ORGANIZING TRANSIT IN SMALL URBAN AND RURAL COMMUNITIES

David G. Ripplinger
Upper Great Plains Transportation Institute
North Dakota State University

July 2012

ACKNOWLEDGEMENTS

This research was sponsored by the Federal Transit Administration, United States Department of Transportation, and conducted by the Small Urban and Rural Transit Center within the Upper Great Plains Transportation Institute at North Dakota State University.

This report was written in partial fulfillment of the requirements for a Ph.D. in Transportation and Logistics at North Dakota State University. The author thanks John Bitzan, Siew Lim, Jim Miller, and Denver Tolliver for their comments and suggestions at various stages of research.

ABSTRACT

The justification of government support of rural transit on the basis of the presence of increasing returns to scale and the most efficient regional organization of transit is investigated. Returns to density, size, and scope at most levels of output were found. Cost subadditivity, where a monopoly firm can provide service at a lower cost than two firms, was found for many, but not all observations. The presence of natural monopoly in rural transit in a strict sense is rejected. The findings and implications are directly applicable to rural transit in North Dakota and should be helpful in informing future federal policy as well as rural transit policy, service design, and operation in other states.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The Federal Government has played an active role in public transportation since the 1960s when it began providing capital subsidies to faltering urban mass transit systems (Wachs 1989). Involvement was expanded to include the funding of operating expenses in 1974, and later to fund transit for different riders and geographic areas including providers of transportation for the elderly and disabled in 1975, and service in rural communities in 1978 (Federal Transit Administration 2007a, 2007b, 2010).

In the past decade, federal spending on transit has increased from $7.9 billion in 2001 to $10.5 billion in 2011 (2011 dollars). During this time, spending on public transportation service in rural areas increased more than 60% from $251 million to $537 million, while total federal spending increased 53%. Meanwhile, spending on six dozen other programs that fund passenger transportation outside of the Federal Transit Administration (FTA) and U.S. Department of Transportation (DOT) totaled in the billions of dollars (GAO 2003).

With the Federal Government currently experiencing record, unsustainable deficits, federal programs are experiencing increased levels of scrutiny from the public and elected officials. While some have proposed sweeping cuts (House Concurrent Resolution 34), justification of federal support for public transportation and the efficiency and effectiveness of federally funded transit agencies have been questioned.

While making changes in funding levels may be politically or financially expedient, economic rationale for government support exists. Government intervention in public transportation is typically justified by three economic arguments (Elgar and Kennedy 2005). First, transit agencies may experience increasing returns to scale and be unable to cover their average costs at the socially desired level of output. Second, negative externalities of automobile transportation including air and noise pollution or congestion may be reduced by subsidizing transit. Finally, government intervention can be used to internalize the positive external impacts of transit - improved accessibility, higher land-use values, and economic agglomeration.

The federal government has the ability to influence the efficiency of transit at the agency and regional level by promulgating and enforcing regulations on which federal funding is conditioned. One unexercised alternative is strict guidance on the configuration of transit within a region. Arrangements that maximize efficiency depend on the cost structure of transit with alternatives including the number of agencies and the type, level, and area of service provided. In a multiproduct setting, the presence of natural monopoly would justify the existence of a regional single agency. In its absence, many combinations would need to be evaluated.

This study investigates the cost structure of the rural transit industry. The intent is to evaluate the justification of government intervention on the basis of the presence of natural monopoly and to evaluate the efficiency of alternative regional transportation configurations and the potential cost savings relative to the current organization of rural transit.

1.1 Rural Transit in the United States

Despite receiving significant federal financial support and playing an important role increasing the mobility of riders and enhancing the vitality of the communities, rural transit is an often overlooked sector. As a result, there is a lack of knowledge of the aggregate scale and scope of rural transit, and of the significant state and federal efforts that have been made to ensure efficiency. Rural transit is unique in its financing and administration, as well as in the services it provides. An understanding of rural transit is necessary to properly frame the issues of government support and efficiency.

1.1.1 Rural Transit Finance and Administration

The Federal Transit Administration sponsors more than 20 major grant programs to support public transportation (GAO 2003). Among these is the Formula Grants for Other than Urbanized Areas commonly referred to as Section 5311 (in reference to its location in the transportation chapter of the United States Code). The intent of the Section 5311 program is to increase access to health care, shopping, education, employment, public services, and recreation for residents of rural areas (defined as those places with populations less than 50,000). Program funds are meant to assist in the development, improvement, maintenance, and use of rural transit, and to support intercity bus transportation.

States play a critical role in administering the Section 5311 program. With the exception of specific requirements identified by FTA, states are provided maximum discretion in managing transit service that is provided by local agencies that receive Section 5311 funds. States develop state-level management plans, develop and apply project funding criteria, monitor local activity, and ensure compliance with federal requirements. In some cases, states may directly operate rural transit service as is the case in Rhode Island, Connecticut, and Delaware. States are required to provide technical assistance, and to ensure that funds are distributed fairly and equitably, that private operators are given an opportunity to participate, that intercity bus transportation service is supported, and that FTA and other federally funded transportation services are coordinated to the maximum extent feasible. In addition to administering the federal program to provide transit to rural areas, states also oversee programs that fund buses and bus facilities (Section 5309), transportation for elderly and disabled persons (Section 5310), job access and reverse commute transportation (Section 5316), and efforts to eliminate barriers facing disabled individuals (Section 5317). Many transit agencies that receive Section 5311 monies also receive funding from these programs.

Federal programs that support capital purchases including Sections 5309, 5310, and 5311 impact the cost structure of transit agencies that receive these funds. As few agencies are able to purchase vehicles without assistance, fleet expansion typically occurs only when federal or state funds are available. While states have selection criteria that prioritize need, in many cases, agencies must wait in queue for their turn to procure new vehicles. At the same time, disposal of vehicles that have not yet met their minimum useful life or have a value more than $5,000 must be cleared with FTA or kept on an agency’s vehicle roster and maintained (Federal Transit Administration 2008).

1.1.2 Rural Transit Service

Transit agencies are a product of the communities they serve, with geography, demography, economics, and history dictating the services they provide and how they are best delivered. By definition, rural areas have small, low-density populations. These attributes impact travel behavior and rural transit service design.

Demand-response transportation, where trips are scheduled at the request of riders, is the predominant type of rural transit service. Long-distance trips to regional centers are often provided as demand-response service. A second type of rural transit service is fixed-route, where regularly scheduled service with timed stops occurs along a defined route. However unlike urban fixed-route service, rural routes may be hundreds of miles in length. Intercity bus service, defined by the use of large over-the-road buses where riders are seated on an elevated passenger deck over a baggage compartment, is the third common rural service type.

While rural transit agencies provide service to the general public, much of their service is tailored to meet the needs of the elderly and disabled, who along with low-income individuals form a group commonly referred to as transportation disadvantaged individuals. This is, in part, a historical remnant, as many early recipients of Section 5311 funds were social service agencies that received Section 5310 funds to purchase vehicles and other federal monies to provide non-transportation-related services to the elderly and disabled. At the same time, members of these groups are among the most likely to be unable to meet their mobility needs independently and are a target market for transit agencies in any community.

Another consequence of the historical joint delivery of rural transit and social services is that many recipient agencies are small and community focused. Transit service was designed to meet local, as opposed to regional transportation needs. Administrators managed many social programs with little training, experience, and, in some cases, interest in transit. These small, multi-faceted social service agencies continue to dominate rural transit to the present day.

1.1.3 The Scale and Scope of Rural Transit

Rural transit is a large, although often overlooked, industry. In 2010, nearly 1,600 organizations received Section 5311 funds (Federal Transit Administration 2012). These agencies provided more than 123 million trips, 570 million vehicle-miles, and 32 million vehicle-hours of service. The capital and operating expenses of providing this service was $412 million and $1.2 billion, respectively. Nationally, operating expenses averaged $10 per trip, $2 per vehicle-mile, and $38 per vehicle-hour of service. This compares to $4 per trip, $8 per vehicle-mile, and more than $107 per vehicle-hour of service for urban systems.

Agencies that received Section 5311 funds operated 23,136 vehicles, ranging in size from cars or minivans to large buses that seat 50 passengers or more. On average, these vehicles were operated for less than four hours and driven less than 70 miles per day.

In 2010, 25% of operating expenses incurred by agencies that received Section 5311 monies were funded by Federal Transit Administration programs, 5% by other federal programs, 19% by states, 25% by local governments, and the remainder by fares and contracts. For urban systems, the federal government supported 8%, states 25%, and localities 28% of operating expenses. In 2010, 89% of rural transit capital expenses were funded by the FTA while state and local sources provided 6% and 4%, respectively.

Federal Transit Administration programs typically require a 50% local match for operating expenses and a 20% local match for capital purchases. This difference in rates results in overcapitalization as factor prices as perceived by transit agencies are skewed (Obeng et al 1997). While the majority of local match for operating and capital purchases is provided by local and state government, transit agencies typically manage these funds as if they are their own.

Rural transit agencies differ fundamentally from their urban counterparts. Rural agencies are usually much smaller in terms of budget, staff, and size. Their fleets typically consist of automobiles and small buses, referred to as cutaways that seat 15 or fewer passengers. However, rural transit agency service areas are often much larger than those served by urban agencies. For example, River Cities Public Transit based in Pierre, South Dakota, serves an area twice the size of the State of Maryland. Rural transit has undergone significant changes in the past decade. Agencies have increased in size and expanded service due to increases in real levels of federal and other funding. Moreover, as the general population has aged and workers and households have been adversely affected by recession and high unemployment, the number of transportation disadvantaged individuals has increased. At the same time, the cost of transportation has risen with increased energy prices. This has impacted individual travel behavior as well as the cost of delivering passenger transportation service.

1.1.4 Transit Efficiency and Effectiveness

The Federal Transit Administration has its own definitions of efficiency and effectiveness distinct from those familiar to economists. In transit, efficiency is defined as the ratio of output produced per unit of input, while effectiveness is the ratio of output consumed per unit of output produced or per unit of input.

The federal government has long supported increased efficiency and effectiveness of public transportation primarily by funding technical assistance and coordinating activities. Consolidation, the creation of a single administrative and operating body, is not mandated by federal transit policy, although it may achieve federal efficiency and effectiveness goals.

Under current federal regulations, states must certify compliance with federal coordination requirements. These include the participation of rural transit providers in the development of locally-developed human services transportation plans and coordination by rural transit agencies with other federally funded transportation providers.

Increased efficiency of coordinating agencies may be realized by grouping riders, providing fewer vehicle-trips, and sharing staff, equipment, and facilities (GAO 1999). Realized efficiencies can result in reduced costs or increased levels and quality of service as cost savings are redirected to provide additional or improved service. However, there are barriers to coordination. These include an unwillingness of individuals and organizations to work together, inflexibility of agencies to modify existing administrative or operational behaviors, and federal requirements (Burkhardt 2004). Additional benefits may result from consolidation. Cook notes a number of benefits of consolidation including more effective regional planning, the ability to address regional transportation problems, development of specialized staff, operational and administrative economies, and improved communication and collaboration with other agencies (2002). However, consolidation presents a number of challenges in addition to those that confront coordination. They include the loss of local control in the design, administration, and operation of service, as well as the potential loss of jobs in some communities.

The potential consolidation of rural transit agencies is affected by the state administration of federal programs that support rural transit. State control of funding and the ability to develop state-level priorities provides state agencies that administer federal transportation programs considerable power over transit agency organization and service delivery. States have de facto control over the entry of new service providers and can force or prevent consolidations regardless of benefits.

The effectiveness and efficiency of public transportation agencies has traditionally been quantified using performance measures, typically the ratios of various operational and financial statistics, as opposed to a more formal economic analysis. Individual performance measures are typically compared within peer groups and are, in some cases, used in funding decisions (Ryus 2010). This analytical method has significant shortcomings, as peer group assignments are often based on endogenous variables with no control for within-group variability.

The benefits of consolidation of rural transit agencies have not been quantified. In the limited cases where consolidation benefits have been identified, qualitative descriptions, rather than economic or other quantitative measures have been presented (e.g. Cook 2002). Furthermore, the economic efficiency of new, expanded, or modified rural transit service is not known. At an even more basic level, there has been no rigorous economic analysis to justify the more than one billion dollars in support provided to rural transit each year.

1.1.5 Transportation Costing Theory and Methods

A firm's production technology contains economic information that can be used to develop industrial and regulatory policy. Knowledge of the presence of economies of scale and scope, input factor and cost elasticities, average and marginal costs, and optimal fleet size can be used to establish pricing policies, determine cost effects of new or modified service, and estimate the impacts of privatization (Karlaftis and McCarthy 2002).

Direct estimation of a production function has traditionally been hindered by endogeneity, as factor input levels are chosen by the firm. This condition can be addressed by estimating a firm's cost function which by duality contains the same information as the production function (Shephard 1953). The transcendental logarithmic (translog) function introduced by Christensen, Jorgenson, and Lau (1970) has become the preferred method for estimating cost functions. The translog function can be thought of as the second-order approximation of an arbitrary cost function. It has a number of benefits over alternative functional forms, such as the quadratic (Lau 1974) or generalized Leontief (Diewert 1971), as it places no a priori restrictions on returns to scale or elasticities of substitution, allows for the imposition of homogeneity restrictions, and requires fewer parameters to be estimated. The translog function utilizes cost-share equations (Shephard 1953) to increase the model's degrees of freedom without increasing the number of parameters, and is estimated using the Seemingly Unrelated Regression (SUR) technique (Zellner 1962).

A multiproduct version of the translog function that accommodates the presence of observations with zero output was introduced by Caves, Christensen, and Tretheway (1980). The Generalized Translog Multiproduct Cost Function (GTMCF) maintains the advantages of the single output translog function and accommodates observations with zero output using a Box-Cox transformation (Box and Cox 1964).

Transportation cost models often include network variables. These include network size and other variables that capture relevant environmental attributes or impact production technology. Network variables include service area, number of stops, length of routes, points served, average length of trip, and speed.

Transit agency outputs can be differentiated into intermediate and final output. Intermediate or supply-side output is the amount of output supplied by a firm. Final or demand-side output is the amount of output that is consumed (Small and Verhoef 2007). Examples of intermediate output in transit include seat-miles and vehicle-hours of service. Final outputs include trips and passenger-miles.

Models that include intermediate output variables measure efficiency, as defined by FTA, but ignore the role and satisfaction of traveler demand. The use of final output variables allows the modeler to measure effectiveness. At the same time, haphazard analysis may result in a well-run agency serving an area with low demand being labeled as ineffective.[1]

Many rural transit agencies provide demand-response and long-distance fixed-route service. As these outputs and production technology are unique, a multiproduct framework appears appropriate.

---

[1] The availability of passenger miles by service type data would allow for the estimation of a cost function that controls for delivery of transit in low-density areas. However, this data is typically unavailable.

Previous studies of urban transit costs have estimated a variety of short- and long-run as well as variable and total cost functions depending on the availability of data, the modeler's assumptions, and the goals of the research. Some transportation cost models have employed 'quasi-fixed' inputs when firms are assumed to be in disequilibria as suggested by Caves, Christensen, and Swanson (1981). Oum and Zhang (1991) introduced a method for scaling capital stock, a quasi-fixed input, using a utilization factor when a firm has excess capital. Given the state's control of capital funding and the difficulty of disposing of unwanted vehicles, it seems unlikely that rural transit operates at long-run equilibrium.

Estimation of a long-run total cost function allows for the estimation of economies of scale and scope. Long-run total cost functions can either be estimated directly or be derived by solving for the optimal level of capital from a short-run total cost function (Keeler 1974). A firm's excess fleet can be calculated as the difference between the actual and optimal fleet size.

Inclusion of network size variables allows for the determination of two distinct types of economies of scale: economies of firm size and economies of density (Caves, Christensen, and Tretheway 1984). Economies of firm size are the cost savings resulting from a proportional increase in output and network size. Economies of density are the cost savings resulting from an increase in output while holding network size constant.

Economies of scope exist when a single firm is able to produce two or more products at a lower cost than multiple firms (Panzar and Willig 1977). Natural monopoly is the condition where a single firm can produce any combination of products at lower cost than any combination of firms. In the multiproduct case, subadditivity is a necessary and sufficient condition for natural monopoly (Evans and Heckman 1984).

1.2 Statement of the Problem

Knowledge of the economic structure of rural transit is necessary to design policies and allocate resources that ensure their efficient service provision.

The delivery of effective, efficient service is fundamental to the success and sustainability of federally funded rural transit programs. Regional and local coordination activities, including those mandated by federal policy or encouraged by state transit administrators, must work within existing organizational, economic, and political frameworks. With limited incentives, penalties, or evidence to support sweeping changes, the likelihood of a thorough overhaul of rural transit service such as consolidation of transportation providers or the transfer of operations from one agency to another is small. Furthermore, few planners have the specialized expertise in transportation economics necessary to address the situation independently.

The economic framework and methodology required to estimate the economic benefits of consolidation and to justify government intervention on the basis of cost structure are well-developed. However, the data required to analyze rural transit agencies are not readily available in most states. The economic costs of introducing, expanding, or modifying rural transit service can be estimated with knowledge of the industry's cost structure that can be found by fitting an appropriately specified and estimated cost function. This will provide information regarding economies of size, density, and scope. It can also be used to test for the presence of natural monopoly conditions and to measure excess capital.

Knowledge of the cost structure of rural transit is valuable to federal and state policymakers, state administrators, transit planners, rural transit agencies and other transportation organization managers, riders, and other regional mobility stakeholders. Each can benefit from an improved understanding of the underlying transportation cost concepts and estimated impacts of reorganization. Furthermore, federal policymakers are able to make changes to national transit policy to appropriately encourage significant reorganization of transit by state administrators.

1.3 Purpose of the Study

This study provides information and a decision-making framework for policymakers and administrators to design and administer polices to ensure that the limited public resources available to rural transit are properly allocated and used. It includes estimates of the benefits of consolidation of rural transit agencies as well as other service alternatives, including the decoupling of demand-response and fixed-route service and the modification of existing capital allocation policies. It also evaluates the justification of government subsidy of rural transit on the basis of its cost structure. The analysis uses data from North Dakota transit agencies that provide service to rural areas. North Dakota was selected because of the availability of necessary data, its large physical size, and the diversity of agencies that provide rural transit. Expansion to other states was considered. However, the high cost of data collection and the large, but undocumented and difficult-to-model state-level decisions that impact rural transit argued against expanding the scope geographically. Given the sensitivity of reorganization of rural transit and the added scrutiny it introduces, it was decided that estimating a cost structure and investigating the implications of reorganization for a single state would be better accepted. However, the concepts presented in the study should be helpful to efforts in other states in that cost estimates will accommodate high-level evaluation of the impacts from reorganization.

The objectives of the study are accomplished by answering the following research questions:

1. Is increased service in an existing service area more efficiently provided by a single existing rural transit agency or by a new one?
2. Is expanded service in a previously unserved area more efficiently provided by an existing single rural transit agency or by creating a second agency?
3. Are demand-response and fixed-response service more efficiently provided by a single firm or should two agencies provide each service exclusively?
4. Do rural transit firms have significant unused vehicle capacity?
5. Is a single regional transit agency always more efficient at providing multimodal service or are there cases where two agencies can provide service more efficiently?
6. Is there economic justification for government support of transit on the basis of increasing returns to scale or natural monopoly?

The study tests the following hypotheses:

1. Rural transit experiences economies of density. At a basic level, bus transit operation entails the movement of a vehicle over a roadway using labor, rolling stock, and fuel. It is plausible that the per-mile cost of such service is relatively constant. However, when planning, administration, and maintenance, which do not vary proportionally with output are considered, the presence of increasing returns to density is reasonable. The presence of increasing returns to density in rural transit infers that it would be more efficient for a single transit agency to provide service in an existing service area, as is common practice, rather than to create a new transit agency to do so.
2. Rural transit experiences economies of firm size. Economies of firm size differ from economies of density in that it considers a proportional increase in service area. There are costs incurred by expanding the geographic area of service including increased

communication costs and deadhead miles where vehicles travel unoccupied to locations where service is provided. However, indivisibilities in administration, facilities, and maintenance are expected to lead to cost savings that more than offset such increases. The presence of increasing returns to firm size infers that it is more efficient for an existing agency to provide increased levels of service in an expanded service area.

3. Rural transit experiences economies of scope. While the services are fundamentally different, rural transit agencies often deliver demand-response and fixed-route service using the same vehicles, drivers, facilities, administrators, and equipment. One exception is that demand-response service requires a scheduler/dispatcher to schedule trips, generate routes, and communicate with vehicle drivers. Recently, some rural transit agencies have implemented technology, computer-aided scheduling and dispatch (CASD) software, to assist with this function. The presence of economies of scope would validate the standard practice of having a single rural transit agency provide demand-response and fixed-route service.
4. Rural transit has excess capacity. Federal programs typically subsidize 80% of the cost of capital purchases and regulate the disposition of vehicles purchased with federal funds. The high cost of operation and the relatively light use of available vehicles, as evident from national statistics suggests that there are an excessive number of vehicles industry wide. The presence of excess capacity would support a revision of federal and state policies relating to capital funding allocation and rules for vehicle disposition.
5. Rural transit is a natural monopoly. The intuition for the presence of natural monopoly follows from the same arguments for economies of scale and scope, although these are not sufficient conditions for natural monopoly. The presence of natural monopoly in rural transit would support the existence of single rural transit agencies as the sole providers of demand-response and fixed-route service and for government subsidy of transit.

1.4 Summary

As the primary financial supporter of rural transit, the Federal Government has an interest in ensuring that service is delivered in an efficient, effective manner. Reorganization of rural transit agencies is one potential method of meeting these goals. Federal support for rural transit is administered by states that are given flexibility in developing and overseeing locally provided rural transit service. However, there is little guidance for reorganization, whether the reorganization is consolidation, assignment of new services, or reassignment of existing service.

Knowledge of rural transit's production technology can be used to estimate the costs of new or modified service and to determine the difference between the optimal and actual fleet size. This can be accomplished using a multiproduct translog cost function.

In this study, the cost structure of rural transit will be estimated, and the implications for consolidating rural transit agencies will be considered. Chapter Two surveys the transportation cost literature with an emphasis on research in the field of transit. Chapter Three describes the ideal and actual data set used, as well as presenting the methodology used to estimate the cost structure of rural transit. Chapter Four describes the results of the analysis and economic measures. Chapter Five explores the results of the analysis and their implications for consolidation and resource allocation.

2. TRANSPORTATION COST ECONOMICS

Determining the presence of natural monopoly in rural transit as a justification for government intervention, and evaluating the efficiency of alternative regional transit service configurations, requires knowledge of rural transit's cost function. Understanding of relevant economic theory and knowledge of model specification and empirical alternatives as well as their strengths and shortcomings is necessary to specify the correct theoretical and econometric model and variables, estimate the model, and calculate the appropriate economic measures.

In this chapter, transportation cost concepts, econometric models, empirical methods, and transit cost studies are reviewed. The intent is to provide the theoretical foundation, empirical structure, and context necessary for estimating the cost structure of rural transit to evaluate the economic justification of government intervention in rural transit and to evaluate different alternatives that ensure the efficiency of individual transit agencies and regional transit systems.

The literature review consists of three sections. The first section reviews economic cost concepts, with in-depth coverage of the mathematical conditions required for natural monopoly. Next, alternative functional forms used for estimating transportation costs are reviewed. Finally, seminal transit cost studies, and those that have considered multiproduct output and regionalization are presented.

2.1 Transportation Cost Concepts

Understanding transportation cost concepts is required to provide a theoretical foundation for analysis, guide the identification of appropriate econometric models, and empirical methods, assist in the interpretation of results and their economic implications, and aid in evaluating policy alternatives. The transportation cost concepts presented are well-accepted. They include concepts from economic cost theory, as well as specialized models for investigating the production of multiple outputs and network economies, where spatial attributes play a crucial role. Concepts related to the joint production of goods include economies of scope and natural monopoly. Network concepts include two versions of economies of scale: economies of firm size and economies of density.

The section begins with a brief review of production theory, a discussion of the dual relationship between production and cost functions, and consideration of alternative cost function frameworks. The topics of excess capacity and quasi-fixed inputs are discussed from a theoretical standpoint. Next, the definitions and equations for calculating economies of firm size, density, and scope are presented. Finally, the mathematical conditions for natural monopoly are reviewed.

2.1.1 Production Theory

McFadden (1978) presents a thorough overview of production and cost theory as well as the concept of duality. Single output firms are assumed to have a production function:

$$y = f(x, z) \tag{1}$$

where y is the level of output, x is the level of inputs, and z are other technological factors. A transformation function is an n-dimensional analog of the production function

$$T(X,Y) = 0 \quad (2)$$

where Y is a vector of outputs and X is a vector of inputs. The transformation function is equal to zero when the maximum amount of Y is produced with a given amount of X.

Direct estimation of production functions has been hindered in the past by endogeneity, as firms simultaneously determine levels of production and inputs. However, this condition is readily addressed by estimating a system of simultaneous equations.

If the production function is monotonically increasing in inputs, continuous, and concave, or if a transformation function is strictly concave, a unique cost function which is dual to the transformation function may be estimated (Shephard 1953). Duality ensures that all the economic information of a well-behaved production function is contained in the cost function.

2.1.2 Cost Function

Cost functions take the form:

$$C = C(W,Y,Z) = \min_X WX \ \ st\, T(Y,X,Z) \quad (3)$$

where C is the minimum cost of producing level of output, Y, given factor prices, W, and technological and environmental variables, Z. The cost function may be short- or long-run depending on the ability of the firm to vary the level of all inputs. A cost function is non-decreasing, homogeneous of degree one, concave, and continuous in factor prices (Varian 1992). Shephard's Lemma states that the first derivative of a firm's cost function with respect to the factor price is the factor demand equation (4). This property proves useful when estimating cost functions using certain flexible functional forms, such as the translog function, that will be discussed in the next section.

$$\frac{\partial C}{\partial w_i} = X_i(W,Y,Z) \quad (4)$$

A firm's factor share, S_i, the proportion of total costs used for a given factor input, i, can be calculated using (5) where w_i is the price of the factor input, x_i is the quantity used, and C is cost.

$$S_i = \frac{w_i x_i}{C} \quad (5)$$

A firm is said to be in the short-run if the quantity of at least one input is held fixed. In the short run, the firm chooses variable inputs to minimize costs, subject to the amount of capital. The firms' short-run cost function shows the minimum cost of producing any output level given the amount of the fixed factor. It includes costs that vary with output (variable costs) and those that do not (fixed costs).

$$SRC = VC(W,Y,Z,K) + \mathrm{r}\mathrm{K} \quad (6)$$

Variable cost functions are short-run considerations and include the cost of all variable inputs used by the firm, but not those associated with fixed inputs. Economic measures, including economies of size and scope, are typically considered to be long-run cost (LRC) concepts. Consequently, they require knowledge of a firm's long-run cost function, which may be directly estimated, or calculated using the short-run cost function.

Deriving the long-run cost function requires determining the optimal level of capital k*. This is done by differentiating equation (6) with respect to capital, K, setting the result equal to zero, and solving for K^*.

$$\frac{\partial SRTC}{\partial K*} = \frac{\partial VC}{\partial K*} + r = 0 \quad (7)$$

The long-run cost function is then found by substituting K^* into the SRC equation (6).

$$LRC = VC'(W, Y, Z, K^*) + r \times K^* \quad (8)$$

2.1.3 Excess Capacity

In the short run, firms are unable to adjust their level of fixed inputs. This can result in excess capacity when capital is underutilized, or firms may have inadequate capacity, where additional capital would result in reduced costs. The case of excess capacity is shown by point A in Figure 21. Here a firm using K' units of capital to produce Y^* units of output is not in long-run equilibrium as the short-run average total cost is greater than the long-run average total cost (Nelson 1989). To be in long-run equilibrium, a firm would need to produce Y^* units of output using K^* units of capital, point B, or Y' units of output using K' units of capital, point C, as these points are located on the long-run average cost curve. The firm's excess capital at point A is the difference between K' and K^*.

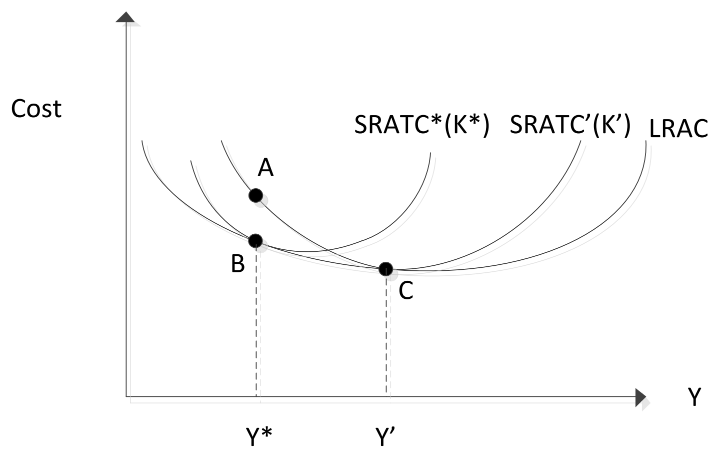

Figure 2.1 Optimal Level of Capital

Caves, Christensen, and Swanson (1981) suggest that variable cost functions be estimated with capital modeled as a quasi-fixed input to remove bias caused by firms' excess capacity. That is, assuming that firms are in long-run equilibrium when they are not may result in bias. Despite the seeming improvement introduced by using this approach, several authors have found that capital has no effect or a positive effect on variable costs. This seems to suggest that the marginal product of capital is negative. Because of this finding, Oum and Zhang (1991) investigate the issue in more detail. Using standard economic theory, Oum and Zhang show that short-run variable costs flatten out once a firm has excess capacity, at points beyond K_0 (Figure 2.2). The intuition is that more capacity increases the productivity of other variable factors when it is used, but has no impact on costs when it is in excess of what is used. This explains a lack of significance in the impact of K on variable costs when some firms in a sample are at excess capacity. Moreover, they also show that fixed cost flattens out once excess capacity is realized. The reason is because the rental cost of capital is still incurred, but use-related depreciation is not.

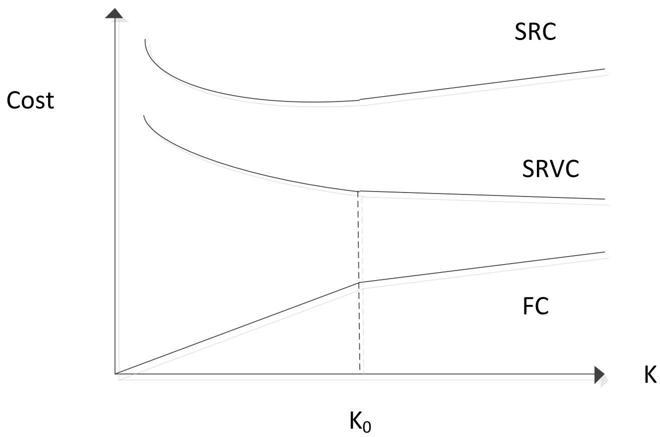

Figure 2.2 Short-Run Variable, Short-Run Total, and Fixed Costs

2.1.4 Economies of Size and Density

Several authors have shown that there are two distinct ways in which transportation firms can expand output. One way is to transport more passengers or goods over a fixed network. Another is to expand the size of the network. Consequently, a properly specified transportation function should include a measure of network size, N.

$$C = C(W, Y, Z, N) \quad (9)$$

Returns to density are defined as the increase in output resulting from a proportional increase in input holding input prices and network size constant. It is calculated as the inverse of the elasticity of cost with respect to output, ε_y,

$$RTD = \frac{1}{\varepsilon_y} \quad (10)$$

Returns to density are said to be increasing, decreasing, or constant if RTD is greater than, less than, or equal to 1.

Returns to size are defined as the increase in output from a proportional increase in input and network size keeping input prices constant. Mathematically, this is represented by equation (11), where ε_N is the elasticity of cost with respect to network size,

$$RTS = \frac{1}{\varepsilon_y + \varepsilon_N} \quad (11)$$

As with returns to density, returns to size are said to be increasing, decreasing, or constant if RTS is greater than, less than, or equal to 1.

Returns to size and density can be calculated from the variable cost function by subtracting the elasticity of variable cost with respect to capital, ε_k,

$$RTD = \frac{1-\varepsilon_k}{\varepsilon_y} \quad (12)$$

$$RTS = \frac{1-\varepsilon_k}{\varepsilon_y + \varepsilon_N} \quad (13)$$

These same calculations can be utilized in a multiproduct framework using the concept of multiproduct economies of scale (Baumol,Willig, and Panzer 1988). Multiproduct economies of scale that describe changes in cost resulting from a proportional change of all outputs can be calculated as

$$S = \frac{C(y)}{\sum_{i=1}^{1} y_i \frac{\partial C}{\partial y_i}} \quad (14)$$

where y_i are unique outputs. Multiproduct economies of scale are said to be increasing, decreasing, or constant if S is greater than, less than, or equal to 1.

2.1.5 Economies of Scope

Economies of scope exist when a single firm is able to produce two or more products at a lower cost than multiple firms (Panzar and Willig 1981). Global economies of scope for the two-product case, which considers the joint production of all outputs, is calculated by

$$SC = \frac{C(y_1,0)+C(0,y_2)}{C(y)} \quad (15)$$

Global economies of scope exist if

$$C(y_1, y_2) < C(y_1, 0) + C(0, y_2) \quad (16)$$

that is, if the cost of jointly producing y_1 and y_2 is less than the cost of disjoint production by two firms.

Product-specific economies of scope measure the proportional increase in cost resulting in the production of all outputs except the m^{th}.

$$SC_m = \frac{C(y^{(m)})+C(y^{(-m)})-C(y)}{C(y)} \quad (17)$$

Product-specific economies of scope exist if SC_m is greater than zero.

2.1.6 Natural Monopoly

Natural monopoly is the condition where a single firm can jointly produce multiple outputs at a lower cost than any combination of firms. In the single product case, the existence of increasing returns to scale is a sufficient condition for a firm to be a natural monopoly. For the multiproduct case, neither economies of scale nor scope are sufficient.

The concept of natural monopoly is fundamental to our investigation of the efficiency of the joint production of fixed-route and demand-response transit service, and to the efficiency of consolidation of smaller systems into regional systems. As will be discussed later in the chapter, nearly all transit cost studies that have considered consolidation have limited their analysis to economies of scale and scope; however, that is not sufficient evidence of the presence of natural monopoly.

Sharkey (1982) and Baumol, Willig, and Panzer (1988) present five separate sufficient conditions for cost subadditivity. That is, any of the five is sufficient. Before reviewing them, the cost concepts on which they rely are presented.

Average incremental cost is the change in total costs resulting from the production of a new type of output, divided by the new output added.

$$AIC_i = \frac{C(y) - C(y_{-i})}{y_i} \quad (18)$$

where y_{-i} is all outputs other than the added ith output y_i.

Ray average cost is the average cost of a bundle of outputs as output is proportionally increased:

$$RAC = \frac{C(ty)}{t} \quad t>1 \quad (19)$$

Ray average costs are declining if

$$\frac{C(y)}{y_i} \geq \frac{C(ty)}{ty_i} \quad \text{for } t>1 \text{ and any } i \quad (20)$$

where i is a bundle of outputs. Ray average costs are the slope of the line intersecting the origin and the cost function at point z in Figure 2.3.

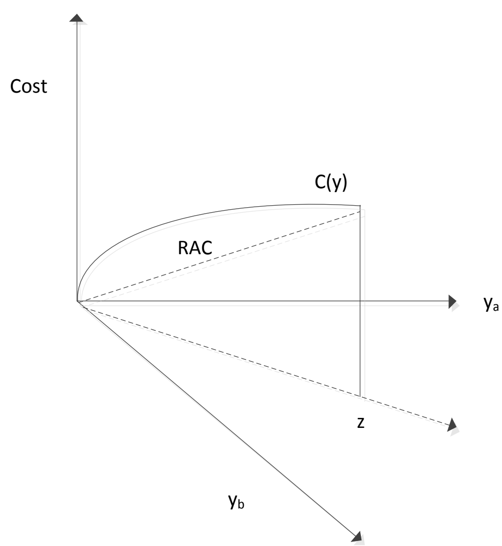

Figure 2.3 Ray-average Cost

Trans-ray convexity is the condition where the cost of producing any combination of two goods is less than the cost of their disjoint production. Mathematically, convexity is defined as

$$C(ty_a + (1-t)y_b) \leq tC(y_a) + (1-t)C(y_b) \quad (21)$$

where $0 \leq t \leq 1$. Trans-ray convexity is presented graphically in Figure 2.4.

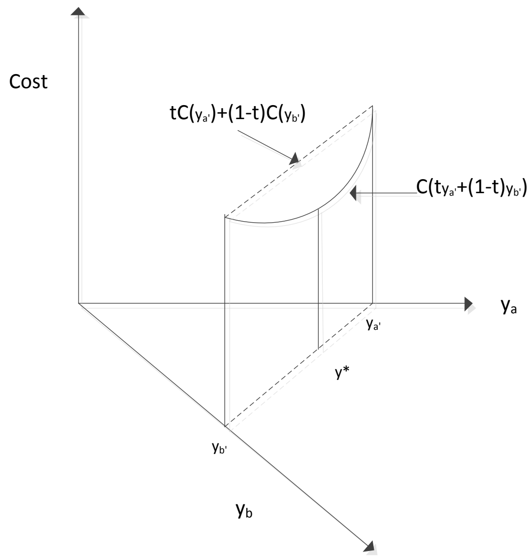

Figure 2.4 Trans-ray Convexity

Trans-ray supportability exists if, for any cross section of the cost function, a hyperplane that supports C(y*) is below any other combination of outputs on the cross section.

Supportability requires that the entire cost surface be supportable for all outputs below y* in order to be supportable at y*. That is, the entire cost surface would be above a hyperplane that passes through the origin and C(y*).

A cost function exhibits *cost complementarity* if the marginal cost, with respect to any output, declines when another output increases. That is

$$\frac{\partial^2 C}{\partial y_i \partial y_j} < 0 \tag{22}$$

Strict cost subadditivity is defined mathematically as:

$$C(\textstyle\sum_{i=1}^{n} y_i) < \sum_{i=1}^{n} C(y_i) \tag{23}$$

where the cost of producing multiple outputs is strictly less than the cost of multiple firms producing them disjointly.

The five separate sufficient conditions for the presence of a natural monopoly are (any one of the five implies subadditivity):

1. Decreasing average incremental costs for each product up to y^* and economies of scope at y^*.
2. Decreasing ray average costs up to y^* and trans-ray convexity of costs up to the hyperplane crossing through y^*.
3. Decreasing ray average costs up to the y* hyperplane and trans-ray supportability at y*.
4. Supportability up to y^*.
5. Strong cost complementarity.

Because each of the sufficient conditions is stronger than the strict subadditivity condition itself, a direct approach is preferred.

2.1.7 Testing Cost Subadditivity

Evans and Heckman introduce a test for subadditivity that allows for testing of the presence of natural monopoly in the multiproduct case. This is accomplished by comparing the cost of joint production of output by a monopoly firm with disjoint production by two hypothetical firms comprising various portions of output for all observations in a sample (1984). Two constraints are imposed on the region of outputs provided by hypothetical firms, limiting the minimum level of output and the ratio of output combinations to those observed in the data.

For the two-product case, Firms A and B produce

$$\tilde{y}_t^a = (\,\Phi\tilde{y}_1^*\,, \omega\tilde{y}_2^*\,) \tag{24}$$
$$\tilde{y}_t^b = (\,(1-\Phi)\tilde{y}_1^*\,, (1-\omega)\tilde{y}_2^*\,) \tag{25}$$

where $0 \leq \Phi \leq 1, 0 \leq \omega \leq 1$ and $\tilde{y}_1^*$ and $\tilde{y}_2^*$ are the levels of output observed in the data.

Mathematically, the test for subadditivity is defined by

$$Max_{(\Phi,\omega)}\, Sub_t(\Phi,\omega) = \left[\tilde{C}_t - \tilde{C}_t^A(\Phi,\omega) - \tilde{C}_t^B(\Phi,\omega)\right]/\tilde{C}_t$$

where

$$\tilde{C}_t^A(\Phi,\omega) = \tilde{C}(\tilde{y}_t^a)$$
$$\tilde{C}_t^B(\Phi,\omega) = \tilde{C}\left(\tilde{y}_t^b\right)$$
$$\tilde{C}_t = \tilde{C}\left(\tilde{y}_t^a + \tilde{y}_t^b\right) = \tilde{C}(\tilde{y}_t) \tag{26}$$

The test compares the cost of producing $\tilde{y}_t^a$, $\tilde{y}_t^b$, and $\tilde{y}_t$ levels of output by Firms A, B, and a monopoly firm by calculating costs across a region by varying the values of Φ and ω. The test considers each observation in the sample. The parameters Φ and ω typically take on the values (.1, .2, .3…, .9) and the long-run cost function is used. If the quantity $Max_{(\Phi,\omega)}\, Sub_t(\Phi,\omega)$ is negative and statistically significantly different than zero, the hypothesis that the cost function is not subadditive is rejected. Evans and Heckman apply the method to a two-product, two-firm case in the telephone industry and rejected the presence of subadditivity and natural monopoly. Shin and Ying (1992) use pooled cross-sectional time series data to examine subadditivity. This allows them to address a shortcoming of the use of highly correlated time series data that makes it difficult to distinguish between economies of scale and technological change. Their test differs from Evans and Heckman's as it does not place restrictions on the admissible region given the greater variability in observed levels of output. They also consider three measures of output. The analysis finds lower costs for the monopolist in only 20 and 38 percent of combinations in 1976 and 1983 and the condition of subadditivity was not met for any observation in the data.

Bitzan tests for the presence of subadditivity to investigate the cost structure of the rail industry and the implications of its structure after reorganization (1999). He conducts the test twice. First, the test models two firms jointly providing service to a fixed network area equal in size to the sample mean. Next, the concept of *overall subadditivity* is tested by allowing the proportion of the network served by the firms to vary. Bitzan models an exhaustive number of combinations of output and network allocations between the two firms for each observation in his sample. The study finds evidence of natural monopolies for railroads over a fixed-sized network, but not in serving multiple markets.

2.2 Multiproduct Econometric Models

General flexible form models, defined as those that do not place restrictions on first and second order derivatives, have dominated transportation cost studies since their introduction more than three decades ago. They quickly replaced the Cobb-Douglas cost function (1928) which imposes restrictions on the structure of costs, such as homogeneity in outputs, constant elasticity of substitution, and relative marginal costs of different outputs not depending on input prices. Early flexible functions, such as the Leontief proposed by Diewert (1971), provided increased flexibility in estimating elasticities of substitution, but still imposed homogeneity in outputs. The quadratic multiproduct cost function introduced by Lau (1974) takes the form

$$C = \alpha_0 + \sum_{i=1}^{m} \alpha_i y_i + \sum_{j=1}^{n} \beta_j w_j + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} \alpha_{ij} y_i y_k + \frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{n} \beta_{jl} w_j w_l + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} \gamma_{ij} y_i w_j \qquad (27)$$

A primary advantage of the quadratic multiproduct cost function is its accommodation of observations where output is zero. The quadratic suffers a number of shortcomings, however. Homogeneity in factor prices may not be imposed without eliminating the flexibility of the model. The quadratic cost function also includes a large number of parameters that must be estimated requiring a relatively large sample.

The multiproduct translog function takes the natural logarithm the dependent and independent variables

$$\ln C = \alpha_0 + \sum_{i=1}^{m} \alpha_i \ln y_i + \sum_{j=1}^{n} \beta_j \ln w_j + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} \alpha_{ik} \ln y_i \ln y_k + \frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{n} \beta_{jl} \ln w_i \ln w_j + \sum_{i=1}^{m}\sum_{j=1}^{n} \gamma_{ij} \ln y_i \ln w_j \qquad (28)$$

However, the multiproduct functional form is unable to accommodate observations with zero output as the natural logarithm of zero does not exist. Caves, Christensen, and Tretheway (1980) remedy this with the Generalized Translog Multiproduct Cost Function (GTMFC). The GTMCF uses a Box-Cox transformation (Box and Cox 1964)

$$\frac{\left(y_i^{\lambda} - 1\right)}{\lambda} \qquad (29)$$

the value of which approaches the natural logarithm as λ approaches zero.

The GMTCF, like the translog function, imposes symmetry,

$$\alpha_{ij} = \alpha_{ji}, \gamma_{ij} = \gamma_{ji} \qquad (30)$$

homogeneity of degree 1,

$$\sum_{i=1}^{n} B_i = 1,\ \sum_{i=1}^{n} \gamma_{ij} = 0\ for\ i = 1,2, \dots n, \\ \sum_{i=1}^{n} \beta_{ij} = 0\ for\ i = 1,2, \dots m, \qquad (31)$$

and utilizes cost-share equations (Shephard 1953)

$$S_i = \frac{w_i x_i}{C} = \frac{\partial lnC}{\partial lnw_i} = \beta_i + \sum_{j=1}^{n} \beta_{ij} \ln w_j + \sum_{j=1}^{m} \gamma_{ij} \ln y_j \qquad (32)$$

to increase the model's degrees of freedom without increasing the number of parameters. The GMTCF is estimated as a system of equations consisting of (28) and cost-share equations for each of the factor inputs (32), less one, to avoid singularity using the Seemingly Unrelated Regression (SUR) technique (Zellner 1962).

The GTMCF has a number of benefits over alternative functional forms, including the quadratic. It places no a priori restrictions on returns to scale or elasticities of substitution. The use of the log metric allows for the imposition of homogeneity restrictions. Including variables divided by their sample mean allows for parameter estimates to be interpreted as elasticities when all values are at the mean levels. The use of cost share equations and imposition of symmetry and homogeneity greatly reduce the number of parameters that must be estimated. However, Pulley and Humphrey (1993) note that the use of the Box-Cox transformation could result in errors when estimating scope economies.

2.3 Transit Cost Studies

While more than 100 transit cost studies have been published in the past three decades, none have focused solely on rural transit in the United States. However, many of the studies are helpful for our purposes; especially as they inform model specification and estimation decisions. Furthermore, results from urban analysis can be used for comparison. Recent research in Europe on transit regionalization is also informative.

In this section, transit cost studies are reviewed. Focus is placed on multiproduct and regional studies that are most related to our project. Empirical considerations including the measurement of transit output measures, calculating the cost of capital, and deriving the long-run cost curve are reviewed. Estimates of transit production technology: economies of density, size, and scope are presented. Finally, multiproduct and regionalization studies are reviewed.

2.3.1 Transit Output Measures

Measures of transit output vary across the cost literature. Few studies describe the reasoning behind their choice of measures. Furthermore, many studies have used robust data sets, such as the National Transit Database (NTD), that include intermediate and final output. Data availability has not been the sole factor in output choice.

There has been limited commentary on the issue of intermediate and final outputs in the transit cost literature. Small and Verhoef (1990) point out that studies interested in production efficiency use intermediate outputs, while those interested in effectiveness use final outputs. Button and O-Donnell (1985) criticize the use of intermediate outputs as they fail to differentiate operating conditions or capture the economic motive to providing service. Using that same reasoning, Hensher explicitly rejects the use of vehicle-kilometers using passenger revenue instead (1988). Fraquelli, Piacenza, and Abrate (2004) discuss the issue and use an aggregate measure of intermediate output introduced by Gagnepain and Ivaldi (2003) - that is, the product of vehicle-kilometers, load capacity, and locations served.

Vehicle distance traveled, an intermediate measure, is the most commonly used transit output measure. Viton (1981), Obeng, Talley, and Colburn (1987), Karlaftis, McCarthy, and Sinha (1999), Karlaftis and McCarthy (2002), use vehicle-mile data from the National Transit Database. Vehicle-miles by mode are used by Viton (1992), Talley (1986), Colburn and Talley (1992) in their multiproduct studies. Farsi, Fetz, and Filippini (2007) use seat-kilometers, the product of vehicle distance traveled by the size of the fleet. Obeng (1985) uses passenger-trips, a final output, as a measure of output. Berechman (1983) uses gross revenue. Harmatuck (2005) uses passenger-trips and passenger-miles.

Some studies have used both intermediate and final outputs when modeling transit costs. Berechman and Guiliano (1984) use vehicle-miles and revenue-passengers. Berechman (1987) uses passenger-trips and vehicle-kilometers.

The choice of output variables when modeling rural transit is a sensitive one. The communities that these agencies serve are diverse with varying populations, population density, travel demands, and geography. Comparing the amount of output consumed across agencies may lead to a belief that a particular agency is ineffective at providing service while in reality its performance is a function of the characteristics of the community it serves. Conversely, the role of government is not to finance the efficient (that is, lowest cost) movement of empty buses, but to improve the mobility of individuals (Federal Transit Administration 2005). Taken to an extreme; however, using effectiveness as the sole guide to allocate and organize resources would mean that some parts of the country would go without transit service due to low demand and challenging geography. So ultimately a balance must be struck, one that is political, not economic in nature.

Returning to economics and model specification, perhaps the most reasonable solution is that supported by Berechman (1993) where intermediate and final outputs are modeled, with a thorough consideration and discussion of the role of demographics, demand, and geography.[2]

2.3.2 Short-Run Capital Stock Coefficients

Most studies that have estimated variable costs (Viton 1981; Obeng 1985, 1987; Karlaftis, McCarthy, and Sinha 1999; Karlaftis and McCarthy 2002; Fraquelli, Piacenza, and Abrate 2004) have estimated a positive coefficient of capital. Authors have justified a number of reasons for this phenomenon. Filippini (1996) conjectures that this is due to multicollinearity between cost and capital measure variables. Windle (1988) argues that the result occurs because agencies do not minimize costs in the long run and consequently employ too much capital. Levaggi (1994) believes that the situation arises from government subsidy of capital purchases. However, none had attempted to address the situation using the method suggested by Oum and Zhang (1991).

2.3.3 The Price of Capital and Derivation of the Long-Run Cost Function

The price of capital is required to estimate a firm's long-run cost function. The price should capture the economic cost of capital, including its opportunity cost and depreciation. However, many transit cost studies have used accounting measures, for example, annual total capital costs divided by fleet size, that do not capture the true economic cost.

[2] In this study, intermediate outputs are used since final output (e.g. passenger miles) are not available.

Viton (1981) calculates the price of capital as the purchase price of rolling stock multiplied by the capital recovery factor,

$$CRF = \frac{r(1+r)^n}{(1+r)^n - 1} \qquad (33)$$

where r is the interest rate and n is the life of the capital good in years. Obeng (1987) uses a weighted arithmetic mean to obtain vehicle capital prices:

$$w_k = w * (D + r) * \exp(-DA) \qquad (34)$$

where w is new bus prices, D is the straight line depreciation rate, A is the average age of the fleet, and r is the interest rate. This measure is similar to Nelson's (1972)

$$P_i = N_i(1 - \mathrm{C})V_{0i}\exp(-\delta A)\delta \qquad (35)$$

where the price, P, in period i is a function of fleet size, N, an adjustment factor, C, for the maximum amount of capital expenses paid by FTA, the price of a new vehicle, V, depreciation, δ, and average fleet age, A.

Keeler estimates the long-run total cost function by solving for the optimal level of capital from a short-run total cost function (1974). This method has been used in the transit cost literature by Viton (1981), Obeng (1985), and Karlaftis, McCarthy, and Sinha (1999). Viton finds that the optimal level of capital is relatively invariant to price (1981). Comparing actual levels of capital to the optimal amount, Viton (1981) estimates that transit agencies have 57% excess capacity while Karlaftis, McCarthy, and Sinha (1999) estimate excess capacity of 42%.

2.3.4 Economies of Scale, Size, and Density

Most published transit cost studies in the past three decades have included a measure of economies of scale. Many have included a network size variable that allows for economies of size and density, types of economies of scale, to be distinguished. As measured in some previous studies, economies of scale are the cost savings resulting from a proportional increase in output when network size is not accounted for in the model. Economies of firm size are the cost savings resulting from a proportional increase in output and network size. Economies of density are the cost savings resulting from an increase in output while holding network size constant. Estimates of economies of scale, size, and density have varied across the literature.

Viton finds decreasing returns to scale in domestic urban transit (1981), while a study modeling short-run variable costs by Karlaftis and McCarthy (2002) finds the measure to be increasing. Berechman and Guiliano (1984) have mixed results for the urban domestic transit industry with decreasing returns to scale when passenger-trips were used as the measure of output, but increasing returns to scale for vehicle-miles. Colburn and Talley (1992) find increasing returns to scale when modeling a single U.S. multiservice transit agency. Fraquelli, Piacenza, Abrate (2004) find that Italian public transit exhibits increasing returns to scale. Berechman (1987) find evidence of short-run economies of scale in the Israeli bus industry. Small British transit agencies show evidence of economies of scale, but diseconomies for larger ones (Button and O'Donnell 1985). De Borger (1984) estimates the economies of scale for Belgian transit, noticing they change dramatically over time. Multiproduct returns to scale are present in Swiss transit agencies that provide rail, bus, and trolley service (Farsi, Fetz, and Filippini 2007). However, only rail service is found to have product-specific returns to scale.

For studies that include a network variable, Obeng (1984) identifies decreasing returns to size for domestic transit, while Windle's analysis (1988) reports constant returns to size. Significant economies of size are present in Swiss transit (Filippini, Maggi, and Prioni 1992, Filippini and Prioni 2003). Cambini, Piacenza, and Vannoni (2007) find increasing returns to size in Italian transit. Spanish transit has short-run economies of size, but not long-run (Matas and Raymond 1998). Ottoz, Fornengo, and Di Giacomo (2008) find increasing returns to size for privately operated bus service providers and decreasing returns to size for public ones.

Studies show that returns to density are, for the most part, increasing (Cambini, Piacenza, and Vannoni 2007, Filippini, Maggi, and Prioni 1992, Filippini and Prioni 2003, Karlaftis 1999, Obeng 1984, Wang Chiang and Chen 2005, Windle 1988). However, Karlaftis, McCarthy, and Sinha (1999) show that increasing returns exist for medium and large transit agencies, but not for small agencies or for those that serve suburban areas. Increasing returns to density exist for privately operated bus service providers and decreasing returns to size are present for public firms (Ottoz, Fornengo, and Di Giacomo 2008). Table 2.1 summarizes the measures of output and economies of scale, density, and size of previous transit cost studies. Note that economies of size and density are types of economies of scale that may be distinguished when a network size variable is included.

2.3.5 Multiproduct Analyses and Consolidation

Cost studies of multiproduct transit agencies are especially helpful as rural transit agencies often provide more than one service. At the same time, studies on transit regionalization or consolidation are illuminative as these changes in structure are often discussed as potential opportunities to increase system efficiency.

A Generalized Translog Multiproduct Cost Function was used by Appelbaum and Berechman (1991) to estimate the costs of the Israeli bus sector that provides traditional bus and specialty services (e.g. tours, charters). They measure output in terms of passenger-trips and vehicle-kilometers for both modes. They find significant economies of scale, but do not explicitly consider economies of scope between the modes.

Viton (1992) uses a quadratic model to estimate the costs of the U.S. urban transit industry including fixed-route, rail, streetcar, trolley bus and demand-response service. He investigates a number of combinations of San Francisco area transit agencies. He finds that economies of scope depend on changes in wage rates. If wage rates remain unchanged, economies exist; however, if wages rise, these economies will disappear or become negative. Product-specific economies of scope exist for all modes except fixed-route. While there were beneficial combinations of existing transit agencies, a merger of all regional providers would raise total system costs.

Table 2.1 Economies of Scale, Size, and Density: Empirical Findings

Author	Location	Function	Output	Economies of Scale	Economies of Density	Economies of Size
Viton (1981)	United States	short-run	vehicle-miles		increasing	increasing
		long-run		decreasing		
Berechman and Guiliano (1984)	United States	long-run	vehicle-miles	increasing		
			passenger-miles	decreasing		
De Borger (1984)	Belgium	short-run	vehicle-miles	increasing (short-run)		
Obeng (1984)	United States	short-run	passenger-miles		increasing	
		long-run	passenger-miles	decreasing		
Button and O'Donnell (1985)	Britain	long-run	passengers, revenue	increasing (small); decreasing (large)		
Berechman (1987)	Israel	short-run	passenger-trips, vehicle-	increasing		
Windle (1988)	United States	long-run	passenger-miles		increasing	constant
Talley (1989)	United States	long-run	vehicle-miles			
Appelbaum and Berechman (1991)	Israel	long-run	passenger trips, vehicle-miles	increasing		
Colburn and Talley (1992)	United States	long-run	vehicle-miles	increasing		
Filippini, Maggi, and Prioni (1992)	Switzerland	long-run	passenger-kilometers, seat-kilometers		increasing	increasing
Viton (1992)	United States	long-run	vehicle-miles	increasing		
Matas and Raymond (1998)	Spain	short-run	vehicle-kilometers		increasing	increasing (short-run);
		long-run				decreasing (long-run)
Karlaftis, McCarthy, and Sinha (1999)	United States	short-run	vehicle-miles		increasing	increasing
		long-run			(large agencies)	
Karlaftis and McCarthy (2002)	United States	short-run	vehicle-miles	increasing		
Filippini and Prioni (2003)	Switzerland	long-run	vehicle-kilometers, seat-		increasing	increasing
Fraquelli, Piacenza, and Abrate (2004)	Italy	short-run	kilomters*stops	increasing		
Wang Chiang and Chen (2005)	Taiwan	long-run	vehicle-kilometers		increasing	increasing
Cambini, Piacenza, and Vannoni (2007)	Italy	short-run	vehicle-kilometers, seat-kilometers, total seats-kilometers		increasing	increasing
Farsi, Fetz, and Filippini (2007)	Switzerland	long-run	seat-kilometers	increasing		
Ottoz, Fornengo, and Di Giacomo (2008)	Italy	short-run	vehicles*kilometers		increasing	increasing

A more thorough consideration of combinations of Bay Area regional providers is found in Viton's 1993 paper. He finds that combinations involving Bay Area Rapid Transit (BART) are usually beneficial. There are significant benefits from a number of consolidations involving two or three agencies. His findings support the formation of larger multimodel transit operations. As not all combinations led to lower costs, thorough evaluation of each alternative is recommended.

Multiservice transit costs are modeled by Talley (1989) using data from a single agency. He finds that external competition or internally provided services can result in competitive downward pressures on unionized labor and wages. The provision of paratransit or contract services can result in lower costs.

Colburn and Talley (1992) investigate the cost structure of an urban multiproduct transit agency that provides bus, elderly and handicapped, van pool, and dial-a-ride service. They use a method introduced by Deller, Chicoine, and Walzer (1988) to determine the presence of cost complementarities by approximating $\partial^2 C/\partial Y_i \partial Y_j$. If that value is negative for $i \neq j$, then economies of scope exist. For a multiproduct translog cost function this was approximated by $\alpha_i\alpha_j+\alpha_{ij} <0$. They found the condition to exist for four of the six services the agency provides. Combinations involving dial-a-ride service had higher costs. The results support the reorganizing of the agency into two firms: one providing dial-a-ride service and another providing bus, elderly and handicapped, and van pool service.

The joint production of three services: fixed-route, trolley and rail in Switzerland is investigated by Farsi, Fetz, and Filippini (2007). They estimate a long-run cost function using the quadratic functional form and found evidence of global economies of scale and scope. They use product-level economies of scope to evaluate the efficiencies resulting from combining individual modes. They find significant product-level economies of scope for all modes at the 1st quartile of output and at the median for combinations involving trolleys.
The cost function of Italian transit agencies as estimated by Fraquelli, Piacenza, and Abrate (2004) finds increasing returns to scale. They use an aggregate measure of output introduced by Gagnepain and Ivaldi (2003): the product of vehicle-kilometers, load capacity, and locations served. Consequently, economies of scope were not measured. However, the authors do compare the cost of joint service to urban and rural areas and find that those that serve both have lower costs supporting the mergers of such agencies.

Cambini, Piacenza, and Vannoni's (2007) analysis calculates that global economies of scope exist for Italian transit agencies and for modal combinations that involve bus service. The study suggests that cost savings can result from mergers of urban and intercity operators. They discuss using bidding as a mechanism to merge agencies.

2.4 Summary

The review of transportation cost concepts, econometric models, and empirical methods provide a theoretical foundation and context for current research. They serve as a basis for the development of the study's methodology, which is presented in the next chapter.

3. METHODS AND DATA

The objective of this study is to evaluate the justification of government support of rural transit on the basis of the presence of natural monopoly and to determine the most efficient regional organization of transit. The previous chapter explained relevant economic theory, discussed alternative multiproduct functional forms, and findings of previous transit cost studies. This chapter explains the empirical methods and describes the data used to estimate the cost structure of rural transit.

A Generalized Translog Multiproduct Cost Function is used to estimate rural transit short-run variable costs. Long-run costs are obtained by finding the optimal level of capital. Next, the presence of economies of size, density, and scope are calculated using existing measures. Finally, the existence of natural monopoly is evaluated by testing for subadditivity.

3.1 Theoretical Model

Rural transit agencies are assumed to be in the short-run as they are unable to easily purchase or dispose of vehicles and have contractual obligations to provide service. Production is assumed to be impacted by environmental and technological conditions and network characteristics. The rural transit transformation function takes the form

$$T(Y, X, Z, N, K) = 0 \qquad (36)$$

where Y is a vector of outputs, X is a vector of variable inputs, Z is a vector of environmental and technological variables, N are network variables, and K is the level of capital.
The short-run rural transit cost function

$$\begin{aligned} C &= C(W, Y, Z, N, K) \\ &= \min_x WX \quad st\; T(Y, X, Z, N, K) = 0 \end{aligned} \qquad (37)$$

identifies the lowest cost, C, of producing output Y, given factor prices, W, environmental and technical variables, Z, the firm's network, N, and capital level, K.

A short-run variable cost function is typically posited to take the form

$$VC = \mathrm{V}C(\mathrm{W}, \mathrm{Y}, Z, N, K) \qquad (38)$$

However, given the presence of excess capacity for many rural transit agencies, the inclusion of K would result in bias, as the effect of K on variable costs is truncated at the point K^0 as discussed in the previous chapter (Oum and Zhang 1991). This situation is remedied by modeling capital use $S = \mu k$ where μ is the observed capital utilization rate.

The short-run variable cost function of transit agencies serving rural communities is

$$VC = \mathrm{V}C(w_l, w_f, w_m, \mathrm{Y}_{dr}, \mathrm{Y}_{fr}\; Z, N, S(K)) \qquad (39)$$

where VC is total operating costs; w_l is the price of labor; w_f is the price of fuel; w_m is the price of maintenance; Y_{dr} is demand-response miles; Y_{fr} is fixed-route miles; Z is technological attributes; N is network size; and S(K) is the level of capital use.

3.2 Empirical Model

The Generalized Translog Multiproduct Cost Function introduced by Caves, Christensen, and Tretheway (1980) is used to model variable costs. This function allows for the inclusion of outputs with zero values, as some transit agencies do not provide demand-response or fixed-route service. The model is defined as:

$$\ln VC = \alpha_0 + \sum_{i=1}^{n} \beta_i \ln w_i + \sum_{j=1}^{m} \alpha_j \ln Y_j + \eta \ln S + \omega \ln N + \sum_{k=1}^{p} \epsilon_o \ln Z_k + \frac{1}{2}\sum_{j=1}^{m}\sum_{l=1}^{m} \zeta_{jl} \ln Y_j \ln Y_l + \frac{1}{2}\sum_{i=1}^{n}\sum_{q=1}^{n} \gamma_{iq} \ln w_i \ln w_q + \frac{1}{2}\theta \ln S \ln S + \frac{1}{2}\iota \ln N \ln N + \frac{1}{2}\sum_{k=1}^{p}\sum_{r=1}^{p} \kappa_{kr} \ln Z_k \ln Z_r + \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_{ij} \ln w_i \ln Y_j + \sum_{j=1}^{m} \psi_j \ln Y_j \ln S + \sum_{j=1}^{m} \nu_j \ln Y_j \ln N + \sum_{j=1}^{m}\sum_{k=1}^{p} \xi_{jk} \ln Y_j \ln Z_k + \sum_{i=1}^{n} o_i \ln w_i \ln S + \sum_{i=1}^{n} \pi_i \ln w_i \ln N + \sum_{i=1}^{n}\sum_{k=1}^{p} \rho_{ik} \ln w_i \ln Z_k + \sigma \ln S \ln N + \sum_{k=1}^{p} \tau_k \ln S \ln Z_k + \sum_{k=1}^{p} \upsilon_k \ln N \ln Z_k \quad (40)$$

where VC is the variable cost associated with producing output, Y_j, using inputs with prices w_i. Factor prices are the price of fuel, labor, and maintenance and materials. Capital use, S, is used to measure the agency's use of the capital. A network attribute, N, service area measured in square miles, is included to accommodate the impacts of the spatial environment and service obligations on production costs. Technology variables, Z_k, include riders per mile, the ratio of elderly and disabled riders to total riders, the year, the average number of seats of agencies vehicles (a proxy for vehicle size), and a dummy variable for transit agencies that employ full-time administrators.[3]

Differentiating (40) with respect to the natural logarithm of factor price w_i, with rearrangement we get

$$\frac{\partial \ln VC}{\partial \ln w_i} = \frac{w_i}{C}\frac{\partial VC}{\partial w_i} = \beta_i + \sum_{j=1}^{m} \lambda_{ij} \ln Y_j + \sum_{q=1}^{n} \gamma_{iq} \ln w_q + o_i \ln S + \pi_i \ln N + \sum_{k=1}^{p} \rho_{ik} \ln Z_k \quad (41)$$

From Shephard's Lemma it is known that the derivative of the cost function with respect to factor price is the factor demand equation

$$\frac{\partial C}{\partial W_i} = X_i(W, Y, Z) \quad (42)$$

Thus from (41) and (42) we have the firm's share equation for factor *i*

$$FS_i = \frac{w_i x_i}{C} = \beta_i + \sum_{j=1}^{m} \lambda_{ij} \ln Y_j + \sum_{q=1}^{n} \gamma_{iq} \ln w_q + o_i \ln S + \pi_i \ln N + \sum_{k=1}^{p} \rho_{ik} \ln Z_k \quad (43)$$

where a firm's factor share, FS_i, is the percent of costs expended on factor *i*.

A similar equation, the ratio of depreciated capital cost to variable cost, is also included (Oum and Zhang 1991).[4]

$$\frac{\delta \mu K + rK}{VC} = \frac{\partial \ln VC}{\partial \ln S} = \eta + \sum_{i=1}^{n} o_i \ln w_i + \sum_{j=1}^{m} \psi_j \ln Y_j + \theta \ln S + \sigma \ln N + \sum_{k=1}^{p} \tau_k \ln Z_k \quad (44)$$

[3] The data set includes the total number of riders per period, but not passenger trip distance. Thus passenger miles cannot be calculated.
[4] Oum and Zhang (1991) suggest that this will increase efficiency.

Here δ is depreciation, a cost associated with capital utilization, and r is rental cost that is accrued to a firm's entire capital stock. The factor share equations are estimated in a system of equations along with the variable cost function.

All variables in (40), with the exception of the dummy variable for the presence of a full-time administrator, are divided by their sample mean. Moreover, the logarithms of all variables are used. Consequently, parameter estimates are the elasticity of cost with respect to the respective variable evaluated at the sample mean of all variables when an agency does not have a full-time director. As not all transit agencies in the sample provide both fixed-route and demand-response service, a Box-Cox transformation with λ=.0001 is used for outputs.

$$y_i^{\lambda=.0001} = \frac{y_i^{.0001}-1}{.0001} \qquad (45)$$

Symmetry (46) and homogeneity of degree 1 in prices (47) are also imposed.

$$\zeta_{jl} = \zeta_{lj}, \gamma_{iq} = \gamma_{qi}, \kappa_{kr} = \kappa_{rk}, \lambda_{ij} = \lambda_{ji}, \xi_{jk} = \xi_{kj}, \rho_{ik} = \rho_{ki} \qquad (46)$$

$$\sum_{i=1}^{n} \beta_i = 1, \quad \sum_{i=1}^{n} \gamma_{iq} = 0 \ for\ q = 1,2,\dots n, \sum_{i=1}^{n} \lambda_{ij} = 0 \ for\ j = 1,2,\dots m, \ \sum_{i=1}^{n} o_i = 0,$$
$$\sum_{i=1}^{n} \pi_i = 0, \sum_{i=1}^{n} \rho_{ik} = 0 \ for\ k = 1,2,\dots p \qquad (47)$$

Estimation of the system of equations of (40), (43), and (44) is done by adding error terms and using Zellner's seemingly unrelated regression (SUR) method. Share equations (43) and capacity utilization equation (44) increase the efficiency of the model and help mitigate concerns of inadequate degrees of freedom. To avoid singularity, the maintenance share equation is deleted prior to estimation.

3.3 Derivation of the Optimal Level of Capital

Short-run total costs are

$$SRC(W, Y, Z, N, K) = VC(W, Y, Z, N, K) + w_k K \qquad (48)$$

where SRC is short-run total cost, VC is variable costs, w_k is the price of capital, and K is the level of capital. There are two components to the price of capital: the rental price, *r*, and the depreciation cost, δ. Rental price depends on the total level of capital while depreciation cost depends on capital use. Moreover, capital utilization is used in the short-run variable cost function as in Oum and Zhang (1991). Thus,

$$SRC(W, Y, Z, N, K) = VC(W, Y, Z, N, S) + \delta S + rK$$
$$= VC(W, Y, Z, N, S) + \delta \mu K + rK \qquad (49)$$

where μ is the capital utilization rate.

The optimal capital stock can be determined by minimizing (49) with respect to K.

$$\frac{\partial SRTC}{\partial K} = \frac{\partial VC}{\partial S}\mu + \delta\mu + r = 0 \qquad (50)$$

The long-run equilibrium state can be found by evaluating (40) at the optimal capital stock with the utilization rate set to 1 (Oum and Zhang 1991).

The solution to (50) contains K* and its logarithm. Consequently, one cannot solve for K* explicitly. However, with knowledge of factor prices and output a simple iterative procedure can be used to solve the equation numerically. That is, K is varied until the equation is approximately 0. Substituting K^* into (40) and adding rK* and δμK* we derive the long-run costs.

Similarly, the optimal level of capital for any transit agency i, K_i^*, is calculated by solving the derivative of the short-run variable cost function with respect to capital evaluated at the agency level of outputs and network size with all other variables, except DIRECTOR, placed at their sample mean. Excess capital is calculated as the difference between the agency's actual capital, K_i, and the optimal level of capital, K_i^*.

Economic theory shows that the long-run cost function can be obtained by inserting K*, the optimal level of capital, into the firm's short-run cost function. Because K* is a function of output, input prices, and technological variables, the long-run cost function shows the explicit relationships between these variables and costs, as capital is freely varied. However, as stated above, an algebraic solution to K* is not available since the solution to (50) contains K* and its logarithm. Thus, by obtaining a numerical value for K* at any output, we are able to obtain minimum costs, but not an algebraic representation of the cost function (that shows how minimum cost varies with output in the long run).

3.4 Measures

Estimated cost functions can be used to calculate a number of economic measures that describe rural transit production technology that can be used to develop industrial and regulatory policy.

3.4.1 Elasticities of Substitution

Factor, own- and cross-price elasticities can be estimated using the Allen-Uzawa partial elasticities of substitution (Uzawa 1962). Partial elasticities of substitution between factors i and j, σ_{ij}, are

$$\sigma_{ij} = \frac{\gamma_{ij}+s_i s_j}{s_i s_j} \qquad \text{i,j=1,2,.....n i}\neq\text{j} \qquad (51)$$

where are γ_{ii} and γ_{ij} are estimated parameters from the generalized translog function and s_i and s_q are cost shares. Own and cross-price elasticities of factor demand, ϵ_{ij}, are

$$\epsilon_{ii} = \frac{\gamma_{ii}+s_i^2-s_i}{s_i} \qquad \text{i=1,2,......n} \qquad (52)$$

$$\epsilon_{ij} = \frac{\gamma_{ij}+s_i s_j}{s_j} \qquad \text{i,j=1,2,.....n i}\neq\text{j} \qquad (53)$$

3.4.2 Economies of Size, Density, and Scope

Estimates of economies of size and density are estimated using parameters from the short-run cost model and measures introduced by Caves, Christensen, and Tretheway (1984). Caves, Christensen, and Swanson (1981) show how returns to density and size can be calculated using the parameter estimates from a short-run variable cost function. Returns to density are defined as the increase in output resulting from a proportional increase in input holding input prices and network area constant. It is calculated as the inverse of the elasticity of cost with respect to output at the mean of all variables

$$RTD = \frac{1-\varepsilon_k}{\varepsilon_y} = \frac{1-\eta}{\alpha_1+\alpha_2} \qquad (54)$$

where $\alpha_1, \alpha_2,$ and ηare estimated parameters from the short-run variable cost function.

Returns to size are defined as the increase in output from a proportional increase in input and network size keeping input prices constant. Mathematically, this is represented by equation (55), where ε_N is the elasticity of cost with respect to network size.

$$RTS = \frac{1-\varepsilon_k}{\varepsilon_y + \varepsilon_N} = \frac{1-\eta}{\alpha_1+\alpha_2+\omega} \qquad (55)$$

where α_1, α_2, η, and ω are estimated parameters from the short-run variable cost function.

Global economies of scope are calculated using the measure described by Panzar and Willig (1981).

$$SC = \frac{C(y_{\mathrm{dr}},0)+C(0,y_{\mathrm{fr}})}{C(y)} \qquad (56)$$

3.4.3 Cost Subadditivity

The presence of natural monopoly in rural transit is evaluated by testing for subadditivity. This method compares the cost of joint production of demand-response and fixed-route services by a single agency to the same level of service produced by two agencies.

Evans and Heckman (1982) impose two constraints on the admissible region, limiting the minimum level of output and the ratio of output combinations to those observed in the data. As zero levels of output for fixed-route and demand-response service are observed in our sample these constraints are ignored.

Our test of the presence of natural monopoly, following Bitzan (1999), is conducted twice. First, output, fixed-route and demand response miles, is allowed to vary while service area is held equal to the sample mean. Next, overall subadditivity is tested by allowing outputs and network size to vary. The test is conducted for all observations in the sample. Total monopoly and two-firm costs are estimated using parameter estimates from the translog cost function assuming each firm employs the optimal amount of capital. All variables other than outputs, network size, and time are placed at their sample means. Simulations are run for each of 41 unique output combinations and each of the 643 observations for the two-output test. A total of 365 simulations are run for each observation when two outputs and network size are considered.

The local test for subadditivity where outputs and service area are allowed to vary is

$$\begin{aligned} &C(q^m) < C(q^a) + C(q^b) \\ &C(q^m) = C(q_1, q_2, q_3) \\ &C(q^a) = C(\lambda q_1, \gamma q_2, \Phi q_3) \\ &C(q^b) = C((1-\lambda)q_1, (1-\gamma)q_2, (1-\Phi)q_3) \\ &\lambda, \gamma, \phi = (0.1, 0.2, \ldots, 0.9) \\ &q_1, q_2, q_3 = \text{demand-response miles, fixed-route miles, and service area} \end{aligned} \qquad (57)$$

The presence of natural monopoly for a large regional transit system is evaluated by testing for subadditivity of outputs and service area at the observed maximums.

Long-run costs are estimated for each observation in the dataset and each of the hypothetical comparison firms. This means that the optimal capital stock is estimated for each observation and each hypothetical comparison firm. This is necessary as the optimal level of capital depends on the amount of output produced. This is done using numerical approximation, the same method used to calculate K* for the sample mean. Given the large number of observations and combinations, an algorithm was programmed to automate the process.

3.5 Data and Variable Construction

The analysis uses data from quarterly and annual reports required and collected by the North Dakota Department of Transportation from 1998-2005. The Quarterly Request for Reimbursement State Aid for Public Transit Program form reports operating costs including wages by job function, fuel costs, contracted services, maintenance costs, miscellaneous, capital, and total costs, as well as ridership and vehicle miles traveled. The Annual Grant Application for State Aid requires agencies to submit information on types and level of service, service area, service fleet, employees, and other transportation services available in their area.

North Dakota was selected because of the availability of data, its large physical size, and the diversity of agencies that provide rural transit. Expansion to other states was considered; however, the high cost of data collection and significant, undocumented and difficult-to-model state-level policies that impact rural transit argued against expanding the scope geographically. Given the sensitivity of reorganization of rural transit and the added scrutiny it attracts, it was decided that estimating a cost structure for a single state was prudent.

Total operating costs (VC) are available from the Quarterly Request for Reimbursement. Demand-response and fixed-route miles traveled (Y_1, Y_2) are determined using the total vehicle miles reported in the Quarterly Request and scheduled fixed-route service in the Annual Grant Application for State Aid. Fixed-route miles are calculated by multiplying route lengths by their quarterly frequency of service. Demand-response miles are calculated as the difference between total miles and estimated fixed-route miles.

The price of fuel is found by dividing fuel cost, reported in the Quarterly Request for Reimbursement, by vehicle miles and multiplying by the fleet's fuel economy. Fleet-wide fuel economy is a function of the fleet mix, which is found using vehicle type included in the Annual Application and vehicle fuel economy available from the American Public Transportation Association's Transit Vehicle Database (2005). The price of labor is found by dividing wages, as reported in the Quarterly Request, by full-time equivalent (FTE) employees, as calculated using the number of full and part-time employees reported in the Annual Application. The price of maintenance is calculated using maintenance costs reported in the Quarterly Request scaled by the number of vehicle-seats contained in the Annual Application. The price of contracted services is not included as no rural transit agencies reported contracting for transportation.

Capital use, S_i, is calculated by multiplying an agency's capital utilization factor, μ_i, times the agency's capital stock, K_i,

$$S_i = \mu_i K_i = \frac{DRM_i + FRM_i}{Recom} K_i \tag{58}$$

The utilization factor is the ratio of demand-response and fixed-response miles per vehicle to the number of miles recommended by the Federal Transit Administration's useful life guidelines which vary by vehicle type (Laver et. al 2007). The useful life for the "Light-duty small bus, cutaways, and modified van" category is used. Fleet size is used as the measure of agency's capital stock. Mileage by vehicle type data was not available. For the most part, there was homogeneity of vehicle types across firms in the sample. Most vehicles were cutaways.

The cost of a new vehicle is calculated using data from the American Public Transportation Association's Transit Vehicle Database. 10-year straight line depreciation is used. The federal funds rate is used as the interest rate.

Network size, N, is the transit agency's service area measured in square miles, as identified in the Annual Application. The data set is unbalanced, as transit agencies entered, exited, or merged during the period. At the same time, some transit agencies did not file reports for each quarter.
Only intermediate outputs are explicitly included in the model, although the variable rides per mile measures the effectiveness of transit service. Passenger-miles data are not available. The database was unable to accommodate the construction of a seat-miles by mode variable as rural transit agencies in North Dakota do not assign their vehicles to exclusive modes nor do they report service type miles by vehicle. Table 3.1 presents the definition and source data for model variables.

Table 3.1 Variable Definition, Units, and Source Data

Variable	Definition	Units	Source Data
Variable costs	quarterly operating costs	dollars	Quarterly report
Fixed-route miles	scheduled fixed-route service	miles	Annual report
Demand-response miles	balance of vehicle miles	miles	Calculated; annual report, quarterly report
Price of fuel	fleet mix-adjusted price of fuel	dollars	Calculated; annual report, quarterly report, APTA transit vehicle database
Price of labor	quarterly price of FTE employee	dollars	Calculated; annual report, quarterly report
Price of maintenance	vehicle-size adjusted price of maintanenance	dollars	Calculated; annual report, quarterly report
Price of capital	vehicle-size adjusted depreciation and rental cost of capital	dollars	Calculated; annual report, quarterly report, APTA Transit Vehicle Database
Capital utilization	ratio of actual vehicle use to Federal use guidelines	percent	Calculated; annual report, quarterly report
Network size	transit agency service area	square miles	Annual report
Director	Agency has full-time director	dummy	Annual report
Elderly-disabled ratio	ratio of elderly and disabled riders to total ridership	percent	Calculated; quarterly report
Rides per mile	rides per vehicle mile	number	Calculated; quarterly report
Seat average	average number of seats per vehicle	number	Calculated; annual report

Descriptive statistics for the transit agency database are presented in Table 3.2.

Table 3.2 Descriptive Statistics

	Minimum	Maximum	Mean
Variable costs	$ 826	$ 211,300	$ 26,555
Demand-response miles	0	159,424	13,998
Fixed-route miles	0	78,765	3,628
Price of labor	$ 3,126	$ 7,968	$ 5,052
Price of fuel	$ 1.10	$ 4.78	$ 2.12
Price of maintenance	0	$ 3,465	$ 2,472
Average Seats	6	30	13
Service Area (sq. miles)	1	10,710	2,584
Elderly-disabled ratio	7%	100%	89%
Rides per mile	42	54,417	5,553
Fleet size	1	19	4.5
Director - 55% of observations were organizations with full-time directors			

2011 dollars

3.6 Summary

A short-run cost model was developed using intermediate outputs, factor prices, capital utilization, network size, technological variables, and time to explain the cost of providing rural transit service. The optimal level of capital is found mathematically taking the first derivative of the short-run total cost function with respect to capital and solving for the optimal level of capital. The amount of excess capital is calculated as the difference between the actual and optimal level of capital. Long-run total costs are calculated by substituting the optimal level of capital into the short-run total cost. The estimated cost function is then used to calculate economies of density, size, and scope. The presence of natural monopoly is evaluated by comparing the cost of monopoly and two-firm production of multiple outputs. The efficiency benefits of regionalization are evaluated using a simulation.

4. RESULTS

The purpose of this study is to evaluate the justification of government support of rural transit on the basis of the presence of natural monopoly and to determine the most efficient regional organization of transit. In the previous chapter, we discussed the specification of the short-run rural transit cost function, the calculation of long-run costs using the optimal level of capital, and presented methods for determining economies of density, size, and scope, and for evaluating the existence of natural monopoly.

This chapter presents empirical results of the analysis using the methods described in the previous chapter. Parameter estimates for the short-run variable cost model, as well as their significance and interpretation are presented. In addition, the optimal level of capital for rural transit firms is calculated and compared to actual levels. The presence of natural monopoly is evaluated by testing for cost subadditivity. The cost efficiency of regionalized transit service is also assessed.

4.1 Rural Transit Variable Cost Model

A four equation system, consisting of the rural transit variable cost function (40), share equations for labor and fuel, and the capital utilization equation, are estimated using Zellner's seemingly unrelated regressions technique (Zellner 1962).

Parameter estimates and standard errors for cost function coefficients are presented in Table 4.1. The system R^2 value is .903 signifying a relatively high level of fit. All first order parameter estimates are of the expected sign and statistically significant at the 1 percent level, with the exception of elderly-disabled ratio which is significant at the 10 percent level.

Table 4.1 Rural Transit Variable Cost Model

Parameter	Estimate	SE
Intercept	0.6593	0.0220 ***
		(29.95)
Demand-response miles	0.5268	0.0305 ***
		(17.26)
Fixed-route Miles	0.2331	0.0194 ***
		(12.00)
Wage	0.6367	0.0062 ***
		(102.67)
Fuel Price	0.0988	0.0025 ***
		(39.64)
Maintenance Price	0.2644	0.0066 ***
		(39.89)
Capital Use	-0.2164	0.0385 ***
		(5.62)
Seat Average	0.4383	0.0877 ***
		(5.00)
Time	-0.2363	0.0228 ***
		(10.34)
Elderly-Disabled Ratio	0.2289	0.1294 *
		(1.77)
Service Area	0.2172	0.0210 ***
		(10.35)
Rides per Mile	0.5069	0.0267 ***
		(19.00)
Director	-0.2585	0.0238 ***
		(10.88)
Wage x Wage	0.1109	0.0028 ***
		(39.48)
Wage x Fuel Price	-0.0529	0.0019 ***
		(28.19)
Wage x Maintenance Price	-0.0580	0.0023 ***
		(25.22)
Wage x Capital Use	-0.0308	0.0032 ***
		(9.64)
Wage x Demand-response Miles	0.0013	0.0004 ***
		(3.48)
Wage x Fixed-route Miles	0.0062	0.0007 ***
		(8.34)
Wage x Time	0.0294	0.0045 ***
		(6.49)
Wage x Rides per Mile	-0.0073	0.0035 **
		(2.06)
Wage x Elderly-Disabled Ratio	-0.0360	0.0134 ***
		(2.69)
Wage x Seat Average	-0.0457	0.0095 ***
		(4.79)

Table 4.1 (Continued)

Parameter	Estimate	SE	
Wage x Director	0.0506	0.0064	***
		(7.86)	
Wage x Service Area	-0.0044	0.0017	***
		(2.63)	
Fuel Price x Fuel Price	0.0577	0.0021	***
		(27.58)	
Fuel Price x Maintenance Price	-0.0048	0.0011	***
		(4.25)	
Fuel Price x Capital Use	0.0288	0.0016	***
		(18.55)	
Fuel Price x Demand-response Miles	-0.0001	0.0001	
		(0.55)	
Fuel Price x Fixed-route Miles	-0.0022	0.0003	***
		(7.35)	
Fuel Price x Time	0.0118	0.0018	***
		(6.44)	
Fuel Price x Rides per Mile	-0.0223	0.0014	***
		(15.39)	
Fuel Price x Elderly-Disabled Ratio	-0.0235	0.0054	***
		(4.35)	
Fuel Price x Seats	0.0155	0.0038	***
		(4.05)	
Fuel Price x Director	0.0144	0.0026	***
		(5.47)	
Fuel Price x Service Area	-0.0052	0.0007	***
		(7.55)	
Maintenance Price x Maintenance Price	0.0628	0.0024	***
		(26.69)	
Maintenance Price x Capital Use	0.0020	0.0033	
		(0.59)	
Maintenance Price x Demand-response Miles	-0.0012	0.0004	***
		(3.07)	
Maintenance Price x Fixed-route Miles	-0.0040	0.0008	***
		(5.09)	
Maintenance Price x Time	-0.0412	0.0048	***
		(8.53)	
Maintenance Price x Rides per Mile	0.0296	0.0038	***
		(7.82)	
Maintenance Price x Elderly-Disabled Ratio	0.0594	0.0142	***
		(4.17)	
Maintenance Price x Seat Average	0.0302	0.0101	***
		(2.99)	
Maintenance Price x Director	-0.0650	0.0068	***
		(9.52)	
Maintenance Price x Service Area	0.0096	0.0018	***
		(5.37)	

Table 4.1 (Continued)

Parameter	Estimate	SE	
Capital Use x Capital Use	-0.0980	0.0164	***
		(5.96)	
Capital Use x Demand-response miles	4.73E-05	2.14E-05	**
		(2.21)	
Capital Use x Fixed-route miles	0.0004	0.0001	***
		(7.75)	
Capital Use x Time	0.0531	0.0112	***
		(4.76)	
Capital Use x Rides per Mile	0.0177	0.0118	
		(1.50)	
Capital Use x Elderly-Disabled Ratio	-0.2086	0.0729	***
		(2.86)	
Capital Use x Seat Average	-0.2488	0.0651	***
		(3.82)	
Capital Use x Director	0.1069	0.0475	**
		(2.25)	
Capital Use x Service Area	-0.1084	0.0096	***
		(11.31)	
Demand-response miles x Demand-response Miles	0.0001	5.53E-06	***
		(17.52)	
Demand-response miles x Fixed-route Miles	-2.95E-05	4.80E-06	***
		(6.16)	
Demand-response miles x Time	0.0034	0.0054	
		(0.63)	
Demand-response miles x Rides per Mile	0.0215	0.0056	***
		(3.84)	
Demand-response miles x Elderly-Disabled Ratio	-0.0448	0.0472	
		(0.95)	
Demand-response miles x Seat Average	0.0781	0.0601	
		(1.30)	
Demand-response miles x Director	0.0599	0.0398	
		(1.50)	
Demand-response miles x Service Area	0.0789	0.0084	***
		(9.40)	
Fixed-route Miles x Fixed-route Miles	0.0002	3.55E-05	***
		(5.97)	
Fixed-route Miles x Time	0.0077	0.0022	***
		(3.49)	
Fixed-route Miles x Rides per Mile	0.0246	0.0033	***
		(7.54)	
Fixed-route Miles x Elderly-Disabled Ratio	-0.0083	0.0102	
		(0.81)	
Fixed-route Miles x Seat Average	2.72E-05	3.97E-05	
		(0.68)	
Fixed-route Miles x Director	-0.0358	0.0057	***
		(6.32)	

Table 4.1 (Continued)

Parameter	Estimate	SE	
Fixed-route Miles x Service Area	0.0128	0.0040	***
		3.23	
Time x Time	-0.1156	0.0272	***
		4.25	
Time x Rides per Mile	-0.0346	0.0134	***
		2.58	
Time x Elderly-Disabled Ratio	-0.0928	0.0516	*
		1.80	
Time x Seat Average	-0.1245	0.0265	***
		4.70	
Time x Director	0.0755	0.0209	***
		3.61	
Time x Service Area	-0.0297	0.0057	***
		5.17	
Rides per Mile x Rides per Mile	0.0547	0.0145	***
		3.78	
Rides per Mile x Elderly-Disabled Ratio	0.0881	0.0602	
		1.46	
Rides per Mile x Seat Average	-0.1440	0.0444	***
		3.24	
Rides per Mile x Director	-0.0900	0.0204	***
		4.41	
Rides per Mile x Service Area	0.0320	0.0079	***
		4.05	
Elderly-Disabled Ratio x Elderly-Disabled Ratio	-0.0229	0.0598	
		0.38	
Elderly-Disabled Ratio x Seat Average	1.0797	0.2217	***
		4.87	
Elderly-Disabled Ratio x Director	0.4334	0.0869	***
		4.99	
Elderly-Disabled Ratio x Service Area	0.0553	0.0360	
		1.54	
Seat Average x Seat Average	-0.5280	0.1835	***
		2.88	
Seat Average x Director	0.3736	0.0746	***
		5.01	
Seat Average x Service Area	-0.0437	0.0275	
		1.59	
Director x Service Area	-0.0337	0.0137	***
		2.46	
Service Area * Service Area	0.0575	0.0090	***
		6.39	

System weighted R^2=.903
Number of observations =693
* Statistically significant at the α=10% level
** Statistically significant at the α=5% level
*** Statistically significant at the α=1% level

The estimated coefficient for demand-response miles, .526, is significantly larger than that for fixed-route miles, .2331, inferring that a one percent increase in demand-response response miles results in a .53 percent increase in costs while a one percent increase in fixed-route miles results in a .23 percent increase in costs for the average firm. This is understandable given higher costs associated with irregular starts and stops and the need for scheduling and dispatch activity to deliver demand-response service. Furthermore, the elasticity for demand-response service is higher as the average firm delivers more demand-response than fixed-route miles. That is, more of the available economies are exhausted due to higher output level. In combination, these elasticities suggest the presence of multiproduct returns to density.

The capital use coefficient is negative, suggesting that capital has positive returns consistent with economic theory. The positive coefficient for seats per bus implies that agencies with larger sized vehicles have higher costs, a phenomenon that may be explained by lower fuel economy. The negative time coefficient can be interpreted as improved technology resulting in lower costs of production over time. The positive coefficient for elderly and disabled riders infers that agencies with relatively high elderly and disabled ridership have higher costs than those that do not. This is expected as elderly and disabled riders often require specialized higher-cost service.

The estimated coefficient for rides per mile is positive. Agencies that provide more rides per mile may have more stops, travel at lower speed, and experience lower fuel economy. Passenger miles, a measure of service use, was not available.

The estimated coefficient for a rural transit agency's service area is positive, suggesting an increase in costs with the area served by a transit agency. This is consistent with the notion that extra costs are associated with a larger service area.

The employment of a full-time director has a negative effect on costs. This may be explained by full-time management resulting in more efficient service. While there is a strong correlation between agency size and the presence of a full-time director, the role of system size is expected to be captured by other agency size variables used in the model.

Second-order coefficient estimates provide a robust picture of the interactions of factors and their impact on the cost of rural transit. The coefficient for demand-response miles x fixed-route miles is negative. This suggests the presence of cost complementarity, meaning that the marginal cost of one output decreases with the more of the other.

Coefficients for interactions between measures of output, demand-response and fixed-route miles, and technological and environmental attributes: seat average, service area, and rides per mile are all positive. That is, as if output and any of the technological and environmental variables increase so do costs. Second-order coefficients for output and elderly-disabled ratio are negative. This infers that at higher levels of output with higher percentages of elderly and disabled riders costs decrease. This is a surprising result, which may be explained by specialization of agencies to provide service to these transportation disadvantaged riders.

4.2 Regularity Conditions

Homogeneity in prices was imposed as a model constraint. Monotonicity is satisfied as input shares for all observations are positive. Concavity of the cost function was checked by computing eigenvalues of the Hessian matrix. The eigenvalues evaluated at the sample means are all negative, satisfying the negative semi-definiteness condition. The condition was also satisfied for more than 95% of observations.

4.3 The Price and Level of Capital

The average price of capital is $2,748 for a typical cutaway transit vehicle. The average observed fleet size of rural transit agencies is 4.6 vehicles. Capital utilization measured as actual agency vehicle miles divided by FTA recommended miles times the agency's fleet size averaged 2.35. The calculated optimal level of capital at the sample mean is 2.7 vehicles. A firm at the sample average has almost twice as much capital as the long-run optimal level.

Excess capital is calculated as the difference between observed and optimal level capital required for a given level of output and service area. Excess capacity for select rural transit agencies is presented in Table 4.2. All of the transit agencies have a fleet that exceeds its optimal size. James River Transit has the greatest excess capacity as a percent of its optimal fleet size.

The increase in short-run total costs can be estimated using the fitted variable cost function and the price and level of capital used. The differences in costs vary greatly ranging from six percent higher short-run total costs for West River Transit and Souris Basin and Transportation to 61% for James River Transit.

Table 4.2 Excess Capital by Transit Agency

	Observed fleet	Optimal fleet	Excess Fleet Capacity (Percent)	Estimated Increase in Total Cost Resulting from Excess Capital
Mean	4.6	2.7	70%	12%
Elder Care	6	4.5	33%	12%
James River Transit	16	6.9	132%	61%
South Central Transit Network	11	6.9	59%	23%
West River Transit	11	9.8	12%	6%
Souris Basin Transportation	12	10.8	11%	6%

1.1. Elasticities of Substitution and Demand

Elasticities of Allen-Uzawa partial elasticities of substitution calculated at the sample mean are presented in Table 4.3.

Table 4.3 Elasticities of Substitution

	Substitution
Wage x Fuel Price	0.16
Wage x Maintenance Price	0.66
Fuel Price x Maintenance Price	0.82

Estimated own and cross-price elasticities of demand are presented in Table 4.4. The cross-price elasticities are all positive indicating that they are substitutes in production.

Table 4.4 Own and Cross-Price Elasticities of Demand

	Wage	Fuel Price	Main. Price
Wage	-0.19	0.10	0.42
Fuel Price	0.02	-0.32	0.08
Main. Price	0.17	0.22	-0.50

4.4 Economies of Size, Density, and Scope

Economies of size and density are calculated using parameter estimates from the variable cost function. The measures evaluated at the sample mean, and first, second (median), and third quartiles are presented in Table 4.5. Increasing returns to density and size exist when estimated using parameters from the short-run variable cost equation at all levels of output and service area, but decrease as agency size increases.

Table 4.5 Economies of Density and Size

	Mean	1st Quartile	Median	3rd Quartile
Returns to Density	1.60	1.64	1.62	1.59
Returns to Size	1.24	1.24	1.16	1.08

Economies of scope calculated at the sample mean is 1.52, indicating that economies of scope are present. Like returns to density and size, economies of scope diminish as agency size increases. Economies of scope of rural transit agencies are presented in Table 4.6.

Table 4.6 Economies of Scope

	Economies of Scope
Mean	1.52
1st Quartile	1.61
Median	1.43
3rd Quartile	1.02

4.5 Subadditivity

The presence is subadditivity is evaluated using the method described in Chapter 4. The results from the analysis for the two-output case, with all other variables at mean levels, are presented in Table 4.7. Monopoly costs were found to be lower than two-firm costs more than 90 percent of the time during each year of the sample. The average percent increase in costs resulting from two-firm operations varied from 14.1 percent in 1999 to 12.4 percent in 2005, the last year of observations. The cost subadditivity condition was met for more than half of the observations in 1998 and just more than 30 percent of the 2005 observations. Subadditivity was only found for single output agencies. For many of the remaining observations monopoly costs were lower for all but a few combinations of output, specifically when two-output firms had a high percentage of output for one mode of service and a low percentage of output for the other. In practice, the likelihood of two firms providing such combinations of service may be unlikely as a single firm would either provide all service, or two firms would provide a single output. Thus, although cost subadditivity was not found in all cases, there is strong support for natural monopoly when outputs are varied with network size held constant. Cost superadditivity, where two-firm costs were lower than monopoly costs for all combinations of output was not found in any case.

The results for evaluating subadditivity in rural transit when expanding output and service area are presented in Table 4.8. About half the simulations result in monopoly costs being lower than two-firm costs, far lower than when varying output only. The average percent increase in costs resulting from two-firm operation are also much lower especially in later years of the sample. The number of observations where cost subadditivity is met remains the same, but again only are found for single output transit agencies. Cost superadditivity, which exists when the cost of all combinations of production are higher for a monopolist than for two firms, is met for some agencies in each year and grows over time. Superadditivity is found for agencies with the largest service area.

Table 4.7 Subadditivity: Two Outputs

			Monopoly Costs Lower than Two-Firm Costs		Percent Increase in Costs Resulting from Two-Firm Operation (across all simulations)	Cost Subadditivity Condition Met		Cost Superadditivity Condition Met	
Year	Number of Observations	Number of Simulations	Number	Percent	Average	Number	Percent	Number	Percent
1998	103	4,223	4,019	95.2	13.8	52	50.5	0	0
1999	112	4,592	4,302	93.7	14.1	61	54.5	0	0
2000	91	3,731	3,531	94.6	13.5	41	45.1	0	0
2001	86	3,526	3,334	94.6	13.9	38	44.2	0	0
2002	88	3,608	3,326	92.2	12.9	37	42.0	0	0
2003	66	2,706	2,538	93.8	13.8	24	36.4	0	0
2004	80	3,280	3,092	94.3	13.5	33	41.3	0	0
2005	66	2,706	2,522	93.2	12.4	20	30.3	0	0

Table 4.8 Subadditivity: Two Outputs and Service Area

			Monopoly Costs Lower than Two-Firm Costs		Percent Increase in Costs Resulting from Two-Firm Operation (across all simulations)	Cost Subadditivity Condition Met		Cost Superadditivity Condition Met	
Year	Number of Observations	Number of Simulations	Number	Percent	Average	Number	Percent	Number	Percent
1998	103	37,595	20,877	55.5	11.7	52	50.5	8	7.8
1999	112	40,880	20,919	51.2	8.3	52	46.4	8	7.1
2000	91	33,215	16,769	50.5	8.8	41	45.1	8	8.8
2001	86	31,390	15,668	49.9	9.9	38	44.2	8	9.3
2002	88	32,120	15,090	47.0	9.8	36	40.9	8	9.1
2003	66	24,090	10,365	43.0	7.1	24	36.4	6	9.1
2004	80	29,200	13,660	46.8	8.1	33	41.3	11	13.8
2005	66	24,090	8,897	36.9	4.2	20	30.3	10	15.2

4.6 Regional Simulations

Unlike some previous studies, our interest in the existence of natural monopoly is not to determine support for preserving or dismantling existing firms, but rather preserving or combining them into a single regional service provider. The framework used to evaluate the presence of subadditivity is now used to conduct simulations that estimate the cost of various scenarios that consist of unique combinations of demand-response miles, fixed-route miles, and service area not previously seen in actual observations. The simulations require identification of the level and geographic area of service. The test for subaddivitiy is used to determine if a single transit agency or two transit agencies can provide service at a lower cost. For this application, the presence of subadditivity would support the merger of existing rural transit agencies into a single firm.

The first scenario considers the provision of 100,000 miles of demand-response service and 25,000 miles of fixed-route service across a service area 8,000 square miles in size. This is roughly one-eighth the size of the state of North Dakota and is equivalent in magnitude to the state's eight Department of Transportation districts and Department of Human Service regions. While individual agencies in the sample did provide levels of service in excess of these values as single-product firms, no multiproduct firms of this size were present. Cost superadditivity was found, meaning that under all combinations two firms were able to provide service at lower cost than a monopoly firm. The second and third scenarios evaluate the cost when only demand-response or fixed-route service is considered, that is the level of output of the other service is always zero. The results are somewhat trivial given our previous knowledge of the presence of subadditivity for single output firms for actual firms. As expected, subadditivity exists for these cases.

Another trio of scenarios evaluates consolidation at a smaller scale. These scenarios include combinations of 50,000 miles of demand-response service and 10,000 miles of fixed-route service delivered across a 2,500 square mile service area. For this combination of service, neither cost subadditivity no cost superadditivity were present. However, monopoly costs were lower for 92 percent of the combinations considered.

Table 4.9 Minimum Cost Regional Service Provision Simulation

Scenario	Demand-Response Miles	Fixed-Route Miles	Service Area (sq. miles)	Monopoly Costs Lower than Two Firm Costs	Cost Subaddivity	Cost Supperaddivity
1	100,000	25,000	8,000	0	No	Yes
2	0	25,000	8,000	100	Yes	No
3	100,000	0	8,000	100	Yes	No
4	50,000	10,000	2,500	92	No	No
5	0	10,000	2,500	100	Yes	No
6	50,000	0	2,500	100	Yes	No

On the other hand, when only fixed-route or demand-response service is provided over this smaller service area, subadditivity is supported.

4.7 Summary

In this chapter, we presented the results of the analysis of the cost structure of rural transit. The analysis shows that most rural transit agencies are significantly overcapitalized. Economies of density and size exist for rural transit agencies. However, economies of size are roughly constant for larger firms. The presence of natural monopoly of multiproduct firms is generally supported in the case of a fixed network equal to the average North Dakota agency. However, when network size is varied, little support for natural monopoly exists. On the other hand, subadditivity was found only with single output agencies. In the next chapter the implications of the findings on policy and transit organization are explored and discussed.

5. FINDINGS AND IMPLICATIONS

The results of the analysis presented in the previous chapter provide a suitable basis to answer the project's research questions. While satisfactory for this purpose, some results are surprising as they differ from project hypotheses and conventional wisdom regarding the economic structure of transit and the efficient design and provision of rural transit. The findings and implications of the study are directly applicable to rural transit in North Dakota and should be helpful in informing future federal policy as well as rural transit policy, service design, and operation in other states. However, as the analysis relied on North Dakota data only, drawing inferences for other states and circumstances should be done with caution. That being said, the study is a solid first contribution in providing the rural transit industry with the type of rigorous economic information needed to guide policy and planning at the federal, state, and regional level.

A primary motivation for the study was the increased scrutiny of federal programs, and spending in general, in recent years. This concern is manifested in the validation of the justification for government involvement in certain activities and the efficiency of government programs. The project's analysis and measures of the economic structure of rural transit provide the information necessary to evaluate the justification of government intervention in public transportation on the basis of the presence of its cost structure as determined by the existence of increasing returns to scale for single service transit agencies or natural monopoly for multi-service agencies. Testing for the presence of subadditivity and regional transit cost simulations can be used to evaluate the efficiency of current and proposed transit agency configurations.

The study is successful in addressing the problem statement that framed the project as it provides knowledge of the economic structure of rural transit necessary to design policies and allocate resources that ensure efficient service provision. The issues of rural transit efficiency and effectiveness are integrally linked to the sustainability of federal support for transit. The study provides both the framework and the information required to estimate the costs of new, expanded, or modified rural transit service. This includes the estimation of economies of size, density, and scope, the presence of natural monopoly, and the amount of excess capital.

The project's primary interest is in estimating the cost structure of rural transit and identifying policy implications resulting from this structure. The fundamental economic considerations are embodied in the project's six research questions.

1. Is increased service in an existing service area more efficiently provided by a single existing rural transit agency or by adding a new one?
2. Is increased service in an expanded area more efficiently provided by a single rural transit agency or by creating a second agency?
3. Are demand-response and fixed-response service most efficiently provided by a single firm or should two agencies provide each service exclusively?
4. Do rural transit firms have significant unused vehicle capacity?
5. Is a single regional transit agency always more efficient at providing multimodal service or are there cases where two agencies can provide service more efficiently?
6. Is there economic justification for government support of transit on the basis of increasing returns to scale or natural monopoly?

The answers to these questions have immediate, practical real-world implications for federal and state transit policies and the design and operation of transit agencies.

The project's research questions adapted into properly-formed hypotheses can be tested using various measures and methods that rely on rural transit's cost function.

1. Rural transit experiences economies of density.
2. Rural transit experiences economies of firm size.
3. Rural transit experiences economies of scope.
4. Rural transit has excess capacity.
5. Rural transit is a natural monopoly.

The evaluation of these hypotheses is presented in the next section along with other study findings. As alluded to previously, not all hypotheses were supported by the analysis.

Our analysis relied on data from North Dakota transit agencies that provide service to rural areas. This situation is both a limitation and strength of the study. This dataset was suitable for our purpose, as it contained the data required to estimate the cost function. There was considerable variability among the agencies that provide rural transit in terms of the levels, combinations, and size of area served. Limiting the analysis to one state removed the challenge of having to appropriately identify and measure relevant policy, environmental, and technological attributes that vary by state. It also protects the analysis from scrutiny, as skeptics might criticize the analysis and its findings on the basis of the inclusion of observations from multiple states even if differences were appropriately modeled.

The analysis estimated a short-run cost model using intermediate outputs: demand-response and fixed-route miles, factor prices, capital utilization as opposed to capital level, network size, technological characteristics, and time. The optimal level of capital was found mathematically by taking the first derivative of the short-run total cost function and solving for capital. The amount of excess capital was calculated as the difference between the actual and optimal level of capital. Long-run costs were obtained by substituting the optimal level of capital into the short-run total cost function. The estimated cost function was then used to calculate economies of density, size, and scope measures. The presence of natural monopoly is evaluated by comparing the cost of joint and disjoint production of multiple outputs. The efficiency benefits of regionalization are evaluated using a simulation.

In this chapter, research findings based on the results presented in Chapter 4 are discussed. The policy implications of the analysis in North Dakota and at the federal level are presented. Next, research implications and opportunities for further research are discussed.

5.1 Findings

Knowledge of the cost structure of transit allows us to test our hypotheses including the presence of natural monopoly. It can also be used to determine the relative efficiency of monopoly and two-firm provision of transit service with varying levels of output and service. In this section, project results are used to evaluate our hypotheses and discuss other relevant findings of the analysis.

The analysis found returns to density at all levels of output. Returns decline as output increases, implying that at very high levels of service, returns to density may not exist. This finding is in agreement with our hypothesis. The presence of returns to density infers that it is more efficient for an existing rural transit agency to provide increased output within its service area than to create a new transit agency to do so.

Returns to size were found at all levels of output. However, returns to size declined with system size. As our interest is the efficiency of larger regional systems which have relatively high levels of output and service area where returns to size do not exist, our hypothesis fails. This implies that it is more efficient for a second agency to provide service in a new, large service than for an existing agency to do so. The actual level of output and service area at which two-firm operation is more efficient can be evaluated by simulation using specific parameters.

Economies of scope were estimated to be present at all levels evaluated. However, these economies have all but disappeared when output levels are equal to the third quartile of values observed in the sample. At even higher levels, economies of scope disappear. As with returns to size, our focus is on evaluating the efficiency of large systems. For high levels of output and service area, diseconomies of scope prevail. This infers that when large regions, and demand-response and fixed-route miles of service are considered it is more efficient for firms to specialize in one mode of service.

Significant amounts of excess capacity were found for all transit agencies evaluated. This is as expected and in agreement with previous studies that have considered urban transit systems. This suggests that short-run costs are higher than would otherwise be the case, as agencies incur expenses to own and maintain relatively little used vehicles. This finding supports revisiting state and federal policies related to capital funding allocation, purchasing, and vehicle disposition.

Cost subadditivity was found only where transit agencies provide a single mode of service. However, some observations did have nearly all combinations with lower costs for monopoly as opposed to two firm service in the two-product case. Moreover, the combinations where monopoly costs were higher are not realistic. These results suggest that at the average service area of North Dakota agencies, single service provision is desirable. In contrast, when service area is expanded, little support for natural monopoly exists. In this case, a number of agencies experience subaddivitity where costs of providing service by two firms is always lower than for a monopoly.

5.2 Policy Implications

The findings of the project have policy implications at the state and federal levels. These implications include guidance on the design and delivery of individual transit agencies which is currently under the purview of states and guidelines on vehicle purchases and disquisition which is both a state and federal issue. We limit our investigation of state-level implications to North Dakota and leave conjecture to other states open to the reader, but caution against drawing strong conclusions as the cost structure may differ from that experienced by the agencies in our sample.

The regionalization of transit agencies is often perceived, if not intended to imply the consolidation of smaller transit agencies into a single transit operating agency. Regionalization is typically considered for delivery of service across relatively large geographic areas, for example the 8,000 square mile area used in the cost simulation. Provision of transit service by a single agency across an area this large on the basis of economic efficiency is not supported by our analysis. This finding does not align with conventional wisdom, where many feel that rural transit is a natural monopoly. We find transit provision across large areas is better served by two agencies, one providing demand-response service the other fixed-route service.

Our analysis focuses solely on economic costs. This is just one, albeit important dimension, of regional or statewide transit planning. It ignores the issue of effectiveness of service which is defined as the amount of transit service consumed relative to inputs or intermediate outputs, demand-response or fixed-route miles. A single agency may be able to employ planners and coordinators who are able to design and operate a system with higher ridership. At the same time, successful coordination of a number of transit agencies in a region may achieve similar outcomes. Consolidation of local transit agencies may cause resentment, loss of local political and financial support, and possibly a decline in ridership.

What may be most politically palatable and economically efficient would be support for regional fixed-route service and as few demand-response systems as possible. Here the fixed-route service would provide service among small communities and regional hubs while demand-response providers would provide service within individual communities. Coordination between the service types would be fundamental to its success in terms of service effectiveness as riders would need to be willing and able to make use of multiple systems to make trips from small communities to large ones.
Despite the presence of returns to density and returns to size for most levels of output and service area, rural transit was not found to be a natural monopoly. Even if this condition is absent, government intervention may have merit in terms of its positive externalities achieved and negative externalities avoided, in addition to the presence of economies of scale. Extensive economies of density suggest that marginal cost pricing will not recover full costs. This is a strong justification for government intervention.

Given the presence of significant levels of excess capital, which aligns with the findings of other studies, federal vehicle purchase and disquisition guidelines should be reviewed and possibly modified. Requiring transit agencies, especially small ones, to keep unused vehicles on their roster is inefficient as they could be put in service elsewhere by other transit agencies, government agencies or government sponsored non-profits, or the private passenger transportation sector. In addition there is a cost to owning these vehicles even if they sit idle. This includes insurance, minimum maintenance, and storage. While individual communities may desire to have their own vehicle, without the requisite amount of service being delivered, it is more efficient for a vehicle to be shared.

5.3 Future Research

The methodology used to fit and measure the cost structure of rural transit served our purposes well and should be amenable for investigation of transit costs in other settings be they urban or rural as well as for other transportation modes. While the methods used are part of the transportation cost literature, together they form an amenable combination for the analysis of relevant cost issues facing multiproduct firms where some firms have zero levels of output. In this section, the use of these tools for transportation costing in general and transit costing specifically are described. Also included is discussion of transit cost issues related to economies of scope and natural monopoly.

The use of capital utilization, as opposed by to the level of capital as proposed by Oum and Zhang achieved its goal of addressing the impact of a quasi-fixed input on the estimation of variable cost functions in transportation. To the knowledge of the author, this is the first use of the method in transit although it has been alluded to in some studies. The suitability of the method makes it a helpful tool for estimating variable cost functions in transportation where a proxy for capital use can be devised which may not always be the case. Previous studies that estimated positive coefficients for capital stock in the past are immediate candidates for its use. Future transit cost studies should consider the utilizing the method as well.

The use of the generalized translog allowed for the estimation of the cost function despite the existence of a number of transit agencies that provide a single mode of service. This issue is of particular importance in rural transit where single modes of service are much more prevalent than urban systems where federal policy mandates the provision of demand-response service to complement fixed-route service. The ability to estimate economies of scope and average incremental costs are necessary to have a full understanding of the cost structure of industries where joint production with occasional zero levels of output are observed.

While the presence of natural monopoly conditions for transit is necessary to determine economic justification for government intervention on the basis of transit’s economic structure, the test for cost subadditivity had not been employed previously, although other weaker measures such as economies of scope have. The application of this test and framework to other rural transit systems or urban transit systems is not only logical, it is vital given the amount of funding that provided to transit in part because of this unvalidated condition.

Knowledge of the rural transit cost function allowed for the simulation of a number of regional service provision scenarios that varied by the level of output and service area. The application of this method was somewhat limited in this paper. We evaluated only six scenarios to demonstrate the capability of cost simulation and to evaluate some general cases. In practice, this framework can be used to quickly and relatively effortlessly evaluate numerous scenarios under the guidance of state, regional, and local transit officials using specific parameters.

The variable cost function model specified two intermediate outputs: demand-response and fixed route miles; and a rides-per-mile variable that incorporated service effectiveness. The specification was the result of data availability as well as the focus on efficiency as opposed to effectiveness. While this met our needs, the remains a significant unanswered question regarding the proper specification of output for transit cost models. This question may be best answered by society through the political process as it decides the goal of transit. Is it to provide the greatest number of trips at the lowest cost? Or it is to provide equitable levels of service across regions regardless of use? Or is it something else or a combination of goals that have different priorities?

The methodology can and should be applied to other states and urban areas to answer economic questions related to transit. Extension of the findings of this paper to other states or to urban areas can only be taken so far. Different policies, practices, technology, and geographies impact the cost of delivering transit service. While the author is comfortable with the assessment of multiproduct rural transit in North Dakota not having cost subadditivity and conjecture that the relationship would hold in other states, that's is not known with certainty. Given the amount of scrutiny that consolidation holds, it would be best to have new state-level analysis conducted in states concerned with the issue. There are many regulatory questions in transit that can be addressed with knowledge of a cost function. A strength of the study is its ability to inform current decisions facing rural transit; concerns about regionalization and consolidation are real and significant. Other states likely have similar concerns. At the same time unique issues that are economic in nature may be addressed using the same method.

A particular application of the method is that of evaluating the economic impacts of decoupling urban fixed-route and demand-response service. Federal regulations require that complementary demand-response service be provided in conjunction with fixed-route service. However, this presents a conundrum as decisions to expand fixed-route service are governed in many respects by an urban transit agencies ability to finance required demand-response service. Would it be less expensive to decouple the service? Would it be more equitable for minimum amounts of urban area wide demand-response service be provided and fixed-route service to be delivered in response to the travel behavior of members of the general public.

5.4 Summary

The purpose of this study is to evaluate the justification of government support of rural transit on the basis of the presence of natural monopoly and to determine the most efficient regional organization of transit. The study found returns to density, size, and scope at most levels of output. Cost subadditivity where a monopoly firm can provide service at a lower cost than two firms was found for many, but not all observations. Consequently, the presence of natural monopoly in rural transit in a strict sense is rejected. A single output firm could be the most efficient provider of service over a large area, but multiple output firms would not. The findings and implications of the study are directly applicable to rural transit in North Dakota and are should be helpful in informing future federal policy as well as rural transit policy, service design, and operation in other states. However, as the analysis was conducted using North Dakota data only, one should draw conclusions for other states and circumstances with caution. The study is first step in providing the rural transit industry with the type of rigorous economic information needed to guide policy and planning at the federal, state, and regional level.

REFERENCES

American Public Transportation Association. 2005. Transit Vehicle Database.

Appelbaum, E., Berechman, J., 1991. Demand conditions, regulations, and the measurement of productivity. Journal of Econometrics 47 (2/3), 379-400.

Baumol, W.J., Willig, R.D., Panzer, J.C., 1988. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt Brace Jovanovich.

Berechman, J., 1983. Cost, economies of scale, and factor demand in bus transport. Journal of Transport Economics and Policy 17 (1), 7-24.

Berechman, J., 1987. Cost structure and production technology in transit: an application to the Israeli bus sector. Regional Science and Urban Economics 17 (4), 519-534.

Berechman, J., 1993. Public Transit Economics and Deregulation Policy. New York: North-Holland.

Berechman, J., Giuliano, G., 1984. Analysis of the cost structure of an urban bus transit property. Transportation Research 18B (4/5), 372-287.

Bitzan, J., 1999. The structure of railroad costs and the benefits/costs of mergers. Research in Transportation Economics 5, 1-52.

Box, G.E.P., Cox, D.E., 1964. An analysis of transformations. Journal of the Royal Statistical Society 26 (2), 211-252.

Burkhardt, J.E., 2004. TCRP Report 101: Toolkit for Rural Community Coordinated Transportation Services. Washington, D.C.: National Academies of Science.

Button, K.J., O'Donnell, K.J., 1985. An examination of the cost structures associated with providing urban bus services in Britain. Scottish Journal of Political Economy 32 (1), 67-81.

Cambini, C., Piacenza, M., Vannoni, D., 2007. Restructuring public transit systems: evidence on cost properties and optimal network configuration. Review of Industrial Organization 31 (3), 183-203.

Caves, D.W., Christensen, L.R., Swanson, J.A., 1981. Productivity growth, scale economies, and capacity utilization in U.S. railroads, 1955-1974. American Economic Review 71 (5), 994-1002.

Caves, D.W., Christensen, L.R., Tretheway, M.W., 1980. Flexible cost functions for multiproduct firms. Review of Economics and Statistics 62 (3), 477-481.

Caves, D.W., Christensen, L.R., Tretheway, M.W., 1984. Economies of density versus economies of scale: why trunk and local service costs differ. RAND Journal of Economics 15 (4), 471-489.

Christensen, L.R., Jorgenson, D.W., Lau, L.J., 1973. Transcendental logarithmic production frontiers. The Review of Economics and Statistics 55 (1), 28-45.

Cobb, C.W., Douglas, P.H., 1928. A theory of production. American Economic Review 18, 138-165.

Colburn, C.B., Talley, W.K., 1992. A firm specific analysis of economies of size in the U.S. urban multiservice transit industry. Transportation Research 26B (3), 195-206.

Cook, T.J., 2002. North Carolina Department of Transportation Research Project 2002-11: Regionalizing Public Transportation Services Final Report. Raleigh, N.C.: Institute of Transportation Research and Education.

De Borger, B.L., 1984. Cost and productivity in regional bus transportation: the Belgian case. Journal of Industrial Economics 33 (1), 37-54.

Deller, S.C., Chicoine, D.L., Walzer, N., 1988. Economies of scale and scope in rural low-volume roads. Review of Economics and Statistics 70 (3), 459-465.

Diewert, W.E., 1971. An application of the Shephard duality theorem: a generalized Leontief production function. Journal of Political Economy 79 (3), 481-507.

Evans, D.S., Heckman, J.J., 1984. A test for subadditivity of the cost function with an application to the Bell system. American Economic Review 74 (4), 615-623.
Elgar, I., Kennedy, C., 2005. Review of optimal transit subsidies: comparison between models. Journal of Urban Planning and Development 131 (71), 71-78.
Executive Order No. 13300, 3 CFR 298. 2004.
Farsi, M., Fetz, A., Filippini, M., 2007. Economies of scale and scope in local public transportation. Journal of Transport Economics and Policy 41 (3), 345-361.
Fazioli, R., Filippini, M., Prioni, P., 1993. Cost-structure and efficiency of local public transport: the case of Emilia Romagna bus companies. International Journal of Transport Economics 20 (3), 305-325.
Federal Transit Administration, 2005. Strategic Research Plan: FY2006-FY2010.
Federal Transit Administration, 2007a. Elderly Individuals and Individuals with Disabilities Program Guidance and Application Instructions. Circular C 9070.1F.
Federal Transit Administration, 2007b. Nonurbanized Area Formula Program Guidance and Grant Application Instructions. Circular C 9040.1F.
Federal Transit Administration, 2008. Grant Management Requirements. Circular C 5010.1D.
Federal Transit Administration. 2009. Rural National Transit Database.
Federal Transit Administration, 2010. Urbanized Area Formula Program: Guidance and Grant Application Instructions. Circular C 9030.1D.
Filippini, M., 1996. Economies of scale and utilization in the Swiss electric power distribution industry. Applied Economics 28, 543-550.
Filippini, M., Maggi, R., Prioni, P., 1992. Inefficiency in a regulated industry: the case of Swiss regional bus companies. Annals of Public and Cooperative Economics 63 (3), 437-454.
Filippini, M., Prioni, P., 2003. The influence of ownership on the cost of bus service provision in Switzerland – an empirical illustration. Applied Economics 35 (6), 683-690.
Fraquelli, G., Piacenza, M., Abrate, G., 2004. Regulating public transit networks: how do urban-intercity diversification and speed-up measures affect firms' cost performance? Annals of Public and Cooperative Economics 75 (2), 167-337.
Gagnepain, P., Ivaldi, M., 2002. Incentive regulatory policies: The case of public transit systems in France. Rand Journal of Economics, 33 (4), 605-629.
Government Accounting Office, 1977. Hindrances to Coordinating Transportation of People Participating in Federally Funded Grant Programs: Volume I. GAO-77-119.
Government Accounting Office, 1999. Transportation Coordination: Benefits and Barriers. GAO/RCED-00-1.
Government Accounting Office, 2003. Transportation Disadvantaged Populations: Some Coordination Efforts Among Programs Providing Transportation Services, but Obstacles Persist. GAO-03-697.
Harmatuck, D.J., 2005. Cost functions and efficiency estimates of Midwest bus transit systems. Transportation Research Record No. 1932, 43-53.
Hensher, D.A., 1988. Productivity in privately owned and operated bus firms in Australia. In: Dogson, J.S., Tophan, N., eds. *Bus Deregulation and Privatisation.* Avebury: Aldershot, pp. 141-170.
House Concurrent Resolution 34. Establishing the budget for the United States Government for fiscal year 2012 and setting forth appropriate budgetary levels for fiscal years 2013 through 2021. April 13, 2011. 112th Congress. 1st Session.
Karlaftis, M., McCarthy, P., 2002. Cost structures of public transit systems: a panel data analysis. Transportation Research Part E: Logistics and Transportation Review 38 (1), 1-18.

Karlaftis, M., McCarthy, P., Sinha, K., 1999. System size and cost structure of transit industry. ASCE Journal of Transportation Engineering 125 (3), 20-215.
Keeler, T.E., 1974. Railroad costs, returns to scale, and excess capacity. Review of Economics and Statistics 56 (2), 201-208.
Lau, L.J., 1974. Comments. In: Intriligator, M.D., Kendrick, D.A., eds. *Frontiers of Qualitative Economics, Volume II*: Amsterdam: North Holland Publishing, pp. 176-199.
Laver, R., Schneck, D., Skorupski, D., Brady, S., Cham, L., 2007. *Useful Life of Transit Buses and Vans*. Washington, D.C.: Federal Transit Administration.
Levaggi, R., 1994. Parametric and non-parametric approach to efficiency: the case of urban transportation Italy. Studi Economici 49 (53), 67-88.
Matas, A., Raymond, J-L., 1998. Technical characteristics and efficiency of urban bus companies: the case of Spain. Transportation 25 (3), 243-263.
McFadden, D.L., 1979. Cost, revenue, and production functions. In: Fuss, M.A., McFadden, D.L. eds. *Production Economic: A Dual Approach to Theory and Applications*. Amsterdam: North-Holland Publishing, pp. 3-109.
Nelson, C., 1972. *An Econometric Model of Urban Bus Transit Operations.* Washington, D.C.: Institute for Defense Analysis Paper.
Nelson, R.A., 1989. On the measurement of capacity utilization. The Journal of Industrial Economics 37 (3), 273-286.
Obeng, K., 1984. The economics of bus transit operations. Logistics and Transportation Review 20 (1), 45-65.
Obeng, K., 1985. Bus transit cost, productivity and factor substitution. Journal of Transport Economics and Policy 19 (2), 183-203.
Obeng, K., Golam-Azam, A.M.H., Sakano, R., 1997. *Modeling Economic Inefficiency Caused by Public Transit Subsidies.* Westport, CT: Praeger.
Obeng, K., Talley, W., Colburn, C., 1987. The effects of subsidies on public transit long-run costs. Journal of the Transportation Research Forum 35, 69-87.
Ottoz, E., Fornengo, G., Di Giacomo, M., 2008. The impact of ownership on the cost of bus service provision: an example from Italy. Applied Economics 40 (19), 337-349.
Oum, T., Zhang, Z., 1991. Utilisation of quasi-fixed inputs and estimations of cost functions. Journal of Transport Economics and Policy 25 (2), 121-134.
Panzar, J.C., Willig, R.D., 1977. Economies of scale in multi-output production. Quarterly Journal of Economics 91 (3), 481-493.
Pulley, L.B., Humphrey, D.B., 1993. The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals. Journal of Business 66 (3), 437-462.
Revis, J.S., 1975. Coordinating Transportation of the Elderly and Handicapped: A State of the Art Report. Washington: Institute of Public Administration.
Ryus, P., 2010. TCRP Report 141: A methodology for performance measurement and peer comparison in the public transportation industry. Washington, D.C.: National Academies of Science.
Sharkey, W.W., 1982. *The Theory of Natural Monopoly.* Cambridge: Cambridge University Press.
Shephard, R.W., 1953. *Cost and Production Functions.* Princeton: Princeton University Press.
Shin, R.T., Ying, J.S., 1992. Unnatural monopolies in local telephone. The RAND Journal of Economics 23 (2), 171-183.
Small, K., Verhoef, E.T., 2007. *The Economics of Urban Transportation.* London: Routledge.

Talley, W.K., 1990. Paratransit services, contracting-out and cost savings for public transit firms: a firm specific analysis. Transportation Planning and Technology 15 (1), 13-25.

Varian, H., 1992. *Microeconomic Analysis.* New York: W.W. Norton & Company.

Viton, P.A., 1981. A translog cost function for urban bus transit. The Journal of Industrial Economics 29 (3), 287-304.

Viton, P.A., 1992. Consolidations of scale and scope in urban transit. Regional Science and Urban Economics 22 (1):25-49.

Viton, P.A., 1993. How big should transit be? Evidence on the benefits of reorganization from the San Francisco Bay Area. Transportation 20 (1), 35-57.

Wachs, M., 1989. U.S. transit subsidy policy: in need of reform. Science 244 (4912), 1545-1549.

Wang Chiang, W.S., Chen, Y., 2005. Cost structure and technological change of local public transport: the Kaohsiung City bus case. Applied Economics 37 (12), 1399-1410.

Windle, R.J., 1988. Transit policy and the cost structure of urban bus transportation. In Dogson, J.S., Tophan, N., (Eds.) *Bus Deregulation and Privatisation.* Avebury: Aldershot, pp. 119-140.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association 57, 348-368.