Examining the Inefficiency of Transit Systems Using Latent Class Stochastic Frontier Models

by Ryoichi Sakano and Kofi Obeng

This paper estimates a single class stochastic cost frontier model that accounts for heterogeneity by including background variables and a latent class model of the same specification. It is found that both the two-sided random errors and one-sided errors (cost inefficiency) are substantially smaller in the latter model than it is in the former, suggesting possible bias in the estimates of inefficiency from the single class model. Further, 58.9%–68.39% of the calculated inefficiencies are due to differences in technology captured by the latent classes. The paper concludes that using background variables only to capture heterogeneity may exaggerate measured inefficiencies in transit systems and suggests the latent class approach as a solution.

INTRODUCTION

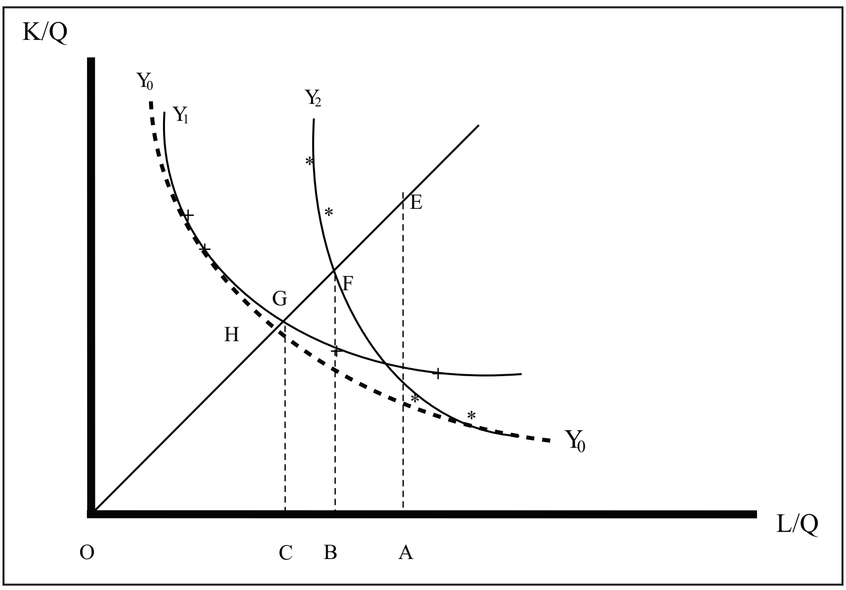
Stochastic frontier models have been used to empirically examine inefficiencies in firms. These models have been applied to cross-sectional data to compare inefficiency in various public transportation studies (De Borger et al. 2002). The earlier models assumed a homogeneous technology for all firms and measured inefficiency as a deviation from the common production or cost frontier. As a result, any heterogeneity in technology among firms that was not taken into account in the estimation was mistakenly included in the measurement of inefficiency, and this resulted in its over-estimation. Because the technology and operating environments of public transit firms vary, this possibly led to incorrect rankings of firms in terms of their inefficiency by skewing "the results towards the most influential group … and introducing size-related bias into the results" (Karlaftis and McCarthy 2002). This heterogeneity bias may lead to erroneous choice of the best practice frontier upon which inefficiency calculations are based and may explain why many studies give diverse estimates of public transit inefficiency.

Additionally, using a common technology frontier assumes firms are alike. If this assumption is violated and firms cluster into classes, then a firm may appear inefficient relative to the common frontier when in fact it would have been relatively less inefficient had the comparison been based upon the frontier of its class. Figure 1 illustrates this heterogeneity problem. In it, the unit isoquant Y_0 is the common frontier and it bounds all points from below. One group of firms denoted by the plus signs clusters on and above the frontier Y_1 and may be characterized as firms using labor-saving technology with relatively high capital productivity. These firms may be using more capital and less labor to produce each unit of output, have extensive networks, operate in high density areas, and operate vehicles of more recent vintage and large capacity. The others denoted by the asterisks may be characterized as firms using capital-saving technology with relatively high labor productivity. They may use more labor and less capital to produce each unit of output, have outdated technology and networks in less dense corridors, and operate smaller, older vehicles. These latter firms cluster on and above the frontier, Y_2 , which adequately describes their technology. If firm "E" is evaluated relative to the common frontier Y_0 , its technical efficiency in terms of inputs required to produce a given level of output will be OH/OE. Alternatively, if it is evaluated relative to the frontier Y_2 for the firms in its class, its technical efficiency is OF/OE. Very clearly, in this example, using a common frontier to evaluate the technical efficiency of heterogeneous firms exaggerates the results. This bias is greater for those firms on the left upper part of the frontier Y_2 than on the right bottom part of frontier Y_2 because the frontier Y_2 is closer to the common frontier Y_0 at the right bottom part. On the other hand, efficiency measures of firms on frontier Y_1 are affected differently. This suggests comparing transit systems based upon the frontier for their class.

To address this heterogeneity problem, some studies that are based upon nonparametric efficiency measures such as data envelopment analysis, which does not involve estimating parameters, use a two-step process. First, they calculate efficiency and second, they use Tobit or similar regression to relate the efficiency so calculated to firm characteristics that include measures of heterogeneity (e.g., Boame 2004). Their main problem is that the calculated efficiency measure has inherent bias because it is based on a common frontier, which does not adequately represent the technology of all firms, and this bias is carried forth to the next stage of regression analysis. Others, particularly those using data envelopment analysis (DEA), add observable heterogeneity factors as inputs since, with a large sample size, the inclusion of these factors will not affect the calculation of technical efficiency unless these factors are perfectly correlated with an existing variable (Nunamaker 1985). The rationale in such studies is that with heterogeneity, considered differences in technology are presumably removed and all the observations will cluster along a common frontier such as Y_0. However, the inclusion of heterogeneity factors still forces all firms to be evaluated with one frontier.

In parametric efficiency or cost studies involving panel data, heterogeneity is addressed in a number of ways. One clusters firms by some characteristics or heterogeneity measures, and after clustering, estimates efficiency for the firms in each cluster (Karlaftis and McCarthy 2002). The clustering approach is very similar to what we have shown in Figure 1 and recognizes that a different technology underlies each cluster. Another uses the fixed or random effects models developed by Schmidt and Sickles (1984) and Pitt and Lee (1981) and the "true" fixed and random effects models proposed by Greene (2005a) to deal with heterogeneity.[1] Others use the Battese and Coelli (1992, 1995) one-step method to estimate efficiency and account for heterogeneity (e.g., Jha and Singh 2001, Dalen and Gomez-Lobo 2003). Still, as in the non-parametric case, others include heterogeneity variables in the frontier specification before estimation.

Figure 1: Firm Heterogeneity and Efficiency



A common strand in the parametric and non-parametric studies is their ad hoc assumption of heterogeneity in technology and the inclusion of a few heterogeneity factors in the frontier. Another, found in the parametric studies, is their assumption that the errors caused by unobservable factors are either time-invariant heterogeneity or time-invariant inefficiency. As Greene (2004) points out, the main problem is that some unobservable factors causing heterogeneity may be counted mistakenly as inefficiency. To solve this problem, Orea and Kumbhakar (2004) suggest a latent class stochastic frontier model which assumes that observations can be sorted into groups or classes, each with a varying technology. This model is flexible enough to include time-varying and time-invariant sources of heterogeneity, and its results provide a more accurate measure of inefficiency than the others. A random parameter stochastic frontier model can also be used, and it is in fact more efficient in estimation than the latent class frontier model (Tsionas 2002). However, it requires distributional assumptions about the parameters, and this makes the latent class model preferable in this paper.[2]

Therefore, the objective of this paper is to argue for a latent class approach in studies of public transit efficiency and to show that using background variables, which describe the characteristics of the transit system and its operating environment to account for heterogeneity, is not enough. The rationale for this objective is the variety of transit systems that we have. Considering bus transit systems, for example, it is possible to distinguish among small, medium, and large systems, each using different types of vehicles (small, medium, and large with various seating configurations) to provide services in heterogeneous environments. For such transit systems, the choice of a single reference technology in efficiency studies may be inappropriate and is a limitation. Recognizing this limitation, this paper estimates 1) a latent class stochastic frontier, and 2) an index frontier that assumes a homogeneous technology for a panel of U.S. urban transit systems. It uses the results from these frontiers to calculate firm-specific cost inefficiency and identify its sources. While the index frontier accounts for heterogeneity by including background variables, the former model does so by including background variables and estimating latent class frontiers. The results show that just including background variables in public transit cost and efficiency studies does not completely account for heterogeneity, and that a substantial portion of the reported inefficiency may be due to heterogeneity. Specifically, a very substantial portion of heterogeneity can be captured by simultaneous inclusion of background variables and estimating latent class models with the results that the inefficiencies are very small.

The benefit of this latter approach is that it provides better estimates of efficiency since it factors out many of those distortions that could result from heterogeneity. Further, it provides information to compare transit systems to their peers based upon similarity in technology. Such peer comparison is common practice in public transit and used by the Federal Transit Administration (FTA) to compare transit systems in its review of their performance for funding purposes. Lastly, estimates of efficiency based upon the latter approach allow transit system management and policy makers to design and target specific programs to improve performance. As far as we know, no study has been done on measuring changes in inefficiency in transit systems over time and across systems that accounts for heterogeneity using both background variables and latent class in cost frontiers together. The results will give some insight into an ongoing debate on the effectiveness of government subsidies in transit systems.

The rest of the paper is divided into seven parts. The literature review is presented first and it is followed by the methodology, data, estimation and results. Next, the latent classes are described in a separate section and the cost inefficiencies in another followed by the results of the homogeneous frontier. The last part of the paper is the conclusion, where the major findings are presented.

LITERATURE REVIEW

Aigner et al. (1977) and later Jondrow et al. (1982) proposed stochastic frontier models with errors composed of a normally distributed error term and a half-normal inefficiency term. That is, the inefficiency term is distributed to the left of the normal curve in the case of a production frontier showing that output is below the frontier, and it is distributed to the right of the normal curve, as in the case of the cost frontier to show that cost exceeds the minimum cost frontier. Since then, many extensions have been made to these models to distinguish between heterogeneity across firms and inefficiency and to properly measure them. For example, Pitt and Lee (1981) introduced the random effects model in which a firm-specific time invariant random effect is considered as inefficiency, and all time-varying and idiosyncratic errors are assumed to be changes over time and heterogeneity. An alternative to the random effects model is the fixed effects stochastic frontier model estimated by Schmidt and Sickles (1984). In this model, each firm's inefficiency is measured as the difference between its fixed and time-invariant constant and that of the most efficient firm. Greene (2005a) critiques this model noting that it is not appropriate in long panels because there may be technological changes that may affect inefficiency and heterogeneity. Even worse, he argues, it is inapplicable where firms face rapid changes in their operating environments. Also, it considers all time-invariant errors as inefficiency, including some time invariant unobserved heterogeneity. To account for these shortcomings, Greene (2005a) proposed the so-called true-fixed and true-random effects models, which differ from the previous fixed and random effects models in two respects. First, the true-fixed and random effects models assume that all fixed or random effects components in a frontier represent heterogeneity in technology, and second, that inefficiency varies with time. Although true fixed and random effects models separate "time invariant" heterogeneity and "time varying" inefficiency, all time-invariant components observed as deviations from the frontier are assumed to be heterogeneity. Further, they assume that any time-varying component represents inefficiency. In general, both heterogeneity and inefficiency may have time-invariant and time-varying components. As a result, Greene (2004) states that fixed effects stochastic frontier models could over or underestimate inefficiency and that the blending of inefficiency and heterogeneity is a modeling problem without a readily available satisfactory solution.

Heterogeneity may also be handled by including it in the inefficiency term. Here, the half-normal error term representing inefficiency is modeled in a functional form so that its mean can vary across firms and over time. This approach is exemplified in the works of Cornwell et al. (1990), Kumbhakar (1990), and Lee and Schmidt (1993) among others. Cornwell et al. (1990) include time and time-squared in their models, an exponent of time and time-squared are found in Kumbhakar (1990), and a set of time dummy variables are found in Lee and Schmidt (1993). Battese and Coelli (1992, 1995), on the other hand, propose a model whose inefficiency term has two components, one a function of time and the other a basic half-normal distribution. This model is generalized in Orea and Kumbhakar (2004) as the product of time-invariant firm effect and any firm-specific variable which may change over time beside the time trend. In particular, Orea and Kumbhakar (2004) considered dummy variables representing a particular type of bank (saving bank) and acquisition of other financial institutions, as well as a time trend. Jha and Singh (2001) used load factor, fleet utilization, network size per bus, and their interactions with the variables in the cost frontier (input price, output, and network size) as well as time trend and squared time trend in their parameterization of inefficiency. Dalen and Gomez-Lobo (2003) included dummy variables in their inefficiency term to reflect yardstick contracts where subsidies are calculated using cost-based standards, subsidy-cap contracts where companies and the government agree on the reductions in subsidies per year, as well as years elapsed on each contract. Picacenza's (2006) parameterization of inefficiency includes a dummy variable to distinguish between types of regulation, average commercial speed as a proxy for network characteristics, and a time trend as well as their interaction terms. More recently, Ahn et al. (2007) used factor analysis to determine which firm-specific time-varying factors to incorporate in their specification of the inefficiency term.

Since heterogeneity comes from differences in the uses of technology and environmental variables, prior clustering of firms followed by estimations of frontiers for the clusters has been used (Kolari and Zardkoohi 1995). Unfortunately, this two-step method may not fully account for heterogeneity among firms because it does not consider all the sources of heterogeneity in its classification. The latent class models of Orea and Kumbhakar (2004) and Greene (2005b) solve this problem and do not involve prior groupings of firms, but simultaneously identify classes based on statistical likelihood and estimate a frontier for each class jointly with the probability of each firm belonging to a particular class. These methods allow technology to vary so that the frontier of each class has a different location as well as a different shape. This may be crucial in industries where firms utilize diverse technologies to reflect differences in their operating environments (e.g., network, regulation, demand). The latent class frontiers have seen limited applications because they are complex to estimate and, until recently, were not readily available in econometric software programs in the public domain.

In transportation, many earlier studies of efficiency estimated stochastic production frontiers by assuming homogeneous technology and did not use the latent class approach. For example, Sakano and Obeng (1995), Sakano et al. (1997), and Obeng et al. (1997) studied inefficiency in urban transit systems using cross-sectional data and examined how federal government subsidies contributed to allocative and technical inefficiencies. They found a large variation in technical inefficiency among transit systems that they attributed to system size and other characteristics. This finding could indicate heterogeneous technology and if so the estimated inefficiencies of some transit systems may be incorrect. Farsi et al. (2006) using transportation data estimated and compared various stochastic frontier models without accounting for the panel nature of their data and did not use the one-step approach. As expected from Greene's (2004, 2005a, 2005b) seminal works, Farsi et al. (2006) found that the estimated technical efficiency was much smaller in the true random effects model than it was in any other model. They attributed this to the assumption in the true random effects model that all-time invariant variation across firms is due to heterogeneity rather than inefficiency. More recently, a number of transportation-related studies have applied the one-step estimation method. Jha and Singh (2001) used it to study a longitudinal panel of nine bus operators in India. Dalen and Gomez-Lobo (2003) applied it to a panel data of Norwegian bus systems, and Picacenza (2006) used it to study a longitudinal panel data of 44 companies.

These reviewed studies show that the direction in stochastic frontier models is to account for heterogeneity and that several models can be used for that purpose. This is important if we are to obtain good information on efficiency. Greene's (2005b) comparison of these models, however, shows they give a wide range of inefficiency estimates. However, in industries where possibilities exist for different firms to use different technologies, as it is in public transit systems, the reviewed literature suggests using a latent class frontier model. Therefore, in the following sections we specify the latent class model used in this study.

METHODOLOGY

Consider a public transit system that uses labor (L), capital (K) and fuel (F) to produce vehicle miles of services (Q). This system faces competitive input markets and pays labor a wage rate of w_L per hour and fuel price of w_F per gallon, and incurs a capital user cost of w_K per vehicle. If the firm minimizes cost subject to a production function constraint then its long run total cost function TC_{it} is $TC_{it} = g(w_{Lit}, w_{Kit}, w_{Fit}, Q_{it}, x_{it}, t)$, where $i = 1,2,3,\ldots, I$ is the number of transit systems, x_{it} is the set of background variables including network size (R), operating subsidies (S_0) and capital subsidies (S_k) all of which may be considered heterogeneity variables, and t is a time variable representing shifts in the cost function over time.[3] For technical improvements this shift is negative, i.e., $\partial TC / \partial t < 0$ showing that the same level of output can be produced cheaply over time. Using the price of labor as the numerator this cost function can be written as,

$$(1)\quad TC_{it} / w_{Lit} = g(w_{Fit} / w_{Lit}, w_{Kit} / w_{Lit}, Q_{it}, x_{it}, t) \exp\{\varepsilon_{it}\}$$

The exponential term is the error and all other terms are already defined. The translog expansion of this function gives the following total cost equation for each transit system (i) in any given period.

$$(2)\quad \ln(TC^*_{it}) = \alpha_0 + \sum_{j=1}^{J} \alpha_j \ln(w^*_{jit}) + \alpha_x (\ln x_{it}) + \sum_{j=1}^{J} \alpha_{jQ} (\ln w^*_{jit})(\ln Q_{it}) + \alpha_t t + 0.5 \sum_{j=1}^{J} \sum_{n=1}^{N} \alpha_{jk} (\ln w^*_{jit})(\ln w^*_{nit}) +$$
$$0.5\alpha_{xx}(\ln x_{it})^2 + \sum_{j=1}^{J} \alpha_{jx} (\ln x_{it})\left(\ln w^*_{jit}\right) + \alpha_{xQ}(\ln x_{it})(\ln Q_{it}) + \alpha_{tx}(\ln x_{it})t +$$
$$\sum_{j=1}^{J} \alpha_{jQ} (\ln w^*_{jit})(\ln Q_{it}) + \sum_{j}^{J} \alpha_{jt} (\ln w^*_{jit})t + 0.5\alpha_{QQ}(\ln Q_{it})^2 + 0.5\alpha_{tt}t^2 + \varepsilon_{it}$$

In equation (2) $w_{jit}^{*} = w_{jit} / w_{Lit}$, $TC_{it}^{*} = TC_{it} / w_{Lit}$, where $j = n = F,K$ and ε_{it} is distributed *i.i.d* without accounting for inefficiency. Among the characteristics that a cost function must exhibit is homogeneity of degree one in input prices. That is, the sum of the first order input price coefficients in the cost functions is one. This characteristic is satisfied in equation (2).

In stochastic frontier cost models, the random error term in (2) has two components. The first is a normally distributed random error term v_{it}, and the second is an inefficiency term, u_{it} distributed half normal as in equation (3) whose value is positive and shows the amount by which actual cost exceeds minimum cost. Both u_{it} and v_{it} are further assumed to be uncorrelated with the other terms in equation (2). Thus,

$$(3) \quad \left.\begin{array}{l} \varepsilon_{it} = v_{it} + u_{it} \\ v_{it} \sim N\left[0, \sigma_v^2\right] \\ u_{it} \sim N^{+}\left[0, \sigma_u^2\right] \end{array}\right\}$$

Recently, several variations of the composed error in equation (3) have been specified in the economics literature to further address heterogeneity issues. One main characteristic of these variations is their inclusion of a firm-specific constant term α_i in equation (2). In fixed effects stochastic frontier models, the inefficiency term u_{it} is assumed to be a set of time-invariant constants, u_i for different firms and it is embodied in the firm-specific time-invariant constant term α_i as, $\alpha_i = u_i + \alpha_0$. Thus, α_0 in equation (2) is replaced by α_i and ε_{it} by v_{it}. With panel data, if there is sufficient "within-firm" variation, the resulting equation can be estimated by least squares. From the estimated coefficients, the inefficiency of each firm is calculated as the difference between that firm's estimated fixed constant and the estimated minimum firm-specific constant from the results, i.e., $\alpha_i - \min(\alpha_i)$. Thus, inefficiency is assessed relative to the least inefficient firm. The problems with this measure, as noted earlier and by Greene (2005), are that it includes firm-level heterogeneity and the influences of omitted variables normally captured by regression constants, and it assumes inefficiency does not vary with time, which may not be true for long panels. These problems may exaggerate the sizes of the inefficiencies calculated for different firms.

To resolve them, the composed error structure in equation (3) can be substituted into equation (2), and then α_i replaced by positive firm-specific constants, θ_i such that,

$$(4) \quad \varepsilon_{it} = \theta_i + v_{it} + u_{it}$$

The resulting equation, called the true fixed-effects model, is then estimated by maximum likelihood methods. The difference between this specification and the previous fixed-effects model is the constant term, which captures time-invariant heterogeneity and omitted variables, not inefficiency. Additionally, inefficiency is now different from heterogeneity; it is time-variant and can be made a function of firm characteristics.

Alternatively, instead of fixed effects models, random effects models can be estimated. In these latter models, the composed error structure in equation (3) is rewritten as,

$$(5) \quad \begin{cases} \varepsilon_{it} = v_{it} + u_{it} + \theta_i \\ \theta_i \sim N^{+}\left[0, \sigma_\theta^2\right] \end{cases}$$

Where, θ_i is now a time invariant random firm-specific coefficient and it is distributed half normal and captures unobserved heterogeneity. In true random effects models θ_i is normally distributed. Instead of being a constant in other applications θ_i is sometimes modeled explicitly as a function of time-invariant firm characteristics.

In all these models, firms are represented by a single technology frontier and the firm-specific constants, whether fixed or random, show individual shifts in the frontier that make each firm different. Lately, this assumption of a single technology has come under scrutiny. It has been argued that firm-level heterogeneity could lead to different technologies. If so, then as argued earlier in this paper, the common frontier may not be the true technology and it may be inappropriate to evaluate inefficiency relative to this frontier. Consequently, a latent class model must be used. To illustrate, assume there are $m = 1,2,3,\ldots, M$ latent classes or groups into which the $i = 1,2,3,\ldots, I$ firms can be intrinsically sorted. Then, equation (2) can be rewritten for each transit system as,

$$(6)\quad \ln(TC_{it}^{*}) = \ln g(Q_{it}, w_{Fit}^{*}, w_{kit}^{*}, x_{it}, t, \alpha_m) + u_{it}\big|_m + v_{it}\big|_m$$

Where, following Greene (2007) $\alpha_m = \alpha + \delta_m$ and δ_m is how much a coefficient in latent class m is different from its mean value. Thus, the subscript m shows there is a different coefficient vector for each of the m latent classes of transit systems, and $u_{it|m}$ and $v_{it|m}$ are respectively the inefficiency and random error terms given the class of each firm. Because the latent class model groups firms by measures of similarity it accounts for heterogeneity.

Using these characteristics, the probability of firm I being in a latent class m is P_{im} and it is given by the expression below,

$$(7)\quad P_{im}(\delta_m) = F\left(\ln TC_{it}, \ln\left(Q_{it}, w_{fit}^{*}, w_{kit}^{*}, x_{it}, t, \alpha\right) + \ln\left(Q_{it}, w_{fit}^{*}, w_{kit}^{*}, x_{it}, t, \delta_m\right)\right)$$

Where, $F(\bullet)$ may be fixed for each class as done in this study or it may be parameterized as a multinomial logit equation. This fixed parameter is estimated jointly with (6) in one step.[4]

Using equation (7), and again following Orea and Kumbhakar (2004), the overall likelihood function (LF) for the latent class model can be written as,

$$(8)\quad \ln LF(\alpha, \phi) = \sum_{i=1}^{I} \ln\left\{\sum_{m=1}^{M} LF_{im}(\alpha_m) P_{im}(\phi_m)\right\}$$

Where, $LF_{im}(\alpha_m)$ is each firm's contribution to the likelihood function and $P_{im}(\phi_m)$ is the probability of being in group m. The estimated parameters from equation (7) can be used to calculate conditional posterior probabilities of each firm belonging to a class as,

$$(9)\quad P(m|i) = LF_{im}(\alpha_m) P_{im}(\phi_m) / \sum_{m=1}^{M} LF_{im}(\alpha_m) P_{im}(\phi_m)$$

The latent class approach to the heterogeneity problem leaves the question of how inefficiency is to be calculated for each firm. Because there is a probability of each firm being in each class, its mean inefficiency is calculated as the sum over all classes of its posterior probability of being in a class $P(m|i)$ times the logarithm of its inefficiency score in that class (E_{im}) (Greene 2002). Thus, the logarithm of firm inefficiency (E_i) is,

$$(10)\quad \ln(E_i) = \sum_{m=1}^{M} P(m|i) \ln(E_{im})$$

DATA

The data for this paper are a random sample of single mode bus transit systems drawn from the U.S. National Transit Database (i.e., NTD Database), which consists of 463 observations for an unbalanced panel of 24 single-mode bus transit systems for the period 1985 to 2004. The data include operating cost, total annual vehicle miles of service used as the output measure, total annual hours worked by labor, gallons of fuel used as a proxy for all non-labor and non-capital inputs, total capital subsidies, total operating subsidies, fleet age, fleet size and transit background data. Labor price is total labor compensation including benefits divided by hours worked.[5] Ideally, fuel price is the total expenditure on fuel divided by gallons of fuel. Since we use fuel as a proxy for all other non-labor inputs, total non-labor operating cost divided by total annual consumption of fuel gives the proxy price of non-labor inputs.[6]

Because we only have information on operating cost, total cost must be calculated. A number of approaches are found in the literature to calculate total costs of transportation systems. Farsi et al. (2007), Karlaftis and McCarthy (2002) and Friedlander and Chiang (1983) calculate capital cost as residual cost by subtracting the costs of labor and fuel from total cost. By dividing this cost by total number of seats offered by an operator's fleet or by fleet size, these authors obtained the price of capital. Friedlander et al. (1981) in estimating a cost function for trucking companies added a 12% opportunity cost for capital to operating cost to obtain total cost and defined capital price as capital service payments (opportunity cost, depreciation, and maintenance of capital items) per unit of capital, where the quantity of capital is carrier operating property net depreciation.[7] Then, they divided total cost into labor costs, fuel expenditures and fuel taxes, purchased transportation, and other expenditures where the latter represents capital cost. In Nadiri and Schankerman (1981), the product of service price of capital and net stock of capital was added to nominal expenditures on labor, materials, rents, supplies, and research to obtain total cost.

Using the latter approach, we define the service price of capital and multiply it by capital stock for each transit system to obtain capital cost and then add it to operating cost to get total cost. Fleet size is a proxy for capital stock and the user cost of each vehicle in any period t is w_K which is calculated using the service price formula $w_{Kt} = P_{Kt}(r_t + d)e^{-d\mu}$. P_{Kt} is the weighted average price of a new bus in a particular year, r_t is the prime rate, d is the straight line depreciation rate assuming bus useful life of 20 years, and μ is average fleet age. This capital price formula is very similar to that in Nadiri and Schankerman (1981) with tax parameters set to zero and a multiplicative factor $e^{-d\mu}$ appended. All costs, prices, and subsidies are in constant 1984 dollars using the consumer price index as a deflator and all variables except time and purchased transportation are in logarithms. Purchased transportation is a binary variable that takes a value of one if a firm purchases transportation from private sector sources, and a zero otherwise.

Table 1 presents summary statistics for the transit systems used in this study.[8] The average transit system operated 390 vehicles and used 2.007 million hours of labor and 3.80 million gallons of fuel to provide 13.843 million vehicle miles of service on 984.86 miles of routes. It paid labor a real wage rate of $15.30 per hour inclusive of benefits, incurred a real capital user cost of $13,420.75 per vehicle, and real total cost of $49.85 million. Additionally, it used vehicles that were on average 8.08 years old. And in real terms the average transit system received $31.14 million and $9.62 million, respectively, in operating and capital subsidies.

ESTIMATION AND RESULTS

Before estimating the equations, two modifications to the data are required. First, because the translog cost model is a Taylor series expansion the point of that expansion must be selected. In this paper, that point is the mean. Therefore, all variables except time are deviations from their mean. Second, because all variables except time are in logarithms, a method of handling zero values in them must be selected. This is the case of operating and capital subsidies, but more so with capital subsidies. Because transit systems do not receive these subsidies every year, zero values for them are relevant observations. Possible approaches to handle zero subsidies are to use the same methods of handling zero outputs in translog cost functions. These include using the Box-Cox transformation (e.g., Caves, Christensen, and Tretheway 1980), substituting very small positive values for zeros (Pulley and Humphrey 1993), using a cost function that is quadratic instead of log-quadratic in the independent variables (Baumol et al. 1982), specifying a composite cost function (Pulley and Braunstein 1992), or a zero-output translog function (Weninger 2003). Of these approaches, the most common exemplified in the work of Berger et al. (1999) and noted by Weninger (2003) is the substitution of small positive values for zeros and it is that used in this paper; a value of one dollar is substituted for a zero subsidy.

Table 1: Transit Systems and Descriptive Statistics

System Name	**System Name**
Anchorage Public Transit	Minneapolis MTC
Providence RI PTA	City of Detroit Department of Transportation
Hartford-Connecticut Transit	Honolulu DOT Service
Metropolitan Bus Authority	Santa Cruz MTD
East Meadow MSBA	Santa Monica Muni Bus
New York Bus Tours Inc.	Alameda-Contra Costa TD
Jackson Heights-Triboro Coach	Santa Barbara MTD
Academy Lines-Leonardo	Los Angele-SCRTD
Hudson Transit Lines-Mahwah	Phoenix Transit System
Suburban Transit Corp	Milwaukee County TS
NJTC-45	Cincinnati-SORTA
Hillsborough Area RTA	Toledo RTA
Descriptive Statistics	
Variable	**Mean**
Operating subsidies	$31,139,874.73
Capital subsidies	$9,620,567.23
Fleet age	8.08
Fuel price	$3.25
Labor wage/hour	$15.30
User price of capital	$13,420.75
Total cost	$49,849,957.03
Vehicle miles of service	13,842,945.49
Right-of-way miles	984.86
Fleet size (Number of buses)	390
Gallons of fuel	3,799,610.49
Labor hours	2,006,899.04

All costs and prices are constant 1984 dollars.

After these modifications, the estimation was performed using the latent class frontier method in LIMDEP (Greene 2007).[9] Because the number of coefficients increases considerably as the number of latent classes increases, it is also important to determine the correct number of classes to use. Following Fraley and Raftery (1998), Orea and Kumbhakar (2004), and Scheiner et al. (2006) the Akaike Information Criteria (AIC) is used to select the correct number of classes. The AIC is of the form, $AIC = -2(\log L - J) / n$ where $\log L$ is log-likelihood, J is the number of coefficients to be estimated, and n is sample size. The values of AIC for one and two class models are, respectively, 0.1206 and -0.7481 and favor a latent class model with two classes. Attempts to estimate a threeclass model failed to produce a solution, indicating it is over-specified (Orea and Kumbhakar 2004) or there is a high degree of within-class multicollinearity (Scheiner et al. 2006).

Columns 3 and 6 of Table 2 show the estimated coefficients of the two latent class models. Most of these coefficients are statistically significant at the commonly accepted probability level of 0.05. The last two rows of the table show that the estimated prior probabilities for membership in latent Class 1 and Class 2 are approximately 0.50, respectively, and these probabilities are highly significant statistically. Thus, half of the transit systems studied fall into latent Class 1 and the other half into latent Class 2. Additionally, the highest class probability of each transit system is one. These results show that the calculated cost inefficiencies are based upon the frontier of each class only. This is because for each transit system, the right-hand-side of Eq. (10) is the weighted sum of its inefficiencies from the class frontiers, where the weights are the probabilities of belonging to a class. In a two latent class system, as the results show, a probability of one for belonging to a class means all the inefficiency measures are based upon the frontier of that class only.

Besides these results, the first order coefficients are consistent with what economic theory predicts them to be. For instance, at the point of approximation, all the elasticities are positive, showing that cost increases with output and that the frontier is concave in input prices. Moreover, because each of the first order coefficients of the price ratios as well as their sum is less than one, the homogeneity of degree one restriction is satisfied in that the first order coefficients of the input prices sum to one. The estimated errors in the latent classes are, respectively, 0.2390 and 0.1460 and are highly significant. Also, the estimated values of the ratio of the two-sided random error to the one-sided error (λ) are 1.2528 and 1.4607, respectively, for latent classes one and two. Being statistically significant, both errors show that most of the unexplained variation in cost is due to two-sided random errors.

LATENT CLASSES

Table 2 also provides some useful information about each latent class. From the first order coefficients the latent Class 1 transit systems are characterized by pure technical growth ($-\partial \ln TC_{it} / \partial \ln t$) of 0.55% as well as a total cost increase of 0.9293% for a 1% increase in output. Test statistic comparing this to constant returns to scale shows no statistically significant difference between them (t statistic = -1.88). Therefore, there are near constant returns to scale in the latent Class 1 transit systems. Also, total cost increases by 0.0303% and 0.0213% for 1% increases in capital and operating subsidies, respectively. These results show that the impacts of operating and capital subsidies on the total cost of these transit systems are quite close. Further, they show that there are economies associated with these subsidies, possibly because they allow transit systems in this group to substitute more for less productive inputs.

Using the characteristics of the latent classes in Table 3, on the average, the latent Class 1 transit systems provide 9.60 million vehicle miles on 766.16 miles of routes at a speed of 16.49 miles per hour and real total cost of $35.04 million. Further, on the average, they pay labor a real wage of $14.75 per hour and buy fuel at $3.47 per gallon[10] while incurring annual capital user cost of $13,403.55 per vehicle. Of these transit systems, 51.54% received operating subsidy averaging about $21.25 million per system and covering more than half their total cost, and $3.51 million per year in capital subsidy. The very large average operating subsidy that some of the transit systems in this class receive could indicate that they may have dedicated tax sources to fund their services and that these sources generate a lot of revenue.

In comparison, column 6 of Table 2 shows that the latent Class 2 transit systems are characterized by pure technical growth of 5.22%. Their cost increases by 0.9830% for 1% increase in output. This increase is not statistically different from what would have been obtained under constant returns to scale (*t statistic = -0.52*). Therefore, just as in the latent Class 1 transit systems latent Class 2 transit systems operate at near constant returns to scale. Additionally, in the latent Class 2 transit systems, total cost increases by 0.0662% for a one percent increase in route miles. A 1% increase in capital subsidy increases total cost in these systems by 0.0473%, while operating subsidies have a statistically insignificant effect on total cost, according to the results in Table 2.

Table 2: Estimated Coefficients of Latent Class Model

Variable	Variable Description	Coefficient	Std. Err.	T-value	Coefficient	Std. Err.	T-value
		Latent Class 1			Latent Class 2		
Constant	Constant	-0.0660	0.0502	-1.3150	-0.2799*	0.0461	-6.0730
$\ln(w_K / w_L)$	Log of relative capital price	0.1335*	0.0583	2.2900	0.1024*	0.0449	2.2810
$\ln(w_F / w_L)$	Log of relative fuel price	0.3102*	0.0540	5.7420	0.4045*	0.0552	7.3310
$\ln(Q)$	Log of output	0.9293*	0.0377	24.6620	0.9830*	0.0331	29.6850
t	Time in years	-0.0055*	0.0025	-2.2080	-0.0522*	0.0036	-14.5750
$\ln(w_K / w_L)\ln(w_F / w_L)$	Log of relative price of capital times relative fuel price	0.3310*	0.1441	2.2980	0.0358	0.1225	0.2920
$0.5[\ln(w_K / w_L)]^2$	One half times the square of log of relative capital price	-0.0324	0.1354	-0.2390	-0.0963	0.2174	-0.4430
$[\ln(w_K / w_L)]\ln Q$	Log of relative price of capital times output	-0.1209	0.1014	-1.1930	-0.0940*	0.0418	-2.2460
$0.5[\ln(w_F / w_L)]^2$	One half times the square of log of relative fuel price	0.1725	0.1592	1.0830	0.2693*	0.1181	2.2810
$[\ln(w_F / w_L)]\ln Q$	Log of relative price of fuel times output	0.1058	0.0886	1.1940	0.1338*	0.0459	2.9140
$0.5[\ln(Q)]^2$	One half times the square of log of output	-0.4670*	0.0977	-4.7800	0.1259*	0.0446	2.8190
$0.5t^2$	One-half times the square of time	-0.0023*	0.0010	-2.4020	-0.0012**	0.0006	-1.8810
$t \ln(Q)$	Time times log of output	0.0233*	0.0040	5.7590	-0.0215*	0.0028	-7.7930
$t \ln(w_F / w_L)$	Time times log of relative fuel price	-0.0177*	0.0068	-2.5990	-0.0014	0.0054	-0.2680
$t \ln(w_K / w_L)$	Time times log of relative capital price	0.0297*	0.0080	3.7130	-0.0126	0.0082	-1.5330
$\ln(S_o)$	Log of operating subsidy	0.0303*	0.0050	6.0900	-0.0170	0.0171	-0.9930
$\ln(S_K)$	Log of capital subsidy	0.0213*	0.0043	5.0060	0.0473*	0.0054	8.8340
$\ln(Q)\ \ln(S_o)$	Log of output times log of operating subsidy	0.0272*	0.0075	3.6070	-0.0745*	0.0179	-4.1610
$[\ln(S_K)]\ln(Q)$		0.0474*	0.0072	6.5430	0.0245*	0.0081	3.0160
$\ln(w_F / w_L)\ \ln(S_o)$	Log of relative fuel price times log of operating subsidy	-0.0322*	0.0102	-3.1480	-0.0673**	0.0368	-1.8270
$\ln(w_F / w_L)\ \ln(S_K)$	Log of relative fuel price times log of capital subsidy	0.0392*	0.0075	5.2430	-0.0123	0.0111	-1.1110
$\ln(w_K / w_L)\ \ln(S_o)$	Log of relative capital price times log of operating subsidy	0.0068	0.0115	0.5950	-0.0172	0.0224	-0.7700
$\ln(w_K / w_L)\ \ln(S_o)$	Log of relative capital price times log of capital subsidy	0.0131*	0.0083	1.5830	0.0401*	0.0107	3.7560
$t \ln(S_o)$	Time times log of operating subsidy	-0.0011*	0.0006	-1.9650	0.0089*	0.0023	3.7860
$t \ln(S_K)$	Time times log of capital subsidy	-0.0019*	0.0005	-3.7290	0.0070*	0.0006	11.1930
$\ln(R)$	Log of network size	-0.0709*	0.0241	-2.9370	0.0662*	0.0208	3.1870
σ	Sigma	0.2390*	0.0273	8.7600	0.1460*	0.0157	9.3300
λ	Lambda	1.2528*	0.4813	2.6030	1.4607*	0.5034	2.9010
Estimated prior probability for Class 1 membership		0.5000*	0.1021	4.8990	0.5000*	0.1021	4.8990
Estimated prior probability for Class 2 membership		0.5000*	0.1021	4.8990	0.5000*	0.1021	4.8990

* Statistically significant at probability of less than 0.05.

Table 3: Characteristics of Latent Classes

Variable	Mean	Standard Deviation.	Minimum	Maximum
LATENT CLASS 1				
Wage rate ($)	14.7504	4.0335	7.6521	25.5398
Capital user cost ($)	13,404	2,566	1,753	21,997
Fuel price ($)	3.4655	1.2159	1.4902	8.9724
Total Cost ($)	35,049,511	32,831,536	6,673,830	128,967,585
Vehicle miles	9,602,925	6,919,388	2,354,600	31,209,800
Route miles	766	769	98	9675
Speed (miles per hour)	16.4858	7.6770	2.1599	57.2637
Average fleet age	7.8651	2.7619	1.8471	16.4600
Capital subsidy ($)	3,510,327	6,419,963	0	33,834,108
Operating Subsidy ($)	21,250,596	28,552,714	0	184,708,922
LATENT CLASS 2				
Wage rate ($)	15.8363	3.2263	7.1813	26.2245
Capital user cost ($)	13,437	2,352	7,356	19,400
Fuel price ($)	3.0410	1.0084	1.2529	7.4379
Total Cost ($)	64,085,980	97,970,495	4,606,234	481,075,315
Vehicle miles	17,921,269	22,063,642	2,236,500	108,215,400
Route miles	1,195	1,100	76	4,933
Speed (miles per hour)	13.6585	2.1155	1.3420	23.2887
Average fleet age	2.0524	0.3391	0.4081	2.8009
Operating subsidy	40,652,020	64,667,813	0	323,661,455
Capital Subsidy	15,497,790	35,424,757	0	268,344,912

Despite this apparent similarity in terms of the effects of subsidies in these transit systems, both latent classes are markedly different in other characteristics. For example, the latent Class 2 transit systems produce the largest output of 17.91 million vehicle miles on 1,195 miles of routes at a lower average speed of 13.66 miles per hour. In providing this service they pay a higher labor wage of $15.83 per hour, incur a yearly capital user cost of $13,437.28 per vehicle, and pay $3.04 per gallon of fuel, all of which result in a much larger total cost of $64.09 million. This high cost, on average, brings about a large amount of $40.65 million per year in operating subsidies, and capital subsidies of $15.50 million. These operating and capital subsidies are received by 97.88% and 91.53% of the transit systems in this class. In addition, these systems operate vehicles that, on average, are 2.05 years old, which is noticeably less than the 7.9 years of the latent Class 1 transit systems.

Very clearly, these characteristics show that the two latent classes are indeed very different. Most of the latent Class 2 transit systems are very large, expensive to operate, and heavily subsidized, compared with the latent Class 1 transit systems. But it cannot be said that the latent Class 2 transit systems consist only of large size transit systems only. This is because the results show that there are small and large transit systems in each class as evident in the wide range between the minimum and maximum values of total cost, vehicle miles, and route miles.

COST INEFFICIENCY ESTIMATES FOR INDIVIDUAL TRANSIT SYSTEMS

Besides the characteristics just described, information about cost inefficiencies in the latent classes can be gleaned from Tables 4 to 6. Table 4 shows cost inefficiency for individual transit systems each year based on the two latent class models. Table 5 shows that, overall, the average estimated cost inefficiency is 6.19%.[11] Among the two latent classes, however, the latent Class 1 transit systems have a higher average inefficiency of 7.49% compared with 4.94% for the latent Class 2 transit systems.[12] In addition, a variation in estimated average cost inefficiency of 1.21% to 25.11% is observed in the latent Class 2 transit systems. Comparatively, a variation in cost inefficiency ranging from 2.27% to 30.02% is observed in the latent Class 1 transit systems as Table 5 shows.

Table 6 and Figure 2 show yearly changes in cost inefficiency in both latent classes. The pattern that emerges here is that periods of increases in inefficiency often are followed by periods in which inefficiency declined. For the latent Class 1 transit systems, except for the brief period between 1986 and 1987 when there was an increase, average inefficiency reduced by 5.25% from 7.82% in 1985 to 7.41% in 1990. After 1991, when inefficiency in this class was highest at 8.66%, it declined by 36.84% to its lowest level of 5.47% in 2000, rose to 8.32% in 2003 to fall again to 7.23% in 2004. A similar trend is depicted by the yearly inefficiencies for the latent Class 2 transit systems. Here, average inefficiency declined from 4.34% in 1987 to 3.78% in 1991 and increased after that to 6.75% in 1995. Between 1995 and 2002, the overall trend was a decline in inefficiency in the latent Class 2 transit systems to 4.01%. This decline, however, was short-lived as it was followed by a 58.35% increase in inefficiency to 6.35% in 2004.

Comparing these results, the cost inefficiencies in both latent classes have a correlation of 0.1884. Also, the inefficiencies are larger in the latent Class 1 transit systems than they are in the latent Class 2 transit systems, especially in the earlier years in our data with the gap between these inefficiencies narrowing in recent years. For example, in 1985, cost inefficiency was 80.2% larger in the latent Class 1 transit systems than it was in the latent Class 2 transit systems. By 2005, this gap had narrowed to 13.86%. This is because while inefficiency was dropping in the latent Class 1 transit systems it was increasing in the latent Class 2 transit systems. Overall, the estimated cost inefficiency in the latent Class 1 transit systems is 51.62% larger than it is in the latent Class 2 transit systems. These findings suggest that there are clear differences in these transit systems in terms of efficiency within their latent classes.

THE HOMOGENEOUS FRONTIER CASE

Another comparison is in terms of what the inefficiencies would have been had heterogeneity not been considered in terms of latent classes. To perform this comparison, an index frontier with the same variables as in the earlier models was estimated without latent classes, and its results are in Table 7. As noted earlier, adding background variables is another way of accounting for heterogeneity in the cost equation while the results reported in the previous two sections account for heterogeneity simultaneously with background variables and latent classes. The results in Table 7 show the inefficiencies for the transit systems studied when only background variables are used to account for heterogeneity. Included in the table are the classes of the transit systems from Table 5. Comparing the results to those reported earlier that considered latent classes, it is observed that the index frontier gives very large inefficiencies. Specifically, without accounting for latent class, the mean inefficiency of the firms falling into latent Class 1 would have been 0.1822 in Table 8, a 143.23% increase. The corresponding increase in inefficiency in the latent Class 2 transit systems would have been 216.40% given their mean inefficiency of 0.1563 in Table 8. Thus, using background variables only to account for heterogeneity exaggerates inefficiency in the latent Class 2 transit systems more than it does in the latent Class 1 transit systems. These results also show that between 58.9% and 68.39% of the calculated inefficiencies in this paper are due to differences in technology as captured by latent class. Therefore, without considering latent classes, the transit

Table 4: Class Prediction and Inefficiency

Prob.	Class	System Name	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
1	1	Anchorage Public Transit	0.0884	0.0492	0.0521	0.0660	0.0317	0.0294	0.0503	0.0662	0.0749	0.1018	0.0905	0.0822
1	2	Providence RI PTA	0.0586	0.0449	0.0459	0.0435	0.0458	0.0682	0.0458	0.0365	0.0368	0.0430	0.0372	0.0341
1	1	Hartford-Connecticut Transit	0.1396	0.1139	0.0973	0.1136	0.1157	0.1236	0.1064	0.0770	0.0679	0.0809	0.0574	0.0803
1	1	Metropolitan Bus Authority	0.1040	0.0745	0.1226	0.0528	0.0557	0.1887	0.2653	0.3002	0.1854	0.0774	0.1220	0.1008
1	1	East Meadow MSBA	0.0710	0.0613	0.0583	0.0531	0.0647	0.0647	0.0675	0.0662	0.0701	0.0786	0.0812	0.0723
1	1	New York Bus Tours Inc.	0.0871	0.0863	0.1005	0.1112	0.1003	0.1066	0.0847	0.0345	0.0591	0.0638	0.0626	0.0538
1	1	Jackson Heights-Triboro Coach	0.0919	0.0948	0.1327	0.1064	0.1082	0.0785	0.0869	0.1748	0.1274	0.1334	0.1371	0.1755
1	1	Academy Lines-Leonardo	0.0323	0.0370	0.0337	0.0278	0.0274	0.0303	0.0346	0.0409	0.0385	0.0397	0.0410	0.0542
1	1	Hudson Transit Lines-Mahwah	0.0227	0.0555	0.0952	0.0876	0.0912	0.0842	0.0642	0.0554	0.0531	0.0498	0.0493	0.0499
1	1	Suburban Transit Corp	0.1058	0.0952	0.0358	0.0352	0.0359	0.0325	0.0729	0.0644	0.0944	0.0795	0.0885	0.1221
1	2	NJTC-45	0.0193	0.0378	0.0977	0.0980	0.0469	0.0474	0.0138	0.043	0.0615	0.1814	0.2511	0.0242
1	1	Hillsborough Area RTA	0.0542	0.0523	0.0605	0.0478	0.0459	0.0334	0.0815	0.0609	0.0789	0.0726	0.0353	0.0629
1	2	Minneapolis MTC	0.0419	0.0425	0.0453	0.0476	0.0518	0.0538	0.0422	0.0446	0.0529	0.0611	0.0736	0.0385
1	2	City of Detroit Department of Transportation	0.0665	0.0416	0.0521	0.0404	0.0413	0.0268	0.0206	0.0358	0.0318	0.0267	0.0305	0.0188
1	2	Honolulu DOT Service	0.0182	0.0191	0.0302	0.0508	0.0257	0.0244	0.0230	0.0371	0.0264	0.0244	0.0261	0.0257
1	2	Santa Cruz MTD	0.0252	0.0319	0.0411	0.0425	0.0396	0.0370	0.0496	0.0516	0.0552	0.0386	0.0496	0.0804
1	1	Santa Monica Muni Bus	0.0768	0.0787	0.0935	0.0883	0.0928	0.0407	0.0696	0.0404	0.0743	0.0973	0.0885	0.0680
1	2	Alameda-Contra Costa TD	0.0365	0.0363	0.0387	0.0290	0.0358	0.0419	0.0377	0.0806	0.1012	0.0609	0.0687	0.0445
1	2	Santa Barbara MTD	0.0396	0.0434	0.0413	0.0484	0.0507	0.0666	0.0397	0.0521	0.0520	0.0333	0.0426	0.0417
1	2	Los Angele-SCRTD	0.0618	0.0719	0.0687	0.0561	0.0594	0.0590	0.0712	0.0738	0.0832	0.0688	0.0683	0.0639
1	1	Phoenix Transit System	0.0649	0.0631	0.0683	0.0472	0.0741	0.0763	0.0554	0.0501	0.0532	0.0710	0.0582	0.0659
1	2	Milwaukee County TS	0.0606	0.0471	0.0265	0.0331	0.0233	0.0237	0.0308	0.0678	0.0340	0.0543	0.0285	0.0583
1	2	Cincinnati-SORTA	0.0382	0.0458	0.0537	0.0514	0.0200	0.0250	0.0316	0.0510	0.0554	0.0703	0.0571	0.0602
1	2	Toledo RTA	0.0548	0.0570	0.0644	0.0468	0.0496	0.0449	0.0474	0.0552	0.0739	0.0511	0.0772	0.0606

Prob.: Estimated highest probability

Table 4 (Continued)

Class	System Name	1997	1998	1999	2000	2001	2002	2003	2004
1	Anchorage Public Transit	0.0649	0.0701	0.0595	0.0392	0.0380	0.0661	0.0602	0.0552
2	Providence RI PTA	0.0421	0.0518	0.0408	0.0519	0.0607	0.0580	0.0494	0.0975
1	Hartford-Connecticut Transit	0.0833	0.0772	0.0444	0.0743	0.0705	0.0704	0.1613	0.0871
1	Metropolitan Bus Authority	0.0705	0.0864	0.1528	0.0656	0.1439			
1	East Meadow MSBA	0.0637	0.0689	0.0389	0.0679	0.0650	0.0827	0.0872	0.0834
1	New York Bus Tours Inc.	0.0477	0.0538	0.1704	0.0517	0.0673	0.0570	0.0641	0.0529
1	Jackson Heights-Triboro Coach	0.1682							
1	Academy Lines-Leonardo	0.0462	0.0361	0.0571	0.0339	0.0519	0.0913	0.1225	0.0903
1	Hudson Transit Lines-Mahwah	0.0525	0.0799	0.0973	0.0602	0.0703	0.0732	0.0983	0.0518
1	Suburban Transit Corp	0.1382	0.0636	0.0795	0.0772	0.0735	0.0706	0.0318	0.0340
2	NJTC-45	0.0883	0.0121	0.0143	0.0833	0.0747	0.0204	0.0315	0.1130
1	Hillsborough Area RTA	0.0286	0.0516	0.0253	0.0241	0.0578	0.0554	0.0606	0.0654
2	Minneapolis MTC	0.0529	0.0672	0.0775	0.0414	0.0246	0.0325	0.0262	0.0529
2	City of Detroit Department of Transportation	0.0170	0.1614	0.0818	0.0345	0.0183	0.0289	0.0487	0.0418
2	Honolulu DOT Service	0.029	0.0222	0.0516	0.0281	0.0368	0.0347	0.0478	0.0537
2	Santa Cruz MTD	0.0603	0.0526	0.0503	0.0495	0.0232	0.0470	0.0338	0.0588
1	Santa Monica Muni Bus	0.0647	0.0628	0.0551	0.0674	0.0533	0.0847	0.0807	0.0585
2	Alameda-Contra Costa TD	0.0583	0.0528	0.0248	0.0198	0.0493	0.0406	0.0700	0.0445
2	Santa Barbara MTD	0.0238	0.0192	0.0276	0.0322	0.0372	0.0294	0.0333	0.0299
2	Los Angele-SCRTD	0.0916	0.0410	0.0898	0.0795	0.0532	0.0487	0.1357	0.1422
1	Phoenix Transit System	0.0575	0.0559	0.0562	0.0514	0.0629	0.0592	0.1150	0.0576
2	Milwaukee County TS	0.0676	0.0387	0.0246	0.0312	0.0214	0.0595	0.0978	0.0462
2	Cincinnati-SORTA	0.0547	0.0443	0.0227	0.0537	0.0245			
2	Toledo RTA	0.0171	0.0233	0.0522	0.0341				

systems studied appear more inefficient than they are. But, as it is clear from the results, a very large portion of this inefficiency is due to the assumption of a homogeneous technology and the inability to account fully for heterogeneity with background variables. Although it could be argued that our finding may be because the chosen background variables do not adequately account for heterogeneity and that perhaps with another set of variables lower inefficiencies would have been obtained from the index frontier, we argue to the contrary that the latent class approach better accounts for inefficiency than just including background variables in the frontier equation. This finding suggests that previous studies may have substantially over-estimated inefficiencies in public transit systems and that perhaps a re-examination of their conclusions is in order.

Table 5: Comparison of Latent Class Inefficiency Estimates Across Transit Systems

Class	Estimated Highest Group Probability	System Name	Mean	Std. Dev.	Minimum	Maximum
		All Transit Systems	**0.0619**	**0.0348**	**0.0121**	**0.3002**
1	1.00	Anchorage Public Transit	0.0618	0.0196	0.0294	0.1018
1	1.00	Hartford-Connecticut Transit	0.0921	0.0289	0.0444	0.1613
1	1.00	Metropolitan Bus Authority	0.1276	0.0720	0.0528	0.3002
1	1.00	East Meadow MSBA	0.0683	0.0112	0.0389	0.0872
1	1.00	New York Bus Tours Inc.	0.0758	0.0311	0.0345	0.1704
1	1.00	Jackson Heights-Triboro Coach	0.1243	0.0334	0.0785	0.1755
1	1.00	Academy Lines-Leonardo	0.0483	0.0250	0.0274	0.1225
1	1.00	Hudson Transit Lines-Mahwah	0.0671	0.0206	0.0227	0.0983
1	1.00	Suburban Transit Corp	0.0715	0.0309	0.0318	0.1382
1	1.00	Hillsborough Area RTA	0.0528	0.0167	0.0241	0.0815
1	1.00	Santa Monica Muni Bus	0.0718	0.0168	0.0404	0.0973
1	1.00	Phoenix Transit System	0.0607	0.0085	0.0472	0.0763
		All Class 1 Transit Systems	**0.0749**	**0.0369**	**0.0227**	**0.3002**
2	1.00	Providence RI PTA	0.0496	0.0144	0.0341	0.0975
2	1.00	NJTC-45	0.0680	0.0608	0.0121	0.2511
2	1.00	Minneapolis MTC	0.0485	0.0140	0.0246	0.0775
2	1.00	City of Detroit Department of Transportation	0.0433	0.0322	0.0170	0.1614
2	1.00	Honolulu DOT Service	0.0318	0.0111	0.0182	0.0537
2	1.00	Santa Cruz MTD	0.0459	0.0131	0.0232	0.0804
2	1.00	Alameda-Contra Costa TD	0.0486	0.0199	0.0198	0.1012
2	1.00	Santa Barbara MTD	0.0392	0.0113	0.0192	0.0666
2	1.00	Los Angele-SCRTD	0.0744	0.0255	0.0410	0.1422
2	1.00	Milwaukee County TS	0.0481	0.0241	0.0214	0.1150
2	1.00	Cincinnati-SORTA	0.0447	0.0150	0.0200	0.0703
2	1.00	Toledo RTA	0.0506	0.0160	0.0171	0.0772
		All Class 2 Transit Systems	**0.0494**	**0.0273**	**0.0121**	**0.2511**

Finally, many environmental and regulatory constraints affect operational efficiency and effectiveness of transit systems. The management of transit systems cannot control these constraints. Our latent class model captures variations in operational constraints among transit firms as well as their technology in heterogeneity, and measures the inefficiency for which management is responsible. It is left to future research to decompose the heterogeneity into its sources such as various government regulations, the operating environment, and technology. This will give further insight to formulate government policies to improve efficiency of transit systems.

CONCLUSION

This study estimates and compares the results of two approaches to account for heterogeneity in estimates of inefficiencies from stochastic cost frontier models using panel data for urban transit systems spanning 1985 to 2004. The first approach, due to Orea and Kumbhakar (2004) and Greene

Table 6: Yearly Variation in Inefficiency Among Latent Classes

	Mean	**Mean**
Year	Latent Class 1	Latent Class 2
1985	0.0782	0.0434
1986	0.0718	0.0433
1987	0.0792	0.0505
1988	0.0698	0.0490
1989	0.0703	0.0408
1990	0.0741	0.0432
1991	0.0866	0.0378
1992	0.0859	0.0524
1993	0.0814	0.0553
1994	0.0788	0.0595
1995	0.0760	0.0675
1996	0.0823	0.0459
1997	0.0738	0.0502
1998	0.0620	0.0493
1999	0.0684	0.0487
2000	0.0547	0.0441
2001	0.0634	0.0403
2002	0.0804	0.0401
2003	0.0832	0.0619
2004	0.0723	0.0635
All	0.0749	0.0494

(2005b), is a latent class stochastic frontier model in which we have also included background variables as additional heterogeneity variables. The second is an index stochastic cost frontier model that accounts for heterogeneity by including background variables in the set of variables. The results show that the estimated coefficients of the latent class cost functions are noticeably different from those of the index stochastic cost frontier. Both the two-sided random errors and one-sided errors (cost inefficiency) are found to be substantially smaller in the latent class models than in the second model. These show that the latent class stochastic frontier specification with background variables included accounts for heterogeneity better than the index frontier and suggest possible bias of the results from the index frontier. This is because the index frontier assumes a homogeneous technology and, despite including background variables in its specification, it does not adequately account for heterogeneity. A comparison of the estimated cost inefficiencies of the individual transit systems from the two specifications shows that accounting for differences in technology with latent classes reduces measured inefficiency. These findings caution researchers and policy makers alike against using models that assume homogeneous technology to evaluate the efficiency of transit systems and as a basis to formulate policies.

A limitation of this paper is that it does not consider contracting, which could reduce the inefficiencies in transit systems. Another is that fuel is used as a proxy for all other inputs and the last is the size of the sample used. These limitations notwithstanding, the purpose of the paper is to demonstrate the usefulness of latent class frontiers in research on efficiency. Future research that addresses these limitations should improve our knowledge of inefficiency in transit systems.

Figure 2: Yearly Variation in Efficiency

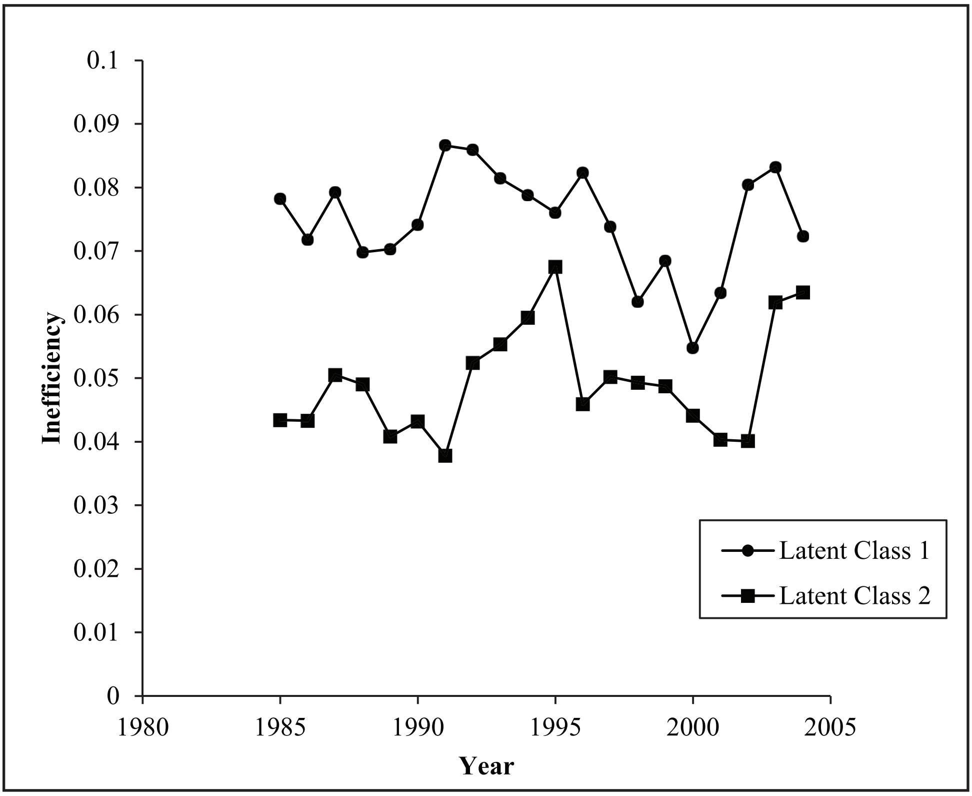


Table 7: Index Equation

Variable		Coefficient	Std. Err	t-value	P[|Z|>z]
Constant|	Constant	-0.1542	0.0264	-5.8470	0.0000
$\ln(w_K / w_L)$	Log of relative capital price	0.1341	0.0480	2.7930	0.0052
$\ln(w_F / w_L)$	Log of relative fuel price	0.2201	0.0412	5.3450	0.0000
$\ln(Q)$	Log of output	0.8037	0.0288	27.9540	0.0000
t	Time in years	-0.0099	0.0022	-4.4000	0.0000
$\ln(w_K / w_L)\ln(w_F / w_L)$	Log of relative price of capital times relative fuel price	0.0884	0.1404	0.6300	0.5288
$0.5[\ln(w_K / w_L)]^2$	One half times the square of log of relative capital price	-0.0630	0.1531	-0.4120	0.6807
$[\ln(w_K / w_L)]\ln Q$	Log of relative price of capital times output	-0.2349	0.0587	-4.0040	0.0001
$0.5[\ln(w_F / w_L)]^2$	One half times the square of log of relative fuel price	0.2043	0.1493	1.3680	0.1712
$[\ln(w_F / w_L)]\ln Q$	Log of relative price of fuel times output	0.0410	0.0573	0.7160	0.4742
$0.5[\ln(Q)]^2$	One half times the square of log of output	-0.1305	0.0338	-3.8580	0.0001
$0.5t^2$	One-half times the square of time	-0.0019	0.0009	-2.2030	0.0276
$t\ \ln(Q)$	Time times log of output	-0.0064	0.0029	-2.2340	0.0255
$t\ \ln(w_F / w_L)$	Time times log of relative fuel price	-0.0056	0.0067	-0.8390	0.4016
$t\ \ln(w_K / w_L)$	Time times log of relative capital price	-0.0091	0.0079	-1.1450	0.2520
$\ln(S_o)$	Log of operating subsidy	0.0109	0.0052	2.1160	0.0344
$\ln(S_K)$	Log of capital subsidy	0.0111	0.0032	3.4510	0.0006
$\ln(Q)\ \ln(S_o)$	Log of output times log of operating subsidy	0.0071	0.0076	0.9320	0.3513
$[\ln(S_K)]\ln(Q)$		0.0359	0.0050	7.2200	0.0000
$\ln(w_F / w_L)\ \ln(S_o)$	Log of relative fuel price times log of operating subsidy	-0.0038	0.0117	-0.3260	0.7442
$\ln(w_F / w_L)\ \ln(S_K)$	Log of relative fuel price times log of capital subsidy	0.0303	0.0075	4.0280	0.0001
$\ln(w_K / w_L)\ \ln(S_o)$	Log of relative capital price times log of operating subsidy	0.0236	0.0139	1.6970	0.0896
$\ln(w_K / w_L)\ \ln(S_o)$	Log of relative capital price times log of capital subsidy	0.0202	0.0076	2.6470	0.0081
$t\ \ln(S_o)$	Time times log of operating subsidy	-0.0005	0.0006	-0.7500	0.4532
$t\ \ln(S_K)$	Time times log of capital subsidy	0.0011	0.0005	2.2780	0.0227
$\ln(R)$	Log of network size	0.1111	0.0209	5.3130	0.0000
σ_v		0.2098			
σ_u		0.2112			
Parameters for the Compound Error					
λ		1.0065	0.1350	7.4570	0.0000
σ		0.2977	0.0005	625.2310	0.0000

Table 8: Inefficiency from Index Frontier

Class	System Name	Inefficiency
1	Anchorage Public Transit	0.1276
1	Hartford-Connecticut Transit	0.1566
1	Metropolitan Bus Authority	0.3197
1	East Meadow MSBA	0.2019
1	New York Bus Tours Inc.	0.1945
1	Jackson Heights-Triboro Coach	0.2394
1	Academy Lines-Leonardo	0.1644
1	Hudson Transit Lines-Mahwah	0.1458
1	Suburban Transit Corp	0.1609
1	Hillsborough Area RTA	0.1134
1	Santa Monica Muni Bus	0.1978
1	Phoenix Transit System	0.1646
	Mean	**0.1822**
2	Providence RI PTA	0.1546
2	NJTC-45	0.2449
2	Minneapolis MTC	0.1442
2	City of Detroit Department of Transportation	0.1232
2	Honolulu DOT Service	0.1232
2	Santa Cruz MTD	0.1596
2	Alameda-Contra Costa TD	0.1377
2	Santa Barbara MTD	0.1487
2	Los Angele-SCRTD	0.1978
2	Milwaukee County TS	0.1699
2	Cincinnati-SORTA	0.1386
2	Toledo RTA	0.1331
	Mean	**0.1563**

Endnotes

1. Greene (2002) notes that the latent class models are also an approximation to the random parameters models.

2. Greene (2002) notes "Results are similar to the random parameters model, but for the same data, the latent class estimator appears to produce a much tighter distribution for u_{it} than the random parameters model." Although theoretically the random parameters models are more efficient, empirically the latent class models may produce better estimation results.

3. In this paper we follow Karlaftis and McCarthy (2002) among others and treat subsidies as exogenous. This exogeneity makes the simplifying assumption that most of the subsidies are lump-sum. Furthermore, though we treat output as exogenous Obeng (2011) shows that it should be treated as endogenous.

4. Before settling on a fixed coefficient for each class we tried parameterizing F(•) with subsidy availability, if or not a transit system purchases transportation as its variables and had no success.

5. In this data, labor hours are equivalent labor times 2080 hours.

6. In calculating the input prices, the cost of purchased transportation is spread among the inputs according to each input's share in cost. Thus, we initially calculated total cost, then allocated the cost of purchased transportation to the inputs according to their shares in cost. After that, we recalculated the input prices as their costs divided by their physical quantities.

7. Friedlander et al. (1981) do not provide a justification for the 12% added to operating cost.

8. All variables that are not binary are transformed into logarithms before they are used in the estimation.

9. The final latent class model estimated uses 23 transit systems because one was identified as outlier.

10. Again note that all non-labor operating costs are allocated to fuel and that fuel is a proxy for materials and fuel and that the fuel price is actually the price of non-labor inputs.

11. To obtain this percentage multiply the entry for inefficiency by 100.

12. Two transit systems in latent Class1 have very high mean inefficiencies and could be considered as outliers. If they are deleted the mean inefficiency of this class is 0.0670 which is 35.63% larger than the mean inefficiency score of the latent Class 2 transit systems.

References

Ahn, S., Y. Lee and P. Schmidt. "Stochastic Frontier Models with Multiple Time-Varying Individual Effects." *Journal of Productivity Analysis* 27, (2007): 1-12.

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

Battese, G. and T. Coelli. "Frontier Production Functions, Technical Efficiency and Panel Data:

With Application to Paddy Farmers in India." *Journal of Productivity Analysis* 3, (1992): 153-169.

Battese, G. and T. Coelli. "A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data." *Empirical Economics* 20, (1995): 325-332.

Baumol, W.J., J.C. Panzar, and R.D. Willig. *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich, Inc. New York, 1982.

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

Berger, A.N., J.C. Demsetz, and P.E. Strahan. "The Consolidation of the Financial Services Industry, Cause, Consequences and Implications For the Future." *Journal of Banking and Finance* 23, (1999):135-195.

Boame, A.K. "The Technical Efficiency of Canadian Urban Transit Systems." *Transportation Research Part E: Logistics and Transportation Review* 50(5), (2004):401-416.

Caves, D. W., L.R. Christensen, and M.W. Tretheway. "Flexible Cost Functions for Multiproduct Firms." *Review of Economics and Statistics* 62(3), (1980):477-481.

Cornwell, C., P. Schmidt, and R. Sickles. "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics* 46, (1990): 185-200.

Dalen, D. and A. Gomez-Lobo. "Yardsticks on the Road: Regulatory Contracts and Cost Efficiency in the Norwegian Bus Industry." *Transportation* 30, (2003): 371-386.

De Borger, B., K. Kerstens, and A. Costa. "Public Transit Performance: What Does One Learn from Frontier Studies?"*Transport Review* 22 (1), (2002): 1-38.

Farsi, M., M. Filippini, and M. Kuenzle. "Cost Efficiency in Regional Bus Companies: An Application of Alternative Stochastic Frontier Models." *Journal of Transportation Economics and Policy* 24, (2006): 95-118.

Farsi M., A. Fetz and M. Filipini. "Economies of Scale and Scope in Local Public Transportation." *Journal of Transport Economics and Policy* 41(3), (2007): 345-361.

Fraley C. and A.E. Raftery. "How Many Clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis." Technical Report No. 329, Department of Statistics, University of Washington, 1998.

Friedlander A.F. and W. Chiang. "Productivity Growth in the Regulated Trucking Industry." *Research in Transportation and Economics* 1, (1983): 369-385.

Friedlander, A.F., R.H. Spady, and S.J. Chiang. "Regulation and the Structure of Technology in the Trucking Industry." T. G. Cowing and R. E. Stevenson eds. *Productivity Measurement in Regulated Industries*. New York: Academic Press (1981): 77-106.

Greene, W. "Alternative Panel Data Estimators for Stochastic Frontier Models." Department of Economics, Stern School of Business, New York University, New York, NY, 2002.

Greene, W. "Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems." *Health Economics* 13, (2004): 959-980.

Greene, W. "Fixed and Random Effects in Stochastic Frontier Models." *Journal of Productivity Analysis* 23, (2005a):7-32.

Greene, W. *LIMDEP Version 9.0*. Econometric Software Inc. Plainview, New York, 2007.

Greene, W. "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model." *Journal of Econometrics* 126, (2005b): 269-303.

Jha, R. and S. Singh. "Small is Efficient: A Frontier Approach to Cost Efficiencies in Indian State Road Transport Undertakings." *International Journal of Transport Economics* 18 (1), (2001): 95-114.

Jondrow, J., K. Lovell, I. Materov, and P. Schmidt. "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *Journal of Econometrics* 19, (1982): 233-238.

Karlaftis, M.G. and P. McCarthy. "Cost Structures of Public Transit Systems: A Panel Data Analysis." *Transportation Research Part E* 38, (2002): 1-18.

Kolari J. and A. Zardkoohi. "Economies of Scale and Scope in Commercial Banks with Different Output Mixes." Texas A&M Working Paper, 1995.

Kumbhakar, S. "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *Journal of Econometrics* 46, (1990): 201-211.

Lee, Y. and P.A. Schmidt. "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency." H. Fried, C. Lovell and P. Schmidt eds. *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford: Oxford University Press (1993).

Nadiri M.I. and M.A. Schankerman. "Production, Technological Change and Productivity in the Bell System." T.G. Cowing and R.E. Stevenson eds. *Productivity Measurement in Regulated Industries.* New York: Academic Press (1981):219-246.

National Transit Database. http://www.ntdprogram.gov/ntdprogram/data.htm (various years).

Nunamaker, T.R. "Using Data Envelopment Analysis to Measure the Efficiency of Non-Profit Organizations: A Critical Evaluation." *Managerial and Decision Economics* 6(1), (1985): 50-58.

Obeng, K. "Indirect Production Function and the Output Effects of Public Transit Subsidies." *Transportation* 38(2), (2011): 191-214.

Obeng, K., G. Azam, and R. Sakano. *Modeling Economic Inefficiency Caused by Public Transit Subsidies.* Praeger: Westport, Connecticut, 1997.

Orea, L. and S. Kumbhakar. "Efficiency Measurement Using a Latent Class Stochastic Frontier Model." *Empirical Economics* 29, (2004):169-183.

Pitt, M. and L. Lee. "The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry." *Journal of Development Economics* 9, (1981): 43-64.

Piacenza, M. "Regulatory Contracts and Cost Efficiency: Stochastic Frontier Evidence from the Italian Local Public Transport." *Journal of Productivity Analysis* 25, (2006): 257-277.

Pulley, L.B. and Y.M. Braunstein. "A Composite Cost Function for Multiproduct Firms with an Application to Economies of Scope in Banking." *Review of Economics and Statistics* 74(2), (1992):221-230.

Pulley, L. and D.B. Humphrey. "The Role of Fixed Costs and Cost Complementarities in Determining Scope Economies and the Cost of Narrow Banking Proposals." *Journal of Business* 66, (1993): 437-62.

Sakano, R. and K. Obeng. "Stochastic Production Possibility Frontier in Urban Transit System: Re-Examination." *The Logistics and Transportation Review* 31(4), (1995):377-392.

Sakano, R., K. Obeng, and G. Azam. *"Subsidies and Inefficiency: Stochastic Frontier Approach." Contemporary Economic Policy* 5(3), (1997):113-127.

Schmidt, P. and R. Sickles. "Production Frontiers with Panel Data." *Journal of Business Statistics* 2, (1984): 367-374.

Scheiner, K.E., C.M. Anderson, and W.C. Horrace. "Estimating Heterogeneous Production in Fisheries." Working paper No. 80, Center for Policy Research, Syracuse University, Syracuse NY. 2006.

Tsionas, M. "Stochastic Frontier Models with Random Coefficients." *Journal of Applied Econometrics* 17, (2002): 127-147.

Weninger, Q. "Estimating Multiproduct Costs When Some Outputs Are Not Produced." *Empirical Economics* 28, (2003):753-765.

Acknowledgment

This research was supported by a summer faculty grant from the School of Business and Economics, North Carolina A&T State University, Greensboro, NC 27411.

Ryoichi Sakano *is associate professor in the Department of Economics and Finance at North Carolina Agricultural and Technical State University. He received his Ph.D. in economics from the University of Alabama and has published in the areas of efficiency and productivity of transit systems, travel behavior, and income distribution in various transportation and economics journals.*

Kofi Obeng *is professor of transportation in the Department of Marketing, Transportation and Supply Chain at North Carolina A&T State University. He received his Ph.D. in City and Regional Planning with a concentration in transportation planning from the University of Pennsylvania and has published on cost, efficiency, productivity and performance, transit subsidy and transit management, and traffic safety.*