 1078-92.

Barrett, S. D. (1991): *Transport Policy in Ireland in the 1990s*. Dublin: Gill and MacMillan.

Barrett, S. D. (1995): "Measuring Rail Productivity in Ireland: A Comment". *The Economic and Social Review*, 26, 323-25.

Battese, G. E. and T. J. Coelli (1992): "Frontier Production Functions, Technical Efficiency and Panel Data, with Application to Paddy Farmers in India". *The Journal of Productivity Analysis*, 3, 153-69.

Bauer, P. W. (1990a): "Decomposing TFP Growth in the Presence of Cost Inefficiency, Nonconstant Returns to Scale, and Technological Progress". *The Journal of Productivity Analysis*, 1, 287-99.

Bauer, P. W. (1990b): "Recent Developments in the Econometric Estimation of Frontiers". *Journal of Econometrics*, 46, 39-56.

Bereskin, C. G. (1996): "Econometric Estimation of the Effects of Deregulation on Railway Productivity Growth". *Transportation Journal*, 35, 34-43.

Bookbinder, J. H. and W. W. Qu (1993): "Comparing the Performance of Major American Railroads". *Transportation Research Forum*, 33, 70-85.

Brand, H. (1974): "Problems of Measuring Railroad Productivity". *Monthly Labor Review*, 97, 26-32.

Brunker, D. (1992): "Total Factor Productivity Growth in Australian National Railways — an Application of Shadow Prices". *Transportation Research B*, 26B, 449-59.

Caves, D. W. and L. R. Christensen (1980): "The Relative Efficiency of Public and Private Firms in a Competitive Environment: the Case of Canadian Railroads". *Journal of Political Economy*, 88, 959-76.

Caves, D. W. and L. R. Christensen (1982): "Verified Statement". Submitted to US Interstate Commerce Commission, Ex Parte 290 (Sub. No. 4), 25 October.

Caves, D. W., L. R. Christensen and W. E. Diewert (1982a): "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers". *Economic Journal*, 92, 73-86.

Caves, D. W., L. R. Christensen and W. E. Diewert (1982b): "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity". *Econometrica*, 50, 1393-1414.

Caves, D. W., L. R. Christensen and J. A. Swanson (1980): "Productivity in US Railroads, 1951-1974". *Bell Journal of Economics*, 11, 166-81.

Caves, D. W., L. R. Christensen and J. Swanson (1981a): "Productivity Growth, Scale Economies, and Capacity Utilization in US Railroads, 1955-74". *American Economic Review*, 71, 994-1002.

Caves, D. W., L. R. Christensen and J. Swanson (1981b): "Economic Performance in Regulated and Unregulated Environments: a Comparison of US and Canadian Railroads". *The Quarterly Journal of Economics*, 26, 559-81.

Caves, D. W., L. R. Christensen, J. Swanson and M. W. Tretheway (1982): "Economic Performance of US and Canadian Railroads: the Significance of Ownership and the Regulatory Environment". In W. T. Stanbury and F. Thompson (eds): *Managing Public Enterprise*. New York: Praeger Publishers.

Caves, D. W., L. R. Christensen, and M. W. Tretheway (1981): "US Trunk Air Carriers, 1972-77: a 'Multilateral Comparison of Total Factor Productivity". In T. G. Cowing and R. E. Stevenson (eds.): *Productivity Measurement in Regulated Industries*. New York: Academic Press.

Charnes, A., W. W. Cooper and E. Rhodes (1978): "Measuring the Efficiency of Decision Making Units". *European Journal of Operational Research*, 2, 429-44.

Christensen, L. R. and D. Jorgenson (1969): "The Measurement of US Real Capital Input, 1929-1967". *Review of Income and Wealth*, 15, 293-320.

Coelli, T. and S. Perelman (1996a): "Efficiency Measurement, Multiple-Output Technologies and Distance Functions: with Applications to European Railways". CREPP 96/05, University of Liège.

Coelli, T. and S. Perelman (1996b): "A Comparison of Parametric and Non-Parametric Distance Functions: with Applications to European Railways". CREPP 96/11, University of Liège.

Coelli, T., D. S. P. Rao and G. E. Battese (1998): *An Introduction to Efficiency and Productivity Analysis.* Boston: Kluwer Academic Publishers.

Compagnie, I., H. Gathon and P. Pestieau (1991): "Autonomy and Performance in Public Enterprises: the Case of Railways and Postal Services". Paper presented at the CIRIEC Seminar *Public versus Private Enterprises: In Search of the Real Issues*, Liège, 4-5 April.

Deakin, B. M. and T. Seward (1969): *Productivity in Transport: a Study of Employment, Capital, Output, Productivity, and Technical Change.* London: Cambridge University Press.

De Borger, B. (1991): "Hedonic versus Homogeneous Output Specification of Railroad Technology: Belgian Railroads 1950-1986". *Transportation Research A*, 25A, 227-38.

De Borger, B. (1992): "Estimating a Multiple-Output Generalized Box-Cox Cost Function: Cost Structure and Productivity Growth in Belgian Railroad Operations, 1950-1986". *European Economics Review*, 36, 1379-98.

Denny, M., M. Fuss and L. Waverman (1981): "The Measurement of Total Factor Productivity in Regulated Industries, with an Application to Canadian Telecommunications". In T. G. Cowing and R. E. Stevenson (eds.): *Productivity Measurement in Regulated Industries.* New York: Academic Press.

Deprins, D. and L. Simar (1989): "Estimating Technical Inefficiencies with Correction for Environmental Conditions with an Application to Railway Companies". *Annals of Public and Cooperative Economics*, 81-101.

Diewert, W. E. (1974): "Application of Duality Theory". In M. D. Intriligator and D. A. Kendrick (eds.): *Frontiers of Quantitative Economics.* Amsterdam: North-Holland.

Diewert, W. E. (1976): "Exact and Superlative Index Numbers". *Journal of Econometrics*, 4, 115-145.

Diewert, W. E. (1980): "Capital and Theory of Productivity Measurement". *American Economic Review*, 70, 260-67.

Diewert, W. E. (1992): "The Measurement of Productivity". *Bulletin of Economic Research*, 44, 163-98.

Diewert, W. E. and T. J. Wales (1987): "Flexible Functional Forms and Gloabal Curvature Conditions". *Econometrica*, 55, 43-68.

Dodgson, J. S. (1985): "A Survey of Recent Developments in the Measurement of Rail Total Factor Productivity". In K. J. Button and D. E. Pitfield (eds.): *International Railway Economics.* Aldershot: Gower.

Dodgson, J. S. (1993): "British Railway Cost Function and Productivity Growth, 1900-1912". *Explorations in Economic History*, 30, 158-81.

Duke, J., D. Litz and L. Usher (1992): "Multifactor Productivity in Railroad Transportation". *Monthly Labor Review*, August, 49-58.

Farrell, M. J. (1957): "The Measurement of Productive Efficiency". *Journal of the Royal Statistical Society*, Series A (General), 120, 253-81.

Filippini, M. and R. Maggi (1991): "Efficiency and Ownership in the Case of the Swiss Private Railways". Paper presented at CIRIEC Seminar *Public versus Private Enterprises: In Search of the Real Issues*, Liege, 4-5 April.

Fishlow, A. (1966): Productivity and Technical Change in the Railroad Sector, 1840-1910". *Studies in Income and Wealth*, 30, 583-646.

Freeman, K. D., T. H. Oum, M. W. Tretheway and W. G. Waters II (1985): "The Total Factor Productivity of the Canadian Class I Railways". *The Logistics and Transportation Review*, 21, 249-76.

Freeman, K. D., T. H. Oum, M. W. Tretheway and W. G. Waters II (1987): "The Growth and Performance of the Canadian Transcontinental Railways 1956-1981". Centre for Transportation Studies, University of British Columbia.

Friedlaender, A., E. R. Berndt, J. S. Chiang, M. Showalter and C. A. Vellturo (1993): "Rail Costs and Capital Adjustments in a Quasi-Regulated Environment". *Journal of Transport Economics and Policy*, 27, 131-52.

Gathon, H.-J. and S. Perelman (1992): "Measuring Technical Efficiency in European Railways: a Panel Data Approach". *The Journal of Productivity Analysis*, 3, 135-51.

Gathon, H.-J. and P. Pestieau (1995): "Decomposing Efficiency into its Managerial and its Regulatory Components: the Case of European Railways". *European Journal of Operational Research*, 80, 500-07.

Gillen, D. W., T. H. Oum and M. W. Tretheway (1990): "Airline Cost Structure and Policy Implications". *Journal of Transport Economics and Policy*, 24, 9-34.

Gollop, F. M. and D. W. Jorgenson (1980): "US Productivity Growth by Industry, 1947-1973". In J. Kendrick and B. Vaccaro (eds): *New Developments in Productivity Measurement*. Chicago: Chicago University Press.

Gollop, F. M. and M. J. Roberts (1981): "The Source of Economic Growth in the US Electric Power Industry". In T. G. Cowing and R. E. Stevenson (eds): *Productivity Measurement in Regulated Industries*. New York: Academic Press.

Gordon, R. J. (1991): "Productivity in the Transportation Sector". NBER Working Paper No. 3815, National Bureau of Economic Research, Cambridge, MA.

Grabowski, R. and S. Mehdian (1990): "Efficiency of the Railroad Industry: a Frontier Production Function Approach". *Quarterly Journal of Business and Economics*, 29, 26-42.

Greene, W. H. (1993): "The Econometric Approach to Efficiency Analysis". In H. O. Fried, C. A. Knox Lovell and S. S. Schmidt (eds): *The Measurment of Productive Efficiency: Techniques and Applications*. Oxford: Oxford University Press.

Grosskopf, S. (1993): "Efficiency and Productivity". In H. O. Fried, C. A. Knox Lovell and S. S. Schmidt (eds.): *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford: Oxford University Press.

Hariton, G. and R. Roy (1979): "Productivity Change in Canadian Air and Rail Transport in the last Two Decades". *Logistics and Transportation Review*, 15, 507-16.

Hensher, D. A., R. Daniels and I. Demellow (1995): "A Comparative Assessment of the Productivity of Australia's Public Rail Systems 1971/72-1991/92". *The Journal of Productivity Analysis*, 6, 201-23.

Hooper, P. (1987): "Productivity Change in Transport: a Survey". *Transport Reviews*, 7, 341-67.

Hulten, C. R. (1973): "Divisia Index Numbers". *Econometrica*, 41, 1017-25.

Jackson, C. (1991): "European Railway Performance 1990". *Railway Gazette International*, 147, 483-84.

Jackson, C. (1992): "Swedes Ahead in European Rail Productivity". *Railway Gazette International*, 148, 479-81.

Jackson, C. (1993): "Passenger Traffic Holds Up as Freight Collapses". *Railway Gazette International*, 149, 499-502.

Jha, R. and S. P. Singh (1994): "Inter-temporal and Cross-section Variations in Technical Efficiency in the Indian Railways". *International Journal of Transport Economics*, 21, 57-73.

Jondrow, J., C. A. K. Lovell, I. Materov and P. Schmidt (1982): "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function". *Journal of Econometrics*, 233-38.

Kalirajan, K. (1990): "On Measuring Economic Efficiency". *Journal of Applied Econometrics*, 5, 75-86.

Kendrick, J. W. (1961): *Productivity Trends in the United States*. National Bureau of Economic Research, New York: Princeton University Press.

Kendrick, J. W. (1973): *Postwar Productivity Trends in the United States, 1948-1966*. National Bureau of Economic Research, New York: Columbia University Press.

Kendrick, J. W. and E. S. Grossman (1980): *Productivity in the United States: Trends and Cycles*. Baltimore: Johns Hopkins University Press.

Kopp, R. and W. E. Diewert (1982): "The Decomposition of Frontier Cost Function Deviations into Measures of Technical and Allocative Efficiency". *Journal of Econometrics*, 19, 319-31.

Kumbhakar, S. (1987): "The Specification of Technical and Allocative Inefficiency in Stochastic Production and Profit Frontiers". *Journal of Econometrics*, 34, 335-48.

Kumbhakar, S. C. (1988a): "Estimation of Input-Specific Technical and Allocative Inefficiency in Stochastic Frontier Models". *Oxford Economic Papers*, 40, 535-59.

Kumbhakar, S. C. (1988b): "On the Estimation of Technical and Allocative Inefficiency using Stochastic Frontier Functions: the Case of US Class I Railroads". *International Economic Review*, 29, 727-43.

Kumbhakar, S. (1989): "Modeling Technical and Allocative Inefficiency in Translog Production Function". *Economics Letters*, 31, 119-24.

Martland, C. D. (1989): "Improving Railroad Productivity: Implications of US Experience for Canadian Railroads". *Transportation Research Forum*, 29, 427-33.

Martland, C. D. (1997): "Sources of Financial Improvement in the US Rail Industry, 1966-1995". *Proceedings*, Transportation Research Forum, Annual Meeting, Montreal.

McFadden, D. (1978): "Cost, Revenue, and Profit Functions". In M. Fuss and D. McFadden (eds): *Production Economics: a Dual Approach to Theory and Applications*. Amsterdam: North-Holland.

McGeehan, H. (1993): " Railway Costs and Productivity Growth: the Case of the Republic of Ireland, 1973-1983". *Journal of Transport Economics and Policy*, 27, 19-32.

McGeehan, H. (1995a): "Measuring Rail Productivity in Ireland: A Note". *The Economic and Social Review*, 26, 315-21.

McGeehan, H. (1995b): "Measuring Rail Productivity in Ireland: A Reply". *The Economic and Social Review*, 26, 327-29.

Meeusen, W. and J. van den Broeck (1977): "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error". *International Economic Review*, 8, 435-44.

Meyer, J. R. and A. L. Morton (1975): "The US Railroad Industry in the Post-World War II Period: a Profile". *Explorations in Economic Research*, 2, 449-501.

Nakajima, T., K. Oda and T. Takahashi (1996): "Total Factor Productivity of the Japanese National Railways Based on the '3T' Activities". KEO Discussion Paper no.45, KEIO University, Tokyo.

Nash, C. (1981): "Government Policy and Rail Transport in Western Europe". *Transport Review 1981*, 1, 225-50.

Nash, C. (1985): "European Railway Comparisons — What Can We Learn?" In K. J. Button and D. E. Pitfield (eds): *International Railway Economics*. Aldershot: Gower Publishing.

Nash, C. and J. Preston (1994): "Railway Performance — How Does Britain Compare?" *Public Money and Management*, 14, 47-53.

Oum, T. H. and M. W. Tretheway (1989): "Hedonic versus General Specifications of the Translog Cost Function". *The Logistics and Transportation Review*, 25, 3-21.

Oum, T. H., M. Tretheway and W. G. Waters (1992): "Concepts, Methods and Purposes of Productivity Measurement in Transportation". *Transportation Research A.*, 26A, 493-505.

Oum, T. H., M. W. Tretheway and Y Zhang (1990): "A Note on Capacity Utilization and Measurement of Scale Economics". *Journal of Business and Economic Statistics*, 9, 119-23.

Oum, T. H. and W. G. Waters (1996): "A Survey of Recent Developments in Transportation Cost Function Research". *The Logistics and Transportation Review*, 32, 423-63.

Oum, T. H. and C. Yu (1992): *An International Comparison of the Economic Efficiency of Passenger Railway Systems*. Research Division Report No. RES 91-21, The Royal Commission on National Passenger Transportation, Ottawa.

Oum, T. H. and C. Yu (1994): "Economic Efficiency of Railways and Implications for Public Policy: a Comparative Study of the OECD Countries' Railways". *Journal of Transport Economics and Policy*, 28, 121-38.

Oum, T. H. and Y. Zhang (1991): "Utilisation of Quasi-Fixed Inputs and Estimation of Cost Functions". *Journal of Transport Economics and Policy*, 25, 121-34.

Oum, T. H. and Y. Zhang (1997): "A Note on Scale Economics in Transport". *Journal of Transport Economics and Policy*, 31, 309-15.

Perelman, S. and P. Pestieau (1988): "Technical Performance in Public Enterprises: a Comparative Study of Railway and Postal Services". *European Economic Review*, 32, 432-41.

Preston, J. M. (1996): "Assessing Public Sector Performance: the Case of Western European Railways". In D. Hensher, J. King and T. H. Oum (eds): *Proceedings of the 7th World Conference on Transport Research, Volume 4: Transport Management*, 343-56.

Reebie Associates (1988): "Railroad Productivity Evaluation: Proposed Measures for RCAF and URCS Applications". Washington DC.

Roy, J. P. and D. Cofsky (1985): "An Empirical Investigation for Canadian Class I Railroads of Both Performance and Industry Cost Structure". *Proceedings of the 20th Canadian Transportation Research Forum Annual Meeting*, Toronto, 766-84.

Schmidt, P. and C. A. K. Lovell (1979): "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers". *Journal of Econometrics*, 9, 343-66.

Schmidt, P. and R. Sickles (1984): "Production Frontiers and Panel Data". *Journal of Business and Economic Statistics*, 2, 367-74.

Schwier, C., J. Jones and A. Pignal (1990): "Canadian Regional Railway Passenger Services: Was There Ever Any Potential For Financial Improvement?" *Proceedings of the 25th Canadian Transportation Research Forum*, Saskatoon, 254-67.

Shephard, R. W. (1953): *Cost and Production Functions*. Princeton: Princeton University Press.

Shephard, R. W. (1970): *Theory of Cost and Production Functions*. Princeton: Princeton University Press.

Simar, L. (1992): "Estimating Efficiencies from Frontier Models with Panel Data: a Comparison of Parametric, Nonparametric and Semi-parametric Methods with Bootstrapping". *The Journal of Productivity Analysis*, 3, 171-203.

Solow, R. A. (1957): "Technical Change and the Aggregate Production Function". *Review of Economics and Statistics*, 39, 312-20.

Steering Committee on National Performance Monitoring of Government Trading Enterprises (1992): *Measuring the Total Factor Productivity of Government Trading Enterprises*. Industry Commission, Belconnen, ACT Australia.

Surface Transportation Board (formerly Interstate Commerce Commission) (1997): *Railroad Cost Recovery Procedures – Productivity Adjustments*. Decision STB Ex Parte 290 (Sub-No. 4), Washington DC.

Thompson, L. S., W. G. Wood and F. Lures (1991): "World Railway Performance Update". *Developing Railways*, 8-16.

Thompson, L. S. and J. M. Fraser (1993): "World Bank Survey of Railway Performance". *Developing Railways*, 5-13.

Tretheway, M. W. and W. G. Waters II (1990a): "Productivity Trends in Canadian and US Railroads". *Proceedings of the 25th Canadian Transportation Research Forum*, Saskatoon, 436-50.

Tretheway, M. W. and W. G. Waters II (1990b): "Productivity Measurement and its Application to the Rail Industry". Paper presented to the Transportation Research Forum Annual Conference, Long Beach, CA.

Tretheway, M. W. and W. G. Waters II (1991): "Railroad Productivity After Staggers". *Proceedings of the Transportation Research Forum Annual Conference*, New Orleans, 521-31.

Tretheway, M. W. and W. G. Waters II (1995a): "Aggregation and Accuracy in Measuring Total Factor Productivity: Evidence from Rail Productivity Studies". *Journal of Transportation Research Forum*, 35, 60-70.

Tretheway, M. W. and W. G. Waters II (1995b): "A Sensitivity Analysis of Total Factor Productivity Measurement: Evidence from Rail Productivity Studies". *World Transport Research* (refereed Proceedings of the World Conference on Transport Research) Volume 4, *Transport Management*. New York: Elsevier.

Tretheway, M. W., W. G. Waters II and A. K. Fok (1997): "The Total Factor Productivity of the Canadian Railways, 1956-1991". *Journal of Transport Economics and Policy*, 31, 93-113.

Tubb, Paul (1977): "Measurement of Railway Operating Performance". Discussion Paper, UBC Centre for Transportation Studies, Vancouver.

Uzawa, H. (1964): "Duality Principles in the Theory of Cost and Production". *International Economic Review*, 5, 216-20.

Waters, W. G. and M. W. Tretheway (1991): "Productivity Adjustment to Price Levels in Regulated Rail Markets: Recent Developments in Canada and the United States". *Journal of Transportation Research Forum*, 32, 172-81.

Wilson, W. W. (1997): "Cost Savings and Productivity in the Railroad Industry". *Journal of Regulatory Economics*, 11, 21-40.

Xu, K., R. Windle, C. Grimm and T. Corsi (1994): "Re-evaluating Returns to Scale in Transport". *Journal of Transport Economics and Policy*, 28, 275-86.

Yu, C. (1995): "A Comparative Study of Alternative Methods for Efficiency Measurement with Applications to the Transportation Industry." Unpublished PhD dissertation, Faculty of Commerce and Business Administration, the University of British Columbia.

Yu, C. (1998): "The Effects of Exogenous Variables in Efficiency Measurement — a Monte Carlo Study". *European Journal of Operational Research*, 105, 569-80.

Date of receipt of final manuscript: May 1998