Expense preference behavior in public transit systems

K. Obeng *

Department of Economics, Transportation/Logistics, North Carolina A&T State University, Greensboro, NC 27410, USA

Received 23 April 1999; received in revised form 28 June 1999; accepted 12 November 1999

Abstract

This paper extends previous works that view transit systems as minimizing their after-subsidy costs. The paper uses the expense preference behavior model in economics and derives first-order conditions for the manager. From the first-order conditions, the paper formally shows that the decomposition of relative price inefficiency between management behavior and subsidies found in the work of Sakano et al. (1997) can be derived from a utility maximizing model, thus placing that decomposition within the shadow price literature. Extensions to the models to calculate expense preference are also presented. The results of the estimated models show that transit systems have expense preference for capital and not labor. This expense preference behavior increases total costs by about 15% and capital subsidies by about 20%. © 2000 Elsevier Science Ltd. All rights reserved.

1. Introduction

The notion of efficiency in transit systems continues to confound research which shows that despite their best efforts transit systems operate inefficiently. Costs continue to rise while revenues have not increased proportionately resulting in increasing levels of deficits that must be met by subsidies. To explain these deficits several models have been proposed including principal-agent relationships and x-inefficiency. While the principal-agent relationship contends that the reason for the higher cost is because of a divergence in the objectives of management and the transit board, x-inefficiency attributes the cost increases to firms adopting objectives that deviate from cost minimization.

Niskanen’s (1968) model of bureaucratic budget maximization provides another explanation for the increasing deficit. In this model, bureaucrats maximize their utilities by producing more...
output or by using more inputs. In each case the result is over budgeting that results in nonoptimal allocation of resources. For example, according to De Alessi (1969) bureaucrats may increase their budgets by adopting capital-intensive operations, while Hayes and Chang (1990) argue that bureaucrats may spend more on visible inputs. According to Hayes and Chang (1990) bureaucrats may over utilize capital as part of their efforts to increase the visibility of their production. This visible input hypothesis does not apply to capital alone, although traditionally that has been the focus. In the context of city management it has been suggested that mayors may hire more than the desirable number of employees to increase their support base and their chances for reelection (Hayes and Chang, 1990). Managers too may spend more on salaries, additional staff, and office furnishings (Hannan and Mavinga, 1980).

A condition for expense preference behavior is the separation of management from ownership or control, i.e., a principal-agent relationship. Another is imperfection in goods and capital markets which is true in public transit systems given the high levels of subsidies in this industry. In public transit systems too, there is a separation between management and control because the managers (agents) are hired employees; the major stakeholders (principals) are the transit boards, city councils, state departments of transportation in some cases, and the public. Because of its high costs, only in rare situations that the principals can monitor the activities of the agents to ensure they follow cost minimization principles. Consequently, the agents may pursue self-interest objectives different from those established by the principals.

Such a principal-agent relationship has been examined recently by Strausz (1997) who showed that the delegation of monitoring to a third party can be profitable, i.e., reduce agency costs. Profit sharing and providing incentives through other compensation schemes, such as employee stock ownership plans (ESOP), also have been suggested as solutions to the principal-agent problem because they can increase work effort (productivity), skills, and information flow. For example, Kruse (1992) showed that ESOP is associated with higher productivity. However, Ma (1994) discounts profits sharing as a solution to this problem because it imposes risks on the agent. Similarly, although it is widely used (e.g., in piece work) a performance contract is often discounted as a solution to the principal-agent problem because employers may raise the performance bar once it is achieved. Carlos and Nicholas (1993) note the importance of organizational culture in reducing agency costs. Here, an organizational culture that focusses on enhancing attachment to the organization, its goals, mission, and values (i.e., affective commitment) is thought to reduce shirking on the parts of employees.

Despite these solutions, principal-agent problems still persist in organizations. Therefore, in public transit systems we should expect expense preference behavior in the form of excessive expenditures on visible inputs such as facilities and equipment because managers pursue self-interest objectives. Indeed the subsidy patterns of the various levels of government in recent years seem to support the capital intensive and visible input hypotheses in public transit systems. Between 1990 and 1994 federal capital subsidies to public transit increased by 23.4% while operating subsidies which are spent on a less visible input increased by 17.79% (Federal Transit Administration, 1996b, p. 1). At the federal level, a factor that contributed indirectly to capital overuse in the past was the incentive tier component of the operating subsidy allocation formula which was inherently biased against noncapital inputs. Obeng and Azam (1995) found that the incentive tier increased the implied prices of noncapital inputs thereby making firms behave as if the prices of these inputs had increased hence buying less. Recent studies of relative price inefficiency in public

transit systems (e.g. Sakano et al., 1997), however, find an overall labor bias, thus casting doubt on the visible input hypothesis as it applies to public transit. These studies were based upon net cost minimization (i.e., total cost net subsidies), or constrained output maximization and, therefore, might not have appropriately accounted for management behavior.

This paper has two general purposes. First, it presents a generalized behavioral model for public transit. Second, it uses this model to investigate expense preference behavior in public transit firms. The paper shows that other behavioral models found in the public transit literature can be derived from this generalized behavioral (or utility maximization) model. Additionally, this model is used to obtain an empirical model that allows for the testing of expense preference behavior. The estimation of this empirical model with US public transit data suggests that firms have expense preference for capital, but not for labor, which is consistent with the common belief in the industry of over-investment in capital.

The rest of the paper is organized as follows. In Section 2, we present a constrained utility maximization model of the bureaucrat and derive a measure of relative price inefficiency from it. Section 3 presents extensions of the basic model of private cost minimization and specifies the empirical model to be estimated, while Sections 4 and 5 deal with the estimation and results, respectively. Finally, Section 6 deals with the conclusions of the paper.

2. A utility maximization model

Allocative inefficiency occurs when firms do not employ their inputs in proportions that lead to cost minimization. That is, for each input pair allocative inefficiency occurs if the marginal rate of technical substitution deviates from the corresponding ratio of input market prices. The amount of the deviation is allocative inefficiency and can be measured by estimating a stochastic cost frontier and its corresponding share equations. Alternatively, as in Good (1992), the deviation can be made a function of organizational characteristics and the first order condition estimated jointly with the underlying production function. The sum of the products of the coefficients of organizational characteristics and the values of these characteristics then measures allocative inefficiency.

The principal-agent problem noted above, rent-seeking, x-inefficiency, and moral hazards are among the reasons why such non cost minimizing behavior might occur. In US public transit systems, there is another reason and that is the presence of operating and capital subsidies. William Vickrey, for example, argued that if a firm receives a subsidy there is no pressure to operate efficiently and management will divert its attention from controlling costs to pleading for the subsidy (Arnott et al., 1994, p. 209). This is particularly true when there are economic rents to be realized from such subsidies, i.e., if the subsidies lead to profits. Even if management were to minimize costs, it is unlikely to minimize its total actual costs, but rather its private costs since these costs are those to be paid from its revenue generating functions (Obeng et al., 1997). Private cost is, in the context of the subsidy-receiving firm, actual total costs less subsidies from all sources, i.e., after-subsidy costs. Minimizing private costs assures after-subsidy profits if the resulting cost is less than passenger revenues. As we show in Table 1 there are many examples of transit systems making after-subsidy profits.
In the context of the shadow price literature, minimizing after-subsidy cost is similar to the hypothesis that firms choose inputs to minimize their shadow costs of given levels of output (Atkinson and Halvorsen, 1980). In fact, the objective of after-subsidy cost minimization leads to allocative distortion as do the shadow pricing results. Particularly, the after-subsidy cost minimization model shows that relative price efficiency deviates from the ratios of actual input prices. Also, the after-subsidy cost minimization objective could be likened to those in models in which firms minimize their shadow prices such as Eakin (1991) and Dor et al. (1997). These similarities imply that we can position the after-subsidy cost minimization results in the general literature of shadow pricing and employ shadow price models to analyze firm behavior.

To do so we consider a typical US public transit system which is the sole provider of transit services in the city from which it receives its franchise. This system is publicly owned and financed, and receives operating subsidies from local, state, and federal government sources. The transit system also receives capital subsidies to cover the costs of capital broadly defined to include property, equipment and vehicle acquisitions, and debt service. Though traditionally we have assumed that such a transit system makes losses, recent changes in funding, particularly the establishment of dedicated funding sources, make such an assumption invalid in some systems; the assumption is true only to the extent that passenger revenues do not cover costs. Table 1, for example, compares the 1996 operating funds (including operating subsidies and passenger revenues) of a random sample of transit systems to their operating costs. The table shows that some transit systems made after-subsidy profits if we consider operating subsidies. In some transit systems these after-subsidy profits were very small, and in others such those in Philadelphia, New

<table>
<thead>
<tr>
<th>Name of transit system</th>
<th>Operating funds (include operating subsidies) in thousands of US dollars</th>
<th>Operating costs in thousands of US dollars</th>
<th>Capital subsidies expended in thousands of US dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix RPTA</td>
<td>7859.7</td>
<td>6928.5</td>
<td>748.6</td>
</tr>
<tr>
<td>Tuscon-Sun Tran</td>
<td>26955.6</td>
<td>26871.4</td>
<td>7768.4</td>
</tr>
<tr>
<td>Denver RTD</td>
<td>165983.0</td>
<td>158715.6</td>
<td>18599.4</td>
</tr>
<tr>
<td>Greater Gridgeport TD</td>
<td>10609.9</td>
<td>10512.9</td>
<td>396.7</td>
</tr>
<tr>
<td>Atlanta-MARTA</td>
<td>243603.2</td>
<td>222496.9</td>
<td>212335.8</td>
</tr>
<tr>
<td>Baltimore Maryland MTA</td>
<td>254109.6</td>
<td>252956.9</td>
<td>140535.9</td>
</tr>
<tr>
<td>Washington-WMATA</td>
<td>666166.7</td>
<td>647309.0</td>
<td>378192.8</td>
</tr>
<tr>
<td>Riverside TA</td>
<td>17364.0</td>
<td>17300.4</td>
<td>6737.1</td>
</tr>
<tr>
<td>St. Petersburg PSTA</td>
<td>25960.7</td>
<td>25952.8</td>
<td>1858.1</td>
</tr>
<tr>
<td>St. Louis MCT</td>
<td>9008.4</td>
<td>6056.7</td>
<td>1923.8</td>
</tr>
<tr>
<td>Louisville TARC</td>
<td>34187.9</td>
<td>34783.3</td>
<td>5921.8</td>
</tr>
<tr>
<td>Springfield PVTA</td>
<td>21068.7</td>
<td>18724.1</td>
<td>18034.5</td>
</tr>
<tr>
<td>Detroit D-DOT</td>
<td>122429.4</td>
<td>122150.7</td>
<td>10814.6</td>
</tr>
<tr>
<td>San Diego Transit</td>
<td>56655.1</td>
<td>56478.1</td>
<td>12790.2</td>
</tr>
<tr>
<td>Denver RTD</td>
<td>16598.3</td>
<td>158715.6</td>
<td>18599.5</td>
</tr>
<tr>
<td>Albany COTA</td>
<td>25878.3</td>
<td>26086.7</td>
<td>1372.7</td>
</tr>
<tr>
<td>NY-MTA-NYCTA</td>
<td>3377827.2</td>
<td>2818836.6</td>
<td>1116870.0</td>
</tr>
<tr>
<td>Philadelphia SEPTA</td>
<td>708299.3</td>
<td>664201.8</td>
<td>253830.5</td>
</tr>
</tbody>
</table>

*Source: Federal Transit Administration (1996a).*
York City, Atlanta and Washington, DC, they were quite large. For the nation as a whole total operating revenues of transit systems were US$17.624 billion in 1996 compared to US$16.302 billion in operating expenses (Federal Transit Administration, 1996c, p. 14) which indicates after-subsidy profit of US$1.322 billion. Table 1 also provides information about the capital subsidies actually expended by each transit system.

Since transit systems are publicly owned and provide social services, their utility maximization problem is akin to that of the public sector bureaucrat who expends some inputs or produces more than the efficient level of output (e.g., Grosskopf and Hayes, 1993). Thus, the expense preference transit manager maximizes a utility function that has output and inputs as its arguments; the manager derives utility from producing more output \( Q \), for example, by serving a large population through overextending services into lower population density areas, and this results in overproduction. This is especially true when the manager perceives the marginal social benefits of this additional production to be more than the associated marginal social costs. Alternatively, the manager employs more inputs \( (x_N, x_G, x_K) \) for labor, fuel and capital, respectively) to produce a given level of output since it is only in doing so that he can justify the size of his budget and the subsidies. The manager who derives utility from output can be effective if the additional output is tailored, for example, to low income areas. Yet, the manager will be cost inefficient because production will exceed the cost minimizing level. This dichotomy between effectiveness and efficiency is recognized but not addressed in this paper. Also, we assume a separable utility function so the cross-effects in the manager’s utility function are ignored although they could be important. This assumption is useful in the present analysis since we are not concerned with the components of marginal utility and the resulting cross-effects.

For ease of analysis and interpretation let \( U(Q, x) \) be the linear utility function of the manager. Furthermore, let the marginal utility of an input be \( v_x \), and the marginal utility of output be \( v_Q \), where both marginal utilities are positive. With these assumptions Eq. (1) shows the utility function of the manager

\[
U(Q, x) = v_Q Q + v_N x_N + v_G x_G + v_K x_K.
\]  

In Eq. (1), \( N, G, \) and \( K \) are labor, fuel and capital, respectively, and we use fuel as a proxy for all inputs that are neither capital nor labor. From this utility function, inefficiency can occur from input overuse as a result of using an inefficient technology or from overproduction.

In maximizing its utility, the manager faces the constraint that his revenues \( B \) from non-subsidy sources must at least equal to his actual total costs net of subsidies. This constraint is shown below and is necessary for the firm to be efficient after receiving the subsidy

\[
B \geq w_N x_N + w_G x_G + w_K x_K - A_o(x_N, x_G) - A_k(x_N, x_K).
\]

The first three terms on the right-hand side represent actual total cost. \( A_o \) is the operating subsidy and is a function of labor, fuel, and materials. The capital subsidy, \( A_k \), is a function of

---

1 The utility function \( U(Q, x) \) is the same as in Grosskopf and Hayes (1993). It has been suggested by a reviewer that a nonlinear utility function or an implicit form would be appropriate for our analysis. We note that they will not change our results.
labor, and capital inputs which for our purposes are measured using fleet size as a proxy. \(^2\) Also, \(Q\) is output in terms of vehicle miles and the entire right-hand-side of this equation is net cost or private cost.

The firm maximizes its utility in Eq. (1) subject to Eq. (2). This constrained utility maximization problem can be analyzed by solving the Lagrangian below:

\[
\max L = v_Q Q + v_N x_N + v_G x_G + v_K x_K + \lambda (B - w_N x_N - w_G x_G - w_K x_K \\
+ A_o(x_N, x_G) + A_k(x_N, x_K)),
\]

where \(\lambda\) is a multiplier and ensures that the net cost constraint is satisfied. The first-order conditions for utility maximization from this equation are below:

\[
\begin{align*}
\partial L / \partial x_G &= v_G - \lambda w_G (1 - \theta_{og} F_{og}) + v_Q \partial \bar{Q} / \partial x_G, \quad (4) \\
\partial L / \partial x_K &= v_K - \lambda w_K (1 - \theta_{kk} F_{kk}) + v_Q \partial \bar{Q} / \partial x_K, \quad (5) \\
\partial L / \partial x_N &= v_N - \lambda w_N (1 - \theta_{on} F_{on} - \theta_{kn} F_{kn} + v_Q \partial \bar{Q} / \partial x_N. \quad (6)
\end{align*}
\]

In Eqs. (4)–(6), \(F\) is the ratio of subsidies to input cost, \(^3\) \(\theta\) the subsidy elasticity of an input and may be obtained from subsidy functions specified to include relevant inputs as some of its arguments, and lowercase \(k\) is a capital subsidy. Setting these equations equal to zero and solving, we have \(^4\)

\[
\begin{align*}
\frac{\partial \bar{Q}(x)}{\partial x_N} &= \left( w_N (1 - \theta_{on} F_{on} - \theta_{kn} F_{kn}) - \frac{v_N}{\lambda} \right) / \left( w_k (1 - \theta_{kk} F_{kk}) - \frac{v_k}{\lambda} \right), \quad (7) \\
\frac{\partial \bar{Q}(x)}{\partial x_K} &= \left( w_G (1 - \theta_{og} F_{og}) - \frac{v_G}{\lambda} \right) / \left( w_k (1 - \theta_{kk} F_{kk}) - \frac{v_k}{\lambda} \right). \quad (8)
\end{align*}
\]

A similar equation can be obtained for labor and fuel by forming the ratio of Eq. (6) to Eq. (4). In Eqs. (7) and 8, the numerators and the denominators on the right-hand side are the shadow or implied prices \((w^s_N, w^s_K, w^s_G)\) of the inputs. In terms of shadow prices we may rewrite the right-hand side of Eq. (7) as \(w^s_N/w^s_K\) and of Eq. (8) as \(w^s_N/w^s_G\). This rewriting shows that the ratios of the shadow prices are not the same as the ratios of the actual input prices \((w_G/w_K, w_N/w_G)\) which result from minimizing actual total cost \(C = \sum_i w_i x_i\). Management behavior such as utility maximization and the presence of subsidies distort the relative prices of the inputs and lead to allocative inefficiency.

\(^2\) The federal government allows capital subsidies to be spent on labor assigned specifically to capital.

\(^3\) For example the share of operating subsidy in labor cost is \(F_{on} = A_{oN}/w_N x_N\) and the elasticity of operating subsidy with respect to labor is \(\theta_{on} = \partial \ln A_{oN} / \partial x_N\). The other elasticities can be derived in a similar manner.

\(^4\) The derivation of Eqs. (7) and (8) can be illustrated with an example. Using Eqs. (5) and (6) and setting them equal to zero and solving we get: \(\partial \bar{Q}/\partial x_K = \lambda w_k (1 - \theta_{kk} F_{kk}) - \lambda k\) and \(\partial \bar{Q}/\partial x_N = \lambda w_N (1 - \theta_{on} F_{on} - \theta_{kn} F_{kn}) - \lambda N\). Forming the ratio of the second expression to the first and dividing through by lambda gives Eq. (7).
If we consider only the first terms in the numerator and the denominator on the right-hand side of Eqs. (7) and (8), we see that capital and operating subsidies distort the relative input prices. That is, the subsidies make the relative input prices different from the ratios of actual input prices. If the subsidies reduce input prices by the same proportion, they will leave the ratios of the shadow prices the same as the ratios of actual input prices. These distortions from operating and capital subsidies have been studied more recently by Obeng (1994), Obeng and Azam (1995, 1997), and Sakano et al. (1997) whose models included equations seven and eight, minus the third terms in the numerators and the denominators.

Taking account of the second terms on the right-hand side of Eqs. (7) and (8) we see that subsidies are not the only factors that cause allocative distortions in transit systems. Management behavior such as utility maximization that causes deviations from cost minimization also distorts input price ratios by reducing the numerator and the denominator in these equations. If management maximizes utility instead of minimizing after-subsidy cost, the marginal utilities in these equations will be positive and there will be further allocative distortions from those prevailing with subsidies. These distortions will depend upon how the third terms affect the sizes of the numerator and the denominator. It is possible that management behavior may or may not affect the distortions in relative price efficiency that occur with subsidies alone.

The first terms in the numerator and the denominator in Eqs. (7) and (8) are the after-subsidy prices $w_G$ and $w_N$ of the inputs if the firm minimizes its actual total costs net of subsidies (Obeng and Azam, 1997). This after-subsidy cost is $C^* = \sum_i w_i^* x_i$, where the superscripted prices are defined in Eq. (9).

$$
\begin{align*}
  w_G^* &= w_G(1 - \theta_{oG} F_{oG}), \\
  w_K^* &= w_K(1 - \theta_{kk} F_{kk}), \\
  w_N^* &= w_N(1 - \theta_{oN} F_{oN} - \theta_{kN} F_{kN}). \\
\end{align*}
$$

A way of expressing Eqs. (7) and (8) to be able to estimate the distortions from management behavior is to assume that for each input utility maximization affects the distortion in its price from subsidies by a factor such as $\phi$. With this assumption, we may rewrite Eq. (8) as below. Similar equations can be derived for the other input pairs

$$
\frac{\partial Q(x)}{\partial x_G} / \frac{\partial Q(x)}{\partial x_K} = \frac{\Phi_G w_G(1 - \theta_{oG} F_{oG})}{\Phi_K w_K(1 - \theta_{kk} F_{kk})} = \frac{\Phi_G w_G^*}{\Phi_K w_K^*}.
$$

Here, the right-hand side is the ratio of the shadow prices ($w_G^*/w_K^*$) of fuel and capital. Rewriting this equation gives the expression for relative cost shares as in Eq. (11)

$$
\frac{x_G w_G}{x_K w_K} = \left[ \frac{\Phi_K(1 - \theta_{kk} F_{kk})}{\Phi_G(1 - \theta_{oG} F_{oG})} \right] \left( \frac{\partial \ln Q}{\partial \ln x_G} / \frac{\partial \ln Q}{\partial \ln x_K} \right),
$$

$\Phi_G$ and $\Phi_K$ are the mean distortions in the prices of fuel and capital, respectively, from utility maximization. Therefore, the ratio $\Phi_K/\Phi_G$ gives the allocative distortion between capital and fuel...
Total relative price inefficiency in Eq. (11) is the term in the large set of brackets, which is $\Phi_K(1 - \theta_{kk}F_{kk})/\Phi_G(1 - \theta_{gg}F_{gg})$. If we take the logarithm of this price inefficiency term, we obtain a decomposition of fuel–capital relative price inefficiency between utility maximization and subsidies as in Eq. (12). Thus

$$E_{KG} = \ln\left(\frac{\Phi_K}{\Phi_G}\right) + \ln\left(\frac{1 - \theta_{kk}F_{kk}}{1 - \theta_{gg}F_{gg}}\right),$$

(12)

$E_{KG}$ is the logarithm of relative price inefficiency. Eq. (12) is equivalent to the decomposition provided by Sakano et al. (1997), and it is a formal derivation of that decomposition. It is also equivalent to the logarithm of the relative price efficiency measure in Fare and Grosskopf (1990), i.e., $(w_K/w_K)/(w_G/w_G)$.

We can use Eq. (12) to test for management preferences for inputs in relative terms. If $E_{KG}$ is positive, fuel is overused relative to capital because the shadow price of fuel is relatively small and makes a firm behave as if its actual price has fallen and thereby buy more fuel. Alternatively, if $E_{KG}$ is negative, capital is overused relative to fuel because in comparison the shadow price of capital is lower than the shadow price of fuel.

Except Sakano et al. (1997), applications of stochastic frontier models to public transit data calculate only the first term in Eq. (12) and do not explicitly account for subsidy influence in the second term. A problem in using the stochastic frontier approach to assess relative price inefficiency is that one must assume a specific distribution of the errors in the underlying production or cost function, and the calculation of inefficiency depends on the chosen probability density function. Thus, the calculated relative price inefficiencies may vary from function to function. Also, if management prefers two inputs, Eq. (12) cannot reveal that information when we compare the relative cost shares of these inputs, so we must find an alternative approach to analyzing such preferences. Ideally, the approach should be flexible enough to incorporate after-subsidy costs so that the influences of management behavior and subsidies on input overuse can be examined simultaneously.

3. Private cost minimization and expense preference

Following our assumption that management maximizes its utility, it will spend more on the inputs for which it has a preference than it is required under both actual cost minimization, and after-subsidy cost minimization. This is because subsidies and utility maximization reduce the perceived prices of the inputs per Eq. (8) and make firms behave as if the prices of these inputs had fallen. Dor et al. (1997) identify two ways in which such expense preference behavior can be tested. The first includes an additional constant term in the expenditure or share equation of the
input thought to be preferred. This constant term is derived from the profit maximization relationship that the price of each input is equal to its marginal revenue product. If in the estimation this term is significant and positive then it verifies the expense preference theory. This approach is referred to as the intercept test in the expense preference literature. In a series of proofs Mester (1989) shows that this test is restrictive and valid only when the production technology is Cobb-Douglas.

The second test involves using a nonlinear argument in the cost function to capture expense preference behavior (Mester, 1989; Dor et al., 1997). Here, the inputs are separated into those which managers do or do not have an interest in over utilizing. Then, a proportional amount is added to the over utilized quantity, and actual total cost and input shares are derived and estimated. This proportion is the expense preference parameter. If in the estimation this parameter is positive and significant it provides evidence of expense preference behavior. A major difficulty with this latter approach is determining which input is thought to be overused. Errors in making this determination can introduce some bias into the analysis.

Broadly, these two approaches fall under separable decision models. We may add that a third approach is the joint decision model where the bureaucrat minimizes costs by behaving as if the prices of some inputs were some smaller proportion of their actual prices and, therefore, buying more of these inputs than is required under cost minimization. A cost function is derived and estimated jointly with share equations and the levels of significance of the proportionality factors or joint expense parameters provide tests of expense preference behavior.

In this paper, we use the nonlinear approach in the separable decision model, and focus on transit systems that receive operating subsidies and capital subsidies. As before, this firm employs labor, fuel, and capital in its production process where fuel is a proxy for itself and materials. The firm favors expenses on capital and labor and does not show expense preference behavior for fuel. Furthermore, the firm minimizes its after-subsidy costs and perceives its input prices as \( w_G \), \( w_N \) and \( w_K \) in Eq. (9). At the price of \( w_G \) the firm buys the quantity \( x_G \) of the input that is not preferred, and at the prices \( w_N \) and \( w_K \) it buys \((1 + z_N)x_N\), and \((1 + z_K)x_K\) of the preferred inputs. Here, the \( z \)'s are the separable expense preference parameters. The market prices of these inputs are \( w_G \), \( w_N \) and \( w_K \) and total resource costs are given by Eq. (13)

\[
C = w_Gx_G + w_Nx_N(1 + z_N) + w_Kx_K(1 + z_K).
\]

From Eq. (13) expense preference adds the \( S_Nz_N + S_Kz_K \) proportion to the minimum actual total cost if we do not consider subsidies. Here, \( S \) is the observed actual share of an input in cost. Because we assume the firm minimizes its after-subsidy costs, we rewrite Eq. (13) as in Eq. (14) utilizing the information in Eq. (9). For each input, we solve it for its actual price and substitute it into Eq. (13). For example, for capital its actual price is: \( w_k = w^*_K/(1 - \theta_{kk}F_{kk}) \). Performing the substitution and noting that the after-subsidy total cost is, \( C^* = w^*_Kx_K + w^*_Nx_N + w^*_Gx_G \), and that the after-subsidy share of an input in cost is, \( S^* = w^*x/C^* \), where the subscripts are suppressed, we have

\[
C = C^* \left[ S^*_G(1 - \theta_{og}F_{og})^{-1} + S^*_K(1 + z_K)(1 - \theta_{kk}F_{kk})^{-1} \right.
\]
\[
+ S^*_N(1 + z_N)(1 - \theta_{kn}F_{kn} - \theta_{kN}F_{kN})^{-1} \right].
\]
The first term represents the cost of the input that is not preferred, and the second and third terms are the costs of the preferred inputs. From this equation, if the firm receives subsidies, expense preference behavior adds the $S_K z_k (1 - \theta_{kk} F_{kk})^{-1} + S_N z_N (1 - \theta_{oN} F_{oN} - \theta_{KN} F_{KN})^{-1}$ proportion to actual total cost. Comparing this increase to $S_N z_N + S_K z_K$ we can ascertain if the cost impact of expense preference worsens with subsidies.

The input share in actual total costs for the not-preferred and preferred inputs are also given by the equations below. These equations are derived from Eq. (14) by dividing each term on its right-hand side by the entire equation. For example, for labor we divide $C^* S^*_N (1 + z_N) (1 - \theta_{oN} F_{oN} - \theta_{KN} F_{KN})^{-1}$ which is the last term in Eq. (14) by the entire right-hand side. Thus, for fuel and labor, respectively, we have,

$$S_G = \frac{S^*_G (1 - \theta_{oG} F_{oG})^{-1}}{S^*_G (1 - \theta_{oG} F_{oG})^{-1} + S^*_k (1 - \theta_{kk} F_{kk})^{-1} (1 + z_k) + S^*_N (1 - \theta_{oN} F_{oN} - \theta_{KN} F_{KN})^{-1} (1 + z_N)},$$

$$S_N = \frac{S^*_N (1 + z_N) (1 - \theta_{oN} F_{oN} - \theta_{KN} F_{KN})^{-1}}{S^*_N (1 - \theta_{oN} F_{oN} - \theta_{KN} F_{KN})^{-1} (1 + z_N) + S^*_G (1 - \theta_{oG} F_{oG})^{-1} + S^*_K (1 - \theta_{kk} F_{kk})^{-1} (1 + z_K)}.$$

The share equation for capital can be derived similarly.

Eqs. (14)–(16) must be estimated together after properly specifying the minimum after-subsidy cost function. A translog function is used to describe the technology of public transit systems. Its advantage is that it is not as restrictive as the Cobb–Douglas technology in terms of elasticities of substitution. Because the firm minimizes its after-subsidy cost, the usual regularity restrictions on cost apply. The translog approximation of the minimum after-subsidy cost function is given by Eq. (17) after imposing symmetry restrictions on the coefficients.

$$\ln C^* = \alpha_0 + \sum_i \alpha_i \ln w^*_i + 0.5 \sum_i \sum_j \alpha_{ij} \ln w^*_i \ln w^*_j + \sum_i \alpha_{iQ} \ln w^*_i \ln Q + 0.5 \alpha_{QQ} (\ln Q)^2. (17)$$

Using Shephard’s lemma, the share of an input in after-subsidy or private cost is the derivative of the logarithm of after-subsidy cost with respect to the derivative of the logarithm of the private input prices. Again, this share of input cost is unobserved and is given by Eq. (18). Homogeneity of degree one in after-subsidy input prices requires that the sum of the second-order price coefficients for each input in the share equations is zero, the sum of the constant terms in the share equations is one, and the sum of the output coefficients in the share equations is zero. Furthermore, because after-subsidy cost must be exactly equal to the difference between actual total cost and subsidies from all sources, the sum of the elasticities of operating subsidies with respect to the inputs must equal to one. This same restriction holds for the elasticities of capital subsidies with respect to inputs. These restrictions are in Eq. (19).
\[
\sum_{n=1}^{n} \alpha_{n} = 1, \quad \sum_{m=1}^{m} \alpha_{mn} = 0, \quad \sum_{i=1}^{i} \theta_{oi} = 1, \quad \sum_{j=1}^{j} \theta_{kj} = 1. \tag{19}
\]

The additional subscripts \((o, k)\) in this equation distinguish operating subsidy and capital subsidies, respectively. Substituting Eqs. (15)–(17) into Eqs. (12)–(14), the resulting equations are jointly estimated in this paper.

4. Data

The data for the estimation come from Obeng and Azam (1997) who graciously provided them for this study. Using, these data, we test the expense preference behavior in this paper. The data consist of the original 439 US bus transit systems that Obeng and Azam (1997) used in their work, and cover the period from 1983 to 1992. These authors provide a detailed description of the data elsewhere (Obeng and Azam, 1997), including the calculation of the user price of capital.

Briefly, the sample is an unbalanced panel of 439 transit systems that operate only fixed route bus services and that reported their data to the Federal Transit Administration (then the Urban Mass Transit Administration) from 1983 to 1992 as part of the annual Section 15 reporting system required by Congress. Because the data form an unbalanced panel, each sample is treated as an independent observation. Labor price is total labor compensation divided by labor hours where labor hours are 2080 h times the equivalent labor reported for each system in the Section 15 data. Fleet size is the measure of capital and its user price each year is calculated by the formula

\[w_k = P_k (r + d) e^{-dE}\]

for each system (Obeng and Azam, 1997). \(P_k\) is the weighted average price of a new bus, \(r\) is an interest rate on a high yield municipal bond for the cities where the transit systems are found, \(d\) is a straight line rate of depreciation assuming a twenty-year life for a vehicle, and \(E\) is the average fleet age. Because transit systems operate different types of buses, the weighted price \(P_k\) is the sum of the products of the number of each type of vehicle in a transit system’s fleet and its price new divided by fleet size. Multiplying the user price of capital \(w_k\) by fleet size gives the total user cost of capital for each transit system. Again, we use fuel as a proxy for itself and materials and assign all nonlabor operating costs to it. Then, we calculate fuel price as this cost divided by the gallons of fuel purchased. Also, we follow past literature on transit cost estimation and use vehicle miles as the measure of produced output and route miles as the measure of network size.

The data used in the estimation did not include all the 439 systems in the original data but a sub-sample. This sub-sample includes only those firms whose data fit the models in this paper during the estimation process. Thus, we do not preselect these systems, instead we allow the program used in the estimation to eliminate those data that are inconsistent with the models and the restrictions. For the systems eliminated, the model of private cost minimization as formulated here may not apply, though expense preference behavior may be exhibited. Also, it is possible that other formulations of the model would be applicable to these systems.

Table 2 shows some descriptive statistics about the data. The largest transit system in the sample operates 650 vehicles, while the average system operates 53 vehicles. Thus, the sample consists of small and medium size transit systems and does not include any of the systems that operate in very large metropolitan areas such as New York City, Philadelphia, and Chicago. The
standard deviations in the table are quite low relative to the mean values of the variables except in the cases of average fare and the ratios of subsidies to input costs. These small standard deviations suggest that the transit systems in the sample have similar characteristics. Large differences between these systems are likely to be observed in terms of fares and how much of their costs are covered by subsidies.

5. Estimation

With the appropriate substitutions made, Eqs. (16)–(18) form a system of equations that we can estimate using the nonlinear seemingly unrelated equation method. This approach uses the actual shares of subsidies in input costs to estimate the $\theta$s. Although reasonable, the estimated subsidy elasticities of inputs may be inconsistent with those that would be obtained if we were to estimate subsidy functions also. We avoid this problem by restricting the values of $\theta$s to those from operating subsidy and capital subsidy equations, and estimating these equations jointly with the cost and share equations.

In the subsidy equations the logarithm of the sum of operating subsidies from all sources, is a function of the logarithms of labor hours ($N$), gallons of fuel ($G$), average fare ($P$), population density ($D$) in terms of population per square mile, and average fleet age ($E$). Comparatively, the logarithms of the capital subsidies from all sources are a function of the logarithms of fleet size ($K$), labor ($N$), population density ($D$), and network size ($R$). The exact forms of these equations are below after imposing the relevant restrictions, $\theta_{oN} + \theta_{oG} = 1$ and $\theta_{kk} + \theta_{kN} = 1$, on the coefficients of the inputs in the subsidy equations.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Sample size</th>
<th>Symbol</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor price (logs)</td>
<td>161</td>
<td>$w_N$</td>
<td>2.4220</td>
<td>0.3522</td>
</tr>
<tr>
<td>Fuel price (logs)</td>
<td>161</td>
<td>$w_G$</td>
<td>1.2427</td>
<td>0.2470</td>
</tr>
<tr>
<td>Capital price (logs)</td>
<td>161</td>
<td>$w_K$</td>
<td>9.5926</td>
<td>0.1626</td>
</tr>
<tr>
<td>Vehicle miles (logs)</td>
<td>161</td>
<td>$Q$</td>
<td>13.7440</td>
<td>0.9448</td>
</tr>
<tr>
<td>Labor hours (logs)</td>
<td>161</td>
<td>$N$</td>
<td>11.7376</td>
<td>0.9460</td>
</tr>
<tr>
<td>Fleet size (logs)</td>
<td>161</td>
<td>$K$</td>
<td>3.4528</td>
<td>0.8971</td>
</tr>
<tr>
<td>Gallons of fuel (logs)</td>
<td>161</td>
<td>$G$</td>
<td>12.3435</td>
<td>1.0442</td>
</tr>
<tr>
<td>Population density (logs)</td>
<td>161</td>
<td>$D$</td>
<td>3.4869</td>
<td>1.0276</td>
</tr>
<tr>
<td>Route miles (logs)</td>
<td>161</td>
<td>$R$</td>
<td>5.0342</td>
<td>0.6760</td>
</tr>
<tr>
<td>Average fleet age (logs)</td>
<td>161</td>
<td>$F$</td>
<td>1.8453</td>
<td>0.4893</td>
</tr>
<tr>
<td>Average fare (logs)</td>
<td>161</td>
<td>$P$</td>
<td>0.0486</td>
<td>0.5214</td>
</tr>
<tr>
<td>Operating subsidy (logs)</td>
<td>161</td>
<td>$A_o$</td>
<td>14.0780</td>
<td>1.0828</td>
</tr>
<tr>
<td>Capital subsidy (logs)</td>
<td>161</td>
<td>$A_k$</td>
<td>11.7558</td>
<td>1.8767</td>
</tr>
<tr>
<td>Share of operating subsidy in labor costs</td>
<td>161</td>
<td>$F_{oN}$</td>
<td>0.9757</td>
<td>0.3077</td>
</tr>
<tr>
<td>Share of capital subsidy in capital costs</td>
<td>161</td>
<td>$F_{kK}$</td>
<td>0.5608</td>
<td>0.6086</td>
</tr>
<tr>
<td>Share of operating subsidy in fuel costs</td>
<td>161</td>
<td>$F_{oG}$</td>
<td>1.6928</td>
<td>0.3749</td>
</tr>
<tr>
<td>Share of capital subsidy in labor costs</td>
<td>161</td>
<td>$F_{kN}$</td>
<td>0.1635</td>
<td>0.1589</td>
</tr>
</tbody>
</table>
\[
\ln(A_o) = \Phi_0 + \theta_{oN} \ln(1 + z_N) + \theta_{oN} \ln(N) + (1 - \theta_{oN}) \ln(G) + \Phi_P \ln(P) + \Phi_D \ln(D) \\
+ \Phi_E \ln(E),
\]

\[
\ln(A_K) = \omega_0 + \theta_{kk} \ln(1 + z_k) + \theta_{kk} \ln(K) + (1 - \theta_{kk}) \ln(N) + \omega_R \ln(R) + \omega_D \ln(D).
\]

In Eqs. (20) and (21), \(\theta_{oN}\) and \(\theta_{kk}\) are, respectively, the elasticity of operating subsidies with respect to labor, and the elasticity of capital subsidies with respect to capital; \(\omega_0\) is the constant term in the capital subsidy equation. Comparatively, \(\theta_{oN} \ln(1 + z_N)\) and \(\theta_{kk} \ln(1 + z_K)\) are the proportional changes in operating subsidies and capital subsidies, respectively, because the bureaucrat expense prefers labor and capital.

Previous works of Obeng and Azam (1995, 1997) justify making the subsidies functions of the variables in Eqs. (20) and (21). Notably, these authors point out that population density captures equity considerations in subsidy allocation, average fleet age accounts for the need for vehicles to be maintained and replaced often thus requiring more operating and capital subsidies, while network size exerts higher demands for capital subsidy for vehicle acquisition because of increased service.

Since a translog function is Taylor’s series expansion of a broad class of functions, we select the mean as the expansion point. However, as Obeng and Azam (1997) note, if we use this point of expansion for the ratios of subsidies to input costs it implies no price distortions from subsidies in the average firm. To avoid this implication we use the actual values of these variables in the estimation. Eqs. (16)–(18), (20) and (21) with the appropriate substitutions made were jointly estimated by the nonlinear seemingly unrelated equations method.

6. Results

Table 3 shows the estimation results and reveals that most of the relevant coefficients are significant at the 0.05 level or better including the subsidy elasticities of the inputs. Also, the adjusted coefficients of determination for the cost function and the subsidy equations are quite reasonable. Not shown in the table are the rather low adjusted coefficients of determination for the labor and capital share equations which are 0.058 and 0.1507, respectively, and are typical. The first-order input coefficients are the mean values of the output elasticities of the inputs if we set subsidies to zero. This is because we noted before that the ratios of subsidies to input costs are not normalized by subtracting their mean values from their observed values.

From these coefficients, without subsidies, i.e., setting subsidies to zero, there are economies of scale in the sampled transit systems. A percentage increase in output increases total costs by 0.5841% according to our results. This can be compared to a 0.6458% increase in total cost from a percentage increase in output when the effects of subsidies are considered. These cost elasticities of output suggest that there are fewer economies of scale to be enjoyed by transit systems when there are subsidies. This is because the subsidies make the average cost curve relatively flat so at each level of output the elasticity of average cost with respect to output is smaller than it was before the subsidies.

With subsidies the shares of labor, fuel, and capital, in total private cost change. Since the subsidies make the private price of an input smaller than its actual price, they affect the quantities
that will be demanded and the cost shares. From the coefficients, the calculated labor share in private cost is 0.5456 which is larger than its share of 0.5232 in total actual cost. Similarly, the share of capital in total private cost is 0.2714 which is larger than its share of 0.1806 in total actual cost. But, the presence of subsidies does not increase the share of fuel in cost; the share of fuel in private costs is 0.1990 when we consider subsidies, and this is lower than the corresponding share of fuel of 0.2962 in total actual costs. This implies that capital and labor are substituted for fuel.

Using these results, it is deducible that the lower after-subsidy prices of labor and capital make firms want to increase the quantities they demand of these inputs proportionately more than they want to increase the quantity of fuel demanded. Since the quantities of all the inputs increase from subsidies, we cannot say that management shows expense preference behavior for a particular input. To do so the expense preference behavior must favor some inputs over others. However, that the quantity of capital demanded increased from subsidies is consistent with the positive and statistically significant expense preference coefficient $z_k$ in Table 3. Therefore, we can attribute

---

### Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficients</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operating subsidy: Adjusted $R - SQ = 0.8939$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant ($\Phi_0$)</td>
<td>$-0.1801^a$</td>
<td>0.0877</td>
</tr>
<tr>
<td>ln($N$)</td>
<td>0.5584$^a$</td>
<td>0.0006</td>
</tr>
<tr>
<td>ln($G$)</td>
<td>0.4416$^a$</td>
<td>0.0006</td>
</tr>
<tr>
<td>ln($P$)</td>
<td>$-0.1420^a$</td>
<td>0.0572</td>
</tr>
<tr>
<td>ln($D$)</td>
<td>0.1169$^a$</td>
<td>0.0290</td>
</tr>
<tr>
<td>ln($F$)</td>
<td>0.1597$^a$</td>
<td>0.0556</td>
</tr>
</tbody>
</table>

| **Cost function: Adjusted $R - SQ = 0.9133$** | | |
| Constant ($\phi_0$) | $-0.2294^a$ | 0.1140 |
| ln($w_N$) | 0.6253$^a$ | 0.0234 |
| ln($w_G$) | 0.2333$^a$ | 0.0294 |
| ln($Q$) | 0.5841$^a$ | 0.0278 |
| ln($w_N$) ln($w_K$) | $-0.1016^a$ | 0.0091 |
| ln($w_N$) ln($Q$) | $-0.0632^a$ | 0.0075 |
| ln($w_K$) ln($Q$) | 0.0078$^a$ | 0.0036 |
| ln($w_K$) ln($w_K$) & ln($Q$) | 0.1139$^a$ | 0.0086 |
| ln($w_G$) ln($Q$) & ln($Q$) | 0.1110$^a$ | 0.0020 |
| ln($N$) ln($Q$)$^2$ | 0.0526$^a$ | 0.0203 |
| ln($R$) | 0.3459$^a$ | 0.0499 |
| $z_N$ | $-0.0045$ | 0.1485 |
| $z_K$ | 0.8122$^a$ | 0.3580 |

| **Capital subsidy: Adjusted $R - SQ = 0.5050$** | | |
| Constant ($\omega_0$) | $-0.8859^a$ | 0.1285 |
| ln($K$) | 0.6476$^a$ | 0.0030 |
| ln($N$) | 0.3524$^a$ | 0.0031 |
| ln($D$) | 0.2026$^a$ | 0.0972 |
| ln($R$) | 0.5644$^a$ | 0.1501 |

* Significant at 0.05 level. The coefficients of fuel in the operating subsidy equation, and labor in the capital subsidy equation are calculated from the restrictions. Note that because the coefficients of the hare equations are implied by the cost equation they are not reported separately.
part of the increase in the quantity of capital demanded from the subsidies to expense preference behavior. But, we cannot attribute the increase in labor to expense preference behavior since $z_N$ is insignificant statistically.

From these results, when transit management maximizes its utility, it shows expense preference behavior for capital but not for labor. This behavior results in a 20.36% (i.e., $0.3424 \times \ln(1 + 0.8122)$) increase in capital subsidies to transit systems. Furthermore, this behavior of transit firms adds to costs. The average increase in actual total costs when transit systems show expense preference for capital is about 14.73%. Comparatively, there is an increase of 29.12% in after-subsidy cost from expense preference behavior. Thus, expense preference behavior exerts severe cost pressures on a firm that minimizes its after-subsidy cost.

The finding of expense preference for capital in transit systems is consistent with the general attitude toward extensive capital investment in the public transit industry. Accounts by Black (1995) clearly show that such preference for capital is pervasive in the public transit industry, particularly when it comes to rail projects. His survey of the literature shows that transit systems prefer capital intensive projects because they give a better public image, they are trendy, and they give transit systems a sense of identity. Also, our findings are consistent with the assertion of Hayes and Chang (1990) that bureaucrats spend more on visible inputs which in our case are capital. As well, we find consistency between our results and that of De Alessi (1969) that bureaucrats spend more on capital-intensive projects.

7. Conclusion

Expense preference behavior in public transit systems has been analyzed in this paper. The method adopted extends previous work particularly those of Obeng and Azam (1997) and Sakano et al. (1997) to include expense preference coefficients, and estimates the resulting set of cost, share, and subsidy equations. The results show that our sample of transit firms has expense preference for capital and not for labor. This expense preference behavior leads to higher demands for capital and capital subsidies. Furthermore, expense preference behavior increases total actual costs by 14.73%, capital subsidies by 20.36%, and after-subsidy costs by 29.12%. These findings are consistent with the common belief in the industry of over investment in capital. Additionally, the paper shows that the after-subsidy cost minimization model in previous works, and the decomposition of relative price inefficiencies between management behaviors and subsidies found in (Sakano et al., 1997) can be derived formally from a utility maximization framework that fits into the shadow price literature. Thus, the paper generalizes the results of previous work and, for this, it makes a contribution to the literature. Although we cast our models within the after-subsidy cost minimization framework, it will be interesting to compare our findings to those from other expense preference behavior models using the same data.

Acknowledgements

I acknowledge the financial support of the Urban Transit Institute at North Carolina A&T State University for the project that resulted in this paper. The author also acknowledges helpful
comments from the anonymous referees. The views expressed herein are the author’s and not that of the Institute.

References


