Cost Function Analysis for Transportation Modes: A Survey of Selective Literature

Rinki Sarkar

Working Paper No. 53
ABSTRACT

Cost functions as applicable to transport industries have evolved from simple analytical constructs to fairly complex structures. Salient features of the transport industry have been gradually incorporated. Theoretical properties of the cost function have been embedded into applied work, using flexible functional forms which place few apriori restrictions on the underlying production technology. Attempts have been made to represent the ‘output’ of a transport firm more accurately. These advances have made the interpretation of results more meaningful in relation to the economic characteristics of transportation industries, providing better tools for efficient formulation of policies towards these industries. This paper traces some of these developments since the mid-fifties. Towards this end, cost studies relating to various land-based transport modes is critically surveyed while focusing mainly on the urban bus transit sector.
1. Introduction

The purpose of this paper is to provide a critical survey of studies conducted for analysing costs in transportation industries since the mid-fifties. While the survey will focus mainly on the urban bus transport sector, cost studies related to other modes such as railroads, trucking and inter-city bus transport operations, which have strengthened conceptual as also methodological developments in appropriately specifying cost functions for the urban bus transit sector, will also be reviewed.

Unearthing the behaviour of 'costs' is crucial for examining economic characteristics of transportation industries. 'Cost functions' which depict the relationship between costs and factors affecting costs have been empirically estimated to throw light on various economic facets of these industries such as, the extent of scale economies, substitution possibilities between factors of production, the elasticity of demand for these factors, possibility of reaping economies of density and scope, the extent of capacity utilisation as well as productivity growth over time.

Thus the structure of costs could have significant bearing on policy issues. For instance, the nature of costs could provide sound guidelines for a sound pricing policy, for determining the optimum levels of subsidy or for prescribing the economically efficient size of an undertaking. Intra-modal cost comparisons could be facilitated, forming a basis for policies designed in accordance with the notion of 'balanced transportation' (Miller, 1970, pp. 32), meant to determine 'that traffic allocation which will satisfy transportation needs of the economy at minimum cost' (Meyer et al, 1959, pp. 16). Broader policy matters in relation to

---

1. Winston (1985, pp. 57-59) provides a brief account of literature on cost analysis of transportation industries before the mid-fifties.

the effects of regulation or deregulation, the question of privatisation and the impact of overall policy changes on efficiency could be assessed more rationally by a cost function analysis, thereby 'introducing a degree of quantification into what has often been essentially, a qualitative series of arguments' (Button, 1985, pp. 9).

Cost functions as applicable to the transport industries have evolved from simple analytical constructs to fairly complex structures. Salient features of the transport industry namely, 'the pervasive presence of government, the spatial nature of the transportation product, the importance of service quality and problems related to the temporal nature of demand' (Winston, 1985, pp. 60) have been gradually incorporated, considerably influencing these analytical developments. More specifically, theoretical properties of the cost function have been embedded into applied work, using flexible functional forms which place few a priori restrictions on the underlying production technology. These advances have made the interpretation of results more meaningful in relation to economic characteristics of transport industries, providing better tools for efficient formulation of policies towards these industries.

Another major area of research relates to the definition of 'output' of a transport firm. A peculiar feature of these industries characterised by network technologies is that output actually represents a bundle of services, varying spatially or temporally across the network. Therefore, units of transport service are not homogenous. Further, the composition and quality of services would affect operating costs. Inability to account for the heterogeneous nature of transport output in the cost function analysis would amount to overlooking important variables that may significantly affect costs. 'The analysis of costs in transportation has therefore gradually come to grips with this heterogeneous nature of transport product by

---

3 These may be thought of as routes of a network with varying service characteristics like frequency, speed, hours of operation etc. (Berechman et al., 1985)

4 Ideally, the extent of disaggregation would depend on the ability to view transport output as a vector 'Y' such that Y = \{Yijk\} where, \(Y_{ijk}\) is the flow of commodity 'k' between origin 'i' and destination 'j' at period 't' (Jara Diaz, 1982, pp. 268). Panzar (1989) asserts that '... if point-to-point transportation movements are viewed as the cost-causative outputs of the firm, a firm operating even a relatively small network must be viewed as producing an astronomical number of products' (pp. 44). Therefore, some degree of aggregations is inevitable, since otherwise a large number of parameters would have to be estimated 'exacerbating problems of multicollinearity' (McFadden et al 1978b pp. 224)
recognising that the foundations of the subject lie in the theory of the multiproduct firm as opposed to the traditional theory of the single output producer' (Winston, 1985, pp. 60).

The survey commences with an overview encompassing types of cost studies used to reflect the nature of costs in the transport sector (Section 2). The emphasis is on statistical cost studies and these will be elaborated in detail in the remaining part of the paper (Sections 3 and 4). The last section summarises our major observations.

2. Types of Cost Studies

Broadly, three approaches have been pursued for empirical measurement of costs in the transportation industries:

i) Accounting Cost Studies
ii) Engineering Cost Studies
iii) Statistical Cost Studies

It may be worthwhile to note that there may be considerable overlap among these approaches and any given study may make use of more than one. (Small, 1992, pp. 52). However, in order to provide conceptual clarity, we examine each approach as per the nomenclature given above.

2.1 Accounting Cost Studies:

These studies involve conventional cost accounting procedures. The cost of a commodity is arrived at by valuing inputs used in its production at their unit market price. For example, the total cost of bus services would be the sum of individual costs of inputs such as drivers' wages, tyres and tubes, fuel oil, repair parts etc. Cost of each input item is found by multiplying the quantities consumed by the respective unit market price of the input. This simple approach has been referred to as the 'causal factor method' (Cherwony et al, 1982, pp. 55)
An alternative accounting approach is the 'cost allocation model', which has been used widely for making intra-modal cost comparisons.\(^5\) This method essentially assumes that it is possible to allocate each input to a specific aggregate measure of output.\(^6\) For instance, wages of drivers would vary with the 'vehicle-hours' measure of output and is allocated to the same, while fuel costs are closely related to the 'vehicle-miles' served and would therefore be allocated to this measure of output. Typically, for a bus transit firm the model would look like,

\[ C = u_h \cdot V_h + u_m \cdot V_m + u_v \cdot P_v \]

where, \(C=\)total cost of bus transit services; \(V_h=\)vehicle hours of operation; \(V_m=\)vehicle miles of operation; \(P_v=\)peak vehicles used\(^7\); \(u_h=\)unit cost per vehicle-hour; \(u_m=\)unit cost per vehicle-mile; \(u_v=\)unit cost per peak vehicle. Unit costs are computed by first allocating expense items to one or more measures of transit service, \(V_h, V_m, \) or \(P_v\) and then, determining total expenses related to each aggregate measure of output, in order to derive the ratio:

\[
\frac{\text{Total expenses related to an aggregate service measure}}{\text{Total amount of that aggregate service}}
\]

Two variations of the cost allocation model have been developed. The 'fixed and variable cost allocation model' modifies the approach outlined above by classifying expense items into fixed and variable costs. While the 'temporal variation cost allocation model' accounts for variation in costs related to peak-period services as against base or non-peak period services.\(^8\)

---

\(^5\) Illustrative examples may be found in Kenneth, S., (1992, pp. 53).

\(^6\) Note, this approach is an attempt to ensure that accounting methodology better approximates economic concepts of relating costs to output measures.

\(^7\) Individual expense items which do not vary with vehicle-hours or vehicle-miles are allocated to this output measure \(P_v\). For example, costs of providing storage facilities; maintenance expenses such as maintenance of buildings, fixtures, shops and garage; overhead expenses such as general office costs, salaries of general office clerks and officials, all of which are related to peak hour vehicle needs. (Ferreri, 1969, pp. 7.)

\(^8\) These variations arise mainly from labor cost differences associated with labor agreement provisions and vehicle cost differentials associated with supporting peak period vehicle requirements. Two cost models would be developed, one for the peak period and the other for the base period.
The main criticism of accounting cost studies is their inherent inability to portray costs of a multiproduct firm, where productive facilities could be jointly used in the output of many different kinds of products and services. Arbitrary methods may have to be resorted to for allocating costs since a one to one correspondence between expense items and outputs may not be easily observable. This 'becomes a matter of guess work' (Meyer, et al, 1959, pp. 27). Lastly, accounting cost studies rely on linear cost functions which are restrictive since these implicitly reflect fixed-factor proportions technology. This is unrealistic as substitution possibilities between input factors cannot be accommodated by these linear specifications.

2.2 Engineering Cost Studies:
This approach relies on the use of engineering information for estimating cost functions. The entire production process is divided into various physical, technical phases of production and the quantities of factor inputs consumed in each phase is determined. These technically optimum input combinations are valued at the prevailing market prices and the total cost of production is the sum of costs related to each phase of production (Koutsoyiannis, 1982, pp. 122). Thus in this method, engineering production functions are converted into cost functions.

Meyer et al (1965) estimated an engineering cost function for the U.S. urban bus transport sector. The relationship developed was:

\[ \text{TBOC} = \$13120U + \$0.30M + \$9000L \]

where, TBOC=total costs; u=total number of vehicles; M=total bus-miles produced and L=lane-miles of exclusive bus-way. The entire production process may be viewed as being split up into three phases and cost of each phase determined - cost of the carriage (vehicles), cost of the carriage-way (lane miles of bus-way) and costs of actual production of bus-services (expressed as bus-miles).

'The underlying logic (of this method)\(^{10}\) is very close to conventional cost accounting' (Meyer et al, 1959, pp. 25). Therefore, this method suffers from similar shortcomings. Firstly, it may

\(^9\) For a listing of cost items used to arrive at unit cost figures see Meyer et al. (1965), pp. 216-217.

\(^{10}\) Parenthesis added.
not always be possible to compartmentalise the production process into distinct phases or operations. These phases may interact and may not be 'additively separable' (Walters, 1963, pp. 43). This creates ambiguities while trying to assess the cost of each phase and subsequently, the total cost of production. Further, allocation of costs in the case of a multiproduct firm can be arbitrary.

2.3 Statistical Cost Studies:
The methodology involved in this approach is to develop statistical relationships between costs and factor affecting costs. We may distinguish two types of statistical cost models.\(^{11}\) While reviewing literature on bus transport costs:

i) Simple Statistical Cost Models

ii) Complex Statistical Cost Models

There is an important distinction between the two types of cost models enumerated above which needs to be highlighted since it forms the basis for the analysis of bus cost models which follows. Simple statistical cost models are devoid of any theoretical base and are directly estimated using simple statistical techniques. 'They lack a solid foundation of economic and transport analysis. Therefore, the conclusions that can be drawn from their results are limited in scope and value' (Berechman, 1983, pp. 8). Complex statistical cost models are however, derivable from the neoclassical theory of production and costs, which centres around the optimal behaviour of decision units. Complex statistical tools are then used to estimate the cost function from empirical data. Developments in duality theory have enhanced the appeal of the neoclassical cost functions (which we refer to as the complex statistical cost model) since information about the underlying production technology can be ascertained from the cost function as well. 'The definition of the (neoclassical)\(^{12}\) cost function as the result of an optimisation, yields strongly mathematical properties and establishes the cost function as a sufficient statistic for all the economically relevant characteristics of the underlying technology' (Fuss et al 1978a, pp. 4)

For the purpose of this review, we have classified statistical bus transport cost studies into two groups. In Section 3, we scrutinise studies based on simple statistical cost models while

\(^{11}\) Our own classification.

\(^{12}\) Parenthesis added.
in Section 4 we critically examine studies based on complex or neoclassical cost models. Cost studies pertaining to other land-based transport modes receive wider coverage while discussing these complex cost configurations.

3. Statistical Cost Studies:

3.1 Simple Statistical Bus Cost Studies

Johnston (1956) developed a purely statistical relationship between costs and output for bus operations in U.K. The purpose of this study was 'to test various hypothesis regarding the nature of cost-output relationship by subjecting empirical data to statistical analysis' (pp. 207). A short run cost function was estimated using three years time series quarterly data relating to a single firm. In order to obtain a picture of how costs varied over a wide range of output, the long run cost function was based on a cross-section data of twenty-four bus companies of different sizes for a single year. The short run linear relationship was:

\[ C = 0.6558 + 0.4433 CM \]

where, \( C \) = operating costs; \( CM \) = car-miles of output. Based on these estimates, the short run average cost function was found to be downward sloping to the right throughout its length, supporting the hypothesis that economies of scale exist (pp. 214). In the long run analysis the dependent variable was 'log (total expenses per car-mile)', while the independent variables were log (car-miles of output); percentage of double-deckers in the fleet; percentage of fleet on fuel oil. Fleet characteristic variables, especially the one which characterised the proportion of double-deckers in the fleet, turned out to have a significant effect on costs. This

13 Plant capacity was expected to be fixed over this period.
14 It was felt that this would remove the influence of extraneous variables such as prices.
15 See Koutsoyiannis (1982, pp.138) for a discussion on the relationship between short-run, long-run analysis and the use of cross-section, time-series data.
16 When the proportionate increase in costs is less than the proportionate increase in output, economies of scale are said to exist. Symbolically, \((\frac{\Delta C}{\Delta Y})<1\), or \(\left[\frac{(\Delta C/\Delta Y)}{(C/Y)}\right]<1\), or \(M.C.<A.C.\) (where, \(C = \) total cost, \(Y = \) output, \(M.C. = \) marginal cost, \(A.C. = \) average costs), which implies that economies of scale exist when average costs are falling.
17 This varied from 5% to 100% of total fleet across the cross section of 24 bus companies.
18 This variable was included since companies used fuel oil as well as motor spirit to run the buses and fuel oil was found to be more fuel efficient (more mileage per gallon) than motor spirit.
simple relationship between unit costs and output exhibited constant returns to scale in the long run (pp. 218).

Miller (1970) broadened the framework of analysis by emphasising the importance of including 'city descriptor' variables into the cost function specification. It was hypothesised that since the 'environmental setting' differed across cities, this would undoubtedly affect bus operations and therefore costs. These differences could not be captured by single output variables such as vehicle miles. To overcome this problem, three 'city descriptor' variables were considered in this study: scheduled speed\textsuperscript{19}; intensity variable\textsuperscript{20} and city age variable\textsuperscript{21}

The model based on a cross-section sample of bus firms across 33 cities in the U.S. was developed as:

\[ C = 0.936X_1 - 0.830X_2 - 0.021X_3 + 0.157X_4 - 3.222X_5 - 0.284X_6 - 6.899X_7 + u. \]

where, \( C = \) costs per vehicle-mile; \( X_1 = \) vehicle-miles per year; \( X_2 = 1/X_1; \) \( X_3 = \) average age of fleet\textsuperscript{22}; \( X_4 = \) wage rate\textsuperscript{23}; \( X_6 = \) intensity and \( X_7 = \) city age. All three 'city descriptor' variables, especially 'scheduled speed' were found to have a significant effect on costs. The conclusion drawn was that 'city environment does play a role in explaining operator costs' (pp. 31).

The only study to delve into the nature of bus transport costs in the Indian context was conducted by Koshal (1970). The need to account for differences in terrain and traffic conditions was stressed, and analysis was carried out separately for a cross-section of three

\textsuperscript{19} A crude measure of 'speed' is used. This is given as: vehicle-miles per vehicle hour. See Braetigam et al (1982) for a formal treatment of the 'speed' variable using engineering information.
\textsuperscript{20} The 'intensity' variable was given as: total vehicle mileage per route-miles served. This variable attempts to capture the effect of level of service (route structure, frequency of service) on costs (pp. 31) which may be perceived as providing some preliminary insight into the concept of 'economies of density' to be developed at a later stage. See 'trucking' cost studies reviewed and Windle (1988).
\textsuperscript{21} The city age variable was incorporated as a dummy variable where, old city=0; new city=1. Concentration of economic activities and high residential density was a common feature of an 'old city' while the 'new city' was characterised by much more uniform dispersal of these activities (see pp. 27-28 for details). These varying urban structural characteristics would influence the operating environment and thereby costs of bus transit systems.
\textsuperscript{22} This fleet characteristic variable is important since it may affect maintenance cost.
\textsuperscript{23} The wage rate variable was found to have a significant effect on costs. Since wage costs account for 50-60% of total costs, this result emphasised the need to devise measures for raising labour productivity.
groups of firms, those operating: i) city routes, ii) long distance routes,\(^{24}\) and iii) mountainous routes. The procedure adopted was to estimate a simple linear bus cost model separately for each cost category in the case of each group, '...to bring out the behaviour and importance of various components of costs on various routes' (pp. 34). Symbolically, the model was:

\[ C_i = a_i + b_i S \]

where, \( C_i = \) 'ith' cost category \((i =\) personal costs, material costs, overhead costs, capital costs and depreciation costs); \( S = \) seat kilometers. Vehicle size varied across the sample groups. Therefore, 'seat-kilometers' as a measure of bus capacity was the preferred output measure to reflect variations in costs associated with the size of the vehicle.\(^{25}\) Results indicated that bus transport operates under constant returns to scale. Marginal costs for each section of the industry was estimated and compared with actual fares. City operators were found to charge fares below marginal costs. These services were incurring losses. Long-distance and mountainous route operators were however charging fares above marginal costs. Intra-modal cost comparisons showed marginal costs of long-distance bus services and railways\(^{26}\) to be somewhat similar. This implied that there was no indication of mode-specific 'cost advantage' in the provision of long distance services.

Amidst a policy proposal to merge a number of municipal transport undertakings in U.K., \textit{Lee and Steedman (1970)}, undertook an empirical study to provide concrete evidence on the relationship between efficiency of bus operations and size of the undertakings.\(^{27}\) Cross-section data relating to 44 municipal bus undertakings was used to build cost functions for each of the cost components\(^{28}\) separately. The independent variables were: annual bus-

\(^{24}\) Some firms operated city as well as long distance services. But since data was not available separately for city and inter-city operations, all services were grouped as inter-city services. Ideally, a multiproduct firm analysis could have been conceivable in terms of a firm providing two products: inter-city and intra-city services. Talley et al. (1986) and Taucher et al. (1983) have developed this approach. See also, Berechman (1983, pp. 22)

\(^{25}\) See Jansson (1980) wherein, 'size of the bus' is an endogenous variable in the cost function analysis.

\(^{26}\) Marginal costs for the rail mode was obtained from related research work by Koshal, ('Statistical Cost Analysis-Indian Railways', PhD. dissertation, University of Rochester, 1967).

\(^{27}\) See Chambers (1988, pp. 23, 24, 69 & 70) for a discussion on the relationship between centralisation, decentralisation and scale economies.

\(^{28}\) These were power costs, traffic operating costs, repair and maintenance costs, management and general expenses, and total working expenses.
mileage; average fleet size; price of fuel; wage rate; percentage of bus-mileage on one man operation; time distribution of demand for bus services; population density; average bus speed. Several linear and log-linear relationships were experimented with. Criteria for selecting the best cost function was highest $R^2,$ such that each independent variable was significant at least at 10% level (pp. 17). Using 'bus-miles' as the measure of output, revealed 'constant returns to scale.' It was noted that 'a different scale effect might have been observed if costs per passenger-mile had been used as the dependent variable' instead of using cost per bus-mile (pp. 27). This was not possible due to limitations of data. Nevertheless, the policy implication of this 'constant returns to scale' result suggested little scope for cost reduction through mergers and amalgamations. However, the extension of 'one-man operation' could significantly reduce costs.

In order to study the effect of 'wage rate' and 'fleet characteristics' on costs, Gray (1972) used cross-section data related to US urban bus transit firms for developing a log-linear cost function. Since the expansion of 'higher-carrying-capacity-bus' services along with enhanced ridership could presumably result in a net decline in vehicular traffic and therefore reduced congestion and pollution levels, 'the omission of (such beneficial) externalities from cost considerations may overstate the real cost of transit operations' (pp. 70). The distinguishing feature of Gray's study is that unlike the studies reviewed so far, the dependent variable,

---

29 These variables were included to account for differences in the composition and quality of service.

Time distribution of demand for bus services = Maximum no. of buses in operation (peak hours)

Average number of buses in operation (non-peak hours)

Highly accentuated peak demand for service would mean higher costs per bus-mile.

30 These variables were included in order to reflect variations in physical and traffic conditions. The vehicle utilisation ratio ($= \frac{\text{Annual Mileage}}{\text{Number of vehicles owned}}$) was used as a proxy for 'speed' due to data constraints.

31 'A log transformation compresses the scale in which variables are measured reducing a ten-fold differences between two numbers to a two-fold difference; (Griliches, 1972, pp. 34, FN. 14).

32 Data on passenger-miles was available in the U.S. much later around 1978. (Windle, 1988).

Berechman et al (1984) while studying the cost structure of ACCTD, California were able to use 'passenger-miles' data to substantiate this idea. They found 'diseconomies of scale' using 'vehicle-miles' as the measure of output but 'economies of scale' using 'passenger-miles' as the measure of output.

33 The wage rate, population density, average speed and the peak demand variable were also found to affect costs significantly.

34 Parenthesis added.

35 Viton (1981), refers to this as the 'full costs' of bus transit. Quantifying these externalities, for viewing transit costs from this broader perspective calls for indepth analysis and research.
'costs' is an aggregate of both 'operating costs' as well as 'capital costs'. The independent variables used to estimate the cost function were: bus-miles of service; hourly wage rate; bus-miles per bus-hour (= mean speed); average fleet age; average seats per bus; dummy variable (=1) if publicly owned (=0) otherwise. Overall constant returns to scale was observed. The 'wage rate' variable was found to have a strong positive effect on costs. This reinforces the findings of Miller (1970, pp. 29) as well as Lee and Steedman (1970, pp. 24). The 'mean speed' variable was likely to have strong negative effects on costs - a conclusion that was also reached by Miller (1970, pp. 31). As regards fleet characteristics, new, larger, unsubsidised buses were costlier.

Studies on urban bus transport costs surveyed so far exemplify the straightforward methodological framework involved. Essentially, the procedure consists of direct statistical estimation of the cost function by intuitively 'itemising potential influences on costs' (Williams, 1979, pp. 210). The linear and log-linear functional forms used are easy to estimate by standard least square regression techniques. Despite these advantages of procedural simplicity, a serious limitation of these studies is that the cost models developed are theoretically deficient, making it difficult to impart economic meaning to the results obtained.

The structure of the cost models impose apriori restrictions on the underlying production technology. (Williams et al, 1981, pp. 263; Berechman et al, 1985 pp. 322; Button et al, 1985, pp. 67). For instance, the linear cost function implies a Leontief technology while a log-linear cost function implies a Cobb-Douglas technology. Any investigation into substitution possibilities between factors of production is therefore ruled out and the hypothesis of a U shaped average costs curve with regions of increasing, constant and decreasing returns to scale cannot be tested. This can be ascertained from Table 1.

36 These costs refer to costs of bus capital, which was arrived at by using the relationship: $P_k = [V(S,A), \delta + \tau]$, where, $P_k =$ cost of bus capital; $S =$ seating capacity, $A =$ age; $\delta =$ depreciation rate; $\tau =$ interest rate and $V(S,A)$, the value of a bus $= V_0(S) e^{-\delta A}$, where, $V_0(S) =$ value of a new bus. This was adjusted for the UMTA capital grant formula. The final version was: $P_k = [n(1-0.67s), V_0(S) e^{-\delta A} \cdot (\delta + \tau)]$, where $n =$ number of buses, and $s =$ proportion of fleet purchased using capital grant.

37 See Chambers (1988, pp. 90-91)
Table 1: Characteristics of Simple Statistical Cost Models

<table>
<thead>
<tr>
<th>FUNCTIONAL FORM</th>
<th>UNDERLYING PRODUCTION STRUCTURE</th>
<th>ELASTICITY OF SUBSTITUTION</th>
<th>ECONOMIES OF SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Leontief</td>
<td>Zero</td>
<td>Either increasing or decreasing or constant returns to scale. (Does vary with output, but in one direction only)</td>
</tr>
<tr>
<td>Log-Linear</td>
<td>Cobb-Douglas</td>
<td>Unity</td>
<td>Constant value. (Does not vary with output)</td>
</tr>
</tbody>
</table>

Thus, inherent properties of linear and log-linear functional forms makes any bus transit cost study based on these models narrow in scope and limited in focus. 'Each of these functional forms places restrictions on the questions that may be asked or the answers that may be given' (Viton, 1981, pp. 288).

The output of a transit firm cannot be described by using aggregate measures such as 'vehicle-miles' or 'seat-kilometers', as defined in these studies reviewed. Units of bus service provided by a firm are not homogenous and differ with respect to the route length, frequency, travel speed as also hours of operation. Failure to incorporate these heterogenous output characteristics would mean a faulty specification. 'The output of a transit firm, whatever the mode, is multidimensional by its very nature.... Since the mix of output and the way in which it is produced affect the firm's costs, it is clearly inappropriate to estimate cost functions by using a single measure of output' (Friedlaender et al, 1981, pp. 16). Crude attempts to account for output heterogeneity have however been made in these studies by the inclusion of technical or fleet characteristic variables, city-descriptor variables, as well as variables which account for differences in terrain and traffic conditions.

Exclusion of factor prices from the cost function analysis as in Johnston (1956) or inclusion of only one factor price such as wage rate as in Miller (1970) and Gray (1972) can be a serious misspecification (Jara Diaz, 1982, pp. 261). Also, estimation of the short-run cost function from time-series data and the long-run cost function from cross-section data as in

---

38 Although this problem is recognised, (See Miller, 1970, pp. 25; Koshal, 1970, pp. 30; Lee and Steedman, 1970, pp. 18) the 'multiproduct' nature of the transport firm has not been incorporated explicitly.
Economies of scale were being witnessed by the larger operators only. The explanation given was that a larger operator could ensure better co-ordination of service schedules and routes vis-a-vis smaller operators, and thereby reduce costs. But the most significant result was the positive elasticity of substitution between capital and maintenance. The policy emphasis of this result was clear. If financial constraints prevented bus operators from expanding fleet strength it would be more economical to expend resources on maintenance of existing fleet.

During the seventies and early eighties, the U.S. and U.K. urban bus transit industry was pumped with subsidies in order to prevent decay and decline as a consequence of significant fall in patronage. Pucher (1982), Button (1985), Button et al. (1985) widened the cost function specification by including 'policy variables' in the analysis. In all the three studies the main thrust was on probing whether subsidies to urban bus service firms had 'simply inflated costs instead of providing more better or cheaper services to transit users' (Pucher, 1982, pp. 51). Pucher's study centred around a cross-section of thirty-four U.S. bus undertakings. Although a simple linear cost function was used, the results indicated that subsidisation regardless of source, may have induced some cost escalation. To test whether subsidies led to reduced efficiency of bus operations in U.K., Button (1985) estimated two cost functions for a cross-section of fifty-five bus operators. The linear version representing a Leontief fixed-factor proportion technology was justified since there was not much scope for substituting labor by other factors, in the bus industry. Button refers to this as the 'technologically unprogressive' nature of the bus industry (pp.9). The Cobb-Douglas cost function was also estimated to make the analysis more general by accommodating some substitution possibilities between inputs. Both cost models revealed the positive effect of subsidies on costs of bus service provision.

45 Source refers to Federal, State or Local government in the U.S.
46 The other policy variables like 'public ownership' and 'public management' (introduced as dummy variables) had the expected positive sign but these were not statistically significant.
47 This may be possible due to better routing, scheduling and change in operating practices (such as move to one-man-operations) in response to input cost changes.
48 It may be noted that there was a cause and effect problem that could not be resolved. It was not very clear whether a high level of subsidy meant laxity in management that led to higher costs as hypothesised in the study, or whether unavoidable high cost operations called for higher levels of subsidy as revenues could not be raised through fares alone. The above conclusion could therefore be misleading. Pucher's analysis (1982, pp. 55) also cautioned about this problem.
The short run cost function\(^{44}\) was:

\[
STC = 0.335 Q^{0.47} P_L^{0.22} S^{-1.80} e^{0.0002D} + 0.09 S
\]

where, \(STC\) = short run total cost function; \(Q\) = bus-miles of output; \(P_L\) = wage rate; \(S\) = fleet size (as a measure of capital); \(D\) = population density. The long run total cost function was obtained by minimising costs with respect to the bus fleet. The estimated relationship was:

\[
LTC = 0.410 Q^{0.17} P_L^{0.08} e^{0.00007D}
\]

where, \(LTC\) = long run total cost. The long run marginal cost was also derived and was found to decline with output. This meant that there were significant long run economies of scale. Costs could be reduced considerably by having a single operator supplying bus services to the entire market (pp. 217). Due to data constraints a translog cost function could not be estimated and a Cobb-Douglas specification was resorted to. For this reason, the study is restrictive in nature, pertaining to the analysis of scale economies alone.

In order to scrutinise 'why urban bus systems continue to experience greatest absolute and relative operating deficits,'... Williams et al. (1981a) estimated a translog cost function for a cross-section of twenty small and medium sized bus operators in Illinois. The sample included bus operators of approximately similar size so as to ensure that the technology used would be more or less homogeneous across firms. 'Bus-miles' was used as the measure of output while the input variables were labor, capital, fuel and maintenance. The total cost function was estimated along with the share equations, using the non-linear Zellner iterative estimation technique. The hypothesis of a homothetic production structure could not be rejected but the Cobb-Douglas technology was strongly rejected. Separability of capital and maintenance from fuel and labour inputs meant that the use of capital and maintenance resources would change in response to their own prices irrespective of labour and fuel prices.

\(^{44}\) The price of fuel variable was dropped as it had an incorrect sign and an insignificant effect. The population density variable was included to reflect the level of congestion. High congestion levels meant slow speeds and therefore high operating costs but this also meant operation at full capacity and therefore declining average costs. The aggregate effect on costs would depend on which effect is more dominating. The unit cost of bus capital \((=0.09)\) was arrived at by assuming an interest rate of 6% and a depreciation rate of 3%.
Johnston (1956), is ambiguous and linked to the problem of 'theoretical weakness' associated with these cost models.

Inspite of these serious lacunae, dearth of a theoretical underpinning, restrictive structure, and an inadequate measure of transit output, these aggregate cost models 'can be more practical than sophisticated cost models when one needs a crude but reasonably reliable estimate of costs for a particular type of movement (Winston, 1985, pp. 65). Further, the poor quality of data base may have served as a constraint, compelling research efforts to be restricted to the estimation of simple cost functions for the study of bus transit costs as in the above cases (Williams, 1979, pp. 210). Complex statistical cost models have evolved as alternative specifications to mitigate some of these shortcomings.

4. Complex Statistical Cost Models

Application of complex statistical cost models to the urban bus transit sector has contributed significantly to a deeper understanding of economic characteristics related to this sector. These models are an outcome of developments in duality theory which unravels the relationship between the neoclassical production and cost functions. The neoclassical cost function, (alternatively referred to as the complex statistical cost model), is the minimum cost of producing a given output level during a given time period, expressed as a function of input prices and output.

Use of flexible functional forms such as the translog form for econometric estimation of the neoclassical cost function is another major advancement for the analysis of bus transit costs. Unlike simple statistical cost models, the translog cost function permits a multiproduct analysis.

---

39 In this context it may be worthwhile to note that a fairly recent study focussing on inter-modal and inter-city comparision of costs and productivity related to major urban transit systems in Europe, (Wunsch, P., 1996) '...departs from current trends and goes back to very simple analysis.' Cost functions suggested by accounting cost studies is resorted to instead of going for flexibility. Therefore, despite data limitations, the effect of important environmental factors affecting costs such as vehicle speed, vehicle capacity and network density can be studied and intermodal comparisions facilitated (pp. 176).

40 Pioneering work on the theory and application of duality in production may be attributed to Daniel Mcfadden while he was at the University of California-Berkeley (see Mcfadden, 1978) and a Berkeley engineering professor named Ronald Shephard (see Shephard, 1976). Objections have been raised regarding the 'cost minimisation' assumption especially in the context of transport firms which tend to be partially or fully regulated.
and does not place apriori restrictions on substitution possibilities, elasticities of factor demand or scale economies, providing greater scope for examining the underlying production structure. Further, parametric restrictions\(^{41}\) can always be imposed to confirm the superiority of the unrestricted translog cost function over other more restricted forms.\(^{42}\)

To summarise, '...in estimating cost functions for the transportation industries one should specify a multi-output cost function in a sufficiently flexible form to test hypothesis concerning the underlying structure of production. Moreover, if there is reason to believe that regulatory or other institutional constraints prevent optimal capacity adjustments, one should estimate a short run variable cost function, which can be used to derive the associated long run total cost function and the underlying production function' (Spady et al., 1976, pp. 3). In the following review of complex bus cost studies an attempt is made to highlight how far these conditions - a multiproduct analysis based on flexible functional form, taking account of excess capacity if any, have been satisfied.

4.1 Complex Statistical Bus Cost Studies.

Amidst a scenario of growing urban congestion, energy crisis and the uncertainty of future energy supplies, necessitating immediate expansion of the public transport system, Williams (1979) attempted to analyse the cost structure of urban bus transport in the U.S. A study of this nature seemed imperative for upgrading the urban bus transit systems which were experiencing high operating deficits due to decline in ridership triggered off mainly by the automobile revolution on city roads.\(^{43}\) This study is based on the neoclassical theory of production and costs. It was observed that there was some time gap between order and delivery of new buses preventing optimal adjustments of fleet size in the short run. Thus, firms were not operating on their long run cost curves. In order to reflect this phenomena, the short run cost function was first estimated and the long run cost function was derived from this short run relationship.

\(^{41}\) See Spady et al. (1976, section VI, pp. 61).
\(^{42}\) These restrictive forms refer to homothetic, homogenous, constant returns to scale or Cobb-Douglas versions.
\(^{43}\) See Meyer et al. (1981, pp. 37-55) for a detail account of problems faced by urban mass transportation in the U.S. during the decade of the seventies and the early eighties.
Using a translog cost function for analysing bus operations in U.K., Button et al. (1985) derived results that were contrary to Pucher (1982) and Button's (1985) findings. The results of this study indicated that subsidy levels had no significant effect on costs. Use of a flexible functional form permitted study of other economic effects as well. Using 'passenger revenue' as a measure of output diseconomies of scale was diagnosed for the larger bus companies. This called for a policy that recommended breaking up of some of the larger bus concerns while merging the smaller sized firms. The Cobb-Douglas structure of production was rejected as well as the hypotheses regarding homothetic technology and input separability. Own price elasticities were small indicating meagre sensitivity to price change. Complementarity between capital and maintenance was explained by stressing that older buses needed more maintenance.

Berechman (1983) and Berechman et al. (1984) rely on the usage of time series data for estimating bus-transit cost functions. Both these studies are based on the premise that cross-section data can pose problems for analysis. Very often a cross-section sample may include firms which differ in the form of ownership, fare structure, type of regulations imposed, distribution of demand over time and space, or even technologies used to produce transit services. Lack of homogeneity in production structure as well as output units, across firms in the sample may thus be evident, being an important cause behind erroneous results. In order to avoid this problem it may be preferable to utilise time series data at the firm level.

Berechman's (1983) study related to bus operations in Israel. Services were not distinguished into inter-urban and intra-urban, since the two types of services had similar trip lengths as well as temporal demand characteristics. The output variable, 'gross revenue' at fixed prices was used as a proxy for passenger trips. Due to lack of adequate data only two factors of production, labour and capital were considered. The results indicated a homothetic production structure. However, the Cobb-Douglas technology was rejected and costs were depicted through the homothetic version of the translog cost function. The elasticity of substitution

---

49 Due to lack of requisite data this measure of output was chosen as a proxy for 'passenger-miles'. Because fares were distance related. The input variables in the analysis were labor, capital and maintenance.

50 Braetigam et al (1982, pp. 274) emphasise this problem while studying railroad costs..... 'Failure to capture firm specific effects can lead to biases in estimated coefficients important to policy prescriptions. There is therefore an important role for econometric analysis of costs at the level of a single firm using time series data...... avoiding the effects of mixing technology.'
between labor and capital was less than zero indicating factor complementarity. This result is theoretically unsound in the context of a two factor model. Exclusion of other crucial important factors from the analysis due to data bottlenecks, may have produced this result. Economies of scale in bus operations was observed. Since the cost analysis was restricted to the Israeli bus sector alone, it was stressed that generalisation of these results to other bus companies would be misleading.

Against a backdrop of rising costs and fiscal austerity Berechman et al (1984) utilised time series data related to a single major Californian bus transit system to examine the nature of costs. It was asserted that results on scale economies would be affected by the output measure used. This was because, while the 'technical' measure of output, 'vehicle-miles', signified capacity of the bus company, the 'demand related' measure, 'passenger-miles' reflected intensity of capacity utilisation. Thus, scale economies based on the 'vehicle-miles' measure would denote variations in costs due to changes in capacity, but use of the 'passenger-miles' measure for determining economies would provide insight on the behaviour of costs associated with changes in the intensity of utilisation of this capacity. The input variables used in the model were labour, capital, fuel and maintenance. Results on substitution possibilities between factors, indicated complementarity between labour and capital. 'This complementarity seems reasonable in the context of the current one bus, one driver technology which characterises most bus service' (pp. 281). The liberal capital subsidy grant policy for purchase of new buses was used to justify substitutability between capital and maintenance. This policy encouraged operators to expand their fleet and cut down on their maintenance requirements. As regards structure of production, the Cobb-Douglas technology

---

51 It was pointed out, '... Inclusion of fuel and repair and maintenance as specific factors may provide more information on the production process especially factor substitution and demand' (pp. 22).
52 Federal operating subsidies were curtailed since the early eighties.
53 The Almeda Contra Costa Transit District (ACCTD).
54 See Berechman et al. (1985, pp. 318-320) for a discussion on the pros and cons of using technical versus demand-related measures of output.
55 The terminology used to highlight this distinction, is referred as 'economies of scale' (cost variations associated with capacity changes) versus 'economies of density (cost variations associated with change in intensity of capacity utilisation). These terms have not been explicitly used in this study.
56 It may be noted, the result is similar to Berechman's (1983) results, but is theoretically more acceptable as the input vector has been expanded to include fuel and maintenance in addition to labor and capital.
was rejected\textsuperscript{57} in favour of a non-homothetic\textsuperscript{58} production structure with linear separability between capital and labour as well as maintenance and fuel. Scale economies varied with output measure used, as hypothesised. Diseconomies of scale resulted when the vehicle-miles measure was utilised. On the other hand, the passenger-miles measure revealed significant scale economies. The policy recommendation which followed from these results was to expand central city services (which generates few miles but many passengers) at the cost of suburban services (which generates many miles and few passengers). Further, unit costs could be reduced if 'vehicle-miles' of output were to be supplied by more than one firm.\textsuperscript{59}

\textit{Viton (1981)} estimated a short run translog cost function, using a cross-section sample of 54 U.S. urban transit firms to verify whether bus transit was the least costly mode while comparing 'full costs'\textsuperscript{60} across modes in an urban setting. The variable factors included in the analysis were labour and fuel while the fixed factor capital was the fleet of buses held. The long run cost function was derived from the short run cost function by minimising costs with respect to the fixed factor.\textsuperscript{61} This study provided considerable insight on the distinction between 'economies of scale' and 'economies of density.' 'Economies of density' was defined as the short run economies arising out of increased production holding the fixed factor constant. These economies were derived from the short run cost function, confirming the presence of 'economies of density.' 'Economies of scale' was referred to as cost advantages observed when all factors of production varied. Results on scale economies obtained from the

\textsuperscript{57} The Cobb-Douglas technology was rejected in the case of most studies, (Williams et al., 1981, Button et al. 1985, Berechman, 1983). This casts serious doubts on the use of a Cobb-Douglas cost function for depicting the nature of costs related to urban bus transit systems.

\textsuperscript{58} This meant, scale economies would vary not only with the output variable but factor prices as well. See Chambers (1988, pp. 39 and pp. 74).

\textsuperscript{59} In keeping with this result the implemented policy of contracting out of transit services to private operators, in the U.S., was a step in the right direction. See Talley (1988) for a descriptive account of the nature of contracting out of transit services as is practised in the U.S. See also Tally et al. (1986) for a theoretical justification for service contracting.

\textsuperscript{60} Full costs = Out of pocket costs + time costs + other external costs. Mohring (1972), Boyd et al. (1978) and Jansson. (1980) include 'user-time' costs in their cost function analysis of bus services. User time constitutes time spent accessing the bus system, waiting for vehicles, possibly transferring between vehicles and getting to final destinations.

\textsuperscript{61} 'Rather than assuming that all inputs adjust instantaneously to their full equilibrium levels, researchers have increasingly adopted a framework that distinguishes variable from quasi-fixed inputs, where the latter adjust only partially to their full equilibrium levels within one time period' (Berndt, 1990, pp. 483) 'Capital' has generally been specified to be the fixed input in most studies. The methodology involves adapting the long run cost function for short run analysis by using the stocks of fixed inputs as arguments of the cost function instead of their prices. The long run cost function is then derived from the short run cost function by optimising with respect to the fixed factor. (See Spady et al., 1976, pp. 8-9).
long run cost function showed the larger firms as operating on the rising portion of the long run average cost curve, therefore witnessing diseconomies of scale and the smaller firms experiencing economies of scale in the long run. The optimal fleet size obtained by differentiating short run total costs with respect to the fleet size variable, when compared with the actual fleet size, testified the existence of considerable excess capacity in the industry.\footnote{See Borts(1956) for a discussion on capacity utilisation as decipherable from the neo-classical average cost curve.} A comparative cost analysis between bus transport and rapid rail transit, proved the superiority of bus transit over the rail mode as marginal costs of bus service was much lower than marginal costs of providing rapid rail transit services.\footnote{Marginal costs of rapid rail transit was obtained from the study of Pozdena, R.J. and Merewitz, L., (1978) 'Estimating Cost Functions for Rail Rapid Transit Properties', (Transportation Research, 12A, pp. 73).}

Using a similar approach\footnote{Unlike Viton's (1981) approach, Keeler used a Cobb-Douglas framework for depicting technology and estimated costs separately for 'freight' and 'passenger' output.} for studying the structure of costs in the U.S. railroad industry, Keeler (1973) found considerable excess capacity measured in terms of excess trackage. It was concluded that 'track abandonments could save substantial amounts of money' (pp. 207). The amount of excess trackage was 20,000 miles and the savings from abandonments was evaluated at 2.5 billion dollars. The corollary of this result was that economies of scale were non-existent while economies of density were very much prevalent (pp. 209).\footnote{Meyer J., observed that 'economies of density' may be a function of historical accident or inheritance. Discovery of at least some of the economies of density in the U.S. railroad industry could be explained by U.S. systems being originally engineered for more optimistic levels of traffic expectations in mind. (Winston, 1985, FN.19).} Harris (1977), in a study of rail freight traffic in the U.S. found significant economies of traffic density in the rail freight industries (pp. 561). The policy implication arising out of this result was to adopt differential pricing, low rates being charged for high density lines and high rates considered more appropriate for low density lines.

Few urban bus transit cost studies have explicitly incorporated the multidimensional nature of transport output. As a prelude to a review of these studies it may be worthwhile to examine developments in multiproduct cost function analysis as applicable to other modes such as railroads, trucking and inter-city passenger transport. Considerable research efforts have been pursued with regard to these related sectors, for strengthening the conceptual base as well as
the methodological framework of a multi-output analysis. These studies are briefly scanned in the next section highlighting the essential aspects.

4.2 Multiproduct Cost Function Analysis: The Case of Railroads, Trucking and Inter-City Transport.

'Railroad studies have stressed commodity differentiation in terms of freight and passenger services' (Jara Diaz, 1982, pp. 262). While, some studies have expanded the output vector to incorporate variables that depict 'output characteristics'. Brown et al. (1979) developed a flexible multiproduct cost function to depict the cost structure of the railroad industry. Railroads were viewed as producers of two types of outputs namely, 'ton-miles' of freight service and 'passenger-miles' of passenger service. Analysis of scale economies for a cross-section of railroad firms revealed that excepting for one railroad, scale economies were significantly positive.66 The multiproduct nature of analysis permitted the study of 'economies of scope' as well. 'Economies of scope' were determined by observing the curvature of the iso-cost contours in output space.68 This was found to be convex, indicating cost economies from specialisation in freight as against passenger services.

The translog cost function for a multiproduct firm is undefined, when one or more of the outputs is zero. In order to surmount this problem, Caves et al (1980) estimated a 'generalised translog multiproduct cost function' by a 'Box-Cox' metric transformation69 of the output variable. Two output measures, 'revenue ton-miles' of freight and 'revenue passenger-miles' as

66 Scale economies for a multiproduct firm is given by: \(1-\sum \frac{\partial \ln C}{\partial \ln Y_i} \), which is unity minus sum of individual output (\(Y_i\)) cost elasticities. If \(1-\sum \frac{\partial \ln C}{\partial \ln Y_i}>0\) it implies 'economies of scale' exist. It may be noted that the expression used in this study to compute multiproduct scale economies, implicitly assumes that the marginal cost of each type of output is identical. The more general expression for multiproduct scale economies is given by: \(1-\sum \frac{Y_i \cdot M_i}{C} \) (Bailey, et al. 1982, pp. 1031).

67 'Economies of scope measures the cost advantages to firms of providing a large no of diversified products as against specialising in the production of a single product' (Bailey et al 1982, pp.1025). Thus the costs of a multiproduct firm would be influenced by the scale as well as the composition of output. Jara Diaz (1982) emphasises that any form of 'output' aggregation partially destroys the possibility of analysing scope though it may not destroy scale analysis (pp. 269).

68 An iso-cost contour in output space is given by: \(dC=(\frac{\partial C}{\partial Y_1} \cdot dY_1 + (\frac{\partial C}{\partial Y_2} \cdot dY_2)=0\), where. \(Y_1\) and \(Y_2\) are two outputs of a firm. Slope of this contour is given by: \((\frac{\partial C}{\partial Y_1})/(\frac{\partial C}{\partial Y_2})=\frac{dY_1}{dY_2}\). Thus, \(\frac{dY_1}{dY_2}\) can be derived. A negative value of \(\frac{dY_1}{dY_2}\) indicates concavity and economies of scope, whereas, a positive value indicates economies of specialisation.

69 If '\(Y_i\)' represents output 'i' then the 'Box-Cox' metric transformation may be given as: \(f_i(Y_i)=((Y_i^{\lambda}-1)/\lambda)\), \(\lambda\neq 0\); \(f_i(Y_i)=\ln Y_i\), \(\lambda=0\) and provided that \(\lambda\) is strictly positive, the Box-Cox metric is well

21
were used for the analysis of a cross-section of fifty-six railroad firms in the U.S. The input variables were labour, fuel and capital. To test the significance of the 'generalised translog cost function', two cost models were developed. First, a cost function was estimated for the whole sample which included firms producing both passenger and freight as well as those producing only freight services. Next, a cost function was estimated for a subset of this sample which constituted firms producing positive levels of freight and passenger services. The subset sample referred to as the 'truncated sample' produced incorrect signs for some of the coefficients as also the multiproduct scale economies which was found to be negative. Use of the entire sample however, produced coefficient estimates which were more robust and scale economies which was more economically meaningful. These results strengthened the significance of the 'generalised' version of the translog cost function and it was concluded that in order to obtain global information on the production and cost structure of a multiproduct firm, it is necessary to include firms which produce only a few of the feasible range of outputs, in the analysis (pp. 478).

The importance of expanding the output vector to include quality of service variables was illustrated by Braetigam et al. (1982) in their study of rail freight costs in the U.S. More specifically, it was stressed that, '....the speed of service is an important determinant of rail costs.....' (pp. 394). Engineering models were used to predict overall average speed before introducing this variable into the cost function specification. Engineering process functions were developed to relate speed of service to known technological parameters of the railroad system. This multidisciplinary approach to improve cost function specification was referred to as a hybrid approach (pp. 394). A translog model was used to estimate this hybrid cost function, using time series data for a single firm:

defined for zero output levels: f(0)= -(1/#). The natural log metric is a limiting case of the Box-Cox metric: lim

Note our earlier discussions on engineering cost studies. 

The overall speed of service is affected by: (i) Line haul movement time (ii) Time spent by cars in classification yards. Process functions were developed separately developed separately for each of these variables using engineering parameters. The overall average speed was given by: L / (L/V_a + T_p) where,
\[ C = C(y, s, p_1, p_2, p_3; k) \]

where, \( C = \) costs; \( y = \) ton-miles of service; \( s = \) overall average speed; \( p_1 = \) price of fuel; \( p_2 = \) price of labor; \( p_3 = \) price of equipment; \( k = \) physical level of fixed factor. The results showed that 'the addition of engineering information improves results in a statistically significant fashion' (pp. 402).

Friedlaender et al. (1992), made attempts to assess the impact of the 1980 'Staggers Act' on the use of capital in the U.S. railroad industry. This Act was framed to permit U.S. freight railroad operators better potentials for adjusting their rates and capital structure. A study of this nature seemed imperative since considerable amount of evidence was found to support the claim that the U.S. railroad industry was characterised by 'substantial capital disequilibrium' prior to the Staggers Act (pp. 131). A rail cost function was estimated, using panel data relating to major 'Class-I' freight railroads in the U.S. for the period 1974-86, to decipher the extent of capital restructuring in response to the act.

The short run variable cost function estimated taking cognisance of the fixed nature of 'ways and structures', included a range of independent variables. Data on 'ton-miles' by commodity type was not available. Thus, 'technological variables such as coal and agricultural tons carried as a percentage of total tons carried were used as proxies to depict the heterogenous nature of rail traffic. This was because specialized equipment was used to handle coal as well as agricultural traffic and costs would vary accordingly. 'Track-miles' and 'average length of haul' were included as measures of the 'network' and its 'utilisation'. These network-based technological variables were specifically used to emphasize the importance of the nature of the rail network while depicting the structure of costs in the

\[ L = \text{average length of haul}; \quad V_s = \text{overall average line haul velocity}; \quad L/V_s = \text{average running time}; \quad T_y = \text{delay at classification yards.} \]

\(^{72}\) Since the railroad is a bridge line that connects with major railroads at each end of its line the average speed of service is determined by the requirements of the major railroads. Therefore 'the speed of service is properly modelled as exogenous to the railroad in question' (pp. 396).

\(^{73}\) In this connection, see our reference to Keeler's (1973) study.

\(^{74}\) Ways and structure capital represents road bed, track, bridges and so on, which are long-lived and treated as a fixed factor in this cost function analysis.
railroad industry. The other independent variables were, prices of fuel, labor, equipment materials and supplies and a time trend variable which was meant to capture any unexplained productivity growth over time. Care was taken to lay down the 'error structures' appropriately. Exogeniety of the output, and output associated 'technological' variables could not be validated. It was felt that these were endogenously determined through the profit-maximising behaviour of the railroads and therefore should be related to demand variables. But such variables do not enter the cost function and the approach adopted was to use appropriate firm-specific demand related variables as instruments in the estimation procedure. Consequently, a 3SLS procedure was used to estimate the system of equations consisting of the cost function and the share equations incorporating these instrumental variables.

Both short-run and long-run returns to 'density' and 'size' were subsequently derived. The results indicate that '...given the large amounts of fixed track and 'ws' capital, there are substantial returns to density. Moreover, if capital is adjusted in an optimal fashion, returns to scale are somewhat larger, indicating that increasing returns is not a transitory phenomenon, but may be an inherent characteristic of the railroads' technology'(pp. 142). A similar result emerges with respect to returns to size both in the short-run as well as in the long run which only strengthens this finding regarding rail technology. An unexpected trend emerged as regards capital utilisation in the U.S. railroad industry. There appeared to be insignificant movement towards optimal adjustment of capital despite the implementation of the 'Staggers Act'. Excess capacity was found to be pervasive in the industry. The magnitude of excess capacity was estimated to range from a low of $8.949 billion in 1974, $ 9.384 billion in 1974 to a high of $16.908 billion in 1984 and $12.124 billion in 1986. One explanation provided for this disconcerting result was that a certain level of ways and structure capital was a

---

75 The error term was decomposed into three components.: (i) firm specific error (ii) an error which exhibits first-order autocorrelation within a given equation and (iii) a normally distributed term that may be contemporaneously correlated across equations only. (See pages 136 and 137 for further details).

76 Returns to scale associated with a given increase in tonnage alone is referred to as 'economies of density'. While returns to size captures the behaviour of costs in response to simultaneous changes in the output of the firm as well as its network.

77 This result was derived by analysing the relationship between the opportunity cost of capital and the firm's shadow value of capital which was obtained as follows. The total cost function may be given as \( C = C^T(y, w, t, zF) + P_F zF \) where, \( C^T \) reflects total cost, \( C^V \) the variable cost function and \( P_F \) represents the opportunity cost of fixed way and structures capital. The equilibrium capital stock is obtained when the opportunity cost of capital equals the firm's shadow value of capital or when \( [\delta C^V/\delta zF] = -P_F \) (pp. 142). In the case of overcapitalisation, the opportunity cost of capital exceeds the firm's shadow value of capital while the converse is true in the case of undercapitalisation.
prerequisite for better service quality such as higher speed and therefore competitiveness of the railroads vis-a-vis the trucking industry. However, this could not fully explain reasons behind the magnitude of capital disequilibrium which was quite high. The policy outcome of this result signaled the need for further incentives towards rationalisation of capital structure in the railroad industry.

'Trucking industries provide a particularly good example of the importance of multiproduct analysis in empirical work' (Bailey et al, 1982, pp. 1033). To represent trucking technology more accurately, the thrust of research work has been to expand the output vector to include operating characteristic such as 'length of haul', 'load size', 'shipment composition' in addition to the physical measure of output 'ton-miles'. Most of these studies infer that costs can be quite sensitive to these service characteristics. These cost functions christened as 'hedonic cost functions' have become common in literature following pioneering work by Spady and Friedlaender (1978). 'If point-to-point transportation movements are viewed as the true cost-causative outputs of the firm, a firm operating even a relatively small network must be viewed as producing an astronomical number of products. The hedonic approach has, in large part, arisen as an attempt to deal with the problem of networks...Use of hedonic cost functions enables the investigator to, in effect, perform often unavoidable aggregation based upon informed judgements about characteristics that are likely to have important impacts upon the costs associated with producing a given aggregate output vector' (Panzar, 1989, pp. 43 & 44).

Koenkar (1977) used a pooled cross-section of trucking firms in the U.S. for estimating a Cobb-Douglas multiproduct cost function. Two shipment characteristics, mean length of haul and weight of loads were incorporated into the analysis. Average costs were significantly affected by these 'service characteristic' variables, declining considerably as 'length of haul' and 'weight of load' were increased.

---

78 Other operating characteristics were controlled by selecting a fairly homogenous sample of firms for analysis.
In order to capture the true relationship between output and costs in the trucking industry, Spady et al. (1978) developed a 'hedonic cost function' which explicitly takes output characteristics into account. This 'quality separable' cost function developed by them may be symbolically expressed as:

\[ C = C \left[ \psi(y, q), w \right] \]

where, \( C = \) costs; \( (y, q) \) = vector of functions that measures effective output; \( y = \) generic measure of output (ton-miles); \( q = \) vector of qualitative attributes; \( q_1 = \) average shipment size, \( q_2 = \) average length of haul, \( q_3 = \% \) of ton-miles shipped in less-than-truck load (LTL) lots, \( q_4 = \) insurance costs per ton-miles, \( q_5 = \) average load; \( w = \) vector of input prices. Two cost functions were separately estimated, a hedonic and a non-hedonic version. Comparison of the two forms using statistical tests confirmed superiority of the hedonic over the non-hedonic cost function. Analysis carried out using conventional measures of output such as 'ton-miles' could therefore 'lead to highly erroneous conclusions about the structure of technology' (pp. 170). Results obtained on 'economies of scale' were illusory. It was emphasised that there was nothing inherent in the structure of technology to indicate economies of scale since the industry was characterised by low capital requirements. These issues were reconciled by noting that the regulatory environment may have conferred scale economies as it permitted larger firms to obtain diverse operating rights vis-a-vis their smaller competitors.

---

79 The work of Spady and Friedlaender is regarded as the 'state-of-art' in estimation of cost function and analysis of scale economies in transportation industries. (JaraDiaz, 1982, pp. 63)

80 The quality separable nature of the cost function implies that the effect of variation in output characteristics \( (q_i) \) on effective output \( (\psi_i) \) and therefore on costs is independent of relative factor prices \( (w_i) \).

81 \( \psi_i(y, q_i) = y \phi(q_i) \). Doubling of \( y \) would mean doubling of effective output \( \psi_i \). Therefore, \( \ln(\psi_i(y, q_i)) = \ln y + \ln \phi(q_i) \).

82 LTL shipments as against TL(truck-load) shipments, need to be consolidated. An LTL shipment is often picked up in a straight truck with other shipments, sorted on to line-haul trailers and transported to its destination city where this terminal process is reversed. TL shipments are shipments of sufficient size to be individually transported. Thus LTL practices involve higher terminal costs of handling. This may be meagre in the case of TL shipments which are picked up and delivered in the same vehicle.

83 As data on break-up of commodities carried, was not available, insurance costs per ton-mile was used since it reflected differences in fragility of shipments and costs of special handling. Therefore, this measure could serve to capture differences in the composition of output.

84 As we have noted, for the hedonic version \( \psi_i(y, q_i) = y \phi(q_i) \), whereas for the non-hedonic version \( \psi_i(y, q_i) = y \) and \( \phi(q_i) = 1 \).
counterparts, encouraging them to increase length of haul, shipment size or cut down share of LTL traffic which consequently led to dramatic cost reductions for these fortunate larger firms. If the smaller firms were also allowed to diversify operating characteristics, just like the larger firms, then these smaller operators could have also secured cost advantages by varying shipment characteristics accordingly. Thus it was concluded that economies of scale in the trucking industry were of a regulatory nature. (Bailey et al, 1982, pp. 1034). Further, economies if any, were not economies of scale, but economies of density, arising out of better utilisation of existing capacity by adjusting operating characteristics.

Friedlaender et al. (1981), refined the framework of the foregoing cost function analysis (Spady et al., 1978) in three respects. Firstly, the quality separable nature of the specification was considered to be excessively restrictive since it implied that service characteristics have no direct effect on factor intensities. As an alternative to the quality separable hedonic specification, it was found to be more suitable to view service characteristics as technological conditions (pp. 42). The modified cost function was:

\[ C = C(y, w, q) \]

where, \( y \)= ton-miles of output; \( w \)= vector of input prices; and \( q \)= vector of service characteristics. Secondly, analysis was carried out separately for specialised commodity carriers and common carriers of general commodities. Regulations as also the institutional set-up differs across these carriers, necessitating separate treatment.\(^85\) Thirdly, it was hypothesised that there might be differences in the structure of trucking technology with respect to geographical regions. Therefore, the cost analysis was further disaggregated by regions, categorising trucking firms into the 'official region' and 'south-west region' in the case of specialised carriers, while, common carriers were classified into the 'official region', 'south-west region' and 'inter-regional' carriers. These methodological improvements over quality separable 'hedonic' cost analysis of Spady et al (1978)'... was uniformly superior in

\(^85\) Carriers of specialised commodities carry full truckloads, use owner operators, and have very rigid route structures and commodity operating rights. In contrast, carriers of general commodities tend to concentrate on LTL carriage, use union labor, and operate over a fairly wide geographic area. For specialised commodity carriers, the vector of shipment characteristics \( t \) included: average load per vehicle average length of haul and insurance costs per ton-mile. For general commodity carriers however, the \( t \) vector constituted: average length of haul, average load per vehicle average shipment size, insurance costs, terminal density and percentage of LTL shipment.
terms of usual statistical criteria' (pp. 42), thus representing trucking technology across carriers and regions more accurately. Results on economies of scale were very similar to Spady et al's (1978) findings. The structure of regulation allowed larger firms to be more profitable than smaller firms since these larger operators could obtain higher utilisation of equipment through diversified operating rights. Therefore, in a deregulated environment these economies would disappear because small as well as large sized firms would then be free to carry any commodity to any place along any route in order to reduce costs. A superior methodological framework of analysis, undoubtedly, enhanced acceptability of model results. Both these variants of the hedonic cost functions are '...quite useful, in the absence of multiproduct data, since they do capture some of the dimensions of heterogeneous output and thus provide partial adjustment for the composition of output. Nevertheless, they are not fully adequate...... (since)..... measure of scale economies based on 'hedonic outputs' are necessarily ambiguous...' not being able to distinguish between the effects of scale and scope on cost. (Bailey et al, 1982, pp. 1034).

Harmatuck (1981) used a cross-section of Class-I motor carriers of general commodities to estimate a multiproduct cost function for a better understanding of trucking technology. This was important for increasing the economic content of policy decisions which was hitherto being guided more by political considerations. It was stressed that, '...development of the cost function specification in terms of distinct truckload and less-than-truckload outputs avoids biases found in single output cost specifications as well as in those multiple output specifications which treat multiple outputs as qualitative variations of a single output index rather than as separate and distinct' (pp. 148) Accordingly, four types of outputs were incorporated into the analysis: \( Y_1 = \) number of shipments (TL); \( Y_2 = \) average weight of shipments (TL); \( Y_3 = \) number of shipments (LTL); \( Y_4 = \) average weight of shipments (LTL); \( Y_5 = \) average length of haul. A generalised translog multiproduct cost function was estimated by a 'Box-Cox' metric transformation of the output variables. 86 Another distinguishing feature of this study was the use of 'activity prices' in the cost function specification as opposed to natural prices. 87 Factor inputs were aggregated into various activities such as '!!nec-

---

86 See also Caves et al., (1980)
87 'On practical grounds activity prices are preferred to natural prices because of the difficulties of properly defining input prices.' (pp.143)
haul', 'pick up and delivery', 'platform handling', and 'billing and collecting' to arrive at the unit price per activity.\textsuperscript{88} All the activity price variables turned out to be quite 'significant'. Results on scale economies indicated diseconomies for larger firms but significant economies for smaller firms in the sample. Economies of scope was also determined by observing the curvature of the iso-cost curve. This indicated considerable economies of scope for larger firms while economies of specialisation for the smaller operators.\textsuperscript{89} The policy implication of these results was quite straightforward. Smaller firms should be permitted to grow in order to reap the cost advantages arising from 'economies of scale', while larger firms should be allowed to diversify their outputs in order to benefit from economies due to scope. These results questioned the actual policy of inhibiting entry into the trucking industry as a whole, and specifically entry into the LTL segment as against the TL segments of this industry.

The way in which trucking traffic is allocated between various links and nodes of the network\textsuperscript{90} may affect costs. The importance of network configuration and utilisation in the analysis of trucking costs has been acknowledged by Wang Chiang et al. (1984), by directly introducing network variables as arguments in the cost function. 'Earlier studies...... failed to reflect the corridor specific nature of trucking traffic' (pp. 267). In addition to shipment characteristics the cost function includes variables that depict the types of corridors utilised by the firms and the structure of the network over which these operate. This is given as:

\[ C = C(\psi, w, N) \]

where, \( \psi = (\psi_1, \psi_2, \ldots, \psi_n) \); \( \psi_i \)\textsuperscript{91} = \( \psi_i(Y_i, q_i) \) the hedonically adjusted output along a generic corridor 'i'; \( Y_i \) = physical output along generic corridor 'i'; \( q_i \) = corridor specific output characteristics; \( w \) = vector of factor prices; \( N \) = vector of network characteristics. Four network

\textsuperscript{88}Unit price per activity= (total activity expenditures / output measures of the activity)

\textsuperscript{89}For the smaller firms, \( d^2y_1/dy_2^2 > 0 \) and for the larger firms, \( d^2y_1/dy_2^2 < 0 \), where, \( y_1 \) = number of shipments (TL) and \( y_3 \) = number of shipments (LTL).

\textsuperscript{90}The whole exercise is to choose various options among different route structures to produce a given O-D pattern. The fact that network shape' has been emphasised only in some airline studies is probably due to the non-constraining nature of the problem in terms of a physical network.(Jara Diaz, 1982, pp. 262)

\textsuperscript{91}\( \psi \) = LTL traffic with length of haul under 250 miles; \( \psi_2 \) = LTL traffic with length of haul of 250-500 miles; \( \psi_3 \) = LTL traffic with length of haul over 500 miles; \( \psi_4 \) = TL traffic. Each ' \( \psi_i \)' is itself a hedonic aggregator function of ton-miles and shipment characteristics, within each type of corridor.
variables were explicitly incorporated. The network configuration variables were: 'terminal density' and 'connectivity' given by the 'gamma index' of network theory. The network utilisation variables were: 'traffic density' given by the 'chi index' of network theory and 'circuity' given by the 'indirect routing index' of network theory. The results emphasised the importance of including network variables, which were found to have a significant influence on costs (pp. 271). Further, economies of network configuration and utilisation, were found to be quite strong. The desire of trucking firms to merge and grow could be explained by the natural advantage which large carriers enjoyed over smaller ones in being able to exploit more fully, the economies of equipment utilisation and traffic flows over the network ensuring a high degree of connectivity implying an ability to perform more direct routing and more balanced service that leads to cost reduction (pp. 276).

A host of studies have been conducted during the nineties to assess the impact of regulatory reform on costs for various sections of the U.S. trucking industry. It has been a well established contention that Federal control of trucking produced inefficiently high rates as well as costs. The Motor Carrier Act of 1980 was aimed at deregulating the trucking industry towards removing some of these distortions. Ying (1990) specifically analyses the underlying effect of deregulation on operating characteristics, factor utilization and the nature of returns to scale of trucking firms. The aim was to observe these changes if any, with respect to common carriers of general freight specializing in less-than-truckload shipments since these appeared to have been subjected to more regulatory restrictions. Three time periods were identified, one representing the pre-deregulation period (the year 1975), the other the year of deregulation (1980) and finally the post-deregulation period (the year 1984). Cost functions

92 The network configuration variables depict the physical structure of the network, given by the operating rights granted and the distribution of terminals. Two network configuration variables in the study were: (i) Connectivity given by the gama index of network theory which measures the degree to which the network is fully connected was measured as:

\[
\text{[Actual number of connected links / Possible maximum number of connected links.]}\]

(ii) Terminal Density which was measured as terminals per ton-mile. The network utilisation variables measure how efficiently the carrier routes its traffic over the network. Two variables were used to depict network utilisation: (i) Traffic Density as given by the chi-index of network theory reflects the concentration of traffic over the network. A low value of this index represents high concentration and the index reaches a maximum value when traffic is evenly distributed. (ii) Circuity as given by the indirect routing index of network theory given as:

\[
\text{Actual number of tons / Number of ton-miles generated if all shipments had been routed directly.}\]

93 Better connectivity, higher terminal density, higher concentration on the network and direct routing opportunities meant lower costs.

94 This implicitly indicates the presence of significant 'economies of density'.

30
were separately estimated for each time period based on a cross-section of trucking firms as it was felt that estimating costs with a uniform specification for years before, at the time of and after deregulation would make the analysis more dynamic and allow better identification of changes in motor carrier technology over time (pp. 997). Given the relatively small capital requirements of the trucking industry, a long run translog cost function of the following hedonic specification was estimated:

\[ C = c(w, y, a) \]

where, \( C \) = long run total cost; \( w \) = vector of factor prices; \( y \) = revenue ton-miles; \( a \) = vector of operating characteristics which together included average length of haul, average shipment size, average load, percentage of less-than-truckload traffic and average cargo loss and damage insurance per dollar cost. These output characteristics were included as an attempt to capture the heterogenous nature of output in the trucking industry. Parameter estimates for the operating characteristics of the representative firm\(^{95}\) seemed to indicate that deregulation had probably enabled motor carriers to utilise their networks and terminal facilities more fully and therefore, to reduce empty backhauls substantially. This was justified by the fact that the cost-reducing impact of ‘average load’ seemed to have been enhanced following deregulation. Specifically, carrying commodities shorter distances over a better utilised network and the ability to get more backhauls may have explained why increasing ‘average length of haul’ and larger ‘average shipment sizes’ continued to decrease costs but not as much. Put together, the same reasons validated why transporting higher value goods and indulging in more ‘less-than-truckload’ operations was not as costly as before. Effects of deregulation on factor-utilisation was derived from the estimated elasticities of factor substitution. Broadly, the results indicated a more efficient factor utilisation in the post-deregulation era. Carriers substituted away from high-priced unionized labor and fuel, increasing the role of purchased transportation as well as capital in the production process after regulatory reform (pp. 1002).

Comparision of scale economies across the three periods indicated that trucking finns may have faced ‘regulatory diseconomies of scale’ rather than ‘regulatory scale economies’ during the pre-deregulation period. This was contrary to the findings of Friedlaender et al. (1981) However, following deregulation, returns to scale ‘...have dramatically changed to strongly

---

\(^{95}\) The representative firm is that firm whose variables have sample mean values.
increasing returns' (pp. 1003). This implied that larger carriers would expand their operations even further in an effort to cut costs.

Under similar contextual conditions, Callan et al. (1992) focused on the post-deregulation cost structure of the comparatively unexplored household goods motor carrier industry in the U.S. Two types of organisational structures are prevalent in this industry: the "van lines" and the "non-van lines". A peculiarity of the household goods carrier industry is that routes tend to be variable, non-repetitive, unpredictable and empty movements may be inevitable under these circumstances. The "van line" structure which represents an inter-firm co-operative arrangement evolved to circumvent these problems and especially to avoid cost-prohibitive empty back hauls. The "non-van lines" operate independently without the support of an agency system. It was hypothesised that the relaxation of government controlled price and entry constraints following deregulation should have enabled van lines to exploit agency-related cost advantages more fully and to price their smaller 'non-van line' counterparts out of the market. However, this phenomenon does not seem to have occurred (pp. 19). Therefore, in this study an attempt was made to unearth factors which would broadly explain how firms of varying sizes and spans of operations could continue to function effectively under the new deregulation era.

The methodology adopted was to empirically estimate cost conditions for detecting cost-differentials across the household motor-carrier industry. The cross-section data set, consisting of van lines and non-van lines, related to the year 1984 to allow for sufficient adjustment period following deregulation. A simple hedonic translog cost specification was adopted output heterogeneity being captured by the network 'link' variable: average length of haul; the network 'node' variable: number of nodes in the network; average load and shipment type. First and foremost, the coefficients of 'average load' and 'shipment type'  

96 More specifically the 'van line' is a complex agency system of many carriers contracted as agents to conduct business under the operating rights of the 'parent' van line which coordinates the activities of the co-operating firms (pp. 19).
97 It was implicitly assumed that the production technology is the same across carriers (FN 7. pp. 25).
98 The general form of the cost function was given as: \( C = C(Q, H, N, L, S, w) \) where, \( C = \) total operating costs, \( Q = \) ton-miles produced, \( H = \) average length of haul, \( N = \) number of nodes in the van line system, \( L = \) average load, \( S = \) shipment type, \( w = \) vector of input prices.
99 Since only the 'van line' is a multi-node system the network node variable has no relevance for the 'non-van lines'.
100 Shipment type was given by the proportion of conventional shipment of household goods as opposed to shipment of unusual commodities.
were found to be very significant '....supporting the contention that these attributes are important determinants of carrier costs' (pp. 26). This result is somewhat similar to that of Ying's (1990) study. Regulatory reform undoubtedly permitted motor carriers to restructure their operations so as to better 'fill up' their carriers and curtail uneconomic empty movements as a consequence. Increasing 'average load' has had a cost reducing effect in the household goods sector of the trucking industry as well. Cost elasticities evaluated with respect to operating characteristics at respective means for each of the carrier group revealed an interesting result. Non-van lines were better able to exploit potential unit cost savings associated with increase in 'average load', 'average length of haul' as also by carrying ordinary household goods rather than unusual commodities which are more expensive to transport (pp. 30). As for the van lines, the main cost advantage appeared to arise from their nodal system of operations. The van-line specific network node elasticity estimate clearly suggested that cost economies associated with number of nodes in their system were yet to be fully exploited. 101 As regards output and size-related economies, each carrier group appeared to operate with some degree of economies of density. 102 However, estimates pertaining to size-related economies revealed a different picture for the two groups of carriers. Van lines were able to monitor their size strategically by expanding both 'output' and 'network' to achieve lower per unit costs. As a result these were found to operate in the region of constant or slightly increasing returns to scale. Non-van lines, on the other hand, had yet to fully exploit available cost savings associated with size increases. On the whole, the results clearly validated how non-van lines could continue to remain competitive despite the changing environment of deregulation which should have significantly increased the role of their larger van line counterparts. Cost advantages could still accrue to these smaller firms by reaping economies of density and size, by enhancing shipment load and length of haul and by targeting their operations towards the transport of ordinary household goods as against carriage of unusual commodities.

---

101 The network node elasticity may be derived from the parameters of the cost function by evaluating the expression: \( \delta \ln C/\delta \ln N \), where \( C \) = costs and \( N \) = number of nodes in the van line agency system. This was estimated to be 0.3063 (pp. 30).

102 'To distinguish between cost effects of the two size dimensions, output and network the literature defines decreasing (increasing) unit costs associated with output increases (holding all else constant including network size) as economies (diseconomies) of density, and those due to simultaneous proportional increases in both output and all measures of network size as economies (diseconomies) of scale' (pp. 23). Therefore, with respect to the cost specification in this study, economies of density would be given by: \( \delta \ln C/\delta \ln Q \), where \( C \) = operating costs and \( Q \) = ton-miles of output produced. While economies of size or scale would be symbolically given as: \( (\delta \ln C/\delta \ln Q + \delta \ln C/\delta \ln H + \delta \ln C/\delta \ln N) \) where, \( H \) = the network link variable, 'average length of haul', \( N \) = network node variable, number of nodes in the system relevant only for van lines.
Hedonic specifications as applicable to the transport sector, are an advancement in cost function analysis. However, Xu et al. (1994) raise queries about the concept of 'returns to scale' under these more complex specifications. For instance, while trying to evaluate returns to scale, cost specifications which include output characteristics and network variables in addition to "primary" measures of output such as 'revenue ton-miles', '......make the definition of what is meant by an increase in output more difficult' (pp. 275). While delving deeper into these conceptual ambiguities, they found that output characteristics were positively correlated with firm size. Therefore, as hitherto assumed, it was erroneous to hold output characteristics fixed while drawing conclusions on the impact of firm size on costs. The main aim of this study was to evolve an appropriate measure of returns to scale consistent with the hedonic cost configuration.

Two systems of equations were separately estimated based on a panel data set of post-deregulation 'less-than-truckload' motor carriers in the U.S. In the first case, the conventional procedure was adopted. A translog cost function was estimated along with the share equations for arriving at a measure of 'returns to size', computed as:

$$\frac{\delta \ln C}{\delta \ln q}$$

where, C=costs and q=primary measure of output, 'ton-miles' produced. Modest diseconomies of scale were detected under the assumption that all other variables, including the two output characteristic variables, were unchanged as firm size increased. A correlation analysis confirmed the positive correlation between size and output characteristic. Therefore, in order to take these interrelationships into account, in the second case the equations were re-estimated by adding the following two equations into the system:

---

103 There may be specific reasons why large firms are able to realise a longer average length of haul and higher average loads. Larger firms may be advantageously placed by virtue of their more extensive geographical coverage, better on-time performance and more sophisticated information systems. These attributes provide the shipper with a more inclusive product which in turn attracts firms with complex, long-haul distribution patterns seeking to minimise the number of carriers they deal with on a regular basis. Better information systems, available to the largest firms because of their cost and complexity, might provide more opportunities for consolidation of freight and in turn raise vehicle load size (pp. 279).

104 A cost function of the following general specification was estimated: $C=C(w, q, y_1, y_2)$ where, $w=vector$ of input prices, $q=ton-miles of output, y_1=average shipment load, y_2=average length of haul.

105 At the sample mean, $\frac{\delta \ln C}{\delta \ln q} = 1.0800$ indicating mild diseconomies (pp. 278).
\[ \ln y = a + \sum b_i (\ln w_i) + c_i (\ln q) + d_i (\ln q)^2 \]

where, \( i = 1 \) for average shipment load; \( i = 2 \) for average length of haul; \( w_i \) = factor prices; \( q \) = ton-miles produced. When the conventional measure was used to derive economies of size based on this new system of equations, larger diseconomies of scale were detected.\(^{106}\) The innovativeness of this study was the computation of the more holistic ‘full cost elasticity’ estimate, manifesting the effects of output as well as output characteristics on costs. Symbolically, this was estimated using the following formulation:

\[ \frac{\delta \ln C}{\delta \ln q} = \frac{\delta \ln C}{\delta \ln y_1} \cdot \frac{\delta \ln y_1}{\delta \ln q} + \frac{\delta \ln C}{\delta \ln y_2} \cdot \frac{\delta \ln y_2}{\delta \ln q} \]

where \( C \) = costs; \( q \) = ton-miles; \( y_1 \) = average shipment load; \( y_2 \) = average length of haul. The full cost elasticity estimate was estimated to be 0.8049 indicating increasing returns to scale. These results indicate that ‘.....larger firms do not have a cost advantage due to size directly, but that size indirectly influences the operating characteristics in such a way that their unit costs are lowered’ (pp. 282). Thus, the empirical results strengthened the significance of the new methodological framework, introduced in this study, for depicting returns to scale under a hedonic cost function specification.

The multiproduct nature of an inter-city passenger transport firm may be conceptualised by disaggregating the output vector into a variety of services. Taucher et al. (1983) adopted this approach for studying costs of inter-city carriers in the U.S. The analysis was motivated by a need to decipher the nature of economies of scale and scope amidst a proposed change in the regulatory environment that was likely to provide greater price flexibility and lesser restrictions on types of services provided by firms. A translog-type functional form was used to estimate this cost function. The output vector for the class I carriers included bus-miles by service types: regular-route, charter and local. For Class II and III carriers the output vector constituted bus-miles by service types: regular route, charter, local and school. Input variables were labor, fuel and capital. The results strongly established cost advantages attributable to the joint production of all types of services.

\(^{106}\) At sample mean values, \( \frac{\delta \ln C}{\delta \ln q} = 1.1327 \) (pp. 280).
This concludes our review of multiproduct cost functions as applicable to railroad, trucking and inter-city transport industries. Considerable experimentation has been carried out, especially with respect to the trucking industry. The expanded framework of analysis provides better scope for treatment of the heterogenous nature of transport output. These developments have undoubtly helped in gaining better and new insight into the structure of technology in these industries. On the whole, the level of conceptual broadening as illustrated by these studies, has paved the way for multiproduct cost function analysis in the case of the urban bus transit sector. This is apparent from the next section.

4.3 Multiproduct Cost Analysis: The Case of Urban Bus Transit.

Few cost studies related to the urban bus transit sector have explicitly incorporated the multi-output nature of transport industries. This is partly because most urban transit firms were considered to be typically ‘single-service’ firms, providing the conventional type fixed-route and fixed-schedule services on city roads. However, with changing contextual factors and better conceptual developments, the somewhat more recent literature attempts to unearth the cost structure of urban bus transit firms under a multiproduct framework of analysis.

While determining the cost structure of the U.S. urban bus transit industry, Windle (1988) utilised a hedonic cost specification by expanding the output vector to include output characteristics in addition to the physical measure of ‘passenger-miles’ A translog cost function was estimated, symbolically given as:

\[ C = C[Y, P_i, Z_i] \]

where, \( C \) = costs; \( Y \) = passenger-miles of output; \( P_i = \) input prices (labour, fuel, capital and materials); \( Z_1 = \) output characteristics; \( Z_1 (\text{speed}) = \frac{\text{total bus-miles}}{\text{total bus-hours}} \); \( Z_2 (\text{average trip length}) = \frac{\text{total passenger-miles}}{\text{total number of passengers}} \); \( Z_3 (\text{average load factor}) = \frac{\text{total passenger miles}}{\text{total bus capacity}} \); \( Z_4 = \text{route-miles} \). The study was

---

\(^{107}\) Higher average speed means, fewer buses and drivers would be needed to provide a given number of bus-miles and therefore lower costs.

\(^{108}\) Longer trip length results in fewer stops and thus reduces all costs associated with carrying an additional passenger.

\(^{109}\) Higher the load factor, fewer bus-miles are needed to achieve a given level of passenger-miles and lower will be the costs per passenger-miles for the bus system.
motivated by a need to resolve apparently conflicting views on return to scale with regard to urban bus transit services. While econometric studies produced results which affirmed the hypothesis of constant returns to scale, policy makers, on the other hand almost universally dealt with the bus industry as if it were a natural monopoly with sizable scale economies. The study illustrated how these conflicting observations could be reconciled by distinguishing 'returns to scale' from 'returns to density'. 'Returns to scale' (RTS) was defined as the effect on costs of an equiproportional increase in both passenger-miles and the network variable depicted via route-miles. This was computed as:

\[ \text{RTS} = \left( \frac{\partial \ln C}{\partial \ln Y} + \frac{\partial \ln C}{\partial \ln Z_4} \right)^{-1} \]

Two measures of 'returns to density' were used. The first measure denoted as 'returns to density' (RTD) was given as:

\[ \text{RTD} = \left( \frac{\partial \ln C}{\partial \ln Y} \right)^{-1} \]

This could be interpreted as the effect on costs resulting from increase in passenger-miles, holding all other factors constant. Next, 'returns to passenger density' (RTP) was given as:

\[ \text{RTP} = \left( \frac{\partial \ln C}{\partial \ln Y} + \frac{\partial \ln C}{\partial \ln Z_3} \right)^{-1} \]

This could be interpreted as the effect on costs of increasing passenger-miles and load factor. Thus, 'returns to density' measures reflects the relationship between unit costs and the extent to which existing capacity is being utilised while the 'returns to scale' measure indicates how costs vary as capacity is expanded. These were evaluated from the estimated translog cost function indicating that the bus industry was operating under constant returns to scale but there were substantial returns to density.\(^{110}\) The policy implication that could be inferred from this result was that cost advantages would accrue by increasing density levels in terms of passenger-miles as well as average load. The result also helped to reconcile the observations of past econometric studies and the notion of policy makers regarding scale economies. While past studies were actually focussing on the 'returns to scale' concept, the underlying concept of 'returns to density' was what the authorities had in mind, while

\(^{110}\) RTS=1; RTD=1.25 and RTP=3.13 (pp. 130).
emphasizing that there were considerable economies in the bus transit industry. Further, the 'constant returns to scale' result meant that there were favourable conditions for deregulating the industry considerably.

Talley et al. (1986) conducted a cost function analysis of a multiservice transit firm in the U.S. in order to devise measures to step up financial receipts. This exercise was considered to be appropriate amidst a scenario of expected decline in subsidy grants. Broadly, the policy thrust was to replace segments of the firm's fixed route, fixed schedule mass transit services by a host of para-transit services\(^{111}\) which were presumed to be less costly. The possibility of contracting out some of these services to private operators was also contemplated. The aim of this analysis was to decipher precise cost implications of these changes. The data related to a single public agency (The Tidewater Transportation District Commission) providing a range of para-transit services as well as conventional mass transit service. Time-series quarterly data between 1979 and 1984 was used for the analysis. The multiproduct cost function estimated was:

\[
C = C [Q_{mt}, Q_{ehm}, Q_{vm}, Q_{dm}, P_f, P_f, V_t, V_o]
\]

where, \(C\) = variable cost; \(Q_{mt}\) = motor bus service miles; \(Q_{ehm}\) = elderly and handicapped paratransit service miles; \(Q_{vm}\) = van-pool paratransit service miles; \(Q_{dm}\) = dial-a-ride paratransit service miles; \(P_l\) = price of labour; \(P_f\) = price of fuel; \(V_t\) = number of motor buses owned; \(V_o\) = number of vehicles other than motor-buses owned. A short run variable cost function was estimated because the fleet strength could not be varied over the study period chosen. A translog form could not be used to estimate the cost function as the number of observations in the data set was not large enough. Thus a Cobb-Douglas specification was resorted to. The results indicated that considerable cost savings could be achieved from the provision of contracted out dial-a-ride service. Further, cost comparisons of alternative types of services clearly showed that it was relatively more expensive to increase motor bus services than to increase any one of the paratransit services. The analysis was concluded by stressing that operating deficits could be reduced by '.... restructuring the transit firm as a para-transit firm...

\(^{111}\) These paratransit services include 'commuter type' services as well as 'demand responsive' services, such as dial-a-ride, van-pool and paratransit services for the elderly and handicapped.
that contracts out service..' (pp. 365). This would be a preferable stand rather than the 'traditional' methods of improving the financial position of the firm by either increasing fares or reducing transit service which had actually turned out to be counter-productive.

In a bid to cut costs and minimise inefficiencies in the U.S. urban public transit industry government directives have stressed diversification of output in this industry as well as the importance of attracting private sector participation. As a result, an important characteristic of this industry has been altered over the years. Firms have expanded their operations to encompass different types of paratransit services rather than the earlier approach of restricting services to the more conventional fixed-route, fixed-schedule transit services. Under these changed circumstances, Colburn et al. (1992) attempted to investigate the long run cost structure of an urban multiservice transit firm in Virginia for determining the nature of economies of size and scope and to infer resulting policy implications. The firm selected for analysis was the same as that in the Talley et al. (1985) study and the general specification of the cost function was similar. But, the study period was wider based on quarterly data from September 1979 to August 1988. This permitted estimation of the more flexible translog cost function a prerequisite for drawing conclusions on the nature of scope and scale economies in the industry.

Conceptually, economies of size or scale for a multiservice firm exists if costs do not increase in the same proportion to the increase in scale (amount) of outputs under the assumption that the composition of outputs remains fixed (pp. 196). Based on parameter estimates of the flexible cost function, aggregate size economies were computed using the following formulation:

$$SL = \frac{1}{(\sum \delta lnC/\delta lnQ_i)}$$

112 The researchers claimed that an investigation of this nature had not been attempted before (pp. 196)
113 But there is a subtle difference Talley et al. (1985) had estimated a short run cost function incorporating a fixed capital input. In this study a long run cost function is estimated and therefore the fixed factor term is replaced by its price as $C=C(w,q)$ where $w$ is the vector of factor prices and $q$ the vector of outputs.
where, $C=$ total cost and $Q_i=$ type of service. Overall economies of size exist if $SL > 1$. As presented in the table below, the results showed that the firm exhibits economies of scale over a wide service range.

<table>
<thead>
<tr>
<th>SERVICE RANGE</th>
<th>$SL = 1/\left(\sum \delta ln C/\delta ln Q_i\right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>90% mean</td>
<td>4.00</td>
</tr>
<tr>
<td>95% mean</td>
<td>3.44</td>
</tr>
<tr>
<td>mean</td>
<td>3.04</td>
</tr>
<tr>
<td>105% mean</td>
<td>2.73</td>
</tr>
<tr>
<td>110% mean</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Source: Colburn et al. (1992, pp. 202)

A multiservice transit firm exhibits ‘economies of scope’ if the cost of providing different types of services jointly in one firm is less than the cost that would have to be incurred if these were produced separately. Symbolically, for a four output case this may be summarised as:

$$C(w, Q_1, Q_2, Q_3, Q_4) < [C(w, Q_1, 0, 0, 0) + C(w, 0, Q_2, 0, 0) + C(w, 0, 0, Q_3, 0) + C(w, 0, 0, 0, Q_4)]$$

It seems intuitively clear that if a multiservice firm exhibits diseconomies of scope, it can be broken down into several specialized firms without any increase in costs in the provision of the given levels of service or even with some cost reduction. The above expression may not be useful for empirical testing of the presence or absence of scope economies. A sufficient condition for economies of scope is the presence of weak cost complementarities between services (Panzar, 1989, pp. 21). The presence of weak complementarities implies that the marginal cost of producing any one service does not increase with increases in the quantity of

---

114 As mentioned earlier, the more general expression representing this concept is given as: $\{C/\sum Q_i (\delta C/\delta Q_i)\}$ (Panzar, 1989, pp. 8). This expression collapses to the one given above under the implicit assumption that marginal costs are identical across products (Bailey et al. 1982, pp. 1031).

115 Cost savings from joint provision of services may arise for many reasons. Common use of inputs such as management, maintenance and accounting facilities may generate cost economies. It has been asserted that inputs of a transportation firm may not be perfectly divisible. Hence for a single service firm excess capacity may result. But if two or more services are produced the problem of excess capacity could be mitigated and services could be provided at a lower cost than the sum of costs to be incurred if the services would have to be provided individually (For more detailed discussions regarding sources of economies of scope, see also Bailey et al., 1982, pp. 1026-1028 for an intuitively appealing presentation and Panzar, 1989, pp.19-21 for a more theoretical exposition along with proofs.)
any other service. A twice differentiable multiproduct cost function exhibits weak cost complementarities if:

$$\delta^2 C(w,Q) / \delta Q_i \delta Q_j \leq 0$$

In the study under consideration, evidence of cost complementarity was not found for all the service combinations. Therefore, economies of scope do not exist for this multiservice bus transit firm. More specifically, cost complementarities were found for service combinations involving conventional transit, elderly and handicapped and van pool services but not for service combinations involving dial-a-ride. The policy implication which could be inferred from these results was that, cost efficiency could be gained by the firm through reorganisation providing these cost complementarity exhibiting services, while another firm specialised in the provision of 'dial-a-ride' services (pp. 205).

The possibility of viewing the bus transit firm as a multi-output producer provides useful and new evidence on the operational and cost characteristics. This is exhibited by the studies just surveyed. Similarly, derivation of productivity indices is another important by-product of this expanded framework of complex statistical cost function analysis. Since the notion of 'performance' and 'accountability' has been deeply associated with the functioning of bus transit corporations around the world, the next section attempts to review studies which have concentrated on tracing productivity changes in the bus transport sector as decipherable from flexible cost functions estimated through econometric analysis.

4.4 Productivity and Cost Function Analysis.

The neoclassical cost function has been used to trace 'total factor productivity' growth in transportation industries, over time. Total factor productivity is defined as the proportionate rate of growth of output minus the proportionate rate of growth of output. This measure can be derived easily from the cost function by incorporating a time trend variable 't' and then deriving the following expression (Denny et al. 1981):

$$\text{TFP}_t = -B^* + (1-c_0)Q^*$$
where, $\text{TFP}_g = \text{total factor productivity growth}$; \( B = -\delta \ln C/\delta t \), which denotes the downward shift of the cost function; \( \epsilon_{eq} = \delta \ln C/\delta \ln Q \), the elasticity of cost with respect to output and \( Q^* = \delta \ln Q/\delta t \). Thus, sources of total factor productivity growth can be decomposed into two effects: one due to shifts in the cost function or 'technical change' and the other solely due to changes in scale of operation. While scale effects may be useful for an assessment of efficiency in exploiting existing technology, the shift phenomenon aids in throwing light on the extent of innovations in the industry such as new knowledge of technology, new managerial organisations, or new policy towards an industry. Most economists tend to identify the 'shift' concept with the notion of productivity growth (Oum et al., 1992, pp. 502).

These conceptual refinements have formed an integral part of recent empirical studies directed towards gauging the nature of productivity growth in transportation industries. Derivation of these indices from flexible cost functions which place few apriori restrictions on the underlying production and cost structure, strengthens reliability of these indicators, especially for providing crucial policy guidelines. Thus, '...the advantage of this approach is that total factor productivity is not an arbitrary measure but derived from economic theory...and consequently superior to the indicators of performance commonly used in transit systems' (Obeng et al., 1992, pp. 449). In fact this approach to the study of productivity trends evolved as an alternative to adhoc but widely used 'single' or 'partial factor productivity' measures, which were considered to provide poor approximations to productivity changes. These partial productivity indices\(^{116}\) expressed as a ratio between an organisation's output and an input (especially labor), it was criticised, failed to reveal causal factors accounting for observed productivity growth and was therefore limited from the point of view of economic analysis. 'Such measures may provide useful information but by their very nature they present only a partial picture and it is usually very dangerous to infer conclusions from them directly without further information' (Dodgson, 1985, pp. 14). Despite these serious shortcomings these non-parametric measures are widely used by transportation firms even today. In this section the review is restricted to studies which analyse productivity changes in the urban bus transport sector, wherein productivity indices are derived from estimated cost functions.

\(^{116}\) See Tomazinis (1975), Meyer et al. (1977), Allen (1979) and Feilding et al. (1985) for use of partial productivity measures in the bus transit sector.
Obeng (1985) and Hensher (1988), estimated translog cost functions to derive measures of total factor productivity. The productivity of an input 'i' was given by the elasticity of output with respect to that input. This partial productivity index with respect to input 'i' is given as:

\[(PPI)_i = \frac{\delta \ln Y}{\delta \ln X_i} \]

where, \(Y\) = output measure; \(X_i\) = measure of the 'i' th input factor. It can be easily shown that \((PPI)_i = S_i/e_{eq}\), where, \(S_i\) (share of the 'i' th input in total costs) = \([\delta \ln C/\delta \ln P_i]\), and \(e_{eq}\) (elasticity of cost with respect to output) = \([\delta \ln C/\delta \ln Y]\). Thus, factor productivities would be affected by substitution possibilities depicted through cost shares and scale economies. The total factor productivity index was expressed as:

\[\text{TFP} = \sum_i (PPI)_i = \sum_i (S_i/e_{eq}) = 1/e_{eq}\]

The shortcoming of this measure is that productivity is explained by scale effects alone. However, based on this measure, Obeng (1985) estimated a short run translog cost function for sixty-two bus transit firms in the U.S., in order to determine how far variation in costs could be explained by factor productivities. The short run analysis indicated that total factor productivity seemed to increase with firm size. This was consistent with the observation of scale economies for larger firms partly attributable to better capacity utilisation. Low substitution possibilities for 'labor' and it's high share in total costs indicated that 'labor' productivity was found to account for a large fraction of changes in total factor productivity. The long run cost function derived from the estimated short run cost function showed disconomies of scale for both small and large firms in the sample. The results indicated low partial productivities with respect to all inputs. Diseconomies observed for larger firms in the long run, meant that other factors outweigh the desirable effects of better capacity utilisation to produce diseconomies. The overall conclusion based on the quantitative results stressed the

---

117 The index of capacity utilisation was obtained by the ratio, \(CU=(\text{passenger-miles/capacity-miles})\). The following relationship was developed: \(ECS=0.1585-2.2020CU\), where, \(ECS=\frac{\delta \ln A.C.}{\delta \ln Y}\) is the elasticity of average costs (A.C.) with respect to output (Y).
importance of improving productivity of all inputs, especially the labor input, for reducing transit costs substantially.

Hensher (1988), used a similar approach to study productivity of private local bus operations in Australia. A cross-section sample of twenty-nine operators was used to estimate a translog cost function, based on the output measure 'passenger revenue' and inputs - 'labor', 'fuel' and 'capital'. The analysis indicated that both 'labor' and 'capital' contribute a like amount to the overall productivity of the firm. (pp. 153). However, labor productivity played a major part in explaining the variations in costs. Comparisons between popular partial productivity measures of system performance, and productivity indices derived from the cost function followed different trends. '......this finding is likely to be controversial in the light of the extensive use currently made of (partial ) performance indicators' (pp. 160).

To study productivity growth trends of a single bus operator in Belgium, Borger (1984), estimated a translog cost function using time series data. The methodology used was similar to that used in the pioneering work of Caves et al. (1981) for evaluating productivity growth in the U.S. railroad industry. Productivity indices were derived on the basis of an implicit production function relationship given by: \( f(Y, L, E, K, t) = 0 \) where, \( Y \) = output, \( L \) = labor, \( E \) = fuel, \( K \) = capital, \( t \) = time trend variable included to account for technological shifts in the production function. More specifically, two measures of productivity growth were constructed. The first measure traced the rate at which output could grow over time with all inputs held at a constant level. The second productivity growth index threw light on the rate at which inputs could be reduced over time with output held at a fixed level. Using the dual variable cost function \( C_v = C(Y, P_L, P_E, K, t) \) it was shown that these indices reduce to:

\[
\begin{align*}
(i) \quad P_1 &= -\left( \frac{\partial \ln C_v}{\partial t} / (\partial \ln C_v / \partial \ln Y) \right) \\
(ii) \quad P_2 &= -\left( \frac{\partial \ln C_v}{\partial t} / (1 - (\partial \ln C_v / \partial \ln K)) \right)
\end{align*}
\]

\(^{118}\) This conclusion was reached by testing the following relationship:
\( \ln C = 21.42 + 8.24 \ln P_{PL} + 0.30 \ln P_{PF} + 0.36 \ln P_{PK} \ldots \ldots \) \( (R^2 = 0.95) \), where, \( C \) = total cost and PPI\(_i\) are the partial productivity indices of the input factors.

\(^{119}\) These measures were, Cost efficiency\( = (\text{total kilometers run/total cost}) \); Labor efficiency\( = (\text{total vehicle hours/number of workers}) \); Vehicle-efficiency\( = (\text{total vehicle kilometers/total peak vehicles}) \); Maintenance efficiency\( = (\text{total vehicle-kilometers/number of maintenance employees}) \).

\(^{120}\) See Borger (1984, pp. 40-41)
These productivity indices were derived from the short run variable cost function related to bus firm in Belgium. The trends in productivity during the study period (1951-1979), revealed that most of the productivity growth had occurred between 1951 and 1960. As anticipated, these results could not be reconciled with those based on the simple partial productivity index (seat kilometers/man-hours of labor used) regularly computed by the bus company. The simple productivity index showed a completely different time path and substantially overestimated productivity growth over the entire study period. Vital policy decisions were being taken on the basis of the partial productivity measures. This could adversely affect policy results, jeopardising efficient operation of bus services by the firm.

Kim (1985), attempted to measure intertemporal efficiency differentials as well as average cost differentials related to the provision bus services in Israel.\(^{121}\)\(^{122}\) A translog cost function based on time-series was estimated. In general terms this cost function was specified as:

\[
C_d = g_d(w_d, Y_d, T_d)
\]

where, \(C=\)total cost; \(w=\)vector of input prices; \(Y=\)output; \(T=\)index of technology and the subscript 'd' denotes the respective time period. The intertemporal cost efficiency measure based on the 'Tornqvist approximation' to account for discrete changes between time periods 'o' and 'd' was given as:

\[
\mu_{od} = (\log C_d - \log C_o) - \left[\frac{1}{2}(\varepsilon_{cy_d} + \varepsilon_{cy_o})(\log Y_d - \log Y_o)\right] - \left[\frac{1}{2} \sum_i(Si_d + Si_o)(\log wi_d - \log wi_o)\right]
\]

where, \(\varepsilon_{cy} = (\delta \ln C / \delta \ln Y)\) the cost elasticity with respect to output; \(Si = \delta \ln C / \delta \ln wi\) is the cost share of input ‘i’. The average cost differential was computed as:

\[
[\log(C_d/Y_d) - \log(C_o/Y_d)] = \left[\frac{1}{2} \sum_i(Si_d + Si_o)(\log wi_d - \log wi_o)\right] + \left[\frac{1}{2}(\varepsilon_{cy_d} + \varepsilon_{cy_o})(\log Y_d - \log Y_o)\right] - [\log Y_d - \log Y_o] + \mu_{od}
\]

\(^{121}\) The data set used in this study was the same as that used to estimate a cost function in Berechman's (1983) study which has been already discussed.

\(^{122}\) See Caves et al. (1980) and Friedlaender et al. (1983) for similar approach used to study productivity growth in the railroad and trucking industry respectively. A more comprehensive approach is developed by Denny et al. (1981) for the study of Canadian telecommunications.
Thus, the average cost differential can be decomposed into three effects: the factor input effect, returns to scale effect and finally the pure efficiency differential effect denoted by ‘$\mu_{ae}$’ (pp. 176). The results suggested increasing cost-efficiency trends almost during the entire study period. Average costs had also declined during the entire study period. Decomposition of the average cost differential index into various factors showed that the labor input had caused an upward shift of average costs to the extent of 11.7%, capital had caused a downshift of 1.9%; scale had caused a reduction of approximately 3.1% and the pure efficiency effect had contributed to a downshift equivalent to about 9.4% of the average cost (pp. 180).

In order to determine the performance of a cross-section of bus transit systems operating in U.S. cities, Obeng et al. (1992), deciphered the nature of total factor productivity growth using the expression discussed above:

$$\text{TFP}_t = B' + (1 - \epsilon_{se})Q'$$

A translog cost function, based on pooled time-series data relating to the period 1983 to 1988 and restricted to bus transit systems having a fleet strength of more than twenty-five vehicles, was estimated to arrive at this index. An interesting result emerged from this analysis. Since the elasticity of cost with respect to output for the representative firm was almost equal to one, it followed that the observed growth in total factor productivity was due to technical change or ‘shifts’ in the cost function and not due to scale effects (pp. 453). The rate of increase in total factor productivity was found to be around 1.1% per year, during the period of analysis. Possible explanations for technical growth were given as capital subsidies, greater microcomputer use, privatisation and contracting. However, it was stressed that the effects of these variables on total factor productivity needed closer examination in future research (pp. 454). These trends were similar to partial productivity trends with regard to ‘labor’ but not that of ‘capital’ or ‘fuel’. It was asserted that policy measures based on partial productivity indices could lead to unforeseen undesirable results.
Hensher (1992), obtained similar results in a bid to identify sources of total factor productivity growth while appraising performance levels related to urban public bus operators in Australia. The 'scale' effect was found to contribute very little to the overall growth of total factor productivity (pp. 439). This analysis was based on two methodological innovations. Firstly, a demand equation was added to the system of equations consisting of the conventional cost function and share equations and the whole system was estimated to derive the productivity index. Symbolically, the system may given as:

\[ C = c(Q_o, w, t) \]

where, \( C \) = total cost; \( Q_o \) = vehicle-kilometers of output produced; \( w \) = vector of input prices and \( t \) = time trend variable,

\[ S_i = \frac{\delta \ln C}{\delta \ln w_i} \]

where, \( S_i \) are the associated share equations and

\[ Q_d = f(Q_o, F, Y, AC) \]

where, \( Q_d \) = quantity demanded in terms of passenger-kilometers; \( F \) = fare level; \( Y \) = income level and \( AC \) = cost of competing mode or cost of travel by car. The model was referred to as the 'market equilibrium' model. This 'market equilibrium' approach it was claimed, provided a more realistic interpretation of the role that demand levels may have on productivity (pp. 436). Secondly, to enhance the practicality of the total productivity index, it was then regressed against a set of operational variables to provide guidance to operators on sources of potential improvement in performance (pp. 436). For instance, 95% of the variation in total factor productivity growth was explained by four operational variables namely, log of 'route-kilometers', 'deficit per passenger', 'passenger-kilometers per vehicle-kilometer' and 'buses per employee'. As is evident, two of these partial measures are demand-side measures. This reinforced the significance of evaluating productivity in the context of demand for bus services and therefore the relevance of incorporating a demand function into the analysis.

Besides the objective of better evaluation of performance levels, these studies specifically demonstrate the superiority of the 'total factor productivity' index as against the commonly
used ‘partial productivity’ indices. These developments in the conceptual as well as methodological framework have been largely facilitated due to the use of complex statistical cost models for depicting the cost structure of bus transit firms.

5. Summary and Conclusions

In this paper an attempt has been made to review and document selective literature which traces the evolution of cost studies applicable to transport industries. While the focus has been on the urban bus transit sector, studies undertaken to analyse the behaviour of costs of other land-based transport modes have been quite comprehensively covered. In each case, an attempt has been made to emphasise the contextual considerations under which these studies were conducted and the manner in which the estimated results were accordingly interpreted. In the process, methodological innovations have been appropriately highlighted. On the whole this exercise aided in throwing considerable light on various facets related to economic characteristics of these industries. This could serve as useful background material for further research pursuits directed towards comprehending the structure of costs and related ‘economic effects’ in transportation industries or even other industries based on ‘network technologies’.

Considerable research efforts have been invested for strengthening the theoretical base of the cost function specification as well as for emphasising the use of flexible functional forms which permits the study of a whole host of economic effects related to the industry. Advances made with respect to the definition of ‘output’ in transport industries have also contributed to a better depiction of the cost structure of the transport industry. These developments would undoubtedly aid in more realistic long-term policy making and thereby ensure efficient provision of transport services.
REFERENCES


Koenskar, R., (1977): 'Optimal Scale and Size Distribution of American Trucking Firms',


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kaushik Basu, Arghya Ghosh,</td>
<td>The Babu and The Boxwallah: Managerial Incentives and Government</td>
</tr>
<tr>
<td>2</td>
<td>M.N. Murty, Ranjan Ray</td>
<td>Optimal Taxation and Resource Transfers in a Federal Nation (February</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1994)</td>
</tr>
<tr>
<td>3</td>
<td>V. Bhaskar, Mushtaq Khan</td>
<td>Privatization and Employment: A Study of The Jute Industry in</td>
</tr>
<tr>
<td></td>
<td></td>
<td>267-273</td>
</tr>
<tr>
<td>4</td>
<td>V. Bhaskar</td>
<td>Distributive Justice and The Control of Global Warming (March 1994)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The North, the South and the Environment: V. Bhaskar and Andrew Glyn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Ed.) Earthscan Publication London, February 1995</td>
</tr>
<tr>
<td>5</td>
<td>Bishnupriya Gupta</td>
<td>The Great Depression and Brazil's Capital Goods Sector: A Re-examine-</td>
</tr>
<tr>
<td>6</td>
<td>Kaushik Basu</td>
<td>Where There Is No Economist: Some Institutional and Legal Prerequisites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of Economic Reform in India (May 1994)</td>
</tr>
<tr>
<td>7</td>
<td>Partha Sen</td>
<td>An Example of Welfare Reducing Tariff Under Monopolistic Competition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(May 1994), Review of International Economics, (forthcoming)</td>
</tr>
<tr>
<td>8</td>
<td>Partha Sen</td>
<td>Environmental Policies and North-South Trade: A Selected Survey of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the Issues (May 1994)</td>
</tr>
<tr>
<td>9</td>
<td>Partha Sen, Arghya Ghosh,</td>
<td>The Possibility of Welfare Gains with Capital Inflows in A Small</td>
</tr>
<tr>
<td></td>
<td>Abheek Barman</td>
<td>Tariff-Ridden Economy (June 1994)</td>
</tr>
<tr>
<td>10</td>
<td>V. Bhaskar</td>
<td>Sustaining Inter-Generational Altruism when Social Memory is Bounded</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(June 1994)</td>
</tr>
<tr>
<td>11</td>
<td>V. Bhaskar</td>
<td>Repeated Games with Almost Perfect Monitoring by Privately Observed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Signals (June 1994)</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>13</td>
<td>Kaushik Basu</td>
<td>The Axiomatic Structure of Knowledge And Perception (July 1994)</td>
</tr>
<tr>
<td>16</td>
<td>Mrinal Datta Chaudhuri</td>
<td>Labour Markets As Social Institutions in India (July 1994)</td>
</tr>
<tr>
<td></td>
<td>S. Subramanian</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>K. Ghosh Dastidar</td>
<td>Debt Financing with Limited Liability and Quantity Competition (August 1994)</td>
</tr>
<tr>
<td>22</td>
<td>K. Ghosh Dastidar</td>
<td>Comparing Cournot and Bertrand in a Homogeneous Product Market (September 1994)</td>
</tr>
<tr>
<td>23</td>
<td>K. Sundaram</td>
<td>On Measuring Shelter Deprivation in India (September 1994)</td>
</tr>
<tr>
<td></td>
<td>S.D. Tendulkar</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Sunil Kanwar</td>
<td>Are Production Risk and Labour Market Risk Covariant? (October 1994)</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>26</td>
<td>Ranjan Ray</td>
<td>The Reform and Design of Commodity Taxes in the presence of Tax Evasion with Illustrative Evidence from India (December 1994)</td>
</tr>
<tr>
<td>27</td>
<td>Wietze Lise</td>
<td>Preservation of the Commons by Pooling Resources, Modelled as a Repeated Game (January 1995)</td>
</tr>
<tr>
<td>29</td>
<td>Jean Drèze, Jackie Loh</td>
<td>Literacy in India and China (May 1995). Economic and Political Weekly, 1995</td>
</tr>
<tr>
<td>30</td>
<td>Partha Sen</td>
<td>Fiscal Policy in a Dynamic Open-Economy New-Keynesian Model (June 1995)</td>
</tr>
<tr>
<td>34</td>
<td>Ajit Mishra</td>
<td>Hierarchies, Incentives and Collusion in a Model of Enforcement (January 1996)</td>
</tr>
<tr>
<td>35</td>
<td>Sunil Kanwar</td>
<td>Does the Dog wag the Tail or the Tail the Dog? Cointegration of Indian Agriculture with Non-Agriculture (February 1996)</td>
</tr>
<tr>
<td>37</td>
<td>Sunil Kanwar</td>
<td>The Demand for Labour in Risky Agriculture (April 1996)</td>
</tr>
<tr>
<td>38</td>
<td>Partha Sen</td>
<td>Dynamic Efficiency in a Two-Sector Overlapping Generations Model (May 1996)</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>39</td>
<td>Partha Sen</td>
<td>Asset Bubbles in a Monopolistic Competitive Macroeconomic Model (June 1996)</td>
</tr>
<tr>
<td>40</td>
<td>Pami Dua, Stephen M. Miller, David J. Smyth</td>
<td>Using Leading Indicators to Forecast US Home Sales in a Bayesian VAR Framework (October 1996)</td>
</tr>
<tr>
<td>41</td>
<td>Pami Dua, David J. Smyth</td>
<td>The Determinants of Consumers' Perceptions of Buying Conditions for Houses (November 1996)</td>
</tr>
<tr>
<td>42</td>
<td>Aditya Bhattacharjea</td>
<td>Optimal Taxation of a Foreign Monopolist with Unknown Costs (January 1997)</td>
</tr>
<tr>
<td>43</td>
<td>M. Datta-Chaudhuri</td>
<td>Legacies of the Independence Movement to the Political Economy of Independent India (April 1997)</td>
</tr>
<tr>
<td>45</td>
<td>Partha Sen</td>
<td>Terms of Trade and Welfare for a Developing Economy with an Imperfectly Competitive Sector (May 1997)</td>
</tr>
<tr>
<td>46</td>
<td>Partha Sen</td>
<td>Tariffs and Welfare in an Imperfectly Competitive Overlapping Generations Model (June 1997)</td>
</tr>
<tr>
<td>47</td>
<td>Pami Dua, Roy Batchelor</td>
<td>Consumer Confidence and the Probability of Recession: A Markov Switching Model (July 1997)</td>
</tr>
<tr>
<td>48</td>
<td>V. Pandit, B. Mukherji</td>
<td>Prices, Profits and Resource Mobilisation in a Capacity Constrained Mixed Economy (August 1997)</td>
</tr>
<tr>
<td>49</td>
<td>Ashwini Deshpande</td>
<td>Loan Pushing and Triadic Relations (September 1997)</td>
</tr>
<tr>
<td>50</td>
<td>Rinki Sarkar</td>
<td>Depicting the Cost Structure of an Urban Bus Transit Firm (September 1997)</td>
</tr>
<tr>
<td>51</td>
<td>Sudhir A. Shah</td>
<td>Existence and Optimality of Mediation Schemes for Games with Communication (November 1997)</td>
</tr>
<tr>
<td>52</td>
<td>V. Pandit</td>
<td>A Note on Data Relating to Prices in India (November 1997)</td>
</tr>
<tr>
<td>53</td>
<td>Rinki Sarkar</td>
<td>Cost Function Analysis for Transportation Modes: A Survey of Selective Literature (December 1997)</td>
</tr>
</tbody>
</table>