# The Structure of Public Transit Costs in the Presence of Multiple Serial Correlation

MATTHEW KARLAFTIS
ADK Consulting Engineers

PATRICK MCCARTHY
KUMARES SINHA
Purdue University

## ABSTRACT

Most studies indicate that public transit systems operate under increasing returns to capital stock utilization and are significantly overcapitalized. Existing flexible form time series analyses, however, fail to correct for serial correlation. In this paper, evidence is presented to show that ignoring multiple serial correlation can have important policy implications. Based on monthly time series data from the Indianapolis Public Transit Corporation, the results indicate that failure to correct for serial correlation significantly affects economies of capital utilization estimates and has potentially important implications for optimal size of the transit fleet.

## INTRODUCTION

Since the 1960s, when most U.S. transit systems were privately owned and operated and received no public financial assistance, subsidies have increased rapidly. Government financial support for public transport has grown, while operating losses and investment needs have also been increasing. The total operating subsidy from all levels of government (local, state, and federal) rose from $318 million in 1970 to $9.27 billion in 1990, a

Patrick McCarthy, Professor of Economics, Purdue University, 1310 Krannert Building, West Lafayette, IN 47907. Email: mccarthy@mgmt.purdue.edu.

near thirtyfold increase in 20 years (Pucher 1995). Similarly, the total capital subsidy from all levels of government rose from $200 million in 1970 to $5.56 billion in 1990, nearly 28 times higher (Pucher 1995). In spite of some evidence that subsidies have led to mild increases in ridership (Cervero 1984, Bly and Oldfield 1986), there is a general consensus that subsidies have had a degrading effect on system efficiency and productivity and have increased operating costs.[1] (Obeng 1985, Kim and Spiegel 1987, Bly and Oldfield 1986).

Several analyses have identified two related effects of continuing capital subsidies on public transit operations. First, public transit systems have an incentive to prematurely replace their capital stock (Frankena 1987). Second, by lowering the unit price of rolling stock, public transit capital subsidies provide public transit managers with an incentive to overcapitalize their systems. Although sparse, the evidence on overcapitalization in public transit is not only consistent with this hypothesis but also indicates that the extent of overcapitalization may be large. Viton (1981), for example, found that public transit systems were overcapitalized by as much as 57%. De Rus (1990) and Obeng (1984, 1985) also found that actual fleet sizes for transit systems are much larger than necessary to produce the current levels of output.

Claiming that their bus fleets are not excessive but rather needed for peak period demands, public transit managers generally argue against the empirical evidence that their systems are overcapitalized. This raises an interesting question. Do public transit agencies routinely overcapitalize their systems by as much as 50%, as existing empirical evidence suggests? Or, are existing empirical models deficient in some way that, if corrected, would produce results on the economic structure that are more consistent with observed behavior?

Of the studies that have analyzed public transit costs, all have used either panel data or time series data. However, although some authors have recognized that serial correlation may be present and could affect the study's results, none of these studies adjusted for serial correlation.[2] As is well known, the presence of serial correlation produces unbiased, consistent, but inefficient parameter estimates; further, variance estimates are biased and inconsistent, thereby invalidating hypothesis tests. The question addressed here is whether the presence of uncorrected, serially correlated errors in a flexible form cost function affects the model's implications on a public transit firm's production technology. To explore this, we developed and estimated a translog cost function model for the city of Indianapolis for a 60-month period, from January 1991 through December 1995. The next section, Methodology, identifies the model and summarizes the data used for the analysis, and the third section presents the estimation results. The final section provides concluding comments.

[1] In a cross-country comparative analysis of public transit systems and subsidies, Pucher (1995) notes "In virtually no other country have transit subsidies been as ineffective as in the United States." (p. 401)

## METHODOLOGY

Assuming one output $y$ (vehicle-miles), three variable inputs $z_i$ (labor, fuel, and maintenance), and one fixed input $k$ (number of buses), equation (1) identifies a public transit firm's short-run translog cost function:

$$\ln VC = \alpha_0 + \alpha_y \ln y + \sum_{i=1}^{3} \alpha_i \ln p_i + \alpha_k \ln k + \sum_{i=1}^{3} \gamma_{iy} \ln p_i \ln y + \sum_{i=1}^{3} \gamma_{ik} \ln p_i \ln k + \frac{1}{2} \sum_{i=1}^{3} \sum_{i=1}^{3} \gamma_{ij} \ln p_i \ln p_j + \gamma_{ky} \ln k \ln y + \frac{1}{2} \gamma_{yy} (\ln y)^2 \ln y + u \qquad (1)$$

[2] Berechman and Giuliano (1984), for example, identified second-order autocorrelation in their quarterly time series analysis of AC Transit in Alameda County, California. However, there was no attempt to correct for the problem. De Borger (1984), Berechman (1987), and Colburn and Talley (1992) also estimated time series models, but neither tested nor corrected for serial correlation. Outside of the public transit literature, Braeutigam et al. (1984) corrected for first-order serial correlation in a study on railroad costs using monthly time series data. While monthly time series data could potentially suffer from first- to twelfth-order serial correlation, the authors did not report the results of the model without correction for autocorrelation or whether they tested for higher order autocorrelation. As a result, it is not possible to assess how seriously autocorrelation affected their results.

where $VC$ is the variable cost of production, $p_i$ is the price of variable input factor $z_i (i = 1, 2, 3)$, $u$ is the disturbance term, and $\alpha_0$, $\alpha_y$, $\alpha_i$ ($i = 1, 2, 3$), $\alpha_k$, and $\gamma_{ij}$ ($i, j = 1, 2, 3, y, k$) are parameters to be estimated. According to equation (1), a transit firm's fleet size is fixed in the short run, implying that the level of variable inputs the firm employs at any given set of prices or output will depend on the level of rolling stock available to the system. The associated share equations (using Shephard's lemma) are:

$$s_i = \frac{p_i z_i}{VC} = \alpha_i + \sum_{j=1}^{J} \gamma_{ij} \ln p_i + \gamma_{iy} \ln y + \\ + \gamma_{ik} \ln k + u_i \quad i = 1,2,3 \tag{2}$$

where $s_i$ ($i = 1, 2, 3$) is the share of input $i$, and $u_i$ ($i = 1, 2, 3$) is the error term for share equation $i$. Equations (1) and (2) constitute a multivariate equation system that can be written more generally as (Berndt 1991):

$$Y_t = X_t\, b + u_t \tag{3}$$

where $Y_t$ is the ($n \times 1$) vector of dependent variables, $X_t$ is the ($n \times m$) vector of independent variables, $b$ is the ($m \times 1$) coefficient vector, $t$ denotes a given time period, and $u_t$ is an ($n \times 1$) vector of random disturbances. If $u_t$ has a first-order stationary univariate autoregressive structure, then

$$u_t = R u_{t-1} + \epsilon_t \\ t = 1, \ldots, T \tag{4}$$

where $R$ is an ($n \times n$) autocovariance matrix, and $\epsilon_t$ is a vector of disturbances with mean zero and constant variance. Combining (4) and (5) results in an equation with uncorrelated disturbances:

$$Y_t = RY_{t-1} + (X_i\ RX_{t-1})b + \epsilon_t \\ t = 1, \ldots, T \tag{5}$$

The usual maximum likelihood estimation methods could be applied to equation (5). However, only $J$–1 equations are independent, due to the constraint that the shares at each observation sum to unity (Berndt and Savin 1975).

Because the $J$ disturbances must sum to zero at each observation, the ($J \times J$) disturbance covariance matrices are singular, and this singularity condition imposes restrictions on the autoregressive process. Violation of these restrictions implies that the maximum likelihood estimates, and associated likelihood ratio test statistics, depend on which share equation is deleted from the system.

In order for the parameter estimates of the multivariate equation system in equation (5) to be invariant to the share equation deleted, the matrix $R$ has to be diagonal, and all the diagonal elements must be equal (Berndt and Savin 1975). Moreover, the autoregressive multivariate system in equation (5) can be generalized to account for higher order autoregressive processes. In particular, for an $M$th-order autoregressive process,

$$u_t = R_1 u_{t-1} + R_2 u_{t-2} + \quad + R_M u_{t-M} + t \\ t = 1, \ldots, M \tag{6}$$

where each $R_i$ ($i = 1, \ldots, M$) is a diagonal matrix whose elements must all be equal for the coefficient estimates to satisfy the invariance property (Berndt and Savin 1975).

## ESTIMATION RESULTS

### Model Estimation

Data for the analysis come from monthly observations for the city of Indianapolis over a five-year period, January 1991 through December 1995.[3] Table 1 reports the translog estimation results for the uncorrected model and for the

[3] The Indianapolis transit system, operated as a public enterprise since 1972, is a medium-sized system that, for the period under study, served an average population of 950,000 with an average fleet of 247 buses. Although the Indianapolis Public Transportation Corporation primarily provides fixed route, fixed schedule services, the data also reflect some demand-responsive services that the city offers. Since demand-responsive services account for less than 3% of the total, their inclusion is not expected to significantly affect the results.

**TABLE 1 Short-Run Translog Cost Functions**[a]

| Dependent Variable: Short-Run Costs ($10^6$) | | (a) Uncorrected model | | (b) Corrected model | |
|---|---|---|---|---|---|
| **Variable** | **Parameter** | **Estimate**[b] | **(t-statistic)** | **Estimate**[b] | **(t-statistic)** |
| Constant term | $\alpha_o$ | 14.135 | (1,084.80) | 14.106 | (280.80) |
| Output (vehicle-miles, units of 105)[c] | $\alpha_y$ | 0.571 | (2.78) | 1.145 | (5.39) |
| Price of labor ($/hour)[d] | $\alpha_l$ | 0.648 | (201.50) | 0.574 | (15.40) |
| Price of maintenance ($/hour)[d] | $\alpha_m$ | 0.295 | (104.30) | 0.350 | (11.60) |
| Number of buses[e] | $\alpha_k$ | 0.380 | (1.38) | 0.191 | (0.59) |
| (Price of labor) • (price of labor) | $\gamma_{\mathrm{ll}}$ | 0.175 | (6.76) | 0.167 | (8.14) |
| (Price of maintenance) • (price of maintenance) | $\gamma_{\mathrm{mm}}$ | 0.177 | (12.40) | 0.189 | (16.50) |
| (Number of buses) • (number of buses) | $\gamma_{\mathrm{kk}}$ | 10.280 | (1.54) | 1.813 | (0.63) |
| (Output coefficient) • (output coefficient) | $\gamma_{\mathrm{yy}}$ | –3.019 | (–0.76) | –0.683 | (–0.50) |
| (Price of labor) • (price of maintenance) | $\gamma_{\mathrm{ml}}$ | –0.157 | (–8.88) | –0.161 | (–11.20) |
| (Price of labor) • (number of buses) | $\gamma_{\mathrm{lk}}$ | –0.067 | (–0.87) | 0.042 | (0.45) |
| (Price of maintenance) • (number of buses) | $\gamma_{\mathrm{mk}}$ | 0.051 | (0.74) | –0.054 | (–0.64) |
| Price of labor output | $\gamma_{\mathrm{ly}}$ | 0.050 | (0.92) | 0.355 | (4.91) |
| Price of maintenance output | $\gamma_{\mathrm{my}}$ | –0.066 | (–1.32) | –0.355 | (5.67) |
| Number of buses output | $\gamma_{\mathrm{ky}}$ | 10.326 | (1.88) | 2.701 | (1.22) |
| Dummy for 1st quarter | $d_{q1}$ | 0.043 | (2.76) | 0.017 | (2.03) |
| Dummy for 2nd quarter | $d_{q2}$ | 0.030 | (2.71) | 0.015 | (2.21) |
| 1st-order autocorrelation, cost equation | $ar_{c1}$ | — | — | 0.551 | (5.06) |
| 2nd-order autocorrelation, cost equation | $ar_{c2}$ | — | — | 0.346 | 3.14) |
| 1st-order autocorrelation, share equations | $ar_{sh1}$ | — | — | 0.659 | (6.74) |
| 2nd-order autocorrelation, share equations | $ar_{sh2}$ | — | — | 0.386 | (3.98) |
| 6th-order autocorrelation, share equations | $ar_{sh6}$ | — | — | –0.058 | (–1.39) |
| $\tilde{R}^2$ (system $R^2$) | | 0.985 | | 0.991 | |

[a] Full information maximum likelihood estimates are invariant to share equation deleted (Berndt 1991, p. 463). The estimation results presented in table 1 normalize on the price of fuel.

[b] The data were collected from the Indianapolis Public Transit Corporation accounting, maintenance, and operations reports for fiscal years 1991–1995. The system's total monthly operating cost, excluding depreciation and amortization of intangibles, measures short-run operating costs. To capture the possibility that the system faces systematic differences in its operating environment during different months, preliminary runs of the model included peak-base vehicle ratio, average speed of service, and age of fleet. In each case, we could not reject the null hypothesis of no effect. We also included a time trend in earlier model runs, but the coefficients for the first- and second-order time trends were not significant. Moreover, the model with time trend variables violated the concavity requirements at two points in the sample. As a result, we excluded the time trend in the final model specification.

[c] Total vehicle-miles provided was selected as the output measure since bus operations are the primary determinant of costs in a transit system (Savage 1997). The output measure also includes "deadhead" miles, that is, miles traveled by revenue vehicles when not in revenue service (not available for passengers). These are a small portion of the total and typically include miles traveled to and from storage and maintenance facilities as well as some training mileage.

[d] As is common in many translog cost function models for public transit, well-defined measures for the input prices do not exist. In this study, similar to methodology of others (e.g., Berechman and Giuliano 1984, Applebaum and Berechman 1993, and Talley and Colburn 1993), we allocate monthly expenses to the various input categories (i.e., labor and maintenance) and then divide the expenses by paid monthly labor hours per category. The monthly price of labor, for example, was estimated by dividing the total labor expenses (including wages, fringe benefits, and pension payments to operators and administrative employees) by the paid labor hours to operators and administrative employees. A similar procedure was followed to determine the price of maintenance. For the price of fuel, we used actual prices since these were available from the monthly reports.

[e] Bus fleet size is the number of buses the system owns and operates during a given month.

[f] In earlier model runs, a full complement of monthly dummy variables was included, but likelihood ratio tests could not reject the null hypothesis that the monthly dummy variables in a quarter were equal.

model corrected for multiple serial correlation.[4] From column (a) in table 1, the system $R^2$ indicates that the model fits the data well. 98.5% of the generalized variance in the dependent variable is "explained" by the variation in the explanatory variables in the system of equations.[5] Further, the estimated function satisfies the necessary neoclassical conditions that the cost function be linear, homogeneous, nondecreasing, and concave in input prices.[6] While linear homogeneity is imposed on the model's parameters (see footnote 4), the estimation results must be checked ex post facto to determine whether the function is nondecreasing and concave in input prices. At every point in the sample, the estimated cost function satisfies each of these latter two conditions.

Indianapolis' mean behavior during the five-year period reveals that the coefficients for price of labor ($\alpha_l$) and price of maintenance ($\alpha_m$), respectively, estimate the share of costs attributed to labor and maintenance at mean production. Although the share equation for fuel was dropped in order to estimate the model, the linear homogeneity conditions identified in footnote 4 imply that the coefficient for price of fuel is $\alpha_f = 1 - \alpha_l - \alpha_m$

The interpretation of the coefficient for output, $\alpha_y$, is somewhat ambiguous. Public transit firms operate with a given amount of rolling capital (e.g., buses) and over a fixed (at least in the short run) network. With data on both rolling stock and network size, the coefficient for output would provide information on economies of traffic density. That is, the coefficient would identify the impact that an increase in output would have on the cost of using a fixed rolling stock over a given network. However, for this model, there were no data available on whether Indianapolis' network size significantly changed over the five-year period. If, in fact, there was little change in network size, then the coefficient for output in table 1 reflects economies of traffic density for a given rolling stock. On the other hand, if there were large changes to the network between 1991 and 1995, the coefficient for output would more appropriately reflect economies of capital utilization. Since in either case rolling stock is held fixed, we shall interpret $\alpha_y$ as a measure of economies of rolling stock utilization but bear in mind that it may also reflect economies of traffic density to the extent that Indianapolis' network underwent little change during the sample period.[7]

In general, the estimation results are consistent with expectations. First, the price coefficients are positive and strongly significant. At mean production, labor and maintenance account for 64.8% and 29.5% of costs. From the linear homogeneity conditions, fuel accounts for 5.7% of costs. Also, relative to the latter part of a year, the results indicate that the transit system experiences higher costs during the winter and spring quarters.[8]

The coefficient for output indicates that the Indianapolis mass transit system operates under

[4] To ensure homogeneity of degree one in variable input prices, given the fixed factor $k$ and output $y$, the following restrictions are imposed on the parameters:

$$\sum_{i=1}^{3} \alpha_i = 1,\ \gamma_{ij} = \gamma_{ji}\ \forall i, j$$

$$\sum_{i=1}^{3} \gamma_{ij} = \sum_{j=1}^{3} \gamma_{ji} = \sum_{i=1}^{3} \gamma_{iy} = \sum_{i=1}^{3} \gamma_{ik} = 0$$

[5] The system $R^2$ reported in the table is computed as (Berndt 1991)

$$\tilde{R}^2 = 1 - \frac{|EE'|}{|y'y|}$$

where $|EE'|$ is the determinant of the residual cross-product matrix of the full model, and $|y'y|$ is the determinant of the residual cross-product matrix of a model in which all slope parameters are simultaneously set to zero. See Berndt (1991) for a discussion of this measure.

[6] The cost function is nondecreasing in input prices if the fitted factor shares are positive at each observation and is concave in input prices if the Hessian matrix based on the fitted factor shares is negative semidefinite.

[7] An issue of interest is whether output is endogeneously determined. That is, do increases in output cause changes in cost, or are both variables endogenously determined? To check for this, we used Granger's (1969) causality test. By running two sets of two regressions, we rejected the hypothesis that "$x$ (output) Granger causes $y$ (cost)" and accepted the hypothesis that "$y$ does not Granger cause $x$" (both F-tests were evaluated at the five percent significance level). As a result, we can say that output "Granger causes cost." In this context, the finding that output is exogenous is not surprising since political motivations, in addition to market forces, often play a large role in the transit services actually provided.

[8] In earlier model runs, a full complement of monthly dummy variables were included, but likelihood ratio tests could not reject the null hypothesis that the monthly dummy variables in a quarter were equal.

increasing returns to rolling stock utilization at mean production level. Holding all else constant, including the size of bus fleet, a 10% increase in output increases short-run variable costs 5.7%.[9] In addition to operating under increasing returns to capital utilization, we see in column (a) of table 1 that the coefficient for number of buses is positive and significant at a 0.10 level (one tail test). A 1% increase in rolling stock raises operating costs 0.38%. As discussed more fully in the final section, this finding suggests that Indianapolis' system is overcapitalized, since a necessary condition for cost minimization is that the coefficient on capital stock be negative. [10]

Column (b) in table 1 provides parameter estimates when the model is adjusted for serial correlation, and we see that this has improved the overall fit of the model.[11] The system $\tilde{R}^2$ increases from 98.5% to 99.1%.[12] Adjusting for first- and second-order autocorrelation in the cost function and first-, second-, and sixth-order autocorrelation in the share equations provided the best model fit. It is important to recall that the autocorrelation coefficients for the share equations ($ar_{sh1}$, $ar_{sh2}$, $ar_{sh6}$) are restricted to being equal across share equations to satisfy the diagonality requirement of the $R_i$ matrices. The estimated pattern of the autocorrelation coefficients indicates a positive first- and second-order correlation and a mild sixth-order correlation for the share equations.

We again see that the price variables have their expected positive signs and are strongly significant. However, there are two interesting differences between columns (a) and (b). First, whereas Indianapolis was estimated to be operating under increasing returns to capital utilization when serial correlation is ignored, we see in column (b) that, when adjusted for serial correlation, Indianapolis' transit system is estimated to be operating under mildly decreasing returns to utilization. However, at normal levels of significance we cannot reject the null hypothesis of constant returns to utilization, suggesting that Indianapolis may be an efficient short-run producer of vehicle-miles.[13]

A second difference relates to the coefficient of the fixed factor, number of buses. Although still positive, the estimated coefficient $\alpha_k$ is lower and, more importantly, not statistically significant at any reasonable level of significance.

## Comparative Analysis of the Models

Table 2 presents test results for two restrictive production technologies, homotheticity (i.e., output changes can be met with constant input ratios), and Cobb-Douglas production technologies (homotheticity plus constant unitary elasticities of substitution between inputs) for the uncorrected model and the model adjusted for serial correlation.[14] Consistent with other studies, a Cobb-Douglas production technology is strongly rejected in each case.[15] However, when serial correlation is present but not corrected, the model accepts the null hypothesis of homotheticity, while the correct-

[9] De Borger (1984) found a similar result using annual time series data from Belgium.

[10] If the coefficient on capital stock is negative, then an increase in a transit system's rolling capital stock will generate savings in variable costs in excess of the unit bus price. A bus system is overcapitalized if the variable cost savings is either less than the unit bus price or actually increases. As noted previously, Frankena (1987) found that capital subsidies increase bus turnover. This suggests that public transit systems that overcapitalize, i.e., inefficiently invest in capital, will have an inefficiently low average fleet age and higher unit operating costs relative to optimal capital investment, which could produce a positive correlation between increased capital and operating expenses in a well-specified model. This does not appear to characterize Indianapolis' system, whose 16–18 year average age of fleet (over the five-year period) is a bit older than the industry average.

[11] To test for serial correlation, the model was initially estimated under the constraint that the $R_i$ matrices (from equation 6) equal zero and then was re-estimated with $R_i$ not equal to zero. The usual likelihood ratio test is based on the sample maximized log-likelihood functions obtained from the previous models (Berndt 1991). The null hypothesis of no autocorrelation was rejected at the 0.01 level.

[12] When adjusted for serial correlation, the cost function was also found to be nondecreasing and concave in input prices at each point in the sample.

[13] The t-statistic for the hypothesis that $\alpha_y = 1$ is 0.68.

[14] Testing for homotheticity is equivalent to testing the null hypothesis that $\gamma_{iy} = 0$ ($i$ = labor, maintenance) versus the alternative hypothesis that at least one of these parameters is nonzero.

[15] A Cobb-Douglas technology characterizes the production of transit trips if we can accept the null hypothesis that $\gamma_{yy} = \gamma_{ij} = \gamma_{iy} = 0 \forall$i. Viton (1981) and Obeng (1984) reject a Cobb-Douglas technology in short-run analyses, while Williams and Hall (1981), Berechman and Giuliano (1984), and de Rus (1990) reject a Cobb-Douglas technolgoy in long-run analyses.

**TABLE 2 Test Statistics for Homotheticity and Cobb-Douglas Production Technologies**

| Null hypothesis | $-2(lnL_R - lnL_U)$[a] | # Restrictions ($n$) | $\chi^2_{.01}(n)$ | Result |
|---|---|---|---|---|
| **Model with no correction for autocorrelation** | | | | |
| Cobb-Douglas | 192.22 | 10 | 23.21 | Rejected |
| Homotheticity | 6.26 | 2 | 9.21 | Not rejected |
| **Model with correction for autocorrelation** | | | | |
| Cobb-Douglas | 227.54 | 10 | 23.21 | Rejected |
| Homotheticity | 67.73 | 2 | 9.21 | Rejected |

[a] $ln\ L_U$ is the sample maximized log-likelihood value without restrictions, and $lnL_R$ is the sample maximized log-likelihood with restrictions.

ed model rejects the hypothesis, implying that the cost function is not separable in output and that changes in a factor's price will not only affect input demand ratios but also the cost elasticity with respect to output.[16]

[16] Berechman and Giuliano (1984), de Borger (1984), Berechman (1987), and de Rus (1990) also found a non-homothetic production structure. However, and similar to the model without correction for autocorrelation, Williams and Dalal (1981), Williams and Hall (1981), and Berechman (1983) could not reject the null hypothesis of a homothetic production structure.

Table 3 presents the own price ($e_{ii}$) and Allen elasticities of substitution ($\sigma_{ij}$) evaluated at the sample mean and by year for both models. As expected, the own price elasticities have the correct negative sign. In both models and consistent with other studies, the elasticities are small for labor and maintenance, while fuel appears to be more elastic when the estimated model corrects for autocorrelation. It is interesting to note that while the own price elasticity of labor and maintenance has remained fairly constant over the period, the demand for fuel has demonstrated a tendency to become more inelastic with the passage of time.

**TABLE 3 Short-Run Own Factor Demand Elasticities and Elasticities of Substitution**

| | | | **Price elasticity**[a] | | | |
|---|---|---|---|---|---|---|
| Period | $e_{ll}$ | $e_{ff}$ | $e_{mm}$ | $\sigma_{lf}$[b] | $\sigma_{lm}$ | $\sigma_{mf}$ |
| **Model without correction for autocorrelation** | | | | | | |
| Mean | –0.081 | –0.267 | –0.103 | 0.515 | 0.178 | –0.224 |
| 1991 | –0.074 | –0.352 | –0.078 | 0.591 | 0.131 | –0.133 |
| 1992 | –0.073 | –0.291 | –0.077 | 0.541 | 0.136 | –0.254 |
| 1993 | –0.084 | –0.275 | –0.098 | 0.514 | 0.173 | –0.196 |
| 1994 | –0.075 | –0.198 | –0.093 | 0.468 | 0.172 | –0.382 |
| 1995 | –0.091 | –0.149 | –0.124 | 0.407 | 0.222 | –0.337 |
| **Model with correction for autocorrelation** | | | | | | |
| Mean | –0.092 | –0.354 | –0.053 | 0.854 | 0.141 | –0.068 |
| 1991 | –0.095 | –0.436 | –0.048 | 0.877 | 0.119 | –0.046 |
| 1992 | –0.087 | –0.389 | –0.039 | 0.865 | 0.119 | –0.065 |
| 1993 | –0.094 | –0.373 | –0.051 | 0.858 | 0.139 | –0.062 |
| 1994 | –0.087 | –0.308 | –0.051 | 0.844 | 0.144 | –0.085 |
| 1995 | –0.099 | –0.266 | –0.076 | 0.828 | 0.185 | –0.084 |

[a] $e$: own price elasticity of demand, $\sigma$: Allen partial elasticity of substitution; where: $l$ = labor, $f$ = fuel, $m$ = maintenance.
[b] Elasticities of substitution are symmetric.

The Allen elasticity results indicate that labor and maintenance and labor and fuel are substitutes for each other, while maintenance and fuel are complements. Further, there is a significant change in the Allen elasticity between maintenance and fuel when the model is corrected for autocorrelation, rising from –0.224 (evaluated at the mean) in the uncorrected model to –0.046 in the corrected model. The latter indicates that maintenance and fuel are weakly complementary.

## DISCUSSION AND CONCLUDING COMMENTS

Although the deficiencies of models that suffer from uncorrected multiple serial correlation are recognized in the public transit cost literature, existing translog cost results have failed to account for serial correlation. Policy implications based on such models are accordingly suspect. When serial correlation is ignored, the translog cost function results for a medium-sized public transit system are consistent with existing literature. However, when corrected for the presence of serial correlation, the estimation results have significantly different implications with respect to a transit system's production technology and possibly its optimal fleet size.

This raises three questions that await further research. First, what are the sources of serial correlation? Consistent with other monthly time series data, the data for this analysis exhibit seasonal and cyclical effects that will generate autocorrelated errors. By testing for various orders of monthly related serial correlation and ultimately adjusting for first- and second-order serial correlation in the cost function and first-, second-, and sixth-order serial correlation in share equations, we find that the models presented likely capture much of the seasonally generated autocorrelation. In particular, the sixth-order correlation coincides with the seasonality of demand for bus services, which is lower during the summer and higher during the winter months (Berechman and Giuliano 1984).

An additional source of autocorrelation is missing information, that is, omitted variables that are serially correlated. The lack of data on network changes, for example, over the five-year period shows up in the error terms. If there were systematic changes in network size (e.g., continual increases) over the 60-month period of this analysis, errors that are serially correlated in the model would be generated.[17]

A final source of serially correlated errors could reflect an adaptive expectations framework in which changes in short-run costs are based on changes in the expected or desired level of the explanatory variables. This produces an empirical specification that is formally identical to a Koyck distributed lag model, which is characterized by serial correlation (Pindyck and Rubinfeld 1976).[18]

Second, we noted above that these findings may have implications for a system's optimal fleet size. Recall the findings reported in table 1 for the model that was not corrected for serial correlation. The coefficient on number of buses, the fixed factor, was positive and significant at a 0.10 level. This finding is consistent with other analyses of public transit costs and suggests that the transit system is overcapitalized. On the other hand, when corrected for serial correlation, the coefficient on number of buses was positive but with a t-statistic well below 1.0, implying that Indianapolis' system is operating efficiently. To illustrate the potential importance of correctly specifying the model, the mean fleet size for the sample period was 227 buses. When uncorrected for serial correlation, optimal fleet size was estimated as 171 buses, indicating substantial (35%) overcapitalization. However, when corrected for serial correlation, optimal fleet size increased to 214 buses, implying considerably less (6%) overcapitalization in the system.[19] The importance of this finding is its implication of

---

[17] To obtain desirable properties of the estimates by adjusting for serial correlation in this instance assumes that the omitted variables and included variables are uncorrelated (Maddala 1977).

[18] The authors would like to thank an anonymous referee for suggesting this.

[19] To obtain the optimal level of the fixed factor, we solved the following equation

$$\frac{\partial CV}{\partial x_F^*} = -p_F$$

where the subscript $F$ refers to the fixed factor (optimal number of buses), $CV$ is short-run total cost, $p_F$ is unit price of the fixed factor, and $x_F$ is the optimal level of the fixed factor. Unfortunately, it is impossible to find a closed form solution for $x_F$, since the equation yields an equation involving both $x_F$ and its logarithm. Using the modified Nelson (1972) approach (Berechman and Giuliano 1984), we calculated a price for $p_F$ and solved for the optimal level of rolling stock through a numerical procedure.

the need for further research to determine more accurately the extent of overcapitalization of current systems.

Last, the vehicle-miles supplied is the output measure used for this analysis and may be overly restrictive. A more general approach would entail a multiple output (e.g., passenger-miles as well as vehicle-miles) translog cost model enabling the estimation of short-run economies of scope as well as economies of capital stock utilization.

Public policy toward deregulation, privatization, and subsidization is only useful when we have accurate information on the underlying structure of transit firms' production and costs. The significant differences identified in this paper imply that ignoring statistical problems such as multiple serial correlation in translog cost function models may have important consequences for public transit policy.

## REFERENCES


Applebaum, E. and J. Berechman. 1991. Demand Conditions, Regulation and the Measurement of Productivity. *Journal of Econometrics* 47:379–400.

Berechman, J. 1987. Cost Structure and Production Technology in Transit. *Regional Science and Urban Economics* 17:519–534.

Berechman, J. 1993. *Public Transit Economics and Deregulation Policy.* North-Holland, Amsterdam.

Berechman, J., and G. Giuliano. 1984. Analysis of the Cost Structure of an Urban Bus Transit Property. *Transportation Research B* 18B, no. 4/5:277–287.

Berndt, E.R. 1991. *The Practice of Econometrics:Classic and Contemporary.* Reading, MA:Addison-Wesley.

Berndt, E.R. and N.E. Savin. 1975. Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances. *Econometrica* 43:937–958.

Bly, P.H. and R.H. Oldfield. 1986. The Effects of Public Transport Subsidies on Demand and Supply. *Transportation Research* 20A, no. 8:415–427.

Braeutigam, R.R., A.F. Daughety, and M.A. Turnquist. 1984. A Firm Specific Analysis of Economies of Density in the U.S. Railroad Industry. *The Journal of Industrial Economics* 33, no. 1:3–17.

Cervero, R. 1984. Cost and Performance Impacts of Transit Subsidy Programs. *Transportation Research A* 18A, no. 5/6:407–413.

Colburn, C. and W. Talley. 1992. A Firm Specific Analysis of Economies of Size in the U.S. Urban Multiservice Transit Industry. *Transportation Research B* 26B, no. 3:195–206.

De Borger. 1984. Cost and Productivity in Regional Bus Transportation:The Belgium Case Study. *Journal of Industrial Economics* 37:35–54.

De Rus, G. 1990. Public Transport Demand in Spain. *Journal of Transport Economics and Policy*, May:189–201.

Frankena, M. 1987. Capital-Based Subsidies, Bureaucratic Monitoring, and Bus Scrapping. *Journal of Urban Economics* 21:180–193.

Granger, C.W.J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37:424–438.

Kim, M. and M. Spiegel. 1987. The Effects of Lump-Sum Subsidies on the Structure of Production and Productivity in Regulated Industries. *Journal of Public Economics* 34: 105–119.

Maddala, G. 1977. *Econometrics.* New York, NY:McGraw-Hill Book Co.

Nelson, G. 1972. An Econometric Model of Urban Transit Operations, Institute of Defense Analyses. Washington, DC, 863.

Obeng, K. 1984. The Economics of Bus Transit Operation. *The Logistics and Transportation Review* 20, no. 1:45–65.

____. 1985. Bus Cost, Productivity and Factor Substitution. *Journal of Transport Economics and Policy* 19, no. 2:183–203.

Pindyck, R. and D. Rubinfeld. 1976. *Econometric Models and Economic Forecasts.* New York, NY:McGraw-Hill Book Co.

Pucher, J. 1995. Urban Passenger Transport in the United States and Europe:A Comparative Analysis of Public Policies—Part 2:Public Transport, Overall Comparisons and Recommendations. *Transport Reviews* 15, no. 3:211–227, 401.

Savage. 1997. Scale Economies in United States Rail Transit Systems. *Transportation Research* 31A:459–478.

Viton, P.A. 1981. On Competition and Product Differentiation in Urban Transportation:The San Francisco Bay Area. *Bell Journal of Economics* 12:362–379.

Williams, M. and A. Dalal. 1981. Estimation of the Elasticity of Factor Substitution in Urban Bus Transportation:A Cost Function Approach. *Journal of Regional Science* 21, no. 2:263–275.

Williams, M. and C. Hall. 1981. Returns to Scale in the United States Intercity Bus Industry. *Regional Science and Urban Economics* 11:573–584.