Quantifying the Technical Efficiency of Canadian Paratransit Systems Using Data Envelopment Analysis Method

by

Jingtao Yang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of

Master of Applied Science
in
Civil Engineering

Waterloo, Ontario, Canada, 2005

Abstract

Paratransit service operators in Canada are under increasing pressure to improve the operational productivity of their services due to increased demand and tightening financial constraints. To achieve this, Paratransit operators need to know their performance as compared to peer systems and the best practices within the industry. This will enable each operator to identify where and how much improvement should be made in order to be on a par with the industry's best practices. Little research effort, however, has been devoted to the issue of how to measure and compare paratransit efficiency in a consistent and systematic manner.

This research focuses on evaluating the level of efficiency of individual paratransit systems in Canada with the specific objective of identifying the most efficient service agencies and the sources of their efficiency. By identifying the most efficient systems along with the influencing factors, it is possible that new service policies and management and operational strategies could be developed for improved resource utilization and quality of services. To achieve this objective, this research applies the analysis methodology called Data Envelopment Analysis (DEA) approach which is a mathematical programming based technique for determining the efficiency of individual systems as compared their peers involving multiple performance measures. Annual operating data from Canadian Urban Transit Association (CUTA) for Canadian paratransit systems of year 2001, 2002 and 2003 are used in this analysis. Regression analysis is performed to identify the possible relationship between the efficiency of a paratransit system and some measurable operating, managerial and other factors which could have an impact on the performance of paratransit systems. The regression analysis also allows for the calculation of confidence intervals and bias for the efficiency scores in order to assess their precision.

Acknowledgements

I would like to thank my supervisor at University of Waterloo, Dr. Liping Fu who has provided insight and has had a significant input to my thesis.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 INTRODUCTION

1.1 The Problem

Paratransit is considered as an important public transportation mode in most municipalities in North America. Traditionally, paratransit provides services to people with special requirements, such as seniors and disabled people. In the past ten years, paratransit system expanded significantly in Canada in order to feed increased demand. It has been predicted that over the next ten to twenty years the demand for paratransit service will be doubled from now due the aging of the baby boomers. Some other factors, such as the desire to reduce transportation induced air pollution, commitment to the Kyoto Protocol to reduce greenhouse gas, and the need to conserve energy, will also have some impacts on paratransit demand as well.

Due to its door-to-door service approach with a fare scheme comparable to regular transit, most paratransit systems in North America rely heavily on subsidization. According to the American Public Transit Association, the total operating expense of paratransit services in the United States exceeded 1.2 billion dollars with only 173 million dollars collected in fares. The Canadian Urban Transit Association reported that the total operating expenses of 50 Canadian paratransit agencies amounted to approximately 150 million dollars, of which only 10% was recovered from fare revenues and the rest 90% was subsidized. The anticipated increase in paratransit demand in the near future will place even greater pressure on paratransit agencies for finding ways to reduce their operating costs and improve their service efficiency.

Due to the complexity of paratransit operations, the efficiency of a paratransit service system is influenced by many factors associated with the overall service delivery system. For example, the operating efficiency of a paratransit system depends on many managerial factors such as service policies (e.g. pickup windows, maximum allowed on-board time, curb-to-curb vs. door-to-door service); fleet mix (e.g. fleet size and vehicle capacity); trip scheduling method (i.e., the extent to which it produces viable and efficient vehicle routes and schedules); dispatch control method (especially in the handling/re-scheduling of late-running trips and making use of capacity in the event of late cancellations and passenger no-shows); driver/dispatch training, etc. Conversely, a paratransit system is also impacted by various local system characteristics that cannot be easily altered by the operating authority such as physical and geographical factors (e.g. size of service area and geographic barriers such as bridges), service type, passenger demand density and opportunities for grouping passengers on trips (i.e., shared rides), and average trip length, etc.

Existence of such diverse range of both controllable and uncontrollable factors – as well as differences in the definitions used to define system efficiency - makes it difficult to directly compare the productivity performance of different service systems and identify the potential productivity problems of a given system. For example, a system with low productivity could simply be because it is operating in an area with a low number of passengers eligible to use a service. Conversely, low productivity could be caused by a poor scheduling system that is not capable of creating efficient ride sharing and utilization of available vehicle capacity. Furthermore, productivity is closely tied to the level of service to be maintained; higher quality of service is usually associated with lower productivity levels. Establishing scheduling parameters that impose high productivity requirements on a system requires very tight timetables with little or no tolerance for real world running time variations or scheduling errors. Such tight schedules inevitably cause some vehicles to run late, ruining their on-time performance. Thus, the imposition of higher productivity through tight scheduling may be illusory. Higher productivity may also be achieved through scheduling policies that encourage or mandate group trips for specific purposes – either through negotiation of start/end times or a differentiating fare structure.

The paratransit industry needs tools and guidelines that can be used to perform "what-if" analysis of performance, identify the important controllable factors that affect efficiency, and provide meaningful comparisons across different paratransit operators. In addition, performance studies can also provide answers to the following questions: who are the best performers in the paratransit industry? What is the maximum achievable efficiency given the demand profile and operating environment? What can be accomplished either through changing service policies or more aggressively pursuing ways to reduce vehicle hours? To address all these questions requires understanding the nature and the causal relationships, as well as the tradeoffs among the factors as independent variables and the elements comprising the productivity calculation.

1.2 The Research Objectives

Different techniques have been applied to study the efficiency of a group of organizations or operating units. Data Envelopment Analysis (DEA) is one of the most widely used approaches because of its sound mathematical basis and non-parametric nature. This research therefore focuses primarily on application of the DEA technique for evaluating the efficiency of paratransit service systems. The research has the following three specific objectives:

- Review existing methodologies for efficiency evaluation and performance benchmarking;
- Assess the suitability of the Data Envelopment Analysis (DEA) approach for evaluating the efficiency of paratransit systems using data from the Canadian paratransit sector;
- Identify the best performers of the Canadian paratransit systems and if possible the factors which are associated with these service systems.

1.3 The Research Methodology

This research proposes to utilize the Data Envelopment Analysis (DEA) method to quantify the efficiency score of individual Canadian paratransit systems. A regression analysis is then applied to the efficiency score in order to find out the possible factors which affect the efficiency of paratransit system. The investigation involves mainly three steps as follows:

- Firstly, data of Canadian paratransit systems is analyzed in order to define the inputs and outputs for evaluating the efficiency. Also, the possible factors that may be associated with the efficiency will be selected for regression analysis;
- Secondly, DEA method is applied to calculate the efficiency score, General Algebraic Modeling System (GAMS) is used as the mathematical programming and optimization tool;
- Lastly, a regression analysis is performed to examine the possible factors which may have effect on the efficiency of a paratransit system.

1.4 Organization of the thesis

This thesis is organized into four chapters:

Chapter 1 introduces the problems to be addressed and the research objectives in this thesis. This chapter also includes general information of research methodology and the structure of the thesis.

Chapter 2 provides background introduction of paratransit and DEA, also some recent studies on paratransit topic and applications of DEA.

Chapter 3 explains how DEA is applied to calculate paratransit system efficiency in this research and further study on the factors that may affect the efficiency score.

Chapter 4 summarizes the results of this study and discusses the possible future research that needs to be studied on paratransit efficiency topic.

CHAPTER 2 LITERATURE REVIEW

This chapter provides a detailed literature review on two topics related to the subject of paratransit efficiency. The first section provides a background description of paratransit service systems, including general definition, classification and system characteristics Section 2 provides an overview of the current issues related to efficiency studying in general and specifically DEA as a tool for efficiency measurement.

2.1 Paratransit as a Specialized Public Transportation Mode

Paratransit, also called demand responsive transit, is a mode of publication transportation that is intended for a group of mobility-challenged population, such as the elderly and the disabled, who have difficult to access regular transit and/or require special equipment or arrangement for their trips. Demand for paratransit service is expected to increase significantly over the ten to 20 years due to the aging of the last baby boomers. Table 2-1 shows the distribution of the Canadian population by different age groups. As it can be observed that 10.6% of Canadians aged from 55-64 will reach to 65-74 in 10 years. And by the year 2025, over 25 percent of the Canadian population will be above age 65. As people age, isolation becomes a serious problem, and access and mobility become increasingly critical needs. For elderly people, affordable and reliable transportation options are therefore essential for them to have:

- The ability to live independently
- Access to medical and social services
- Contact with the outside world

- The feeling of belonging to the community

Table 2-1 Canadian population by age group

	Canada	Male	Female	**Canada**	Male	Female
Age group	Persons (thousands)			% of total of each group		
Total	**31,946.3**	**15,816.5**	**16,129.8**	**100.0**	**100.0**	**100.0**
0–4	**1,695.9**	868.0	827.9	**5.3**	5.5	5.1
5–9	**1,915.2**	980.4	934.8	**6.0**	6.2	5.8
10–14	**2,117.2**	1,085.1	1,032.0	**6.6**	6.9	6.4
15–19	**2,125.9**	1,090.4	1,035.5	**6.7**	6.9	6.4
20–24	**2,223.2**	1,136.4	1,086.8	**7.0**	7.2	6.7
25–29	**2,157.7**	1,092.7	1,065.0	**6.8**	6.9	6.6
30–34	**2,223.4**	1,121.5	1,101.9	**7.0**	7.1	6.8
35–39	**2,408.4**	1,210.4	1,198.0	**7.5**	7.7	7.4
40–44	**2,745.9**	1,379.1	1,366.7	**8.6**	8.7	8.5
45–49	**2,570.1**	1,280.3	1,289.8	**8.0**	8.1	8.0
50–54	**2,235.8**	1,106.6	1,129.2	**7.0**	7.0	7.0
55–59	**1,929.5**	955.5	973.9	**6.0**	6.0	6.0
60–64	**1,457.2**	714.6	742.6	**4.6**	4.5	4.6
65–69	**1,170.9**	563.3	607.6	**3.7**	3.6	3.8
70–74	**1,042.0**	487.5	554.6	**3.3**	3.1	3.4
75–79	**849.1**	367.3	481.8	**2.7**	2.3	3.0
80–84	**611.3**	234.6	376.7	**1.9**	1.5	2.3
85–89	**305.0**	100.2	204.7	**1.0**	0.6	1.3
90 and over	**162.7**	42.6	120.1	**0.5**	0.3	0.7
Note: Population as of July 1.						
Source: Statistics Canada, CANSIM, table 051-0001.						
Last modified: 2004-08-10.						

2.1.1 Definition of Paratransit

“Paratransit” means “subsidiary, assistant or alongside of” transit. It includes all public and private mass transportation provided in the spectrum between private automobile travel and conventional, fixed route, fixed schedule bus and rail transit. Paratransit can extensively include carpools and vanpools (shared ride modes), limousines, public autos (station cars), charter bus, shuttle bus, exclusive and shared-ride taxicab, livery vehicles, and bus transportation operating on flexible routes and flexible schedules in response to individual requests for service, the latter called “demand responsive” or “dial-a-ride” services. – Lave and Mathias 2000

The original definition of paratransit was provided by Kirby (1974) as follows: any forms of intra-urban passenger transportation which are available to the public, and distinct from conventional transit (scheduled bus and rail) and can operate over the highway and transit system.

A modernized definition introduced by R. E. Lave and R. G. Mathias is as follows: Any form of ground, passenger transportation that is demand responsive, that requires the passenger to place a request with the service provider (by hail, telephone, or other electronic means), and that operates with flexible routes and/or flexible schedules tailored to the passengers’ trips.

2.1.2 Types of Paratransit

Since the term “paratransit” was first introduced in 1970’s, there have been many regulatory, social, and technological changes. Although these changes have had an impact on paratransit, basic paratransit service modes have remained the same, and mainly include (Lave and Mathias, 2000):

- Commercially viable general public transportation
- Commercially viable target market transportation

- Subsidized general public transportation
- Subsidized target market transportation
- Subsidized medical care recipients
- Exclusive ride taxi
- Shared ride taxi
- Children's services
- General public dial-a-ride
- Fixed route feeder services
- Night owl service substituting for fixed route
- Service routes
- People with disabilities (in the United States, called "ADA Complementary Paratransit")
- Senior services
- Low income workers (in the United States, called "Welfare-to-work" or "access to jobs")
- Other health and human service agency clients

2.1.3 History of Paratransit

Paratransit services first appeared in 1970s as a personalized transportation model. It quickly gained popularity as a possible means of serving low-density residential areas. The availabilities of computerized control make it possible to combine trips to achieve higher levels of productivity. The anticipation of an efficient dial-a-ride brings interest in paratransit. At the beginning, the productivities did not go up to the levels that were contemplated by the forecast analysis due to expensive large mainframe computers that were required by the computer software (Wilson et al., 197?). However, paratransit services found its place in serving persons with disabilities and older persons, also as the transit service of choice in many small towns and parts of larger cities serving all passengers.

New technologies make paratransit services easier to manage and operate, more affordable to public and private transit service providers. For example, Intelligent Transportation Systems (ITS) technologies, such as Automatic Vehicle Location (AVL) and in-vehicle data terminals, are now being applied to paratransit operations. AVL technology linked with schedule information makes it possible to predict estimated times of arrival. Paratransit operators can more promptly inform their customers on the status of their rides.   This can not only help provide much better customer confidence in the service but also improve inter-vehicle and inter-modal transfers. Additionally, the Inter-vehicle data terminals can provide real-time assistance to drivers on their runs and driving directions.

ITS technologies are being introduced to improve the efficiency, productivity and reliability of paratransit systems in Canada. Emerging ITS technologies of special use to elderly and disabled travelers, such as on-board replication of maps and signs, pre-trip electronic route planning, traffic information broadcast systems, on-board navigation systems, safety warning systems, is currently in use in many bus services such as those in

Hull and Halifax (Sesto Vespa,1995). These allow the control centre to monitor operation of the system and to take action to minimize unreliability; they can also give buses priority at traffic signals. The same technology allows the display of the name of the next stop on-board the bus and, at a bus stop, the number and destination of the next few buses and the waiting times until they arrive. The information displayed at bus stops can also be announced audibly. For example, two Canadian ITS projects were recently started to develop systems to broadcast real-time transit information to hand-held display units or home-based computers. These systems would help elderly people by minimizing the time they need to wait at bus stops in inclement weather.

As an important issue, scheduling and dispatching are studied to improve paratransit performance. Dispatching software was first developed for taxi systems, and has recently been used in paratransit systems. For example, Trapeze Software Inc. provides a general paratransit dispatch system, which has been found to give about 8 - 30 percent increase in productivity. It is also being used to increase productivity by allocating trips that are expensive by paratransit to back-up taxis.

Legislation also played a very important role to implement the paratransit. For example, in the United States public transit must be provided to those who could not use fixed route transit due to their limited mobility, which is called Americans with Disabilities Act (ADA). Another example, now that transit organizations in the US have been required to offer paratransit service, new relationships have been established between fixed route and paratransit services. Any decisions about when and where to offer fixed route services now affects when and where complementary paratransit is provided since paratransit must be provided in the same areas as is fixed route transit. Therefore, if a route is extended into the rural area, the complementary paratransit service area must expand as well. If a weekend service is added, complementary paratransit also must be added. If fixed routes are eliminated, the agency must decide whether to continue serving paratransit passengers in those areas or to drop the service. As paratransit began claiming a larger percentage of the overall transit budget, transit agencies have been forced to look for new and creative ways to balance service and satisfy the needs of passengers using all modes of public transportation.

2.1.4 Paratransit Economics

The design of paratransit services mostly focuses on striking an optimal balance between the efficiency (productivity) of the service, and the quality of the service from the users' point of view (Lave and Mathias, 2000). Depending on the service policy, the quality of a paratransit service could vary significantly, ranging from the most costly exclusive-ride taxi service (one person rides at a time), to the shared ride system (multiple persons share a same vehicle and usually have to rider longer than is needed as compared to an exclusive ride). For a shared ride system, the more passengers assigned to a vehicle at the same time, the better efficiency could be achieved, as it minimizes the total vehicle miles traveled and the number of vehicles required. On the other hand, too many passengers assigned to a same vehicle may lower the quality of the service, as it increases the average ride time and the variability of promised pickup and arrival times. The quality of the service may also be sacrificed by running the system faster and spending less time assisting passengers. Most of the cases, the scheduler/dispatchers will try to maximize the number of trips served by each vehicle hourly with the constraints that the minimum service levels maintained in terms of the longest ride times allowable and the maximum lateness for a promised pickup or arrival. In reality, as most paratransit services are subsidized by public funds, providers are constantly pressured to compete for funds and keep costs down. Other systems with more group trips to agency programs may enjoy higher productivities. Moreover, because of the relative inefficiency of paratransit in relationship to fixed route transit, it is nearly impossible to offer competitive wages to paratransit drivers. They may make half of what their fixed route colleagues make driving for the same transit agency.

The cost of providing the service is another important factor that need be considered in the design of paratransit services. For example, higher salaries will attract a staff with better qualifications and result in lower turnover. Costs for training and incentives will likewise improve service and lower turnover. Expenditures for technology and vehicles may enhance service delivery. If budgets remain the same and ridership increases, productivity may increase out of necessity, but quality may suffer.

The relationship of costs, productivity and service quality has been conceptualized as a triangular relationship and called the “big three” of paratransit service. Sometimes these three may conflict with one another; other times they may complement one another. For example, if a system is very effective, providing high quality customer service and assistance, it may not be as cost-efficient because each passenger is receiving additional attention which may slow down the system and drive up costs overall. On the other hand, customers who routinely receive effective, high quality service also could benefit from higher productivity because of good scheduling and dispatching, which can result in overall cost savings. How to achieve balance between these three is a constant challenge that paratransit and transit managers have to face everyday.

2.2 Efficiency

Efficiency is one of the most popular topic when study urban transit. By simple definition, efficiency is the measure of how much output generated compared to the input. Table 2-2 shows some indicators used to measure efficiency in urban transit studies.

Table 2-2 Common Efficiency Measures of urban transit

Efficiency Measures	
Efficiency Measure	Efficiency Indicator
Cost Efficiency	Cost per km/mile
	Cost per hour
	Cost per vehicle
	Ridership per expense
Cost-Effectiveness	Cost per passenger trip
	Revenue per passenger trip
	Ridership per expense
Service Utilization Efficiency	Passenger trips per km/mile
	Passenger trips per hour
	Passenger trips per capital
Vehicle Utilization Efficiency	Km/miles per vehicle
Labor Productivity	Passenger trips per employee
	Vehicle miles per employee
Coverage	Vehicle km/miles per capital
	Vehicle km/miles per service

There are also some other measurements for urban transit efficiency study. Generally, efficiency can be categorized into two groups:

- Technical Efficiency. Technical efficiency means producing maximum output with given inputs; or equivalently, using minimum inputs to produce a given output.

- Economic Efficiency. Economic efficiency measures producing maximum value of output with given value of inputs; or equivalently, using minimum value of inputs to produce a given value of output.

Technical efficiency is measured by the relationship between the physical quantities of output, and economic efficiency is measured by the relationship between the value of the output and the value of the input. Using technical efficiency, there is always relative efficiency score. When we call a system inefficient, we are claiming that we could achieve the desired output with less input, or that the input employed could produce more of the output desired. When examine the economic efficiency, the value of output over the value of input can get an absolute efficiency score.

Economic efficiency can help to examine profitability for an investment better than technical efficiency. Since the purpose of this study is to improve the operation productivity of Canadian paratransit systems, the research will focuses on the technical efficiency and economic efficiency will not be discussed. The following section provides a detailed introduction on technical efficiency.

2.2.1 Technical Efficiency Measurement

There are two main approaches used to measure technical efficiency, parametric and non-parametric frontier approaches. The parametric/ econometric frontier approach (Aigner and Chu, 1968; Aigner et al., 1977; Meeusen and van den Broeck, 1977) specifies a functional form for the cost, profit, or production relationship among inputs, outputs, and environmental factors, and allows for random error. Both the inefficiencies and the random errors are assumed to be orthogonal to the input, output, or environmental variables specified in the estimating equation. The sensitivity of efficiency estimates to misspecification has been demonstrated using Monte Carlo simulations, where both the true functional form of the technology and the distribution of efficiency across observations are known (Gong and Sickles 1992; Banker, Gadh, and Gorr 1993). From these studies, researchers find out that parametric approaches are best applied to industries with well-defined technologies to minimize the risk of misspecification. For industries with imprecise technologies, such as the service sector, non-parametric approaches are more flexible and could be more desirable to use (Charnes, Cooper, Rhodes 1978).

DEA is a non-parametric frontier approach, which begins with Farrell (1957) who drew upon the work of Debreu (1951) and Koopmans (1951) to define a simple measure of efficiency that could account for multiple inputs. The DEA frontier is formed as the piecewise linear combinations that connect the set of best practice observations, yielding a convex production possibilities set. DEA does not require the explicit specification of the form of the underlying production relationship. As a non-parametric approach, however, DEA do not allow for random error. If random error exists, measured efficiency may be confounded with these random deviations from the true efficiency frontier. As well, statistical inference and hypothesis tests cannot be conducted for the estimated efficiency scores. Bootstrap methods may be used to resolve some of these problems. A detailed discussion on DEA is provided in the following section.

2.2.2 Data Envelopment Analysis (DEA) Method

DEA is commonly used to evaluate the efficiency of a group of producers (also called decision making units or DMUs) such organizations, firms, departments or operating units. A typical statistical approach evaluates producers relative to an average producer. In contrast, DEA is an extreme point method which compares each producer with only the "best" producers. A fundamental assumption behind an extreme point method is that if a given producer, A, is capable of producing Y(A) units of output with X(A) inputs, then other producers should also be able to do the same if they were to operate efficiently. Similarly, if producer B is capable of producing Y(B) units of output with X(B) inputs, then other producers should also be capable of the same production schedule. Producers A, B, and others can then be combined to form a composite producer with composite inputs and composite outputs. Since this composite producer does not necessarily exist, it is sometimes called a virtual producer.

The heart of the DEA technique lies in finding the "best" virtual producer for each real producer. If the virtual producer is better than the original producer by either making more output with the same input or making the same output with less input then the original producer is *inefficient*. Some of the subtleties of DEA are introduced in the various ways that producers A and B can be scaled up or down and combined.

The procedure of finding the best virtual producer can be formulated as a linear program. Assume there are data on k inputs (denoted by the vector x_i) and m outputs (denoted by the vector y_i) on each of N firms or decision-making units (DMUs). The k*n input matrix, X, and the m*n output matrix, Y, represent the data of all N DMUs. The purpose of DEA is to construct a non-parametric envelopment frontier over the data points such that all observed points lie on or below the production frontier. For the simple example of an industry where one output is produced using two inputs, it can be visualized as a number of intersecting planes forming a tight cover over a scatter of points in two-dimensional space. To measure technical efficiency, one has to solve the following linear programming problem for each DMU$_j$, j = 1; . . . ;N (Charnes et al., 1978; Fare et al., 1985):

$$Max\theta = \frac{\sum_{r=1}^{m} u_r y_{r0}}{\sum_{i=1}^{k} v_i x_{i0}} \tag{2-1}$$

$$\text{Subject to } \frac{\sum_{r=1}^{m} u_r y_{rj}}{\sum_{i=1}^{k} v_i x_{ij}} \le 1 \quad \text{for each DMU in the Sample,}$$

$$j = 1,\ldots,N \quad u_r > 0, v_i > 0$$

where m is the number of outputs; u_r is the weight of output r; y_{r0} is the amount of output r produced by the DMU evaluated; k is the number of inputs; v_i is the weight of input i; and x_{i0} is the amount of input i used by the DMU. The value of θ obtained will be the efficiency score for the ith DMU. It will satisfy $\theta \le 1$, with a value of 1 indicating a point on the frontier and hence a technically efficient DMU, according to the Farrell (1957). Note that the linear programming problem must be solved N times, once for each DMU in the sample. A value of θ is then obtained for each DMU.

Due to its non-parametric feature and its ability to combine multiple inputs and outputs, DEA has been found to be a powerful tool when used appropriately. A few of the characteristics that make it powerful are:

- DEA can handle multiple input and multiple output models.
- It doesn’t require an assumption of a functional form relating inputs to outputs.
- DMUs are directly compared against a peer or combination of peers.
- Inputs and outputs can have very different units. For example, X1 could be in units of trips taken and X2 could be bus fare of monthly pass.

The same features that make DEA a powerful tool can also create problems. The following limitations must be considered when choosing whether or not to use DEA (Tim Anderson 1996).

- Since DEA is an extreme point technique, noise (even symmetrical noise with zero mean) such as measurement error can cause significant problems.
- DEA is good at estimating “relative” efficiency of a DMU but it converges very slowly to “absolute” efficiency. In other words, it can tell you how well you are doing compared to your peers but not compared to a “theoretical maximum.”
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research.
- Since a standard formulation of DEA creates a separate linear program for each DMU, large problems can be computationally intensive.
- The standard DEA approach has the disadvantage that it cannot distinguish between changes in relative efficiency brought about by movements towards or away from the efficiency frontier in a given year and shifts in this frontier over time.

- The DEA method assigns mathematically optimal weights to all inputs and outputs being considered. It empirically derives the weights so the maximum weight is placed on those favorable variables and the minimum weight is placed on the unfavorable variables. The underlying assumption of this method is that it is equally acceptable to specialize in producing any output or consuming any input. In many cases this kind of free specialization without weight restrictions is not acceptable or desirable and may lead to highly unreliable conclusions.

2.2.3 The DEA Bootstrap Estimator

An important issue in DEA applications is whether or not a limited set of efficiency estimates from DEA can truly uncover the sources of technical inefficiencies of the evaluated firms. That is, what are the factors that had contributed to the low level of efficiency of a firm? This issue is relevant to managers who wish to focus their efforts to increase efficiency where improvements are most needed. The real world applications of DEA estimators offer no guidance to the statistical inference problem since only point estimates of efficiency are obtained from the estimators. As well, the DEA method is non-parametric and hence traditional parametric inference is impossible. An approach to address this problem is to bootstrap the DEA estimator. The application of the bootstrap to DEA estimators is rather recent and under extensive development. It is worth noting that as yet there is no consensus on a single bootstrap methodology for the DEA estimators.

Bootstrap method was first introduced by Bradley Efron 1979. It has quickly become a popular and powerful statistical tool used to address estimation problems. The problem solved by Bootstrap is mainly an estimation problem. Considering a random sample $X = (X_1, X_2, ...X_n)$ from a population with an unknown distribution F, the goal is to estimate the sampling distribution of some pre-specified random variable R(X,F), based on the real data set x, where $x = (x_1, x_{2,} ...x_n)$ denotes the observed realization of $X = (X_1, X_2, ...X_n)$. The Bootstrap method includes the following steps:

- Construct the sample probability distribution $\hat{F}$, assigning probability 1/n at each point in the observed sample: $x_1, x_{2,} ...x_n$.
- Draw a random sample of size n with replacement from $\hat{F}$ while $\hat{F}$ is fixed at its observed value. This sample $X^* = (X_1^*, X_2^*, ..., X_n^*)$ is defined as the Bootstrap sample $X_i^* = x_i^*, X_i^* \sim_{ind} \hat{F}, i = 1,2,...,n.$.
- The distribution of the random variable R(X,F) is approximated by the bootstrap distribution of $R^* = R(X^*, \hat{F})$.

Behind this principle, the core idea is that given $X = x, \hat{F}$ is the central point of F among all the likely F's, and then R^* should be close enough to R. In theory, when $\hat{F} = F$, it must be the case that $R^* = R$. Theoretically, R^* can be calculated after x is observed. Based on well established facts, the bootstrap has been shown to work satisfactorily in many estimation problems, such as the estimate of the variance of the sample median and confidence intervals. It was also used to estimate the distribution of the regression coefficients when the error term's distributions independently follow some unknown distribution.

The bootstrap method was first applied to non-parametric frontier analysis by Simar (1992). Shortly after that, Atkinson and Wilson (1995) used bootstrap method to construct the confidence interval for the means of the DEA efficiency scores, and to analyze the sensitivity of the DEA efficiency scores related to the variations of the estimated frontier. Nowadays, the bootstrap method has been well recognized as a powerful tool to address the statistic aspects of DEA.

The DEA bootstrap algorithm used in this study is adopted from Simar and Wilson (1998) and Lothgren (1998). It is based on the data generating process (DGP) where the inputs are assumed given by random radial deviations off the isoquant of the input set. Formally, each input in the sample of input–output observations x = {(xi; yi); i=1; . . . n, } is specified as

$$(x_i, y_i) = (x_i^f / \vartheta_i, y_i)$$

where $x_i^f \in IsoqL(y_i)$ is the unobservable frontier input for firm i. the true efficiency measures are drawn from the same distribution, that is, $\theta_i \sim F_\theta, i = 1,...,n$. This DGP represents the idea that, conditioned on the output and the input proportions, the stochastic elements in the production process are completely represented by the random input efficiency measures.

The main idea in the bootstrap simulation is to imitate the DGP. The procedure for the algorithm in each re-sample is as follows. Conditioned on observed output and input proportions, the re-sample data are constructed in two steps. First, the frontier inputs are estimated and bootstrap pseudo-inputs are generated by replicating the DGP in using the estimated frontier inputs and pseudo-efficiencies drawn from some estimate of the distribution, F_θ. The algorithm makes use of a smoothed re-sampling procedure, based on consistency argument (Silverman and Young, 1987). Second, the bootstrapped efficiency estimate is obtained by evaluating the distance from the original input relative to the bootstrap estimate of the frontier.

Following is a brief description of the bootstrap algorithm used in this study. First, the input–output vectors are transformed using the original efficiency estimates, $\{\hat{\theta}_{in}\}$. Second, a kernel smoothing of the empirical distribution of the original efficiency estimate is utilized to generate a smoothed re-sample of pseudo-efficiencies, γ_i^*. The smoothing procedure is based on the reflection method described by Silverman (1986). The reflection method consists of two steps:

1. A small perturbation is added to δ_i^*, the non-smoothed re-sample drawn independently with replacement from the empirical distribution of the original estimates of technical efficiency, $\{\hat{\theta}_{in}\}$.

2. A correction of the re-sampled sequence is applied. A perturbation $h\varepsilon_i^*$ is added to δ_i^*, where h denotes the bandwidth parameter and ε_i^* is drawn from a standard normal distribution to generate the smoothed pseudo-efficiency $\hat{\delta}_i^*$.

An important issue in the application of the smoothing procedure is the choice of the bandwidth parameter h. Several approaches to select the bandwidth h are discussed in Silverman (1986). In this study, we use a simple automatic robust bandwidth selection rule for univariate data proposed by Silverman (1986). The final smoothed re-sampled efficiencies, denoted γ_i^*, are obtained by correcting $\hat{\delta}_i^*$ using the average of the re-sampled original efficiencies. This correction guarantees that the re-sampled efficiency (asymptotically) has the same first two moments as the original efficiency estimates $\{\hat{\theta}_{in}\}$.

Third, the smoothed pseudo-efficiencies, $\hat{\delta}_i^*$, are used to generate the bootstrap pseudo-data set, (x_i^*, y_i^*). Fourth, the pseudo-data set and the DEA linear program are used to estimate the bootstrap efficiencies. The bootstrap efficiency estimate for the ith firm is evaluated as the efficiency of the original input xi relative to the bootstrapped isoquant of the input set. Finally, the second to fourth steps are repeated B times to create a set of B firm-specific bootstrapped efficiency estimates $\theta_{in}^{*b}, i = 1,2,...,n, b = 1,2,...,B$, where n is the number of decision-making units (DMUs).

As noted earlier, the bootstrap allows us to estimate bias and to conduct inference on the DEA efficiency scores. The bootstrap estimate of the DEA estimator bias is given by $\hat{bias}_i^{*} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_{in}^{*b} - \hat{\theta}_{in}$ where the first term on the right hand side is the mean of the bootstrapped efficiency score, and the second term is the original efficiency scores.

The simplest and most straightforward method to obtain confidence intervals is the percentile method. The percentile method is based on the empirical distribution function of bootstrapped efficiencies $\theta_{in}^{*b}, b = 1,2,...,B$, defined as $\hat{F}_i(s) = \frac{1}{B}\sum_{b=i}^{B} I(\hat{\theta}_{in}^{*b} \leq s)$, for any real value s, where $I(\hat{\theta}_{in}^{*b})$ denotes the standard indicator function (Efron and Tibshirani, 1993). A (1-2a) (equal-tail) confidence interval for the true technical efficiency h for the ith firm is given by $(\hat{\theta}_{in}^{*(\alpha)}, \hat{\theta}_{in}^{*(1-\alpha)})$ where $\hat{\theta}_{in}^{*(\alpha)}$ in is the αth quantile of $\hat{F}_i$, that is, $\hat{\theta}_{in}^{*(\alpha)} = \hat{F}_i^{-1}(\alpha)$. The quantiles of $\hat{F}_i$ are given by the $[(B+1)\alpha]th$ and the $[(B+1)(1-\alpha)]th$ ordered values of $\theta_{in}^{*b}, b = 1,2,...,B$, respectively, where γ denotes the integer part of any real value γ .

The DEA estimators are biased in small samples. Simar and Wilson (1998) present a simple and direct approach to bias-correct the percentile intervals above using a simple additive bias correction. The bias-corrected firm-specific (1-2a) (equal-tail) confidence intervals are simply obtained by shifting the bounds in the intervals above by the factor $2*\hat{bias}_i^{*}$ as $(\hat{\theta}_{in}^{*(\alpha)} - 2*\hat{bias}_i^{*}, \hat{\theta}_{in}^{*(1-\alpha)} - 2*\hat{bias}_i^{*})$. The use of correction factor $2*\hat{bias}_i^{*}$ is due to the fact that this correction centers the empirical bootstrap distribution on the bias-corrected estimate $\tilde{\theta}_{in} = \hat{\theta}_{in} - \widehat{bias}_i^{\,*}$. We use the above bootstrap algorithm to estimate bootstrap DEA efficiency scores for our sample of paratransit systems. As well, bias in the original efficiency scores and bias-corrected confidence intervals are estimated for the paratransit systems.

CHAPTER 3 MODELING & ANALYSIS

This chapter describes how to apply DEA to calculate the efficiency score of paratransit systems, and the analysis of the factors which may be associated with the efficiency score. The process starts from understanding and analysis of the data. After defining the inputs and outputs from the source data, a GAMS program is developed to help calculate the efficiency of Canadian paratransit systems. A regression analysis is then followed to find out possible factors which may affect the efficiency score.

3.1 Description of Data

The data used in this research is provided by Canadian Urban Transit Association (CUTA) as part of the annual publication entitled *Summary of Canadian Transit Statistics Operating Data*. The published data includes annual operating statistics and trends of all Canadian urban transit systems. The data used in this analysis are paratransit-specific and cover the period from January 1, 2001 to December 31, 2003.

The data in the annual publication includes both publicly and privately operated transit systems providing conventional and specialized transit services to urban municipalities in Canada. Since transit services in Canada are not subsized by the federal government, there is no uniform information requirement and all data is submitted to CUTA on a voluntary basis. As a result, the data element definitions and accounting procedures employed by individual systems may vary considerably. Also, fare structure, service policies, subsidy levels and the local operating environment may vary from system to system and from province to province. Therefore, cautions must be taken in comparing the performance of different transit systems. That is also the reason why this research focuses only on technical efficiency, leaving out the issue of economic efficiency. Another issue is that economic conditions, demographic trends, development activities and differences in urban spatial structure in Canadian urban areas can cloud comparisons of economic efficiency among different transit agencies.

The data contains general information for both conventional transit services and specialized transit services. The data for specialized transit services is used to analyze the efficiency of paratransit systems. It should be noted that the database does not include information on community bus services and some private non-profit paratransit services (for example, transportation service provided by Canadian Cancer Society). Table 3-1 provides a summary of all paratransit system in Canada for the analysis period. A glossary of the terms used in the table is provided in Appendix D.

Table 3-1 Summary of Comparison in Canadian Paratransit 2000 to 2003

	2000	2001	2002	2003
Number of transit systems reporting	58	60	60	60
Total Vehicle Kilometers	52,465,836	52,524,934	55,649,453	55,753,517
Total Vehicle Hours	2,704,238	2,804,652	2,894,866	2,917,468
Passenger Boardings	10,870,147	11,126,423	11,640,015	11,792,766
Total Direct Operating Expenses $	185,447,066	197,224,952	215,068,952	231,337,741
Total Operating Revenues $	17,595,185	18,631,352	19,740,612	20,449,070
Non-Accessible Cars	193	214	173	360
Accessible Vans / Minivans	716	677	627	794
Small Buses	575	713	827	726
Total Employees	2,350	2,361	2,388	2,472

3.2 Description of Inputs and Outputs

The first step in a DEA is to identify the inputs and outputs that can be potentially used to define the efficiency of a paratransit system. Unlike many other industries where output (e.g. consumer products) is a clearly identifiable entity, the output of a paratransit agency (or transit agency in general) can be quantified in various ways. The basic reason for this difference is that the ‘‘output’’ of a paratransit system is service that cannot be stored for future use. If a paratransit vehicle runs during a time period at half capacity, the system cannot store the other half in its inventory. Once service is produced, it ceases to exist regardless of whether it is consumed. This has led to two separate measures of paratransit output: vehicle-kilometers (often referred to as ‘‘produced output type’’) and passenger-kilometers or passenger boardings (often referred to as ‘‘consumed output type’’).

The output measure in this study is revenue vehicle kilometers. By definition it is the total kilometers traveled for total fare passengers carried, or the total service supplied to fare passengers. The use of Revenue Vehicle Kilometers implicitly avoids the empty travel problem (for example, deadheading, training, roadtests, maintenance or any auxiliary passenger services), in contrast to vehicle kilometers that account for all distances traveled by the paratransit systems.

Transit systems most frequently use three input quantities, namely labor, fuel, and capital to produce output:

- Labor is measured as the total equivalent number of full-time employees hired in providing the paratransit service, including operators, maintenance, and administrative personnel, A part-time employee could be counted as half a full-time employee;
- Fuel is usually measured by the total annual amount of fuel used by the system (in liters). However, the CUTA database includes only the annual fuel expenses. we therefore use fuel expenses as a measure of fuel consumption; and
- Capital is the total number of revenue vehicles owned or leased by the system which are actively used or available for use in revenue service.

3.3 Efficiency Modeling

With the inputs and outputs identified in the previous sections, the basic DEA model for a given paratransit system can be formulated as follows:

$$max. \quad \theta = \frac{u_1 Km}{v_1 Vehicle + v_2 Fuel + v_3 Employee} \qquad (3\text{-}1)$$

$$s.t. \quad \frac{u_1 Km_j}{v_1 Vehicle_j + v_2 Fuel_j + v_3 Employee_j} \le 1 \quad \text{for all paratransit systems}$$

$$u_1 > 0\,;\; v_1,...v_3 > 0\,;\; j = 1,2,...,32.$$

where Km_j is the total **Revenue Vehicle Kilometers** provided by paratransit system j; *Vehicles*$_j$ is the total number of vehicles used in service, $Fuel_j$ the total fuel expenses and $Employee_j$ the total number of employees hired by paratransit system j. ϑ is the efficiency of the paratransit system under study.

The software package GAMS® (General Algebraic Modeling System, GAMS Development Corp.) was used to solve the formulated linear programming problems (refer to APPENDIX A for the GAMS program coded for solving the problem). Table 3-2 provides the solution results indicating the efficiency of individual paratransit systems in Canada for each year. From these results, the following observations can be made;

- The technical efficiency of Canadian paratransit systems varies significantly across systems with values ranging from 0.164 to 1.000. The average efficiency of all systems is 0.687, 0.725 and 0.684 for year 2001, 2002, and 2003, respectively. The variation over the three years is quite consistent with a standard deviation of around 0.29.
- Among all of the systems, the paratransit systems operated by the city of Regina (C20) and Woolwich, Welleley & Wilmot (C31) had consistently outperformed other systems (100% efficient) over the three year period. These systems are the best performers that other paratransit systems may consider as a benchmark for improving their efficiency. This is because the efficiency score is a measure of "relative" efficiency on how well or badly a paratransit system is operated as compared to the most efficient ones.
- In terms of change in efficiency score over the three years, there were 2 system, the city of Oakville and the city of Sarnia, that had experienced noticeable increase in efficiency (over 20% increase per year). There were also 3 systems, the city of Brandon and the city of Prince Albert and the Rocky View District, whose whole efficiency scores had decreased significantly (approximately -15% per year). It would be valuable to find out, e.g. through a survey, what actions had been taken by these systems over these years that had lead to the dramatic changes in their technical efficiency.
- The paratransit service with the lowest efficiency score is offered by the city of Cornwall with an efficiency score of between 0.164-0.232. The system in the city of Windsor is ranked the second worst with an efficiency score of between 0.349-

0.458. Again, it would be interesting to examine the particular environments and service management methods associated with these two cities.

Table 3-2 Efficiency Scores of Canadian Paratransit Systems (2001-2003)

Community		2001	2002	2003	% change 2002-2001	%change 2002-2003
AJAX-PICKERING	C1	0.601	0.635	0.654	5.66%	2.99%
BRANDON	C2	0.532	0.470	0.349	-11.65%	-25.74%
BURLINGTON	C3	0.855	0.633	0.655	-25.96%	3.48%
CALGARY	C4	0.883	1.000	0.984	13.25%	-1.60%
CAMBRIDGE	C5	0.841	1.000	0.751	18.91%	-24.90%
CORNWALL	C6	0.164	0.168	0.232	2.44%	38.10%
GUELPH	C7	0.530	0.524	0.562	-1.13%	7.25%
HALIFAX	C8	0.671	0.631	0.552	-5.96%	-12.52%
HAMILTON	C9	0.798	0.736	0.772	-7.77%	4.89%
KITCHENER-WATERLOO	C10	0.777	0.760	0.754	-2.19%	-0.79%
MEDICINE HAT	C11	0.411	0.541	0.524	31.63%	-3.14%
MONTREAL	C12	0.658	0.777	0.754	18.09%	-2.96%
NIAGARA FALLS	C13	0.619	0.758	0.613	22.46%	-19.13%
NORTH BAY	C14	0.643	0.843	0.711	31.10%	-15.66%
OAKVILLE	C15	0.666	0.824	1.000	23.72%	21.36%
PEEL	C16	0.598	0.680	0.609	13.71%	-10.44%
PETERBOROUGH	C17	0.684	0.716	0.665	4.68%	-7.12%
PRINCE ALBERT	C18	1.000	0.504	0.481	-49.60%	-4.56%
QUEBEC CITY	C19	0.464	0.497	0.721	7.11%	45.07%
REGINA	C20	1.000	1.000	1.000	0.00%	0.00%
ROCKY VIEW DISTRICT	C21	1.000	0.729	0.629	-27.10%	-13.72%
SARNIA	C22	0.644	0.782	0.895	21.43%	14.45%
SAULT STE MARIE	C23	0.539	0.645	0.571	19.67%	-11.47%
SHERBROOKE	C24	0.617	0.730	0.744	18.31%	1.92%
ST. CATHARINES	C25	0.790	1.000	0.849	26.58%	-15.10%
THUNDER BAY	C26	0.660	0.708	0.745	7.27%	5.23%
TORONTO	C27	0.767	0.916	0.905	19.43%	-1.20%
VANCOUVER	C28	0.664	0.799	0.718	20.33%	-10.14%
WELLAND	C29	0.861	0.757	0.719	-12.08%	-5.02%
WINDSOR	C30	0.349	0.458	0.425	31.23%	-7.21%
WOOLWICH, WELLESLEY, & WILMOT	C31	1.000	1.000	1.000	0.00%	0.00%
YORK REGION	C32	0.704	1.000	0.349	42.05%	-65.10%
Average		0.6872	0.7257	0.6841	0.0799	-0.0352
Std.Dev.		0.1939	0.1955	0.1941	0.1958	0.1892
Minimum		0.164	0.168	0.232	-0.496	-0.651
Maximum		1.000	1.000	1.000	0.420	0.451

When comparing the efficiency score with the ratio of revenue to cost, which is a measure of economic efficiency, as shown in Figure 3-1, 3-2 ,3-3 and 3-4, it can be found that technical efficiency and revenue-to-cost ratio do not correlate with each other. And there is no curve that can be fitted to the data set in Figure 3-4. It shows that higher efficiency does not necessarily means higher profitability. And a higher ratio of revenue to cost does not necessarily means more service provided with certain amount of inputs (vehicles, labor, fuel, etc.). The difference between the technical efficiency and economic efficiency suggests the different use of them.

Figure 3-1 Technical Efficiency vs. Revenue/Cost 2001

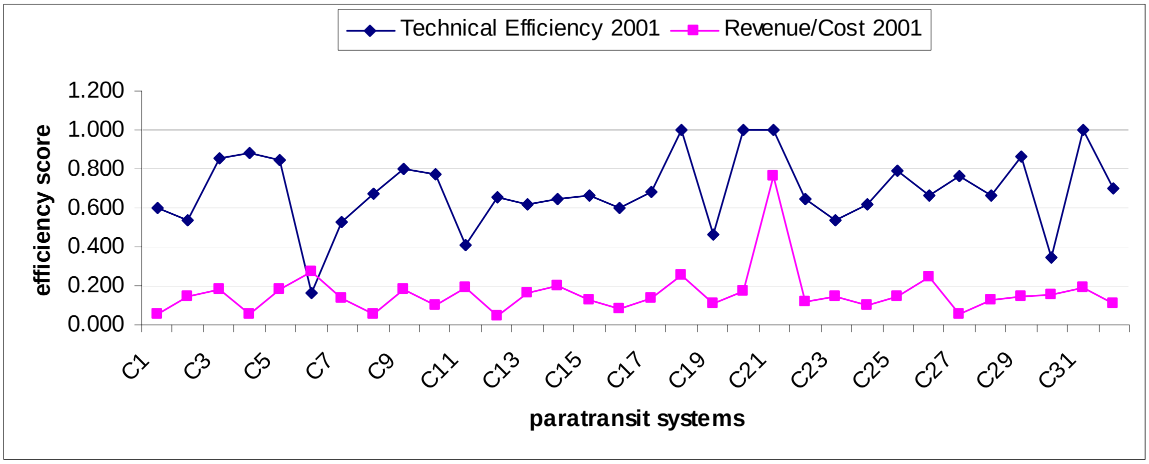


Figure 3-2 Technical Efficiency vs. Revenue/Cost 2002

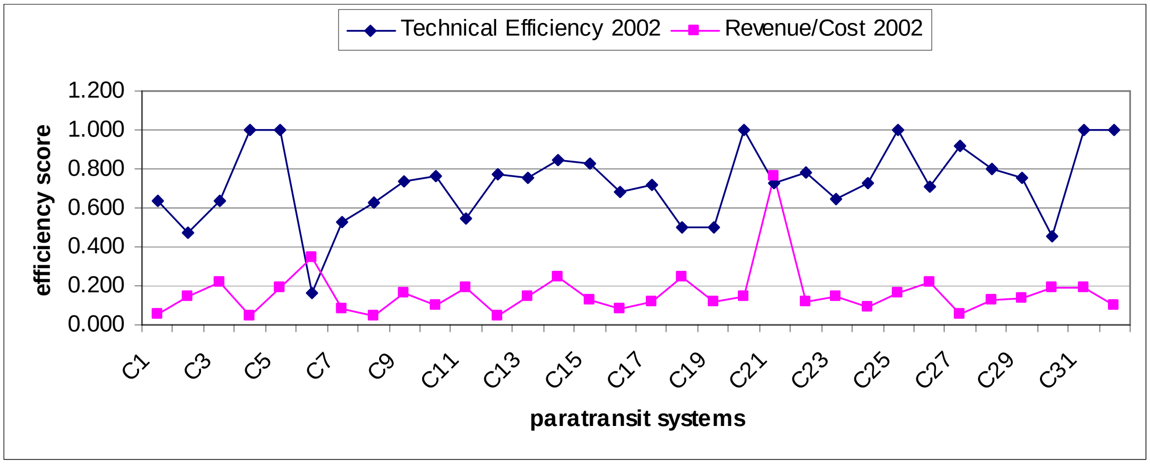


Figure 3-3 Technical Efficiency vs. Revenue/Cost 2003

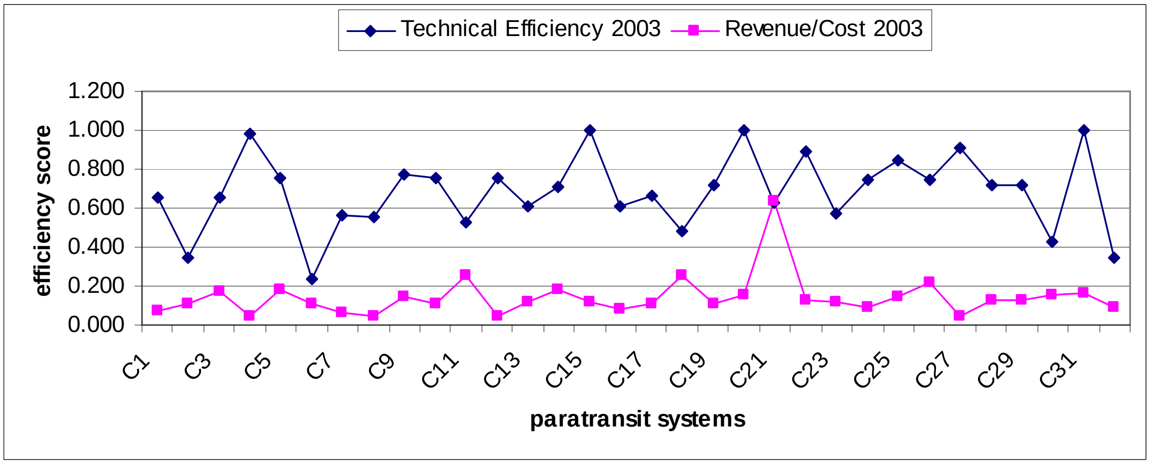
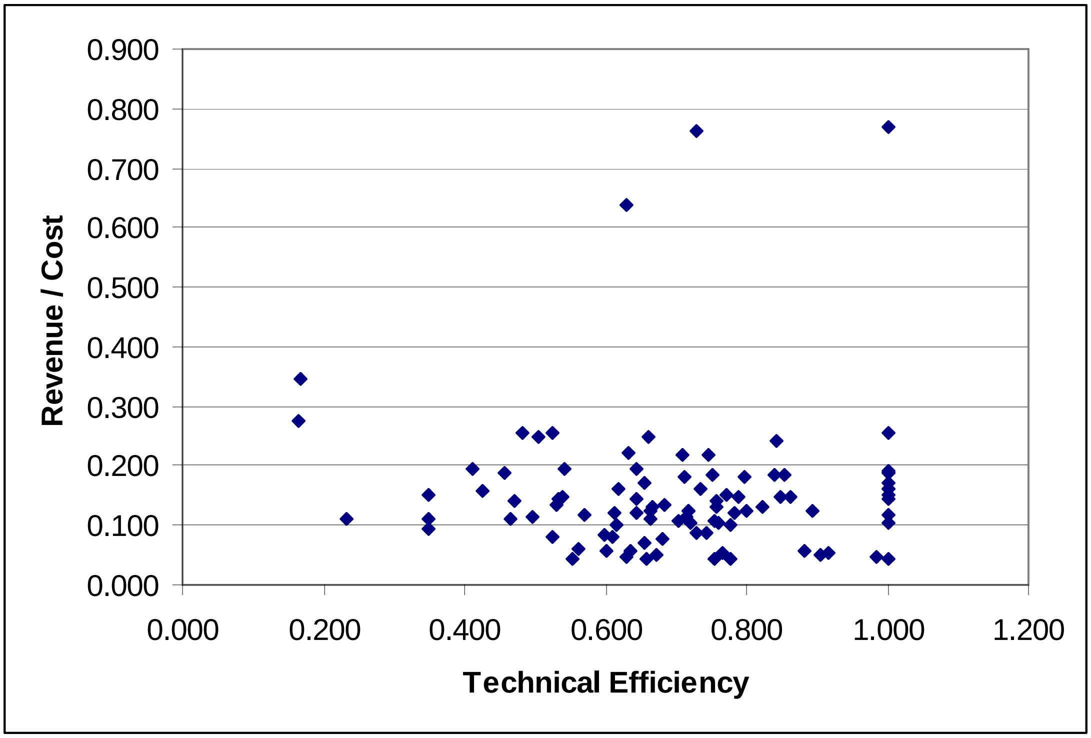


Figure 3-4 Curve fitting Technical Efficiency & Revenue/Cost 2001-2003



3.4 Analysis of Factors Influencing Efficiency

As shown in the previous section, there is a significant variation in technical efficiency across systems. To identify the sources of efficiency or inefficiency associated with these systems, it is necessary to conduct an analysis on the relationship between the efficiency scores of individual systems and their characteristics. The system characteristics are *environment variables that* describe factors that could influence the efficiency of a paratransit system, where such factors are not traditional inputs (or outputs) and are assumed not under the control of the manager. The objective of this section is to find out whether or not and how various environment factors had affected the efficiency of individual paratransit systems. In particular, we are interested in the following three hypotheses:

- Is the level of automation in scheduling a factor influencing the efficiency of a paratransit system?
- Is the average speed a factor influencing the efficiency of a paratransit system?
- Is demand (e.g. density) a factor influencing the efficiency of a paratransit system?

There are several ways in which environmental variables can be linked to efficiency score in a DEA study. We use the two-stage method proposed by Coelli et al., (1998).

The two-stage method involves solving a DEA problem in the first-stage analysis, which includes only the conventional inputs and outputs. In the second-stage, the efficiency scores from the first-stage are regressed on the environmental variables. The sign of the coefficients of the environmental variables indicates the direction of the influence, and standard hypotheses tests can be used to assess the strength of the relationship.

The advantages of the two-stage method include the following: it can accommodate more than one variable; it can include both categorical and continuous variables; it does not make prior assumptions regarding the direction of the influence of the variables; and it is easy to calibrate.

According to previous studies: Chilingerian (1995)-Evaluating physician efficiency in hospitals: a multivariate analysis of best practices; Gillena and Lall (1997)-Airport Performance Measurement: Data Envelope Analysis and Frontier Production Functions; Linna, Nordblad and Koivu-Technical and Cost Efficiency of Oral Health Care Provision in Finnish Health Centers; the disadvantage of the two-stage method is that if the variables used in the first-stage are highly correlated with the second-stage variables, then the result is likely to be biased. The Tobit regression (Tobin, 1958) is used in this case since it can account for censored data. A simple form of the Tobit model is

$$y_t^* = x_t(\beta) + u_t, \quad u_t \sim NID(0, \sigma^2) \tag{3-2}$$

$$y_t = y_t^* \text{ if } y_t^* > 0$$

$$y_t = 0 \text{ if } y_t^* \leq 0;$$

where β is a vector of estimated parameters, y_t is the limited dependent variable, x_t is a vector of independent variables, and u_t is a random, normally and independently distributed error term. y_t^* is a latent variable that is observed only when it is positive.

When the latent variable is negative, zero is observed instead. The Tobit formulation is appropriate since the DEA scores fall between 0 and 1, making it a limited dependent variable. The Tobit model is usually estimated by maximum likelihood.

Following is a brief discussion of the environmental variables that are included in the regression analysis based on the microeconomic foundations of urban transit production and availability of data:

Schedule is a variable used to represent the level of automation in generating schedules and runs by a paratransit service provider. It is expected that higher level of automation may generate more efficient routes and thus lead to higher technical efficiency. In this study, a manual scheduling process is rated 4, a partially computerized scheduling method is rated 2, and a fully automated scheduling is rated 1.

Average Speed is defined as the average ratio of travel distance to travel time for all trips in the service area covered by a paratransit system. Average speed is expected to have a high impact on the efficiency of a system. The higher the average speed in a service area is, the higher the number of trips that can be covered within a given time period and thus the lower the output such as travel cost per kilometer and the vehicle operating cost. An urban transit system with higher average speeds is one where buses stop less often. Less number of stops could also reduce maintenance requirements as well. Thus, higher average speeds should correlate with higher levels of technical efficiency, and vice versa.

User Area Density represents the number of users per unit area. It is easy to understand that the closer all users live together, the more concentrated the pickup/delivery stops and the shorter the trip length. Thus the density of users in an area should be considered when efficiency is examined. In order to include this factor, data of **Service Area** and **Total number of Users** are required.

It is assumed that the efficiency score of a system is linearly related to the variable of scheduling method (Schedule), demand density (Density) and average vehicle speed (Speed) as follows:

$$efficiency_i^* = b_0 + b_1 \times Schedule_i + b_2 \times Density_i + b_3 \times Speed_i + \varepsilon_i \qquad (3\text{-}3)$$

$$efficiency_i = efficiency_i^* \text{ if } efficiency_i^* > 0$$

$$efficiency_i = 0 \text{ if } efficiency_i^* \leq 0$$

where ε is an error term and bj are model coefficients to be calibrated. For the research questions posed above, the following hypothesis concerning the regression coefficients are to be tested:

- $H_0^1 : b_1 = 0$, vs. $H_0^1 : b_1 \neq 0$
- $H_0^2 : b_2 = 0$, vs. $H_0^2 : b_2 \neq 0$
- $H_0^3 : b_3 = 0$, vs. $H_0^3 : b_3 \neq 0$

After obtaining the efficiency score from DEA Model, the least-squares estimate $\hat{b}_j$ for $b_j, j = 1,2,...,32$ is calculated by fitting the linear OLS regression model within the observed sample with the following steps:

- Construct the sample probability distribution $\hat{F}$ by assigning probability of 1/32 to each DMU of the observed 32 paratransit systems.
- Take 2000 random samples of size 32 with replacement from the observed sample of 32 paratransit systems. These samples are the Bootstrap samples.
- Run the DEA model for each Bootstrap sample.
- Within each Bootstrap sample, fit the following linear OLS regression model:

  $$\hat{efficiency}_{ki} = \hat{b}_{k0} + \hat{b}_{k1} \times Schedule_{ki} + \hat{b}_{k2} \times Density_{ki} + \hat{b}_{k3} \times Speed_{ki}$$

  for $i = 1,2,...,32$; $k = 1,2,...,2000$

  where $\hat{efficiency}_{ki}$ is the DEA efficiency score for paratransit system i in Bootstrap Sample k, and $\hat{b}_{kj}$, $j = 1,...,3$ are the Bootstrap replications for $\hat{b}_j$ in Bootstrap Sample k.

- Estimate the standard error $se(\hat{b}_j)$ by the sample standard deviation of the c Bootstrap replications:

  $$\hat{se}(\hat{b}_j) = \left\{\frac{\sum_{k=1}^{c}(\hat{b}_{kj} - \bar{b}_j)^2}{c-1}\right\}^{\frac{1}{2}} \qquad (3\text{-}4)$$

  where

$$\bar{b}_j = \frac{\sum_{k=1}^{c} \hat{b}_{kj}}{c}, \ c = 100,200,...,2000, \ j = 1,...,3$$

The last step is to calculate the test statistic *t* according to $t = \frac{\hat{b}_j}{\hat{se}_c(\hat{b}_j)}$, and then compare the calculated *t* to the critical value $t_{0.025}$ from the student *t* distribution with degree of freedom equal to $100 - 4 - 1 = 95$. If $|t| > t_{0.025}$, reject the null hypothesis $H_0 : b_j = 0$, in favor of $H_0 : b_j \neq 0$, and conclude that the jth factor influences the efficiency of paratransit system at $\alpha = 0.05$ significant level. Otherwise, the null hypothesis $H_0 : b_j = 0$ is tenable and we cannot reject the null hypothesis that the jth factor does not influence the hospitals' efficiency at $\alpha = 0.05$ significant level. The Bootstrap procedure was again coded in GAMS as shown in APPENDIX B. Table 3-3 shows the results of calibration.

Table 3-3. Results of linear regression analysis

OLS Regression	Estimates	Std.Error	t Value	p Value
Schedule	-0.003	0.021	-0.122	0.903
Density	0.004	0.001	2.973	0.004
Speed	0.020	0.003	7.264	1.19E-10
Const	0.233	0.067	3.476	0.0078
Bootsrtap Method				
Schedule	-0.003	0.028	-0.091	0.927
Density	0.004	0.002	2.254	0.027
Speed	0.020	0.004	5.515	3.17E-7
Const	0.233	0.086	2.715	0.008

Results from both OLS regression and Bootstrap regression indicate that both **Users Density** and the **Average Speed** had a statistically significant impact on the technical efficiency of a paratransit system. The positive coefficients associated with these two variables make intuitive sense as they suggest positive correlation between efficiency and demand density and average speed. The variable **Schedule** was found to be statistically insignificant. Initially we thought that this result might be caused by the assignment of numerical values to the different level of automation which did not reflect the difference in efficiency level induced by using different scheduling tools in reality. Subsequently, we tried to model the level of automation as a categorical variable. The result was however the same, that is, there is no significant difference in efficiency between systems with different scheduling methods. While this finding is somehow consistent with some empirical results of several past studies, it does not necessarily conclude that level of automation in scheduling has no contribution to system's efficiency at all. A closer examination of the change in system efficiency as related to change in scheduling method has resulted in mixed results. As indicated in Table 3-4, ten paratransit systems changed their *scheduling tool* over the study period (2001-2003). Three cities including Oakville, Sherbrooke and Windsor changed their scheduling method from *Partially Computerized* to *Fully Computerized* in 2002, resulting higher efficiency score. City of Guelph changed form *Fully Computerized* to *Partially Computerized* and the efficiency score decreased slightly. For the city of Regina, its technical efficiency remained at 100% although its scheduling tool was improved. However, the rest five systems introduced new scheduling tools but experienced lower level of efficiency. It is unknown whether change in scheduling tool is related to the efficiency from analysis of these changes. Further study is therefore required to confirm the effectiveness of scheduling method on the efficiency of a paratransit system.

Table 3-4 Scheduling tool changes vs. Efficiency changes

	Efficiency			Schedule		
Community	2001	2002	2003	2001	2002	2003
AJAX-PICKERING	0.600	0.593		2	1	
GUELPH	0.526	0.504		1	2	
HALIFAX	0.618	0.606		2	1	
NIAGARA FALLS		0.655	0.544		2	1
OAKVILLE	0.664	0.730		2	1	
PETERBOROUGH	0.680	0.666		4	1	
REGINA	1.000	1.000		2	1	
Rocky View District	1.000	0.654		4	2	
SHERBROOKE	0.612	0.703		2	1	
WINDSOR	0.349	0.407		2	1	

After remove the insignificant factor Schedule, the recalibrated efficiency model is:

$$Efficiency = 0.230 + 0.004 \times User\_Density + 0.020 \times Average\_Speed \qquad (3\text{-}5)$$

R Square = 0.3743, t Value = 7.4982

The regression analysis result of efficiency model without Schedule is listed below in Table 3-5.

Table 3-5. Results of linear regression without Schedule

OLS Regression	Estimates	Std.Error	t Value	p Value
Density	0.004	0.001	3.036	0.003
Speed	0.020	0.003	7.264	8.430E-11
Const	0.230	0.060	3.834	2.292E-4
Bootsrtap Method				
Density	0.004	0.002	2.301	0.024
Speed	0.020	0.004	5.591	2.244E-7
Const	0.230	0.076	3.036	0.003

Based on this efficiency model, the following recommendations are made to improve paratransit efficiency:

- locate paratransit garage closer to the higher user density area;
- adopt new types of vehicles to accelerate the passengers loading/unloading and thus increase average vehicle speed, and
- provide training to the drivers to improve their familiarity of the service and area and skills in handling specialized vehicles, etc.

CHAPTER 4 CONCLUSION

This chapter summarizes the major findings and conclusions of the study. It also provides a discussion on some future research options on the subject of paratransit system efficiency.

4.1 Major Findings and Conclusions

This research has applied the Data Envelopment Analysis (DEA) approach to the problem of determining the technical efficiency of paratransit systems. Three years of operating data from 32 Canadian paratransit agencies were used in this analysis. The following is a list of the major findings and conclusions that can be drown with respect to the efficiency of Canadian paratransit systems and the factors that can contribute to the improvement of paratransit efficiency:

1. DEA was found to be effective and relatively easy to use for quantifying the technical efficiency of paratransit systems. Based on the case study of Canadian paratransit systems, it was found that efficiency score was quite sensitive to systems with a wide spread of variation. Large variation in efficiency estimates facilitates the investigation of factors contributing to the efficiency of individual systems.

2. To identify the factors that influence the technical efficiency of paratransit systems in Canada, linear regression analysis was conducted to relate technical efficiency of paratransit systems to various environmental factors. It was found that demand density

and average vehicle speed had a significant impact on paratransit efficiency. Paratransit agency managers can predict their relative efficiency score (comparing with other paratransit systems) by the function:

$$Efficiency = 0.230 + 0.004 \times User\_Density + 0.020 \times Average\_Speed$$

This equation indicates that higher efficiency is associated with higher demand density and higher average speed. Please be noted that the paratransit agencies should not improve their efficiency by encouraging their drivers to drive faster. Instead, agencies should improve their efficiency by locating their station closer to the area of higher user density, routing by shortest time and providing training to the drivers and adopting new types of vehicles for faster passengers loading/unloading. The analysis however could not confirm the effect of the scheduling method used by a paratransit agency in generating daily service runs. This result is somehow counterintuitive and needs to be verified with further research.

3. The analysis does not show any significant difference between OLS regression method and the Bootstrap method. This may be attributed to the simplicity of the model structure with a small number of independent variables. Further research is needed to examine the advantage of the Bootstrap approach.

4.2 Future Research

This research is limited in a number of aspects due to limited availability of operating data. Future research is needed and should focus on the following areas:

- It would be valuable to conduct a survey of paratransit systems to obtain more accurate and detailed information on inputs, outputs, environmental factors, service management factors and distinctive operating practices, etc.
- It is necessary to perform more extensive analysis of the sources of efficiency/inefficiency and influencing factors;
- Future efforts should also devoted to the development of guidelines that paratransit agencies can use to improve their service performance;
- It is important of investigate the effects of other independent variables on paratransit efficiency, such as peak-base ratio of demand or fleet dedicated service verse non-dedicated service ratio, whether or not the employees are unionized, etc.
- The technical efficiency calculated in this study is based on the standard DEA method without placing any restrictions on the input and output weights. Incorporation of reasonable weights restrictions into this DEA model may generate more reliable results.
- Economic efficiency of a paratransit system is also of critical importance to the transit industry and the future development of paratransit services. However the technical efficiency score calculated using DEA is a “relative” score. It only represents how well or badly a firm is operated comparing with its peers, and it is not necessarily related to economic efficiency of a system. The technical efficiency analysis can be used as a basis for evaluating economical efficiency.

REFERENCE

ADAVantage, Inc., Cost Effective Delivery of Paratransit Services Summary Report, American Public Transit Association, Washington, D.C. (September 1996).

ADA Paratransit Plan for the WMATA Region, Washington Area Metropolitan Transit Authority, Washington, D.C. (July 1992).

Andrle, J. Stephen, (1999) “Highlights of the Transit Capacity and Quality of Service Manual”, TCRP Research Results Digests #35

Ball, W. (1995) APTS: A Consideration of Technology Costs and Benefits, University of South Florida, Tampa, Florida.

Balog, N. John, (1996) “Attracting Paratransit Patrons to Fixed-Route Services”, TCRP Report 24

Bower, D. J. (1991) "Automated Paratransit Routing and Scheduling Using a Highway Network Model." Transportation Research Record, No. 1292, pp. 15-23: ill., maps.

Colquhoun, Dave, (2001) “Successful Integration of Calgary Transit into the Community”, paper from APTA Bus and Paratransit Conference, Calgary, Alberta, May 2001.

Competitive Contracting of Transit Services, The Mackinac Center for Public Policy/Reason Foundation, Midland, Michigan, 17pp.

Canadian Urban Transit Association (CUTA), (2001), “Investing in Transit: Canada at the Crossroads” http://www.cutaactu.ca/pdf/issue_paper.pdf

CUTA (2002), “Promoting Better Health Through Public Transit Use”, http://www.cutaactu.ca/pdf/issue_paper_2.pdf

CUTA (2002), “Public Transit and our Quality of Life: Building Better Communities”, http://www.cutaactu.ca/pdf/IssuePaper3ENG.pdf

CUTA (2003), “Profile of the Public Transit Industry in Canada”, http://www.cutaactu.ca/pdf/issue_4.pdf

CUTA (2003), “Better Access Through Mobility: Getting Ahead with Public Transit”, http://www.cutaactu.ca/pdf/IssuePaper6ENG.pdf

CUTA (2003), “Transit's Leading Edge: Innovations in Service and Technology”, http://www.cutaactu.ca/pdf/IssuePaper7ENG_new3.pdf

CUTA (2004), “Transit's Next Generation: Working with Canada's Youth”, http://www.cutaactu.ca/pdf/IssuePaper8ENG.pdf

CUTA (2004), “Transit-Oriented Development: Smart Growth in Action”, http://www.cutaactu.ca/pdf/IssuePaper9ENG.pdf

CUTA (2004), “Bus Rapid Transit: A Canadian Perspective” http://www.cutaactu.ca/pdf/IssuePaper10ENG.pdf

CUTA (2005), “Public Transit and Small Communities”, http://www.cutaactu.ca/pdf/Issue11_feb3.pdf

CUTA (2005), “Transit in Canada: An Industry on the Move”, http://www.cutaactu.ca/pdf/IssuePaper12E.pdf

CUTA (2005), “Special Edition Issue Paper: Towards Kyoto: Public Transit's Role” http://www.cutaactu.ca/pdf/IssuePaperSpec1ENG.pdf

Davies, P., C. Hill, et al. (1991) Assessment of Advanced Technologies for Transit and Rideshare Applications, NCTRP Final Report 60-1A.

Dial, R. B. (1994) ADART: Autonomous Dial-a-Ride Transit, Federal Transit Administration, Washington, DC, 31 pp.

Drachman Institute, Service Routes, Route Deviation, and General Public Paratransit in Urban, Suburban, and Rural Transit Systems. Federal Transit Administration Report A226-7000, U.S. Department of Transportation, Washington, D.C.

Colquhoun, Dave, May 2001. Successful Integration of Calgary Transit into the Community, paper from APTA Bus and Paratransit Conference, Calgary, Alberta, May 2001.

Fielding, G. J. (1975) "Demand-Responsive Transit: An Overview of Integrated Systems." Paratransit Special Report 164, Transportation Research Board, pp. 49-51.

Fu, L., (1999), “On-Line and Off-Line Routing and Scheduling of Dial-a-Ride Paratransit Vehicles”. Computer-Aided Civil and Infrastructure Engineering 14, 1999, pp. 309-319

Fu, L., (1999), “Improving paratransit scheduling by accounting for dynamic and stochastic variations in travel time”, Transportation Research Record, n 1666, 1999, p 74-81

Fu, L., (2000), “Estimation of time-dependent, stochastic route travel times using artificial neural networks”, Transportation Planning and Technology, v 24, n 1, 2000, p 25-48

Fu, L., (2001), "Scheduling Dial-a-Ride Paratransit under Time-Varying, Stochastic Congestion", Transportation Research, Part B, Vol. 36, 6, July, 2002, pp. 485-506.

Fu, L., (2002), “Potential effects of automatic vehicle location and computer-aided dispatch technology on paratransit performance: A simulation study” Transportation Research Record 1760, Transportation Research Board, National Research Council, Washington, D.C., pp. 107-113,

Fu, L., (2002), “A Simulation Model for Evaluating Intelligent Paratransit Systems” Transportation Research, Part A, Vol. 36, pp.291-307,

Fu, L., (2002), “Planning and Design of Flexible Transit Services”, Transportation Research Record 1791, Transportation Research Board, National Research Council, Washington, D.C., pp. 59-66.

Fu, L., (2003), “An Analytical Model for Paratransit Capacity and Quality of Service Analysis”, TRB, Washington, DC

Fu, L., (2004), “Fleet Size and Mix Optimization for Paratransit Services” Transportation Research Record, National Research Council, Washington, D.C., January 10-14

Gray, B. H., (1989), Urban Public Transportation Glossary, Transportation Research Board, Washington, DC, 74 pp.

Jones, W. (1995) “ITS Technologies in Public Transit: Deployment and Benefits”, ITS.

Kirby, R. F., et al (1974) “Paratransit: Neglected Options for Urban Mobility”, Washington, DC.

King, D. R., “Low-Floor Transit Buses”, TCRP Report 15.

Koffman, David, (2004), “Operational Experiences with Flexible Transit Services”, TCRP Report 53.

Lave, R. E. (1996) “A Handbook for Acquiring Demand-Responsive Transit Software”, Transportation Research Board, TCRP Report 18, Washington, DC.

Lave, R. E. and R. G. Mathias (2000) “State of the Art of Paratransit”, Transportation Research Board, Millennium Papers, Washington, DC.

Little, D., (1995) “Impact of Radio Frequency Reframing on Transit Communications” TCRP Reports 11

Martin, L., “How to Compare Costs Between In-House and Contracted Services”, the Mackinac Center/Reason Foundation, Midland, Michigan.

Miller, E. J., et al. (1982) "Computer Application in Paratransit--Reservation, Scheduling and Dispatching: A Case Study." RTAC Forum, Vol. 5, No. 2, pp. 62-73.

National Council on Disability, “Equality of Opportunity: The Making of the Americans with Disabilities Act”, Washington, D.C. (July 1997).

NEXTEA: Section by Section Analysis, Government Affairs Management Associates, Washington, D.C. (1997).

Public Transit: Meeting the Mobility Needs of Elderly and Disabled Persons, American Public Transit Association, Washington, D.C. (October 1985).

Ridesharing and Transportation for the Disadvantaged, in TRR 1170, Transportation Research Board, National Research Council, Washington, D.C. (1988).

Rosalyn, M. Simon, (1997) “Integrating Americans with Disabilities Act Paratransit Service and Health and Human Service Transportation”, TCRP Report 10.

Rosalyn, M. Simon, (1998), “Paratransit Contracting and Service Delivery Methods”, TCRP Report 31.

Stone, J. R. (1992) "Technological Advances in the Taxi and Paratransit Industries: Computer Dispatch and Mobility Management Lead the Way." Taxi and Livery Management, Vol. 4, No. 3), pp. 16-23.

Stone, J. R., A. M. Nalevanko, G. Gilbert, (1993) The Computer Dispatch and Scheduling in the Taxi and Paratransit Industries: An Application of Advanced Public Transportation System, Transportation Research Board Annual Meeting, 14 pages.

Stone, J. R., A. M. Nalevanko, and J. Tsai (1993) "Assessment of Software for Computerized Paratransit Operations." Transportation Research Record, No. 1378, Transportation Research Board, Washington, DC, p. 109.

Teal, R. F. (1993) "Implications of Technological Developments for Demand Responsive Transit." 32 Transportation Research Record, No. 1390, Transportation Research Board, pp. 33-42

Thatcher, Russell, (1995), “Transit Operations for Individuals with Disabilities”, TCRP Report 9.

Transport Canada. “Appendix A: Intelligent Transportation Systems: A Primer.” An Intelligent Transportation Systems (ITS) Plan for Canada: En Route to Intelligent Mobility. November 1999. www.its-sti.gc.ca/en/download/its_plan.htm

Transportation Research Board (1974) Demand-Responsive Transportation, Special Report 147, Washington DC.

Transportation Research Board (1976) Paratransit, Special Report 164 (1976) edited by Sandra Rosenbloom, Washington, DC.

U.S. Congress (1990) Americans with Disabilities Act (ADA), 49 CFR Parts 37 and 38, Serial 12101.

Weiner, Richard, (1998) “ADA Paratransit Eligibility Certification Practices”, TCRP Synthesis Reports #30

Wilson, N. H. M. (1975) "Coordination and Control of Paratransit Services." Paratransit, Special Report 164, Transportation Research Board, pp. 174-178.

APPENDIX A

DEA PROGRAMMING IN GAMS

```
sets  i         "DMU’s" /C1*C29/
      j         ’inputs and outputs’ /Km,  Vehicle, Fuel, Employee/
      inp(j)   ’inputs’ /Vehicle, Fuel, Employee /
      outp(j) ’outputs’ /Km/
;

Table data(i,j)

INPUT THE DATA SET

parameter
x0(inp) ’inputs of DMU j0’
y0(outp) ’outputs of DMU j0’
x(inp,i) ’inputs of DMU i’
y(outp,i) ’outputs of DMU i’
;

positive variables
v(inp) ’input weights’
u(outp) ’output weights’
;

 variable
 eff ’efficiency’
 ;

 equations
   objective ’objective function: maximize efficiency’
   normalize ’normalize input weights’
   limit(i) "limit other DMU’s efficiency";

   objective.. eff =e= sum(outp, u(outp)*y0(outp));

   normalize.. sum(inp, v(inp)*x0(inp)) =e= 1;

   limit(i).. sum(outp, u(outp)*y(outp,i)) =l= sum(inp, v(inp)*x(inp,i));
```

```
model dea /objective, normalize, limit/;

alias (i,iter);

x(inp,i) = data(i,inp);
y(outp,i) = data(i,outp);
parameter efficiency(i) ’efficiency of each DMU’;

loop(iter,
x0(inp) = x(inp, iter);
y0(outp) = y(outp, iter);
solve dea using lp maximizing eff;
abort$(dea.modelstat<>1) "LP was not optimal";
efficiency(iter) = eff.l;
);

display efficiency;
```

APPENIDIX B

BOOTSTRAP ALGORITHM PROGRAMMING IN GAMS

```
set e(j) 'explanatory variables' /Schedule, Density, Speed, CONST/;

INPUT THE DATA SET

alias(e,ee,eee);
parameter XX(e,ee) 'matrix (X^TX)';
          XX(e,ee) = sum(i,data(i,e)*data(i,ee));

parameter Xy(e) 'X^Ty';
          Xy(e) = sum(i, data(i,e)*efficiency(i));

parameter ident(e,ee) 'Identity matrix';
          ident(e,e)=1;

variable
  invXX(e,ee) 'matrix inv(X^TX)'
  dummy
;

equation
  invert(e,ee)
  edummy
;

invert(e,ee).. sum(eee, XX(e,eee)*invXX(eee,ee)) =e= ident(e,ee);
edummy.. dummy=e=0;
model matinv /invert,edummy/;
matinv.solprint=2;
matinv.solvelink=2;
solve matinv using lp minimizing dummy;

parameter b(e);
b(e) = sum(ee, invXX.l(e,ee)*Xy(ee));
```

```
parameter resid(i) 'residuals';
              resid(i) =  efficiency(i) - sum(e,b(e)*data(i,e));

scalar      rss 'residual sum of squares';
            rss = sum(i, sqr(resid(i)));

scalar     df 'degrees of freedom';
           df = card(i)-card(e);

scalar     sigma_squared 'variance of estimate';
           sigma_squared = rss/df;

parameter variance(e,ee);
              variance(e,ee) = sigma_squared*invXX.l(e,ee);

parameter se(e) 'standard error';
              se(e) = sqrt(variance(e,e));

parameter tval(e) "t statistic";
              tval(e) = b(e)/se(e);

parameter pval(e) "p-values";
              pval(e) = betareg( df / (df+sqr(tval(e))), df/2, 0.5);

parameter ols(e,*);
              ols(e,'estimates') = b(e);
              ols(e,'std.error') = se(e);
              ols(e,'t value') = tval(e);
              ols(e,'p value') = pval(e);

display
  "------------------------------------ OLS MODEL ------------------------",
  ols;

set s 'sample' /sample1*sample2000/;
```

```
parameter bs(s,i) 'bootstrap sample';
            bs(s,i) = trunc( uniform(1,card(i)+0.999999999) );

loop((s,i),
  abort$(bs(s,i)<1) "Check bs for entries < 1";
  abort$(bs(s,i)>card(i)) "Check bs for entries > card(i)";
);

alias(i,ii);

set mapbs(s,i,ii);
    mapbs(s,i,ii)$(bs(s,i) = ord(ii)) = yes;

loop((s,i),
 abort$(sum(mapbs(s,i,ii),1)<>1) "mapbs is not unique";
);

parameter data_sample(i,j);

parameter sb(s,e) 'b(e) for each sample s';

loop(s,

   data_sample(i,j) = sum(mapbs(s,i,ii),data(ii,j));
   x(inp,i) = data_sample(i,inp);
   y(outp,i) = data_sample(i,outp);

   solve dea using lp maximizing eff;
   abort$(dea.modelstat<>1) "LP was not optimal";

   XX(e,ee) = sum(i,data_sample(i,e)*data_sample(i,ee));
   Xy(e) = sum(i, data_sample(i,e)*efficiency(i));
   solve matinv using lp minimizing dummy;
   sb(s,e) = sum(ee, invXX.l(e,ee)*Xy(ee));

);
```

```
parameter bbar(e) "Averaged estimates";
            bbar(e) = sum(s, sb(s,e)) / card(s);

parameter sehat(e) "Standard errors of bootstrap algorithm";
            sehat(e) = sqrt(sum(s, sqr(sb(s,e)-bbar(e)))/(card(s)-1));

parameter tbootstrap(e) "t statistic for bootstrap";
            tbootstrap(e) = b(e)/sehat(e);

scalar dfbootstrap 'degrees of freedom';
       dfbootstrap = card(i) - (card(e) - 1) - 1;

parameter pbootstrap(e) "p-values for bootstrap";
            pbootstrap(e) = betareg( dfbootstrap / (dfbootstrap+sqr(tbootstrap(e))),
dfbootstrap/2, 0.5);

parameter bootstrap(e,*);
            bootstrap(e,'estimates') = b(e);
            bootstrap(e,'std.error') = sehat(e);
            bootstrap(e,'t value') = tbootstrap(e);
            bootstrap(e,'p value') = pbootstrap(e);


display
  "----------------------------------- BOOTSTRAP MODEL ------------------------",
  bootstrap;
```

APPENIDIX C

SUMMARY OF INPUT AND OUT PUT INFORMATION

OF CANADIAN PARATRANSIT SYSTEM 2001-2003

ar 2001		RVK	Vehicle	Fuel_Expen.	Employee	Schedule	U_Density	Ave-Speed
AJAX-PICKERING	C1	280772	10	52,557	12.0	2	2.371	16.63
BRANDON	C2	37810	3	7,345	2.0	2	2.147	11.03
BURLINGTON	C3	189700	7	19,776	9.0	1	19.427	18.53
CALGARY	C4	6645965	133	928,208	221.0	1	12.872	22.80
CAMBRIDGE	C5	125875	4	15,016	4.5	1	9.929	21.69
CORNWALL	C6	51026	7	30,503	10.0	1	15.591	6.57
GUELPH	C7	172593	6	40,000	9.0	1	11.114	19.37
HALIFAX	C8	726000	17	205,071	29.5	2	7.980	18.07
HAMILTON	C9	2855844	71	417,337	98.0	2	13.713	20.64
KITCHENER-WATERLOO	C10	656244	15	99,719	31.5	1	18.640	19.64
MEDICINE HAT	C11	250000	12	90,000	13.0	1	50.000	12.76
MONTREAL	C12	2936096	89	468,497	220.0	2	27.628	15.61
NIAGARA FALLS	C13	106545	4	17,557	5.0	2	3.217	20.03
NORTH BAY	C14	160933	5	26,406	8.5	1	11.161	17.34
OAKVILLE	C15	157771	5	27,095	6.0	2	7.255	20.21
PEEL	C16	1001655	28	225,842	45.0	1	2.577	20.66
PETERBOROUGH	C17	270800	7	46,900	13.0	4	19.792	16.41
PRINCE ALBERT	C18	310000	7	27,623	11.5	4	15.221	25.83
QUEBEC CITY	C19	551609	20	149,315	36.0	2	5.798	15.80
REGINA	C20	1034433	29	191,045	6.5	2	40.932	17.75
ROCKY VIEW DISTRICT(AIRDRIE)	C21	265966	6	29,882	6.0	4	0.222	39.04
SARNIA	C22	151412	5	25,856	6.5	2	5.523	16.74
SAULT STE MARIE	C23	163383	7	32,168	8.0	1	7.606	12.97
SHERBROOKE	C24	381740	11	76,932	17.5	2	12.154	15.93
ST. CATHARINES	C25	334358	10	46,021	10.5	1	14.402	19.72
THUNDER BAY	C26	576139	18	99,410	23.5	1	4.954	21.06
TORONTO	C27	8820130	226	1,284,581	358.0	1	38.278	21.02
VANCOUVER	C28	7033264	251	1,098,007	297.0	2	13.889	16.23
WELLAND	C29	90903	2	11,724	4.5	1	11.360	18.66
WINDSOR	C30	186965	13	54,711	15.5	2	18.657	12.89
WOOLWICH, WELLESLEY, & WILMOT	C31	205643	3	28,215	6.0	2	0.940	29.32
YORK REGION	C32	369767	17	54,000	15.0	2	3.478	15.85
Average		1159416.91	32.75	185229	48.72	1.75	13.40	18.65
Std.Dev.		2199097	60.8112	322341	90.7242	0.8799	11.7463	5.6575
Minimum		37810	2	7345	2	1.0	0.222	6.57
Maximum		8820130	251	1284581	358	4.0	50.000	39.04

Year 2002		RVK	Vehicle	Fuel_Expen.	Employee	Schedule	U_Density	Ave-Speed
AJAX-PICKERING	C1	310130	11	65161	11.5	1	1.150	17.61
BRANDON	C2	28749	3	7009	2.0	2	3.726	13.10
BURLINGTON	C3	225894	7	43565	11.5	1	18.630	19.87
CALGARY	C4	6963339	134	851353	216.0	1	15.596	23.69
CAMBRIDGE	C5	123650	4	14120	4.5	1	11.622	21.99
CORNWALL	C6	59436	8	47799	8.0	1	16.126	7.99
GUELPH	C7	191450	7	50000	9.0	2	13.477	19.34
HALIFAX	C8	726,000	17	211360	30.5	1	5.924	21.63
HAMILTON	C9	2,639,902	56	492000	98.0	2	13.457	20.52
KITCHENER-WATERLOO	C10	712271	16	121052	32.5	1	19.921	20.88
MEDICINE HAT	C11	274500	9	85500	11.5	1	12.427	14.37
MONTREAL	C12	3030502	89	466496	225.0	2	28.552	15.96
NIAGARA FALLS	C13	118540	4	18404	5.5	2	3.307	19.84
NORTH BAY	C14	188765	5	26865	7.5	1	13.483	16.84
OAKVILLE	C15	160840	5	23106	6.0	1	8.484	20.25
PEEL	C16	1007078	39	173525	49.5	1	2.365	20.03
PETERBOROUGH	C17	283300	8	48100	13.5	1	24.768	17.60
PRINCE ALBERT	C18	141971	7	33321	11.0	4	15.601	13.08
QUEBEC CITY	C19	586368	22	146992	36.0	2	5.975	15.08
REGINA	C20	1045711	31	164693	7.5	1	38.860	17.40
ROCKY VIEW DISTRICT	C21	178585	8	31000	6.0	2	0.219	43.66
SARNIA	C22	166160	5	26400	6.0	2	6.276	16.69
SAULT STE MARIE	C23	164836	8	29264	8.0	1	8.215	12.93
SHERBROOKE	C24	445000	11	76814	19.0	1	6.733	17.12
ST. CATHARINES	C25	336439	10	38553	10.5	1	15.309	21.46
THUNDER BAY	C26	601267	18	106827	24.0	1	8.092	21.98
TORONTO	C27	9318159	233	1216400	371.0	1	43.743	21.62
VANCOUVER	C28	7034000	262	1009312	300.0	2	13.889	15.89
WELLAND	C29	91205	2	15736	4.5	1	11.186	18.77
WINDSOR	C30	207676	12	54097	13.5	1	19.395	13.59
WOOLWICH, WELLESLEY, & WILMOT	C31	216640	3	30200	6.0	2	0.992	30.41
YORK REGION	C32	300200	17	33000	25.5	2	3.564	12.93
Average		1183705	33.47	179938	49.70	1.44	12.85	18.88
Std.Dev.		2278922	62.5594	302404	92.3398	0.6690	10.2089	6.1626
Minimum		28749	2	7009	2	1.0	0.219	7.99
Maximum		9318159	262	1216400	371	4.0	43.743	43.66

Year 2003		RVK	Vehicle	Fuel_Expen.	Employee	Schedule	U_Density	Ave-Speed
AJAX-PICKERING	C1	316916	11	67936	9.5	1	1.505	20.19
BRANDON	C2	29290	3	9189	3.5	2	5.720	12.61
BURLINGTON	C3	218350	7	43215	11.5	1	19.427	19.16
CALGARY	C4	6487251	133	884682	187.0	1	16.547	24.30
CAMBRIDGE	C5	123796	4	21010	4.5	1	16.826	24.01
CORNWALL	C6	78442	7	45695	9.0	1	21.780	8.79
GUELPH	C7	200103	6	50000	11.5	2	14.932	19.32
HALIFAX	C8	705089	18	201419	33.0	1	6.080	20.85
HAMILTON	C9	2642914	61	471900	105.0	2	7.787	19.39
KITCHENER-WATERLOO	C10	750416	16	141827	33.5	1	21.537	21.62
MEDICINE HAT	C11	272100	9	84500	11.5	1	9.526	14.14
MONTREAL	C12	2988535	86	507218	222.0	2	30.135	15.79
NIAGARA FALLS	C13	112073	5	20973	6.0	1	3.184	18.29
NORTH BAY	C14	176245	5	32677	8.0	1	13.713	17.91
OAKVILLE	C15	173548	5	19014	6.0	1	14.144	21.33
PEEL	C16	1084110	213	201092	56.5	1	2.038	19.49
PETERBOROUGH	C17	270700	8	54200	13.5	1	25.216	16.21
PRINCE ALBERT	C18	147888	7	38490	11.5	4	16.717	13.75
QUEBEC CITY	C19	880282	20	172904	39.0	2	6.578	20.15
REGINA	C20	1066856	32	150432	8.5	1	39.215	16.85
ROCKY VIEW DISTRICT	C21	190149	8	39512	6.5	2	0.133	42.69
SARNIA	C22	196866	5	27553	7.0	2	7.681	16.70
SAULT STE MARIE	C23	143762	8	28431	8.0	1	8.640	12.81
SHERBROOKE	C24	442995	11	81144	19.0	1	7.888	16.54
ST. CATHARINES	C25	326391	10	45189	12.5	1	17.031	20.87
THUNDER BAY	C26	630631	16	118708	21.5	1	10.154	21.11
TORONTO	C27	9377716	249	1262600	378.0	1	49.406	21.22
VANCOUVER	C28	7365936	266	1232756	313.0	2	14.444	15.60
WELLAND	C29	81165	2	15588	4.5	1	11.349	18.26
WINDSOR	C30	195457	12	54033	15.0	1	18.159	13.11
WOOLWICH, WELLESLEY, & WILMOT	C31	224993	3	33634	6.0	2	1.075	28.18
YORK REGION	C32	292637	25	95480	26.0	2	2.967	13.07
Average		1193550	39.72	195406	50.23	1.41	13.80	18.88
Std.Dev.		2273214	71.9351	329835	92.5480	0.6652	10.9848	5.9090
Minimum		29290	2	9189	4	1.0	0.133	8.79
Maximum		9377716	266	1262600	378	4.0	49.406	42.69

APPENIDIX D

Glossary of terms in the source data

Registrants: includes wheelchair users / scooter users, or ambulatory persons registered to use the service, as well as other eligible users, as local defined, e.g. attendants, companions, and other registrants (which include persons with temporary disabilities).

Dedicated Service: Service provided in vehicles exclusively dedicated to the transport of persons with disabilities (e.g. vans, small buses). This service can be provided internally or under contract.

Non-dedicated Service (e.g. taxis): Service available to persons with disabilities provided by non-excusive vehicles, typically taxis under contract.

Taxis Scrip/User Subsidy Program: provides eligible users with coupons that are purchased at a discount and can be used on participating taxi services.

Passenger Data: Passenger data is broken down by type of service (dedicated, non-dedicated/taxi) and by type of user (ambulatory, wheelchair / scooter users. attendants / companions, and other).

Donations/Other for Operating or Capital: Cash donations from corporations, service clubs, private individuals, and other to defray operating or capital expenses.

For specialized transit service (paratransit system), the data information includes:

Total Users: Registrants living within the built-up area provided with paratransit service.

Revenue Vehicle Kilometers/Hours: Annual vehicle kilometers traveled / annual vehicle hours operated by active revenue vehicles (Non-Accessible Cars, Accessible Vans /

Minivans, etc.) in paratransit passenger revenue service, including dedicated and non-dedicated service.

Operations Expenses: includes salaries, wages and benefits (for operators, inspectors, dispatchers, schedulers, management, etc.), uniforms, vehicle licenses and registration, fleet insurance premiums, purchased services (by private contract operators or other municipalities), and net of recoveries or rebates.

Fuel Expenses: for revenue vehicles only, includes gasoline, diesel, propane, natural gas and electric power (for vehicles only, where applicable; does not include electric power for buildings or other uses), and net of recoveries or rebates.

Vehicle Maintenance Expenses: Includes salaries, wages and benefits (for mechanics, servicemen, stores, foremen, management, etc.) parts, materials, supplies, purchased services, and net of recoveries or rebates.

Plant Maintenance Expenses: Includes salaries, wages and benefits (for security, janitorial, tradesmen, supervisors, management, etc.), utilities, parts, materials, supplies, purchased services, shelter maintenance, municipal or property tax, and net of recoveries or rebates.

General and Administration Expenses: Other direct operating expenses not covered above, including salaries, wages and benefits (for general manager's office, planning, marketing, human resources, finance, etc.), liability expenses (other than fleet insurance premiums), advertising, promotion, office supplies, telephone, and net of recoveries or rebates.

Total Operating Expenses: includes administration expenses, internal operations expenses for dedicated service, contract expenses for dedicated services, contract expenses for non-dedicated / taxi services, expenses for the taxi scrip / fare reimbursement program, internal maintenance expenses, and fuel expenses.

Total Vehicles: includes station wagon / sedans/ modified vans, small buses, and other purpose-built vehicles, excludes taxis.

Total Employees: All employees, including operators, mechanics (vehicle maintenance only), other transportation operators, other vehicle maintenance, plant maintenance, and general and administration.

Regular Service Passenger Revenues: From regular passenger services for which the fare system applies, including cash, tickets, passes, etc.

Other Operating Revenues: Revenues from school boards, charters, interest, membership, municipal or other contracts, space rentals, asset disposals, etc.

Total Operating Revenues: Regular service passenger revenues plus other operating revenues.

Peak Vehicles: Maximum number of revenue vehicles required for the weekday a.m. or p.m. peak period, whichever is greater, including scheduled, non-scheduled and auxiliary services.

Base Vehicles: Minimum number of revenue vehicles required for the weekday midday period, including scheduled, non-scheduled and auxiliary services.

Schedule: Tools or method used by paratransit systems to dispatch or schedule for revenue operating, including Manual, Partially Computerized and Fully Automated.

Registrants/Capital: Total Registrants divided by Service Area Population,

Passengers/Registrants: Total Passengers divided by Total Registrants,

Passengers/Hours (Dedicated Service): Dedicated Service Passengers divided by Revenue Vehicle Hours (Dedicated Service),

Kms/Passengers (Dedicated Service): Dedicated Service Revenue Vehicle Kilometers divided by Dedicated Service Passengers,

Average Speed (Dedicated Service): Revenue Vehicle Kilometers (Dedicated Service) divided by Revenue Vehicle Hours (Dedicated Service),

R/C Ratio: Total Operating Revenues divided by Total Operating Expenses,

Net Operating Cost / Capital: [Total Operating Expenses – Total Operating Revenues] divided by Service Area Population,

Total Expense / Passengers: Total Operating Expenses divided by Total Passengers,

Total Expenses / Eligible Passengers: Total Operating Expenses Divided by Eligible Passengers,

Transp. Exp. / Passengers (Dedicated Service): [Dedicated Contract + Internal + Maintenance + Fuel Expense] divided by Passengers (Dedicated Service),

Transp. Exp. / Passengers (Non-Dedicated Service): [Non-Dedicated Contract + Taxi Scrip Program Expenses] divided by Passengers (Non-Dedicated Service),

Transp. Exp. / Hours (Dedicated Service): [Dedicated Contract + Internal + Maintenance + Fuel Expense] divided by Total Vehicle Hours (Dedicated Service).