Report Number CA-11-0026-2

INDICATORS PETE FIELDING

INDICATORS AND PEER GROUPS FOR TRANSIT PERFORMANCE ANALYSIS

Gordon J. Fielding
Mary E. Brenner
Olivia de la Rocha
Timlynn T. Babitsky
Katherine Faust

Institute of Transportation Studies
and
School of Social Sciences
University of California, Irvine
Irvine, California 92717

JANUARY 1984

FINAL REPORT

Document is available to the public through the
National Technical Information Service,
Springfield, Virginia 22161

PREPARED FOR
U.S. DEPARTMENT OF TRANSPORTATION
URBAN MASS TRANSPORTATION ADMINISTRATION
Office of Technology Assistance
University Research and Training Program
Washington, D.C. 20590

Notice

eport Number CA-11-0026-2

INDICATORS AND PEER GROUPS FOR TRANSIT PERFORMANCE ANALYSIS

Gordon J. Fielding
Mary E. Brenner
Olivia de la Rocha
Timlynn T. Babitsky
Katherine Faust

Institute of Transportation Studies
and
School of Social Sciences
University of California, Irvine
Irvine, California 92717

JANUARY 1984

FINAL REPORT

Document is available to the public through the
National Technical Information Service,
Springfield, Virginia 22161

PREPARED FOR
U.S. DEPARTMENT OF TRANSPORTATION
URBAN MASS TRANSPORTATION ADMINISTRATION
Office of Technology Assistance
University Research and Training Program
Washington, D.C. 20590

Notice

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
CA-11-0026-2		

4. Title and Subtitle

INDICATORS AND PEER GROUPS FOR TRANSIT PERFORMANCE ANALYSIS

5. Report Date

January 1984

6. Performing Organization Code

7. Author(s)

Gordon J. Fielding, Mary E. Brenner, Olivia de la Rocha, Timlynn T. Babitsky and Katherine Faust

8. Performing Organization Report No.

9. Performing Organization Name and Address

Institute of Transportation Studies
School of Social Sciences
University of California
Irvine, California 92717

10. Work Unit No. (TRAIS)

CA-11-0026

11. Contract or Grant No.

CA-11-0026

12. Sponsoring Agency Name and Address

U.S. Department of Transportation
Urban Mass Transportation Administration
Office of Management and Demonstrations
Washington, D.C. 20590

13. Type of Report and Period Covered

Final Report
July 1982-
January 1984

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Data from the second year (1979-80) of the Section 15 statistics are used, first to test the validity of a small set of performance indicators for fixed-route bus operations, and second to define relatively homogeneous groups of operators (peer groups) that can be compared. Agencies operating 304 bus systems are included. Rail operations were excluded, as were exclusive, demand-responsive operations. The second year data is both more complete and accurate than that reported for the inaugural year. However, data from the magnetic tape had to be reorganized before it could be used with any of the major statistical software packages.

A large set of performance variables are analysed with factor analysis to establish seven dimensions of transit performance. Seven marker indicators were chosen rather than the nine proposed in previous research.

Cluster analysis is used to create a typology for transit based upon characteristics of operations that are available in the Section 15 statistics. Agency size (measured by total vechicle miles and number of peak vehicles operated), peak to base demand and average bus speed are used to create 12 peer groups.

Results from this research confirm the validity of using a small set of indicators to represent dimensions of transit performance. They will also allow meaningful comparisons between similar systems.

17. Key Words	18. Distribution Statement
Bus transportation, Performance measures, Section 15 Reporting System, Data analysis and collection, Peer Groups.	Available to the public through the National Technical Information Service Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

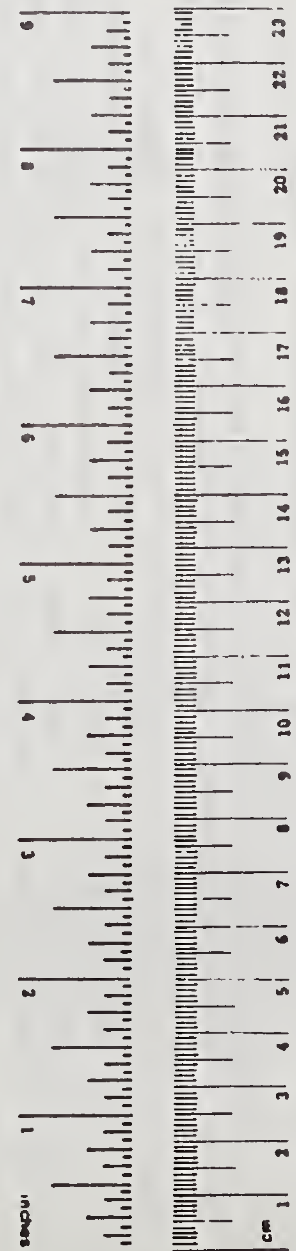
Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

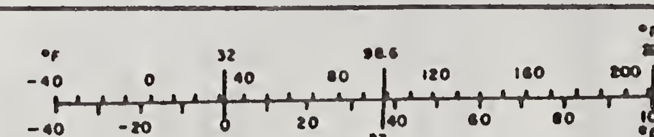
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



ACKNOWLEDGMENTS

In accord with the purposes of the UMTA University Research and Training Program, we have had a high level of student participation in all phases of this research. For several, this was their first exposure to research in transportation. Nancy Minear provided assistance in data analysis and presentation throughout the project, and Kevin Carrigan assisted with the management applications of results. Aki Nogawa and William Billet helped with data preparation.

We would like to acknowledge the assistance received from the Institute of Transportation Studies: to Esther Frank for secretarial assistance; to Gen Giuliano and June Kurata for financial management and to Lyn Long for assistance with bibliographic and research reviews. Jim Ashurst and Mike Slagley of the U.C.I. Computing Facility provided programming assistance. Manuscript typing was completed by the Word Processing Center, School of Social Sciences.

Encouragement and cooperation from UMTA has also affected the direction and quality of the research. Mr. Philip G. Hughes and Norman Ensrud served as technical monitors, and they, together with Michael Jacobs, Transportation Systems Center, helped to obtain access to the Section 15 data. Judy Z. Meade, UMTA University Research and Training Program, served as contract monitor and cheerfully assisted with the administrative budget issues.

G. J. (Pete) Fielding

January, 1984

EXECUTIVE SUMMARY

This research used the second year (1979-80) of the Section 15 statistics, first to test the validity of a small set of performance indicators for fixed-route bus operations, and second to define relatively homogeneous groups of operators (peer groups) that can be compared. Agencies operating 304 bus systems were included. Rail operations were excluded, as were exclusive, demand-responsive operations.

Data Preparation

Chapter 1 reviews the data and the methods used to correct problems and reformat data for statistical analysis. The second year data are both more complete and accurate than that reported for the inaugural year. However, data from the magnetic tape had to be reorganized and validated before they could be used with any of the major statistical software packages.

Performance Indicators

Chapter 2 analyzes a large set of performance variables in conjunction with factor analysis to establish seven dimensions of transit performance. Seven marker indicators were chosen rather than the nine proposed in previous research. The seven marker variables best representing the performance concepts are:

- (RVH/OEXP) Revenue Vehicle Hour per Operating Expense
- (TPAS/RVH) Unlinked Passenger Trips per Revenue Vehicle Hour
- (OREV/OEXP) Operating Revenue per Operating Expense
- (TVH/EMP) Total Vehicle Hours per Total Employees
- (TVM/PVEH) Total Vehicle Miles per Peak Vehicle
- (TVM/MNT) Total Vehicle Miles per Maintenance Employee
- (TVM/ACC) Total Vehicle Miles per Accident

Peer Group Typology

Chapter 3 describes the use of cluster analysis to create a typology for transit based upon characteristics of operations that are available in the Section 15 statistics. Agency size (measured by total vehicle miles and number of peak vehicles operated), peak to base demand and average bus speed are used to create twelve peer groups. This new typology updates and supercedes the typology briefly described in 1982. Each peer group was shown to be distinct in its operating characteristics through statistical analyses and descriptively as follows:

The private bus companies in Peer Group 1 stand out because of their extremely high average speed. They are the smallest in size and the lowest in peak to base ratios relative to other peer groups.

Peer Group 2 consists of transit providers primarily located in small urban areas or suburban areas across the United States with populations under 500,000. They are small (1 to 46 peak vehicles), fast (17 to 22 miles per hour) and have average peak to base ratios.

Although Peer Group 3 is a cross-national group, Southwestern systems are disproportionately represented. While a few systems are in the suburban fringes of major urban areas, most are in small cities or towns. These systems are small (2 to 74 peak vehicles) with low peak to base ratios (1.0 to 1.15) and above average speeds.

Peer Group 4 draws from all parts of the country despite its small size. These systems serve small cities with suburban characteristics. Systems in Peer Group 4 have a high average speed (15.9 to 16.8 miles per revenue vehicle hour) and they tend to be small (fewer than 50 peak vehicles) with low peak to base ratios. Their speed is consistent with their suburban locations.

Peer Group 5 is unusual in that nearly half of its members are private bus companies in the urban New York City area. Most of the rest are small Midwestern city agencies. The systems in this group are distinguished by their very low speeds. They are slightly below average in size, and average in peak to base ratios.

Peer Group 6 draws systems from most regions of the United States but with a particular emphasis on the Midwest and South central regions. While a few medium sized cities are included in this group, many of the systems serve small towns or somewhat rural areas; three-quarters of these systems are in areas with populations under 250,000. Systems in this peer group range in size, but are generally below average in number of peak vehicles. They have low peak to base ratios.

Members of the largest peer group, Peer Group 7, are found in all parts of the United States. They primarily serve small cities and large towns (77,000 to 500,000), although a number are in metropolitan New York. Systems in this peer group are average in size and speed, but above average in peak to base ratios.

Peer Group 8 has primarily Midwestern and Eastern small to medium-sized cities, although a few of its members are from the outer suburban sections of New

York and Chicago. It differs from other peer groups in its high average peak to base ratio (all above 2.3). Systems in this peer group range widely in speed and size, though there are no systems over 400 peak vehicles in this group.

Systems in Peer Group 9 are all from the Southwestern areas of the United States. They predominate in suburban, low density areas with populations between .5 and 1.5 million. Systems in this peer group are above average in size and speed, and about average in their peak to base ratios.

Transit systems in Peer Group 10 are all public agencies in large urban areas (1 to 3 million), in most regions of the United States except the Northeast. These systems have an above average number of peak vehicles (260 to 506) and usually below average speeds, with a wide range of peak to base ratios. Peer Group 10 is similar to Peer Group 11, though the systems are smaller on average and have slightly lower peak to base ratios.

Peer Group 11 includes public transit agencies in major urban areas (1.4 to 16 million) in all regions of the United States. They have a high number of peak vehicles (666 to 1573) and are second in size only to Peer Group 12. These systems are above average in peak to base ratio and are average in speed.

The transit agencies in Peer Group 12 are the major public transit providers in the three largest urban areas of the United States. All three have over 1900 peak vehicles. They are one of the two slowest groups of systems, and they have slightly above average peak to base ratios.

A peer group typology based upon performance characteristics was also devised. However its usefulness is limited because its structure is less clear than the previous typology and it has limited applications in performance evaluation.

Peer Groups and Performance

Chapter 4 describes the performance of each peer group and analyzes the relationships between all 12 peer groups and the seven performance indicators. Much more information is presented on each group so as to clarify the distinguishing characteristics in terms of operating conditions and performance.

Performance profiles were constructed for each peer group by comparing the peer group's average performance on each of the seven performance indicators to the national average for each indicator. Graphical representations of the profiles revealed that each peer group has a distinct pattern of performance across the seven indicators. For instance, Peer Groups one and twelve are both very high and very low on some indicators. But their relative strengths and weaknesses are quite distinct. Other peer groups, such as Peer Groups four and seven, are much closer to the national averages in their performance. It must be emphasized that comparing peer group performance to the national average is used as a descriptive device and is not intended to suggest that these are norms for the transit industry. Each peer group must have its own set of standards.

Statistical analyses were also done to show that the peer groups are significantly different on each performance indicator. Analysis of variance revealed that performance does vary across peer groups for all seven performance indicators. However the results were slightly less significant for revenue generation and maintenance efficiency.

The performance indicators were also examined to see if they adequately discriminated between peer groups both in terms of average performance and the range of values. It was found that each peer group did have its own unique range of values on most of the indicators reflecting important and practical differences between transit systems operating in different circumstances. Some cautions are given for use of certain performance indicators because performance is so varied within some peer groups or the indicator is more valid for certain types of transit systems.

The final analysis in Chapter Four addresses the question of how the four operating characteristics used to form the peer groups relate to performance. Multiple correlation analyses and a comparison of performance profiles showed that size of a transit system, as measured by the number of peak vehicles, was the most important variable in predicting differences in performance. However both speed and peak to base ratio made significant contributions to accounting for differences in performance for some of the performance indicators. Labor efficiency was most strongly related to speed and vehicle efficiency was most strongly related to the peak to base ratio. Revenue generation was not significantly related to any of the operating characteristics although together they predicted a significant amount of the variance between systems.

Use of the Results

Although the results of previous research are already in use, the current research will confirm the validity of using a small set of indicators and encourage meaningful comparison between similar systems.

California is already requiring only five indicators based upon previous research. Other states including Florida, Iowa, Michigan and Pennsylvania have used these same performance concepts to develop performance monitoring and reporting requirements. As a result of this research, they will be able to use the Section 15 data with confidence.

Improved utilization of Section 15 data at the transit agency level promises even more beneficial results. Using the preliminary results of this research, the Orange County Transit District, California, the Transit Department of Seattle METRO and the New York M.T.A. have revised their management information systems to provide quarterly reports representing the major dimensions of performance based upon the Section 15 data format. It will be important to study these results in future years, as well as to examine the consequences for agencies like the Washington D.C. Metropolitan Area Transit Authority that use a much larger list of performance indicators.

A related report was prepared for UMTA to assist in the preparation of the report to Congress on the status of the nation's urban public transportation. Both FY 1980 and FY 1981 data were reported for the seven performance indicators identified by this research. The results were reported as nationally aggregated statistics as well as by the twelve peer groups. These results will allow the Secretary of Transportation to report, not only aggregate changes in American transit, but also changes by national peer group.

Results have already been used in management training. A full day is devoted to the use of Section 15 data in analyzing transit performance at the Transit Managerial Effectiveness Program that is offered by the UMTA University Center for Transit Research and Training located at Irvine. Transit managers become familiar with performance analysis, learn statistical concepts and gain experience with computers by using Section 15 data sets developed in this research.

Another result from the research has been the independent assessment of the Section 15, federal data submission requirement. It is essential that it be continued and the accuracy of information reported be improved. It is also recommended that revisions be made in the current requirements and that more data be requested on the operating environment of each agency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION AND DATA ASSESSMENT

INTRODUCTION

Research on transit performance is important for policy analysis. It allows federal, state and local agencies to outline objectives for transit, to experiment with new forms of service and to monitor performance. It can also assist transit managers: using a small set of performance indicators, they can reliably track overall performance of an agency and evaluate one agency against others with similar operating characteristics.

Evaluation is neglected in public policy research. A great deal of attention is devoted to the formulation of government programs and their implementation. However, few researchers have been concerned with measurement and analysis of programs once they have been implemented. Elinor Ostrom's research on police services is an exception.[1] The present report expands evaluation research by defining measurement and analysis structures for American public transit.

The goal has been to provide a systematic procedure for judging the merits of public transit programs. This is accomplished through constructing a model of transit performance and showing that a small set of seven indicators can reliably represent the major dimensions of transit performance. And further, that these indicators can be used to evaluate individual systems within "peer groups" of similar systems defined by inherent operating characteristics.

Judging the merits of public transit programs is seldom value-free. But, by using the model outlined by this research, analysts can be assured that they are encompassing the major dimensions. Experimentation with different kinds of transit provision and with new methods for satisfying demand is necessary in an industry that has grown rapidly during the past two decades while being responsive to changing policy objectives.[2] Costs per vehicle hour have been rising faster than

---

[1] Elinor Ostrom et al., Community organization and the provision of police services. (Beverly Hills, Calif.: Sage Professional Papers in Administrative and Policy Studies, 1973.)

[2] Gordon J. Fielding, Changing objectives for American transit (Parts I and II), Transport Reviews, 1983, 3 (3 and 4), 287-299, 341-362.

inflation, employee productivity has declined and passengers per revenue vehicle hour have remained static.

Policy and evaluation objectives for public transit are similar to those in other social services: to determine whether service is being produced efficiently and used effectively. As Alice M. Rivlin suggests in Systematic Thinking for Social Action: "Unless we begin searching for improvements and experimenting with them in a systematic way, it is hard to see how we will make much progress in increasing the effectiveness of our social services."[3]

Current State Use of Section 15 Data

A number of states have already undertaken the task of identifying the important dimensions of transit performance. Site visits and questionnaire surveys were used in Florida to define eight operational goals. Quantitative measures derived from the Section 15 data base were then identified for each of the goals, and standard value ranges established.[4]. Special software was developed in Iowa so that both urban and rural transit systems could easily be compared using the categories defined by the UMTA Section 15 requirements. The software system was tested and then used to develop performance standards for use in performance audits.[5] Michigan has also developed an evaluation system using Section 15 data to promote the efficient and effective use of state funding for transit.[6]

While all of these state applications demonstrate the usefulness of Section 15 data for performance evaluation, much less effort has been spent to develop a nationwide system of performance evaluation. Within-state analyses suffer because

---

[3]Alice M. Rivlin, Systematic thinking for social action. (Washington: The Brookings Institution, 1971), p. 119.

[4]Post, Buckley, Schuh and Jerrigan, Inc., Florida transit system performance measures and standards. (Tallahasse, FL: Florida Department of Transportation, Public Transportation Operations Division, 1979.)

[5]Iowa Department of Transportation, Uniform data management system: System development and testing. Report No. DOT-I-81-2. (Ames, Iowa: Iowa Department of Transportation, October 1980.)

[6]James M. Holec, Dianne S. Schwager, and Angel Fandalian, Use of federal Section 15 data in transit performance: Michigan program. Transportation Research Record No. 765. (Washington, D.C.: Transportation Research Board, 1980), pp. 36-38.

of the relatively few number of transit systems in any one state. With a larger number of transit systems, it is possible to better establish the reliability and internal validity of specific performance indicators.

Finding a set of "peer" transit systems in an intra-state comparative analysis is also difficult. Few states have more than a few large urban transit systems and some states have only a limited number of transit providers of any size. The establishment of peer groups across state boundaries allows for the construction of reasonably sized peer groups while controlling for the large variation in operating environments. Current within-state peer groups typically vary in limited ways--by size, mode or urban-rural location. A set of peer groups for the nation can capture finer distinctions in other aspects of the operating environment.

The research described in this report updates previous work on a nationwide approach to performance evaluation.

Links to Previous Research

The first year of Section 15 reported statistics (FY 1979)[7] was used by Anderson and Fielding[8] to test the performance concept model developed by Fielding, Glauthier, and Lave.[9] Nine dimensions of performance, developed from 60 measures, were used to develop a performance index which could be applied to individual transit properties. Although the results of this previous research have been widely used, the Principal Investigator was concerned over the validity of the results because of the limitations of the inaugural data from the Section 15

---

[7]U.S. Department of Transportation, Transportation Systems Center, National urban mass transportation statistics: First annual report Section 15 reporting system: Transit financial and operating data reported for fiscal years ending between July 1, 1978 and June 30, 1979. Report No. UMTA-MA-06-0107-81-1. (Washington, D.C.: U.S. Government Printing Office, May 1981.)

[8]Shirley C. Anderson and Gordon J. Fielding, Comparative analysis of transit performance. Final report No. UMTA-CA-11-0020-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January 1982.) (NTIS No. PB 82-196478).

[9]Gordon J. Fielding, Roy E. Glauthier, and Charles A. Lave, Development of performance indicators for transit. Final report No. UMTA-CA-11-0014-78-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, December 1977.) (NTIS No. PB 278 678).

requirements. The Transportation Systems Center (TSC) compiled the first annual report based on FY 1979 data but warned that "care should be taken in the application and use of the data as presented."[10] Although there was extensive checking and editing, reporting deficiencies and erroneous data remained in the FY 1979 data tape supplied to the UCI research team. Omission of important data was common and in other instances obviously erroneous data remained uncorrected. Where possible the errors were corrected or entries deleted so that the data set (UCI data set) used for the research differed from the TSC data set. Replication of the results published in 1982 was a principal reason for conducting the current study.

The second year data (FY 1980) became available in July 1982 and was both more complete and more carefully verified by TSC.[11] Therefore, research was proposed which would:

1. Use the FY 1980 data to test the validity of the set of performance indicators developed from the FY 1979 data for fixed-route bus operators, and
2. Define relatively homogeneous groups of operators (peer groups) that could be compared in terms of performance.

Agencies operating 304 bus systems were included in the study. Rail operations were excluded. This includes the exclusive operators like the Bay Area Rapid Transit District and rail operation statistics which are reported by mixed-mode operators like the Chicago Transit Authority. Bus operating statistics for mixed-mode operators were included. Exclusive, demand-responsive operators were excluded.

Chapters in this report respond to the two objectives. The latter portion of this chapter reviews the data and the methods used to correct problems and to reformat data for statistical analysis. Chapter 2 reports the analysis of a large set of

---

[10] U.S. Department of Transportation, Transportation Systems Center, National urban mass transportation statistics: First annual report Section 15 reporting system: Transit financial and operating data reported for fiscal years ending between July 1, 1978 and June 30, 1979. Report No. UMTA-MA-06-0107-81-1. (Washington, D.C.: U.S. Government Printing Office, May 1981), p. vi.

[11] TSC supplied the data as a magnetic tape divided into 62 data files. Although this same data was used to prepare the Second Annual Report of National urban mass transportation statistics, 1982, the format is quite different. This is discussed later in this chapter.

performance variables in conjunction with factor analysis to establish seven dimensions of transit performance. The chapter also reports tests of the validity of using a small set of indicators to represent these dimensions. Seven marker indicators were chosen rather than the nine proposed in the 1982 report. Hence this chapter revises the previous results.

Chapter 3 describes the creation of a typology for transit systems using characteristics of transit operations that are available in the Section 15 statistics. Agency size, peak-to-base ratio and average bus speed are used to create 12 peer groups. This new typology updates and supercedes the typology briefly described in 1982. Chapter 4 describes the performance profile of each group and analyzes the relationships between all 12 peer groups and the seven performance indicators.

Chapters 1-3 are somewhat technical. Chapter 4 summarizes the achievements of this research in less technical terms.

Use of the Results

Although the results of previous research are already in use, the current research will confirm the validity of using a small set of indicators and encourage meaningful comparison between similar systems.

California is already requiring only five indicators based upon the 1977 research.[12] Other states including Florida, Iowa, Michigan and Pennsylvania have used these same performance concepts to develop performance monitoring and reporting requirements.[13] As a result of this research, they will be able to use the Section 15 data with confidence, and change the weights of the dimensions to emphasize either efficiency or effectiveness attributes.

Improved utilization of Section 15 data at the individual transit property level promises even more beneficial results. Using the preliminary results of this research, the Orange County Transit District, California, and the Transit

---

[12]California. Business and Transportation Agency. Transportation Development Act: Statutes as amended and related sections of the California Administrative Code as adopted by the Secretary of the Business and Transportation Agency. Report No. DMT-032. (Sacramento, Calif.: California Department of Transportation, Division of Mass Transportation, February 1978.)

[13]James H. Miller, The use of performance-based methodologies for the allocation of transit operating funds, Traffic Quarterly, October 1980, 34(4), 555-585.

Department of Seattle METRO have revised their management information systems to provide monthly and quarterly reports representing the major dimensions of performance based upon the Section 15 data format. It will be important to study these results in future years, as well as to examine the consequences for agencies like the Washington D.C. Metropolitan Area Transit Authority that use a much larger list of performance indicators.

A related report[14] was prepared for UMTA to assist in the preparation of the report to Congress on the status of the nation's urban public transportation. Both FY 1980 and FY 1981 data were reported for the seven performance indicators identified by this research. The results were reported as nationally aggregated statistics as well as by the 12 peer groups. These results will allow the Secretary of Transportation to report, not only aggregate changes in American transit, but also changes by national peer group. As more years of Section 15 data become available, it will be possible to do more longitudinal studies of changes in different types of transit providers as represented by the peer groups.

Results have already been used in management training. A full day is devoted to the use of Section 15 data in analyzing transit performance at the Transit Managerial Effectiveness Program that is offered by the UMTA University Center for Transit Research and Training located at Irvine. Transit managers become familiar with performance analysis, learn statistical concepts and gain experience with computers by using Section 15 data sets developed in this research.

Another result from the research has been the independent assessment of the Section 15 federal data submission requirement. The research results demonstrate the usefulness of the requirement. It is essential that it be continued and accuracy improved. It is also recommended that revisions be made in the current requirements and that more data be requested on the operating environment for each property. Suggestions have been made by the research team to the UMTA Section 15 Reporting System Advisory Committee appointed by the U.S. Secretary of Transportation in 1983 on sampling of passenger statistics, deletion of "road call" data and improved definition for accidents statistics.

---

[14]Gordon J. Fielding and Katherine Faust, Dimensions of bus performance for peer groups of transit agencies in fiscal years 1980 and 1981 using Section 15 data. Working Paper No. 83-5. (Irvine, Calif.: University of California, Institute of Transportation Studies, December 1983.)

DATA AVAILABILITY

Statistics reported in compliance with Section 15 of the Urban Mass Transportation Act of 1964, as amended, were used as the source for all data used for analysis since the purpose of this project was to develop a system of transit evaluation that would be applicable and available throughout the United States. Data for fiscal year 1980 (the second year of data) were obtained from the Transportation Systems Center (TSC), Cambridge, Massachusetts. The data were supplied in the form of a magnetic tape divided into 62 data files, roughly corresponding to the reporting forms filled out by transit systems. The data tape differs from the published version of the statistics[15] in that it is more comprehensive and organized in a more complex way.

Although the Section 15 tape is the most extensive and uniform set of data available for transit at the current time, there are major problems which must be overcome. The magnetic tape must be substantially reorganized before it can be used with any of the major statistical software packages. The tape has a complex organization because it includes information from four different reporting levels (R, A, B, C) and for seven different modes of transit. Any single transit system will be reporting at only one reporting level and for some sub-set of modes. Further, some individual items on the forms are irrelevant to a specific transit system. Rather than leaving blank space for the irrelevant items or coding them as missing information, TSC has employed a hierarchical coding scheme which allows for the economical and methodical coding of different sub-sets of information for each transit system. The following section of this chapter examines this problem in detail and demonstrates how the hierarchical structure was converted to a statistical structure. The data themselves must be carefully scrutinized for validity and reliability.[16] Since few researchers have worked with the Section 15 data tape, the data reorganization and validation procedures are described in some detail

---

[15]U.S. Department of Transportation, Urban Mass Transportation Administration, National Urban Mass Transportation Statistics: Second Annual Report, Section 15 Reporting System. (Washington, D.C.: U.S. Department of Transportation, July 1982.)

[16]Beginning with the FY 1981 data, TSC started to validate the data with methods similar to the ones reported here. Thus the quality of data provided will improve with each successive year. However the same organizaton of the data tape will be used at least until FY 1982.

in this report. Further information can be obtained from the technical report completed as part of the grant.[17]

Certain kinds of information are not available from Section 15 data. Detailed information on the service areas of specific transit systems is not included. Thus it is not possible to examine how well individual transit systems serve specific sub-populations (e.g., elderly, low income, transit dependent) or how geographical features of service areas (urban density, mean temperatures) affect service demand and cost. Organizational features of transit systems such as whether the labor force is unionized, the ownership is public or the primary service is commuter oriented are also unavailable. An earlier project[18] tried to use published sources of information to supplement the Section 15 data but found that these sources seldom had units of analysis that were comparable to the service areas of transit systems.

It is possible to combine Section 15 data with census data by examining an urban area as defined by census standards (e.g., SMSA) as the unit of analysis and aggregating the information for all transit systems that serve that urban area. Vaziri and Deacon[19] have demonstrated how performance evaluation can be done in this way. However, this project wished to provide results that would be useful to the managers of individual transit systems. So an urban area analysis would be relevant to only a few large transit systems whose service area corresponded to the urban area. It was decided to work with individual transit systems using primarily Section 15 data. These data are most readily available to transit managers, not only in magnetic tape form but in published form and on diskettes for microcomputers.

Preparation of the data was done in three major phases--reorganization of the data into a format suitable for statistical analysis, calculation and validation of data values, and evaluation of the quality and properties of specific variables.

---

[17]Gordon J. Fielding, Mary E. Brenner, and Olivia de la Rocha, Using Section 15 data for transit performance analysis. Interim report No. UMTA-CA-11-0026-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January 1983.)

[18]Shirley C. Anderson and Gordon J. Fielding, Comparative analysis of transit performance. Final report No. UMTA-CA-11-0020-82-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January 1982.)

[19]Manoucher Vaziri and John A. Deacon, Application of Section 15 and census data to transit decision making. Final report No. UMTA-KY-11-0002-83. (Springfield, Va.: National Technical Information Service, 1983.)

DATA REORGANIZATION

One of the most important problems to be overcome in working with the Section 15 tape was reorganizing the data into a form suitable for statistical analysis. In the following sections the circumstances making reorganization of the files necessary and the steps required to accomplish the reorganization are explained. After a brief discussion of the problems connected with defining electronic data for statistical analysis, it will be shown how the tape files diverge from a conventional statistical format. The discussion will then focus on the reorganization process both conceptually and from the point of view of programming. It will conclude with remarks on other data processing problems encountered in managing the tape and a summary of the files and variables reorganized.

Defining Data for Statistical Analysis

Several background concepts are useful in understanding the nature of the organizational problems. While all the numbers in a data file are organized in rows and columns, the meaning of the numbers is not inherent in the row and column organization. It must be conveyed to the computer by the programmer. The system or scheme used by the programmer to give meaning to the array of numbers is the logical organization.

The specification of the logical organization is laid out in a document called a codebook. In a codebook the meaning of data is defined by the way the numbers are organized into sets of columns. A large number like $4,000,000 takes up 7 columns, for example. The assigned sets of columns are called fields, and each unique set of information items filling up the fields is called a record.

Table 1-1 shows a codebook from TSC's documentation. According to the codebook, columns one through four of the number array have been reserved for Transit System ID. Columns five through twelve are reserved for the fiscal year end date for the system which is identified in columns one through four. Column thirteen is assigned to the mode code. And so it goes. With the help of this scheme the computer can be informed about the meaning of the data by the way fields in the block of numbers are assigned. This process is called formatting.

By formatting, fields are named so that any number found in that space by the computer can be presumed to have the assigned meaning. The computer can then interpret each record it encounters in the data array by the same standard. There is

TABLE 1-1. A SAMPLE CODEBOOK

COLUMN	NAME	TYPE	DESCRIPTION
1 - 4	TRSID	INTEGER	TRANSIT SYSTEM ID
5 - 12	FY	DATE	FISCAL YEAR
13 - 13	MODE	INTEGER	MODE CODE
14 - 15	EMCOD	INTEGER	EMPLOYEE CLASS CODE
16 - 21	OLABR	REAL	OPERATING LABOR
22 - 27	CLABR	REAL	CAPITAL LABOR

some flexibility in the way data may be formatted, and there may be more than one, meaningful logical organization for the same data file.

Two additional concepts fill out the data definition problem for statistical anaysis. Statistical procedures operate by making systematic comparisons among objects. The objects are compared on those attributes which have been measured in some way. For example, in later analyses transit systems are compared on such attributes as size of fleet and speed.

In statistical data files the most important organizational units are cases (objects) and variables (attributes). A case may be thought of as the full collection of information items defined in the codebook for a single transit agency. If some defined item is missing, the statistical case is incomplete, and a place-holding code must be inserted to fill it out.

A variable, like a case, is a statistical concept. When all cases have been measured on a given attribute, the resulting collection of values is organized in a list called a variable. Statistical procedures compare these lists and depend on the fact that cases always appear in the same order. Once again if no place-holder resides in the position of a missing item, the order is disturbed and statistical results are rendered meaningless.

In formatting data to be read by a statistical package like SPSS[20] or BMDP,[21] the electronic definition of cases and variables depends on there being a

---

[20]Norman H. Nie, C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner and Dale H. Bent, SPSS: Statistical package for the social sciences. (New York: McGraw-Hill, Inc., 1975.)

[21]W. J. Dixon, Ed., BMDP Statistical software 1981. (Los Angeles: University of California Press, 1981.)

uniform amount of information for every case and every variable. Each case must have values or stand-in values for every variable, and every variable must have values or stand-in values for every case.

The Divergence of Tape Files from Conventional Statistical Format

The circumstances leading to the need to reorganize the tape files arise in the way the structure of the files is closely linked to the reporting forms. This close linkage leads to two problems in formatting the data files for statistical purposes. First, information whose presence was predicted by the reporting forms was absent in the actual data requiring the insertion of place-holding values. Second, preparation of the data in the same format as the reporting forms required the design of a new codebook before the data could be read for statistical purposes. Form 404, Transit System Employee Count Schedule, can be used to illustrate how both problems arise.

Missing records. Figure 1-1 shows Form 404 and the information submitted by one transit system as it is recorded in a data file on the tape. A comparison of the form and the data shows the first three fields TRANSIT SYSTEM ID, FISCAL YEAR ENDED and MODE coming from the top of the form and repeating on every record in the data. The next two fields, EMPLOYEE CLASSIFICATION (EC) and OPERATING LABOR (OLABR) are taken from the "Employee Classification" and "Operating Labor" sections of the form. (Information about capital labor is omitted from the example.) The Figure shows a one-to-one correspondence between the numbers assigned to employee categories on the form (11, 12, 13, etc.) and the values under EC in the data. However, the one-to-one correspondence is not quite complete. If Form 404 were used to construct a codebook which acted as the logical organization for the data appearing in Figure 1-1, then there would be a discrepancy between what the logical organization predicts and what actually appears in the data file. There is no record appearing for category 22, Maintenance Support Personnel, in the data file.

This circumstance violates the requirement of statistical software packages for complete information in all variables and cases, and some entry to stand in for the missing category 22 must be supplied or the data will not be correctly read. Until a stand-in value (or dummy record) is inserted, the information cannot be said to form a complete, statistically analyzable case. Therefore, all such instances of "missing" information had to be remedied before statistical analysis could commence.

FIGURE 1-1. CORRESPONDENCE BETWEEN REPORTING SYSTEM FORMS AND THE ORGANIZATION OF TSC DATA FILES

Form No. 404

TRANSIT SYSTEM EMPLOYEE COUNT SCHEDULE

Transit System ID [1056] Level [R]

Fiscal Year Ended [06] [30] [80] Mode motorbus Code [1]

EMPLOYEE CLASSIFICATION	OPERATING LABOR
11. Transportation Executive, Professional and Supervisory Personnel	4.5
12. Transportation Support Personnel	2.5
13. Revenue Vehicle Operators	47.8
21. Maintenance Executive, Professional and Supervisory Personnel	2.3
22. Maintenance Support Personnel	--
23. Revenue Vehicle Maintenance Mechanics	5.6
24. Other Maintenance Mechanics	.5
25. Vehicle Servicing Personnel	2.6
31. General Administration Executive, Professional and Supervisory Personnel	1.0
32. General Administration Support Personnel	2.3
00. TOTAL TRANSIT SYSTEM EMPLOYEES	67.1

ID	FY	M	EC	OLABR
1056	19800630	1	11	4.5000
1056	19800630	1	12	2.5000
1056	19800630	1	13	47.800
1056	19800630	1	21	2.3000
1056	19800630	1	23	5.6000
1056	19800630	1	24	.50000
1056	19800630	1	25	2.6000
1056	19800630	1	31	1.0000
1056	19800630	1	32	2.3000
1056	19800630	1	00	67.100

ID=ID NUMBER
FY=FISCAL YR END DATE
M=MODE
EC=EMPLOYEE CODE
OLABR=OPERATING LABOR (CAPITAL LABOR VALUES OMITTED)

Designing a new codebook. A second important consequence of the correspondence between the data and the forms is the way values are compared, i.e., which values are making up the variables. Again Figure I-1 is used to illustrate. In a statistical routine, the OLABR value of 4.5000 cannot be compared to the OLABR value of 2.5000 beneath it, as would usually be the case in a file ordered for statistical analysis. Instead, the 4.5000 must be compared to another value, not shown in Figure I-1, which has an EC of 11, but a different ID number. OLABR, therefore, is not one variable but eleven variables (the number of employee classifications) collected together in one field. Without some new way of defining the data, the statistical routine would compare the number of Revenue Vehicle Operators to the number of Vehicle Servicing Personnel in the same system, when the need is to compare the number of one system's Revenue Vehicle Operators to the number of Revenue Vehicle Operators employed by another system.

Informing the computer of this relationship between the values in the OLABR field requires devising a new logical organization or codebook for the data to replace that found in the TSC documentation. Figure I-2 shows most of the TSC codebook for Form 404 in its original and revised forms. For statistical purposes, OLABR in Codebook I is too general a category to qualify as a variable. Instead the eleven variables embedded in the OLABR field require the new definition given them in Codebook II. Additional comparison of the two codebooks uncovers another important difference. Codebook I "reads" or formats only one line of Form 404 at a time, and in that sense it has no inbuilt way of defining a statistical case. No higher level of organization clustering records together to form a case exists. Codebook II, on the other hand, reads the whole form, defines the individual lines on the form as variables, and the whole form as a statistical case. Eleven separate records under the old scheme are clustered together as a case in the new one.

The kind of organization found in the TSC tape files is common, economical, and often used as input to management information systems using customized software. File organization of this kind is referred to as hierarchical ordering by computer scientists.

By way of summary, then, two major problems motivated the data reorganization: (1) the absence of stand-in values for missing records; and (2) hierarchical ordering of data. In the following section the strategy used for solving these problems is described.

FIGURE 1-2. A COMPARISON OF ORIGINAL AND REVISED CODEBOOKS USED IN FORMATTING TAPE FILES

CODEBOOK I. UNREVISED TSC DOCUMENTATION

COLUMN	NAME	TYPE	DESCRIPTION
1 - 4	TRSID	INTEGER	TRANSIT SYSTEM ID
5 - 12	FY	DATE	FISCAL YEAR
13 - 13	MODE	INTEGER	MODE CODE
14 - 15	EMCOD	INTEGER	EMPLOYEE CLASS CODE
16 - 21	OLABR	REAL	OPERATING LABOR

CODEBOOK II. REVISED TSC DOCUMENTATION

CARD	COLUMN	NAME	DESCRIPTION
1	1 - 4	TRSID1	TRANSIT SYSEM ID FOR CARD 1
	5 - 12	FY	FISCAL YEAR
	13 - 13	MODE	MODE CODE
	14 - 15	EMCOD	EMPLOYEE CLASS CODE
	16 - 21	TNSEXOL	TRANS. EXEC., PROF., AND SUPP. OP LABR.
2	1 - 4	TRSID2	TRANSIT SYSTEM ID FOR CARD 2
	5 - 15	OMITTED	
	16 - 21	TNSSPOL	TRANS SUPP PERSONNEL OP LABOR
3	1 - 4	TRSID3	TRANSIT SYSTEM ID FOR CARD 3
	5 - 15	OMITTED	
	16 - 21	RVEHOPOL	REVENUE VEHICLE OPERATORS OP LABOR
⋮	⋮		⋮
11	1 - 4	TRSID11	TRANSIT SYSTEM ID CARD 11
	5 - 15	OMITTED	
	16 - 21	TOTEMPOL	TOTAL TRANS SYS EMPLOYEES OP LABOR

Implementing Reorganization

The main goals of reorganization were to supply stand-in values for missing records and to reformat instances of hierarchical ordering, i.e., where several variables had been grouped together in one field. A hypothetical example of the transformations resulting from reorganizing is shown in Figure 1-3.

Data File I illustrates how the problems discussed above appear in the data. In Data File I under the field SYSTEM ID, there is no information present for system number 1003, and systems 1002 and 1004 appear to have only half the information they need. This example illustrates that in the actual data both whole and partial cases are missing.

The second problem, hierarchical ordering, can also be seen in Data File I. The field WAGES contains six different variables, but the values in the fields MODE and EMPLOYEE CATEGORY must be used to find these variables. For example, the first WAGES value, 500, has a MODE value of 1 and an EMPLOYEE CATEGORY value of 0. These values indicate that the first 500 of WAGES is for motor bus drivers' wages. Hence, the only other value it can be compared to is WAGES of 650, six lines down in case 1002, which also has a MODE of 1 and EMPLOYEE CATEGORY of 0. There are six WAGES variables possible because in addition to the MODE and EMPLOYEE CATEGORY combination of 1 and 0 there are also the combinations of 1 and 1 or 1 and 2, etc. Because there are two values of MODE and three values of EMPLOYEE CATEGORY, it takes two times three, or six, combinations to exhaust all pairs possible and identify all six variables. Because the values in MODE and EMPLOYEE CATEGORY are required to distinguish among the six variables clustered in the WAGES field, they are referred to by the functional term "hierarchical ordering variables."

Data File II in Figure 1-3 illustrates how reorganization transforms the data. In this file the six WAGES variables each have their own separate fields. The information in MODE and EMPLOYEE CATEGORY from File I has been incorporated into the new logical organization of File II. Therefore, they disappear from File II. Data File II also has full sets of information (complete cases) for all transit system ID numbers represented, although missing value codes of 999 had to be inserted to make this possible. For example, even though system 1002 has no trolley buses, stand-in values of 999 were inserted in the three trolley bus variables in this case.

FIGURE 1-3. HYPOTHETICAL DATA FILE BEFORE AND AFTER REORGANIZATION

DATA FILE I. HIERARCHICAL ORGANIZATION

SYSTEM ID	MODE	EMPLOYEE CATEGORY	WAGES
1001	1	0	500
1001	1	1	600
1001	1	2	600
1001	2	0	400
1001	2	1	700
1001	2	2	700
1002	1	0	650
1002	1	1	600
1002	1	2	700
1004	2	0	700
1004	2	1	000
1004	2	2	000

MODE
1 = MOTOR BUS
2 = TROLLEY BUS

EMPLOYEE CATEGORY
0 = DRIVER
1 = MAINTENANCE
2 = ADMINISTRATION

DATA FILE II. STATISTICAL ORGANIZATION

SYSTEM ID	MTRBUS DRIVER WAGES	MTRBUS MAINT WAGES	MTRBUS ADMIN WAGES	TRBUS DRIVER WAGES	TRBUS MAINT WAGES	TRBUS ADMIN WAGES
1001	500	600	600	400	700	700
1002	650	600	700	999	999	999
1003	999	999	999	999	999	999
1004	999	999	999	700	000	000

999 = MISSING VALUE CODE

In general, the basic reorganization steps can be reduced to four:

1. Data were read as single records and unwanted information was eliminated.
2. The positions in the retained data needing stand-in values were located.
3. The stand-in values were inserted.
4. The data were formatted with a new logical organization (codebook) which considered all the records belonging to a single transit system as a statistical case.

Although useful at a general level, it is misleading to represent the reorganization process as four steps. Discussion of the programming procedures required to implement the reorganization is more suggestive of the actual scope of the task.

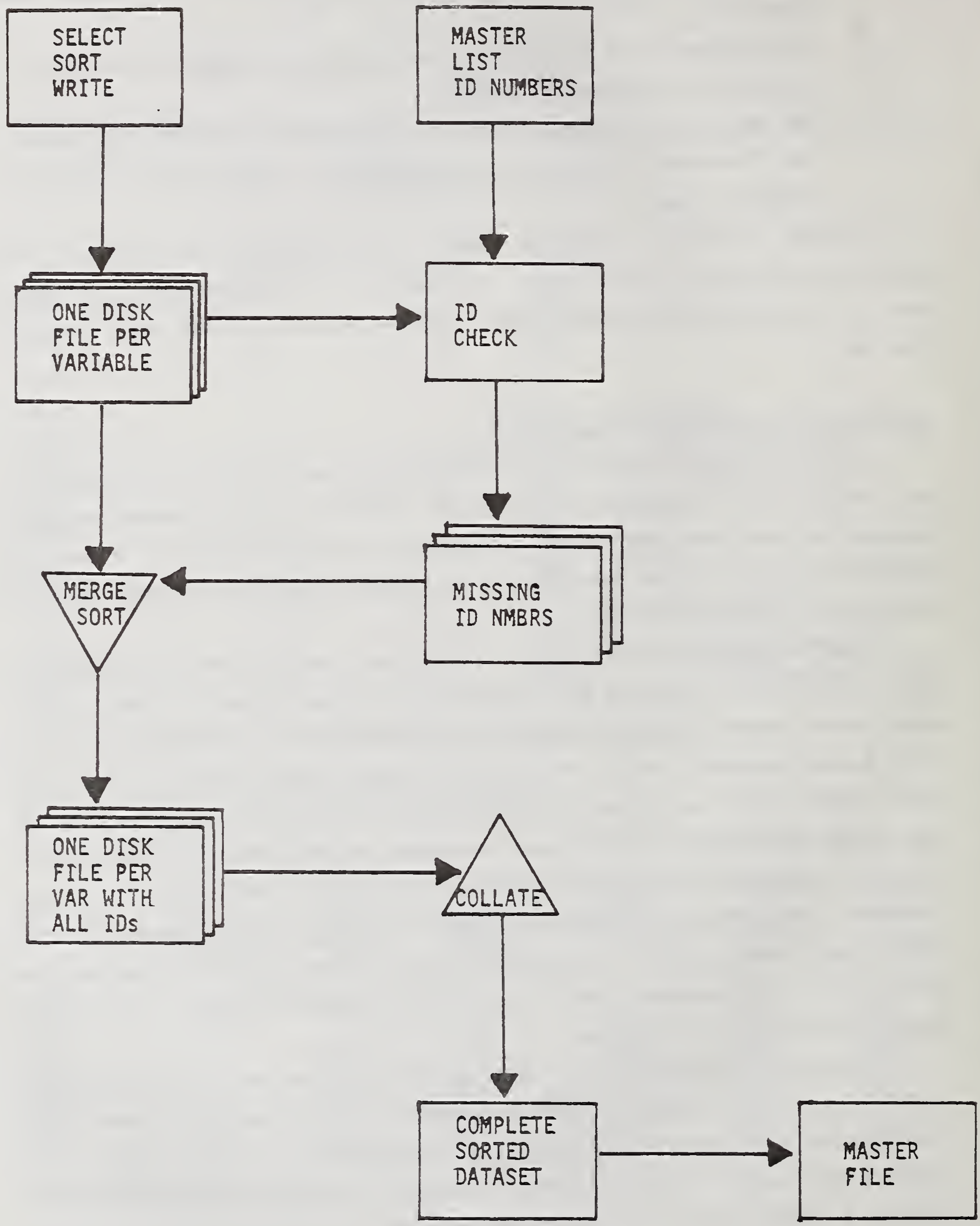
<u>Programming Required for Reorganization</u>

The basic strategy adopted was to process one data file from the tape at a time, selecting the variables that would be required for the projected analysis, reorganizing them, and adding them cumulatively to a master data file. Figure 1-4 gives a summary of the programming procedures required to reorganize the information in a single data file.

Three reorganizing requirements were met with the first set of programming steps. Since only a subset of the information available in each file was actually used, it was economically advantageous to eliminate all but necessary information from future processing steps. Hence the first step was to selectively read only those records which were to be retained. Next, since succeeding steps depended on the transit systems' data being in uniform order, the retained data were sorted in ascending numerical order by transit system ID number. Finally, since it was likely that each variable being processed was missing a different set of needed stand-in values, each variable and its accompanying set of transit system ID numbers was written out to a separate disk file. At this juncture, Step 1 of the basic reorganizaton process is complete.

The second step involved the identification of the systems who required the insertion of stand-in values or "dummy records" for the variables of interest. The identification step was accomplished by comparing the ID numbers present for a variable to a master list of ID numbers. The result of this step was another disk file containing the ID numbers missing for that variable.

FIGURE 1-4. SUMMARY OF PROGRAMMING PROCEDURES



The insertion of the stand-in values, the next move in reorganization, required two steps, referred to collectively as a Merge/Sort routine. First, two data files, one containing ID numbers and variable values and another containing the now identified missing IDs, were merged together electronically. Then, because in the merging the required ascending numerical order is no longer preserved, the merged files were re-sorted. When these two steps were complete, a data file with a complete set of ID numbers resulted, although IDs which were identified by comparison to the master list were dummy records having blanks in the variable value position. (In a later routine a missing value designation, -9, was inserted in the blank.)

When all the variables of interest originating in the same tape file have been processed to this point, they are reunited in a single data file by a collating routine. This step results in a complete, sorted data set containing all the variables. At this juncture, three of the four basic reorganization steps have been accomplished.

In a final step, the complete, sorted dataset is added to the master file. At this stage each transit system has a uniform number of records. This manufactured uniformity is what allowed the imposition of the new logical organization, in actuality a new formatting scheme, which identified for the computer the several variables embedded in a single field. When all the variables required for analysis had been processed in this way, the reorganization step was complete.

Each data file handled presented special characteristics which required special treatment. Not all files required as many steps as described while others required many more. Four major variations in the data reorganization procedure emerged in practice and are discussed in detail in the Technical Report.[22]

The DECsystem-10 conversion of SPSS was used for nearly all programming steps. Two FORTRAN programs, one for identifying missing ID numbers and another for collating variable files, were also required.

One other problem of note arose in the handling of the Section 15 variables. This problem concerned the inability of single precision software (such as the DECsystem-10 conversion of SPSS) to handle field widths exceeding eight columns. The eleven-column wide variables found in the expense files, for example, set up a variety of problems and barriers that had to be circumvented. SPSS cannot write

---

[22]Gordon J. Fielding, Mary E. Brenner, and Olivia de la Rocha, Using Section 15 data for transit performance analysis, Interim report No. UMTA-CA-11-0026-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January 1983.), pp. 11-20.

out a single numeric field wider than eight columns and warns of distortions in accuracy when reading variables exceeding that limit, although tests have shown those distortions to be minor. Binary and alphanumeric formatting are temporary remedies to the problem, but never solve it. The use of double-precision software would simplify the handling of Section 15 data, and we recommend its use where possible.

In all, twenty-three separate data files and 147 variables were prepared by the reorganization sequence discussed above. Appendix A summarizes the files accessed and variables retrieved.

DATA PREPARATION

Once the data were reorganized, additional data preparation was required before analysis could commence. There were three steps to preparing the data: calculating basic variables, identifying and flagging missing information and validating existing data.

Calculating Basic Variables

The Section 15 database contained a wealth of information which was too detailed for our purpose. It was necessary to aggregate many small pieces of information into more comprehensive variables which contained only information about the motorbus mode and which were applicable to an entire year's operation. The building blocks for this process are listed in Appendix A. The final sets of variables are listed in Table 2-1 and in the text of Chapter 3. The remainder of this section outlines the major steps used to calculate the variables used in the analyses.

The information about transit employees was summarized into broader categories. Ten employee categories are reported in Section 15: three in vehicle operations (i.e., supervisors, revenue vehicle operators and support personnel), five in maintenance and two in general administration. These ten categories are further subdivided into capital labor and operating labor. Analysis for this project required only the number of vehicle operators, the number of maintenance employees and the number of administrative employees. The first step in creating these variables was to add together operating and capital employees since we were not interested in this distinction. At this point, the number of revenue vehicle operators was ready for use. The number of maintenance employees was calculated by adding together the five categories of maintenance employees. The number of administrators was calculated by adding together the supervisory personnel in vehicle operations and maintenance to the two categories of administrative personnel.

Other variables which underwent a similar aggregation process were the total number of accidents (combining all categories of collision and non-collision accidents), total amounts of subsidies (combining local, state and federal) and the miles of line used on bus routes (combining mixed right-of-way and one-way directional).

The number of peak and mid-day vehicle variables were created from several sources. Although this information can be directly reported on Form 406, transit systems with a peak to base ratio of one were not required to report their numbers of vehicles for different time periods. For systems missing this information, the number of vehicles was calculated by substituting information from the number of vehicle opertors scheduled for weekdays or the number of vehicles operating on an average weekday. To further assure that this was done only for systems with a peak to base ratio of one, other sources of published information were cross-checked, including APTA reports,[23,24] other Section 15 reports[25] and state reports[26] to validate peak to base ratios.

The data on service supplied by a transit agency and service consumed by passengers underwent a special calculation to annualize them. While the Section 15 reporting system requires that all financial data be reported for a complete fiscal year, information on service variables such as unlinked passenger trips and revenue vehicle hours was collected by a sampling procedure and reported for an "average

---

[23]American Public Transit Association, Operating statistics report 1981: Transit system operating statistics for calendar/fiscal year 1980. (Washington, D.C.: American Public Transit Association, October 1981.)

[24]American Public Transit Association, Operating statistics report 1980: Transit system operating statistics for calendar/fiscal year 1979. (Washington, D.C.: American Public Transit Association, October 1980.)

[25]U.S. Department of Transportation, Transportation Systems Center, National urban mass transportation statistics: 1981 Section 15 report. Report No. UMTA-MA-06-0107-83-1. (Springfield, Va.: National Technical Information Service, November 1982.)

[26]State of California. Office of the Controller. Financial transactions concerning transit operators and non-transit claimants under Transportation Development Act: Annual report for fiscal year 1980-1981. (Sacramento, Calif.: State of California, Office of the Controller, 1982.)

weekday," an "average Saturday" and "average Sunday." This information was com-bined using a formula which annualized it so that it was comparable to the financial data. The formula allowed for 253 weekdays, 53 Saturdays, 52 Sundays and 7 holidays (also calculated as Sundays) with each of these numbers multiplied by the given values for average weekdays, Saturdays and Sundays. This is the same formula used by TSC in the Annual Report. However, for these data the formula was combined with a validation process so a few values differ from those in the Annual Report.

A series of calculations were also needed to disaggregate data so that it applied only to the motorbus mode. Revenue and subsidy information are reported in Section 15 for the entire transit system, not by mode. In addition, multimodal systems have the option of reporting expenses as joint expenses between modes, and a few systems report most of their expenses in this way. A series of weighting formulas were designed which allowed assignment of revenues or joint expenses to specific modes. For example, a proportion of passenger revenue was assigned to the motor bus mode by multiplying the system's total passenger revenues by the ratio of motor bus passengers to total passengers. Although the resulting values are only estimates, they are an improvement on the distortions caused by using overly-large figures or dropping the multi-modal system (32% of the systems reporting in 1980) from the analysis. Appendix B summarizes which variables were weighted and the procedure used. All later analyses were done twice, with weighted and unweighted variables to assure that the results were not an artifact of the weighting procedure.

Detection of Missing Data

The second phase of preparing data for analysis was detection of cases having missing data and which, therefore, needed to be eliminated from further analysis. A database prepared for statistical analysis will usually have a special symbol such as -9 which indicates that information is missing. However, the Section 15 data tape had no such special symbol. Cases with missing data had either a zero or blank. Since there can be "real" zeroes (e.g., a system may have no local subsidies), it was necessary to differentiate "real" zeroes from missing data zeroes. Thus a missing data symbol had to be inserted during the process of calculating the variables. It was possible to detect the missing data problems by considering the logical properties of specific variables, by comparing a variable to other information in the data base and by comparing the Section 15 data to other sources of information.

For some variables, detecting missing data was straightforward and quite logical. For instance, a transit system which had no operating expenses was assumed to have a missing data problem. Other such variables were revenue vehicle drivers, operating subsidies, total vehicle hours, total wages, etc.

But most variables required more judgment on the part of the project staff. It is possible for a transit system to have zero accidents for a given fiscal year, but this is unlikely for large systems. Other transit systems of similar size to the one reporting zero accidents were examined to see if zero was a possible number. A cross-year comparison of reported accidents supplied further evidence on which to base a decision. It was decided for this project that any systems with more than ten revenue vehicles could not have zero accidents, and a missing data symbol was inserted for these systems. Smaller systems were then judged individually--taking into account the number of peak vehicles required (a better measurement of size than revenue vehicles), their safety record in other years as reported in Section 15 Reports or APTA reports and the performance of like-sized systems.

Some judgments about missing data involved making decisions about whether a concept was adequately measured by a combination of several different variables. For instance, vehicle maintenance could be supplied by employees on the transit agency payroll or by contract with other organizations. Thus if a system reported zero maintenance employees, the system was expected to have zero maintenance wages reported but a substantial expenditure for services indicated under either the maintenance function or general administration. In the absence of wages and service expenses, a missing data symbol was used to indicate that maintenance expenses were missing.

For some other variables, the decision was more complex because a zero value could be a real value or it could be an indication of a problem. The example of total vehicle miles will make this clear. Total vehicle miles, as noted above, is constructed from three variables--average weekday miles, Saturday miles and Sunday miles. If weekday miles were zero, it was assumed that information was missing. However, many systems do not offer weekend service, so a zero for Saturday or Sunday miles might be real or might be an indication of a problem. Since this information was based upon a time consuming sampling procedure, there was a definite possibility that a transit system failed to collect this information, and thus had a missing data problem. The Section 15 data tape included information about the service schedule of each system. Therefore, it was possible to determine if a system offered Sunday service or not, and thus whether it had a missing data problem or not.

The problem of missing data received detailed attention because it is an inevitable problem in a data base as complex as the one mandated by Section 15. Over 300 different systems must learn to interpret and fill out numerous forms--ranging from 17 pages for a small, single mode system to 90 pages for a large, multi-modal system. Since 1980 was only the second year in which this information was reported, some systems were still in the process of instituting accounting systems compatible with Section 15 requirements.

Data Validation

The final phase of data preparation consisted of cross-checking the data for validity. Errors could enter the database in many ways--misinterpretation by a transit system of what number should be reported, miscalculation of totals, and key punching errors as data are prepared for the computer. Four major methods were used to validate the data: recomputation of totals, comparisons of redundant information, comparisons of related information and comparison to feasible value ranges. An example of each of these methods with specific variables will be given.

The total number of employees reported for each system was compared to the sum of the separate categories. In about ten cases, the totals differed by more than could be accounted for by rounding errors. In most cases the differences were apparently caused by keypunching errors (e.g., reversal of digits) or simple miscalculations. For these cases, reported totals were replaced by the recalculated totals and cross-checks made with the Annual Reports. Revenue, subsidy and expense totals were also checked.

Much of the financial data was reported in several different places. For instance, the Revenue Summary Schedule (Form 201) summarized the information on the Revenue Subsidiary Schedule, (Form 203). Total operating expenses were also reported in two different places on the magnetic tape. A simple comparison of these numbers revealed a few differences and the correct number was identified by the other validation methods.

Different variables in the database are sometimes different measures of the same thing. For instance, employee counts and employee wages are two different measures of labor utilization. If a transit system has a large number of vehicle operators, it must have a proportionately large amount of vehicle operator wages.

However caution must be used in some of these comparisons. Maintenance employee counts and maintenance wages were sub-divided into distinct, non-comparable sub-groups, so only the totals were comparable.

The final method of identifying mistakes was to look for values that lay outside an expected range for that specific variable. This method worked best for measures that were combinations of two variables such as miles per hour or cost per passenger. Miles per hour (speed) has an expected range of about 5 miles per hour (dense urban areas) to 30 miles per hour (commuter service). Any system that fell outside this range or was in the wrong part of the range for the kind of service it offered probably had a mistake in either its measure of miles or hours.

A variable such as cost per passenger was a little more difficult to work with since inflation and difference in fiscal years caused the feasible range to change over time and the boundaries of a feasible range were indefinite. In this instance all cases were examined which lay more than three standard deviations from the mean as well as the largest and smallest cases. While some of these outliers had apparent, real causes, such as extremely long trip lengths, others were so different from the norm that they were obviously wrong. In these cases we looked for the correct values in other parts of the database, or in other sources. If a correction was impossible, incorrect values were designated as missing.

DATA EVALUATION

Once the data were in a form ready for statistical analysis, it was necessary to select the best variables for the ensuing analyses. Once the variables were chosen, it was then necessary to evaluate the distributional characteristics of each variable in order to select the appropriate statistical technique. Finally, the sample of transit systems with sufficient data to enter into the analyses had to be carefully described in terms of how well they represented the entire set of transit systems that were included in the Section 15 reporting system for FY 1980.

Evaluation of Variables

Some kinds of variables were more likely to have missing data than others. In FY 1980 the most complete data were available for economic variables such as operating expenses and passenger revenue (Table 1-2).

TABLE 1-2. THE DISTRIBUTION OF MISSING DATA IN SELECTED TRANSIT VARIABLES

Variable	% missing values out of 304
Passenger Revenue	0.7%
Total Operating Expense	2.0%
Total Employees	2.6%
Total Vehicle Miles	8.2%
Unlinked Passenger Trips	18.1%
Passenger Miles	24.0%

The most incomplete information was for passenger measures such as unlinked passenger trips and passenger miles. Since service utilization is a major facet of transit performance, it was necessary to keep some measure of this concept although it would result in more systems being excluded from the analysis. Unlinked passenger trips were chosen because this variable was the most complete measure of utilization and seemed less prone to measurement error.

Other variables seemed relatively complete but the validity checks described in the earlier section revealed that there were problems with the values reported. In many instances there was not enough information to cross-check the values or to correct them. The following variables had severe enough problems that they were eliminated from the final analyses:

1. Active vehicles: The number of vehicles actually used to provide service during fiscal year was obtained from the Revenue Vehicle Inventory. However, the information in the Revenue Vehicle Inventory was incomplete and had numerous mistakes. Thus information was not available for each vehicle in the fleet of some systems. Other systems had more active vehicles than they actually owned. Numerous mistakes in the designation of mode resulted in vans being considered motor buses and vice versa. Since about a third of the transit systems had major errors on this item, it was not used in the analyses.

2. Fuel: There were four different fuels reported in use with motor buses: diesel fuel, gasoline, bunker fuel and liquid natural gas. A number of transit systems had combinations of fuels. A major problem was that bunker fuel is not normally considered a motor bus fuel so several systems had a coding error on fuel. It was also hard to compare the efficiency of different fuels. In addition, for those

systems using several kinds of fuel, there was no way to allocate miles to one fuel or another. Although mileage information should have been available on the Revenue Vehicle Inventory, problems with that data precluded its use in this instance. Thus measures of fuel efficiency were eliminated from the final analysis.

3. Subsidies: Transit systems did not use consistent definitions for designating whether particular subsidies were from state, local or federal sources. For instance, in California half of the transit systems called the subsidies from a particular source local funding, while others called it state funding. Since subsidy programs vary from state to state, there was no comparability in these definitions. For the analysis, all sources of subsidies have been combined and only measures of total subsidies used.

4. Miles of line (route miles): The definitions used by the Section 15 reporting systems for measuring miles of line were confusing and interpreted differently by different systems. The excessive variance in reported miles made this variable unreliable for statistical analysis and it was eliminated.

5. Roadcalls: As with miles of line, the definitions for roadcalls were not consistent across transit systems and this variable was not used.

6. Maintenance expenses: Since transit systems can do maintenance in-house or through purchased services, variables relating to maintenance must be used with caution. Transit Systems with no reported maintenance expenses were eliminated. Those remaining in the analysis have a variety of maintenance arrangements and any one measure of maintenance efficiency may not be sufficient to represent the situation for all systems.

Evaluation of the Distribution of Variables

Most of the basic variables were not normally distributed and thus certain kinds of statistical analyses had to be used with caution. There were many more small transit systems (less than 25 vehicles) than large, so the distributions were very peaked at the small end of the scale for variables which reflect the size of a transit system, such as number of peak vehicles, operating expenses and subsidies. There were also a few very large transit systems, such as New York, Chicago and Los Angeles, that were so much larger than the others that they were outliers causing the distribution of variables to be very skewed. These systems were too important in terms of the amount of transit they provide to eliminate them from the analysis. Thus statistical methods had to be chosen which minimized the influence of these outliers or the variables had to be transformed so that they met more of the assumptions of the statistical techniques used. Both of these approaches were used and are described in later chapters where relevant. More information on the distributional properties of specific variables is available in another paper.[27]

Evaluation of the Sample

The cases with insufficient data to be included in the analyses were not randomly distributed throughout the sample of transit systems included in the Section 15 data base. The missing data situation was particularly acute for small systems--those with fewer than 25 revenue vehicles. Thirty per cent of these systems were missing information on passenger trips and six per cent on expenses. Although the analyses still included substantial data on smaller systems, since over one third of the systems reporting data fall into this size category, generalizations to all small systems must be made cautiously.

Since the results of this project are being compared to an earlier project,[28] it was necessary to examine the issue of how stable the sample of transit systems was between FY 1979 and FY 1980. About 18% of the transit systems differed between the two years--with some cases dropping out and new ones entering into the sample. This changeover was most common among the smallest systems. So cross-year comparisons must be made with caution, particularly for the smallest systems.

CONCLUSION

Data used in the analysis of transit performance were based upon the Section 15 data but are not identical with those reported on the TSC tape or in the Annual Report for FY 1980. Obvious errors have been corrected and missing entries have been designated with the -9 symbol. Some agencies were eliminated because either there were too many missing items or obviously inconsistent values could not be verified. The data base was reorganized into a format suitable for statistical analysis in preparation for the factor analysis reported in the next chapter.

---

[27]Gordon J. Fielding, Mary E. Brenner, and Olivia de la Rocha, Using Section 15 data: Adapting and evaluating the magnetic tape version for statistical analysis. Working paper no. 83-6. (Irvine, Calif.: University of California, Institute of Transportation Studies, December 1983.)

[28]Shirley C. Anderson and Gordon J. Fielding, Comparative Analysis of transit performance. Final report No. UMTA-CA-11-0020-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January 1982.) (NTIS No. PB 82-196478.)

CHAPTER 2

IDENTIFYING KEY PERFORMANCE INDICATORS

INTRODUCTION

Section 15 of the Urban Mass Transportation Act of 1964, as amended, has provided for the collection of a unique set of comparable transit statistics by requiring all urban transit applicants for operating assistance to provide a uniform set of information about their transit systems. The first year of Section 15 reported statistics [FY 1978-79] was used by Anderson and Fielding[1] to test the performance concept model developed by Fielding, and Glauthier and Lave[2]. A set of nine performance indicators was selected, representing the three dimensions of transit performance. However, serious questions were raised about the validity and completeness of the first year's data. Only 98 agencies out of 311 could be used in the final factor analysis. The rest were dropped because of missing and imprecisely reported data. Other questions were raised by reviewers about the validity of the indicators selected based upon a single factor analysis solution. Although the results had not previously been satisfying, the method of using factor analysis to identify clusters of variables and performance indicators held promise. If the data set could be improved, using the data techniques described in Chapter 1, then more rigorous factor analytic solutions could be applied on different versions of the data to test the validity of the performance model.

Following the data cleaning and verifying routines outlined in Chapter 1, data from the second year of reported statistics [FY 1980] were analyzed. Chapter 2 addresses two issues. It replicates the methods and compares results to the first year [FY 1979] statistical analysis. Secondly, a thesis is advanced that there exists a highly consistent set of performance concepts relevant to fixed route transit operators and a small, unique subset of performance indicators that are useful for performance evaluation by individual transit managers for systems of all sizes.

---

[1] Shirley C. Anderson and Gordon J. Fielding, Comparative analysis of transit performance. Final report No. UMTA-CA-11-0020-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January, 1982.) (NTIS No. PB82-196478.)

[2] Gordon J. Fielding, Roy E. Glauthier and Charles A. Lave, Performance indicators for transit management. Transportation, 1978, 7(4), 365-379.

Results from the analyses undertaken here are compared to previous research and suggestions are offered for the use of the seven key performance indicators identified as being the most useful for cross-sectional analysis.

Emphasis is given to describing the sequence of steps used to explore the thesis that a highly consistent set of performance concepts exists and that they can be represented by a small, unique set of performance indicators. Results from previous research have been controversial.[3] Therefore, we have endeavored to explain how:

- performance indicators were selected and calculated in alternative ways to minimize bias
- different methods of factor analysis were used to explore the structure of performance concepts
- tests were used to verify the structure of performance concepts.
- seven performance indicators were identified as being the most useful for cross-system analysis.

PERFORMANCE EVALUATION USING SECTION 15 DATA

Section 15 data has been crucial to the analysis; it is only through the use of a nationwide set of comparable data that identification of globally-oriented performance indicators can be assessed. A wide variety of Section 15 statistics was evaluated as performance indicators. Three categories of statistics--service inputs, service outputs and service consumption--provided the framework to organize the much larger set of data.

Figure 2-1 portrays the organizing framework developed in the Fielding, et al. performance concept model. Cost-efficiency indicators measure service inputs (labor, capital, fuel) to the amount of service produced (service outputs: vehicle hours, vehicle miles, capacity miles, service reliability). Cost-effectiveness indicators measure the level of service consumption (passengers, passenger miles, operating revenue) against service inputs. Finally, service-effectiveness indicators measure the extent to which service outputs are consumed.

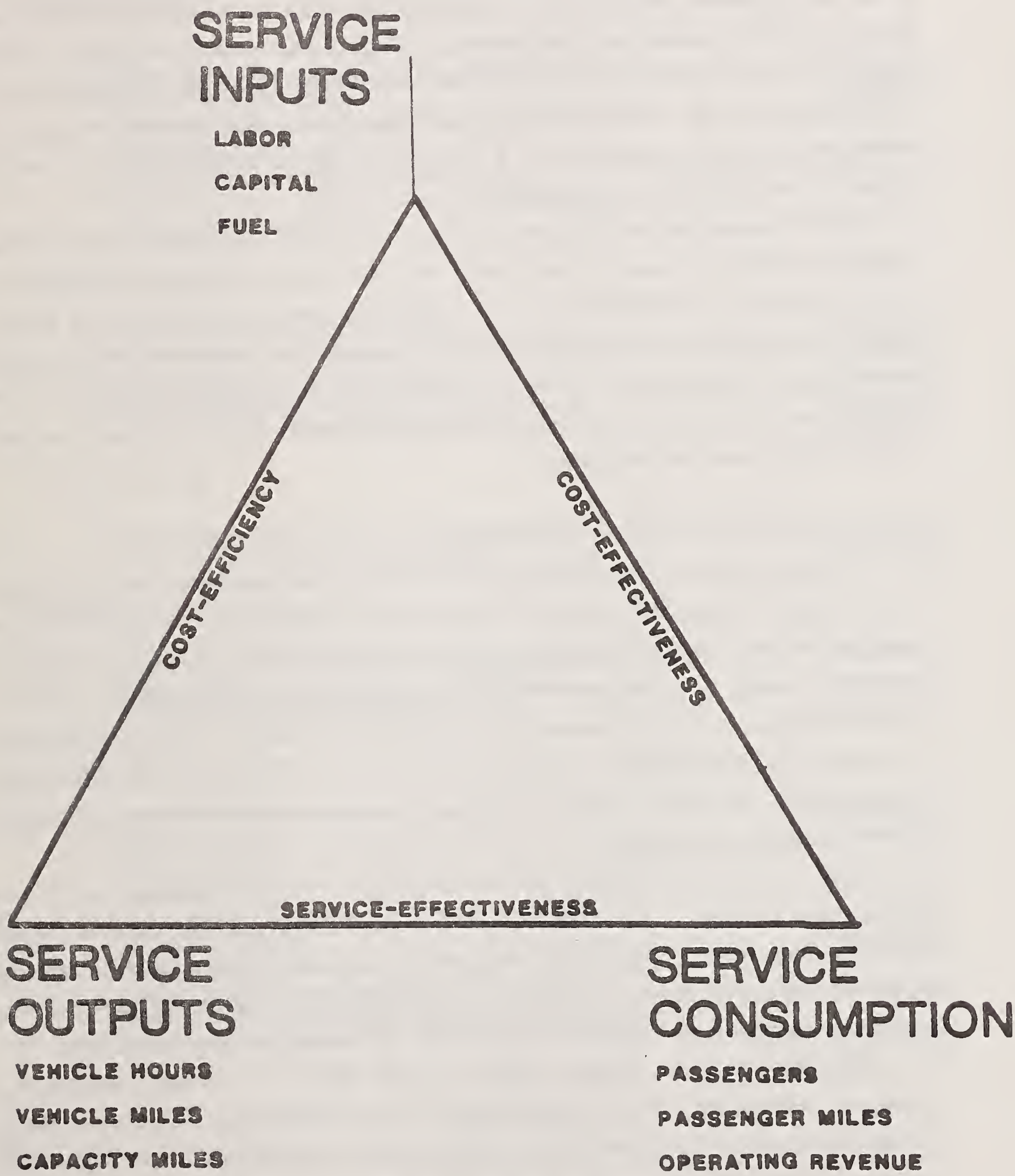
The overriding goal of this research was to identify those key performance statistics: 1) that provide transit analysts with the most salient performance

---

[3]T. A. Patton, Transit performance indicators. Transportation Systems Center Staff Study #SS-67-0.3-01. (Washington: U.S. Department of Transportation, 1983.)

FIGURE 2-1. FRAMEWORK FOR TRANSIT

PERFORMANCE CONCEPTS



information and 2) that target information which is equally valid for each transit agency and thus for cross-system analysis.

One result of the analyses that follow was the identification of a small, unique set of key performance indicators that met the overriding goal of this research. Seven performance variables from a much larger data set were identified. These can be used to assess the performance of any fixed route, motor bus, transit system. A minimum of three of the seven variables will provide key information on cost efficiency, cost effectiveness and service effectiveness. Further, all seven of these performance indicators and a parallel set of "alternates" can be used for cross-system comparisons with peers.

The following sections describe how Section 15 data was used to identify these seven performance indicators, and how they were rigorously tested to ensure their validity for use. The main focus of this research has been to provide transit analysts with a set of easily accessible statistics with which to do individual and peer group comparisons of performance. The second goal was to evaluate the validity of the earlier analysis conducted on the FY 1979 data. The body of this chapter explains how both goals were accomplished.

SELECTING PERFORMANCE INDICATORS

A wide variety of performance indicator ratios was available from the Section 15 data base. In selecting the set of performance indicators to be used for further analysis, the data included variables that would relate to the conceptual model i.e., those that would best represent the three categories of performance concepts--cost-efficiency, cost-effectiveness and service-effectiveness. Particular attention was given to the availability and reliability of the data from which the ratios would be calculated. As noted, some of the Section 15 data variables were more complete or more reliable than others.

Table 2-1 lists the initial set of forty-eight variables selected for further multivariate analysis. The variables are organized under the performance concept to which they relate. This set of forty-eight variables in most cases (other than passenger data) represent the most complete, generally reliable and non-redundant performance indicators available in the current (FY 1980) Section 15 data set.

Variables based on revenue capacity miles were not included because of a detected inconsistency in the measurement of that variable across systems. Ratios based on population data were not included because available population information reflected total urban population rather than service area population. Otherwise,

TABLE 2-1. PERFORMANCE INDICATORS BY CONCEPT

COST EFFICIENCY MEASURES	
Labor Efficiency	
Vehicle Hours per Employee	TVH/EMP
Revenue Vehicle Hours per Operating Employee Hour	RVH/OEMP
Vehicle Miles per Employee	TVM/EMP
Peak Vehicles per Executive, Professional and Supervisory Employees	PVEH/ADM
Peak Vehicles per Operating Personnel	PVEH/OP
Peak Vehicles per Maintenance, Support and Servicing Personnel	PVEH/MNT
Vehicle Efficiency	
Vehicle Hours per Active Vehicle	TVH/AVEH
Vehicle Hours per Peak Vehicle Requirement	TVH/PVEH
Vehicle Miles per Active Vehicle	TVM/AVEH
Vehicle Miles per Peak Vehicle Requirement	TVM/PVEH
Revenue Vehicle Miles per Vehicle Miles	RVM/TVM
Fuel Efficiency	
Revenue Vehicle Miles per Gallon Diesel	RVM/FUEL
Vehicle Miles (Bus) per Gallon Diesel	TVM/FUEL
Maintenance Efficiency	
Total Vehicle Miles per Maintenance Expense	TVM/MEXP
Vehicle Miles per Maintenance Employee	TVM/MNT
1,000,000 Vehicle Miles per Roadcall	TVM/RCAL
Output per Dollar Cost	
Revenue Vehicle Hours per Operating Expense	RVH/OEXP
Vehicle Miles per Operating Expense	TVM/OEXP
Revenue Vehicle Hours per Total Labor and Fringe Expenses	RVH/TWG
Revenue Vehicle Hours per Operations Labor and Fringe Expenses	RVH/OWAG
Revenue Vehicle Hours per Vehicle Maintenance Labor and Fringe Expenses	RVH/VMWG
Revenue Vehicle Hours per Administrative Labor and Fringe Expenses	RVH/ADWG

TABLE 2-1. (continued)

SERVICE EFFECTIVENESS MEASURES	
Utilization of Service	
Passenger Trips per Revenue Vehicle Hours	TPAS/RVH
Passenger Trips per Revenue Vehicle Mile	TPAS/RVH
Passenger Trips per Peak Vehicle TPAS/PVH	
Passenger Miles per Passenger	PASM/TPS
Operating Safety	
1,000,000 Vehicle Miles per Accident	TVM/ACC
Revenue Vehicle Hours per Accident	RVH/ACC
Revenue Generation	
Passenger Revenue per Peak Vehicle	REV/PVEH
Passenger Revenue per Revenue Vehicle Hour	REV/RVH
Operating Revenue per Revenue Vehicle Hour	OREV/RVH
Passenger Revenue per Passenger	REV/TPAS
Public Assistance	
Revenue Vehicle Hours per Local Capital and Operating Assistance	RVH/LSUB
Revenue Vehicle Hours per State Capital and Operating Assistance	RVH/SSUB
Revenue Vehicle Hours per Total Operating Assistance	RVH/OSUB
Revenue Vehicle Hours per Total Capital and Operating Assistance	RVH/TSUB
Passengers per Local Operating Assistance	TPAS/LOA
Passengers per Total Operating and Capital Assistance	TPAS/TSUB
Passenger Revenue per Total Operating and Capital Assistance	REV/TSUB
Passenger Revenue per Total Operating Assistance	REV/OSUB
Passengers per Total Operating Assistance	PAS/OSUB
COST EFFECTIVENESS MEASURES	
Service Consumption per Expense	
Passengers per Operating Expense	PAS/OEXP
Passenger Miles per Operating Expense	PASM/OEX
Passengers per Total Labor and Fringe Benefits	PAS/TWAG
Passengers per Gallon Diesel Fuel	PAS/FUEL
Passenger Miles per Total Expense	PASM/TEX
Revenue Generation per Expense	
Ratio Operating Revenue to Operating Expense	OREV/OEXP
Ratio Total Revenue to Total Expense	TREV/TEX

performance indicator ratios comparable to the 1979 data analyses were selected for use. This facilitated comparison with previous results and identification of shifts due to the better-collected, cleaner and more complete data.

Missing Data Effects

Missing values encountered at any point in the computation of basic and ratio variables and during the multivariate statistical procedures cause a "snowball" effect of missing information to occur. The assumption in the computation and analysis procedures is that every case has information for all of the variables. This problem and solutions used to address it were discussed in Chapter 1 under Data Reorganization. If any case is missing even one piece of information it is thrown out of the computations and subsequent analyses. The missing values problem has a cumulative effect as cases are dropped from the analysis. Thus, from a total of 304 transit systems running fixed route, motor bus service, only two-thirds of the cases--198 systems--had enough information available for use in the analyses. However, this is a vast improvement over the 98 systems which could be used from the FY 1979 data.

Distribution of the Data

One of the first tasks for exploring the data set was to search for extreme outliers and to remove them from the analysis. Extreme outliers could force the analysis to focus on the inflated variance due to the presence of an outlier, rather than the more true-to-data variance present across the range of the other cases.

The next task was to check the univariate descriptive statistics for each of the selected performance indicator ratios to evaluate the distribution of the case values across the variable range. Most commonly used bivariate and multivariate procedures assume a normal-like distribution of the case values in each variable.

As noted in Chapter 1, the large proportion of small systems and the presence of a few very large transit systems affected the distribution of data in most of the basic variables. Two descriptive statistics that provide information on how far a variable deviates from a normal-like distribution of values are skewness and kurtosis. For a normal distribution of data, both skewness and kurtosis equal zero; for each statistic the further from zero the value, the less normal-like is the data distribution. The less normal-like the distribution, the more questionable the statistical results.

Included as Appendix C is a listing of relevant descriptive statistics for each of the forty-eight performance indicators selected for further analysis. The skewness and kurtosis values for the list of forty-eight variables ranged from -5.212 to 16.098 and from 1.373 to 263.908 respectively, indicating that the distributions were far from normal. The proposed multivariate procedures to be used on the performance indicator data set were considered relatively "robust," i.e., valid even under deviations from normality. Robustness is of greatest concern when using inferential statistical techniques. However, even descriptive techniques, like the ones used here, could be affected by highly skewed data. As the goal of this research was to provide a highly reliable set of consistent analytical findings that could serve as a benchmark for cross-year comparisons, it was important to begin with a set of data that had a minimum of distributional problems.

To counter any possible bias in the analyses and to provide a comparable set of more normally distributed performance indicator variables, the base 10 logarithms of the forty-eight performance indicators were calculated. Logarithms preserve the essential data structure of the variables from which they arise while shifting the distribution of the data to a more normally shaped, i.e., less skewed, curve[4]. This provided two sets of comparable data--the forty-eight performance indicator ratios calculated from the Section 15 data and a set of forty-eight logarithm variables calculated from these.

In developing the strongest set of data on which to base analytical findings, a second question arose. As mentioned in Chapter 1, revenue data is reported as a total for the whole system; it is not broken down by mode when more than one mode exists. It had also been necessary to use total subsidy information. A third set of performance ratios was developed using basic variable data, subsidy information, and revenue statistics that were weighted to eliminate revenue from modes other than bus transit. Then, a full set of forty-eight base 10 logarithms was calculated on the weighted data, again, to provide a less skewed data distribution.

As a result of the cleaning, verifying and grooming, four somewhat different sets of performance indicator data were available: a) ratios from reported data, b) logs of reported data variables, c) ratios from the weighted reported data, d) logs of the weighted data variables. As noted, Appendix C includes descriptive statistics

---

[4]J. B. Kruskal, Transformations of data. International encyclopedia of the social sciences. David L. Sills, Ed., Vol. 15. (New York: Macmillan Co., 1968), pp. 182-192.

for each of the four sets of data. The purpose for developing these four sets was to ensure that when final results from multivariate analyses were reported, most contingencies for possible bias in the data had been addressed. Consistent results across the four data sets would provide evidence that a stable performance concept structure had been found in the data.

EXPLORATORY ANALYSES

Multivariate analyses were used to search for a highly consistent set of performance concepts relevant to fixed route transit and for a small, unique subset of conveniently useable performance indicators. Factor analysis is ideal for detecting the most salient features of a set of data and for determining those few key variables with which a whole range of information can be represented. The prime objective in this research was to search for the minimum amount of data necessary to convey the maximum amount of performance information. Parsimony and consistency were the key criteria; factor analysis was the most efficient means.

Factor Analysis Defined

The most distinctive characteristic of factor analysis is its ability to reduce a large set of data to a smaller set of "components" or "factors" which portray the underlying structure of relationships among a set of variables. Based upon the correlation patterns of a large number of variables, the objective of the factor analytic technique is to group together those variables which are highly correlated with each other. The analyst then interprets each factor according to the variables belonging to the group. The idea is to summarize many variables by using a few representative factors. Appendix D portrays the correlation matrix for the variables from the weighted reported data--the more correct of the two raw data sets.

There are two main types of factor analyses, principal components analysis and inferential or "classical" factor analysis. The former works from the assumption that the entire population of cases--not a sample--is being analyzed. Analytical solutions describe the data at hand and the relationships among the variables as represented in the input data. Inferential factor analysis, however, adjusts analytical solutions to make predictions about a larger, ideal population. Because the entire population of motor bus systems was represented in the data, and because no sampling technique had been used to select systems for analysis, principal components factor analysis was used.

The basic factor analysis model assumes that in any set of variables, there exist two main types of variation or variance: variance commonly shared by all the variables in the set and variance unique to each individual variable. Commonly shared variance contributes to the intercorrelations of variables. The patterns of intercorrelations are used to group variables into a smaller number of factors. The number of factors necessary to portray this underlying data structure depends on how much more commonly shared variance continues to be detected with the addition of each new factor. The order in which the factors emerge from the data is important. The first factor accounts for the largest portion of shared variance in the data. With each successive factor, less and less of the shared variance is accounted for. At the point where little more explained variance is detected, the procedure halts and the factor structure is considered complete.

Factor analysis not only provides information on the number of factors underlying the data, it also determines which variables grouped on a particular factor are most highly related or representative of the identified factor. The <u>factor loading</u> of each variable on the respective factors can be interpreted as the correlation of the variable with the factor; high factor loadings represent high correlations.

In performing any factor analysis, there are several problem areas that could exist in the data and obscure the underlying data structure[5]:

1) Two variables carry highly redundant information (collinearity). A correlation coefficient of .98 or larger between two variables would show that either variable could be used to present nearly the same information.
2) A variable loads across several factors equally well (poorly defined structure in the variable). When a variable portrays a pattern of factor loadings that are either almost equal, or are high across several factors, the variable does not contribute to defining the underlying structure of the data set.
3) One factor has all or most of the variables weighting heavily on it (poorly defined structure in the data set). Such a factor then becomes a complex "catch-all" category for data, and the underlying concepts of the data become obscured.

---

[5] A. L. Comrey, <u>A first course in factor analysis</u>. (New York: Academic Press, 1973), pp. 189-197.

The first exploratory factor analysis was begun with the most complete set of performance indicator ratios available in the Section 15 data i.e., the forty-eight performance indicators selected for analyses. It remained necessary to assess how well these variables measured the target information and how relevant the indicators were for cross-system analysis. The next task involved determining from the set of forty-eight variables, which subset of variables provided the best cross-sectional measures and best defined the structure in the data while testing the data for the three possible contaminating problems listed above. After each exploratory factor analysis was performed, the resulting factor loading matrix was evaluated.

As four parallel sets of performance indicators were available, the same type of exploratory factor analysis was carried out on each set. Finding similar results across the four data sets would signal detection of the consistency in the data which would point to the "true" underlying structure in the variables.

Variable Elimination

In the first exploratory analysis on the full set of forty-eight performance indicators, a total of 128 cases were included in the analysis. As mentioned earlier, factor analysis will drop from the analysis every case missing any piece of information. Because the missing values were scattered throughout the forty-eight variable set, the snowball effect of missing data across a set of variables, had eliminated nearly two-thirds of the cases from the analysis. Thus, in the next exploratory pass through the data, it was decided to eliminate from further analyses, those variables that compounded the missing data problem and those that were still somewhat questionable as to the quality and comparability of reported information.

Fuel related variables (RVM/FUEL, TVM/FUEL) were omitted because with four different types of fuel listed for motor bus operations it was difficult to validly compare fuel efficiency across systems. Local and state subsidy related variables (e.g., RVH/LSUB, RVH/SSUB) were removed because definitions of local versus state subsidies were inconsistent. Capital subsidy variables were omitted because they can greatly shift from year to year.

The passenger miles (PASM) variable was missing from almost 20% of the cases. To increase the number of cases entering into the analysis, variables based on PASM (e.g., PASM/OEX, PASM/TPS) were eliminated from the data set.

Variables related to active vehicle counts were also removed because about a third of the cases have a problem of some sort. A distinction was not always made between school buses, charter buses and other motor buses. Some cases listed more active vehicles than total vehicles and vehicle inventories were incomplete for some companies.

The variable RVM/TVM was eliminated because sixty-five of the cases had revenue vehicle miles equal to total vehicle miles, a strong indication of a definitional problem, which greatly inflated the kurtosis value of the variable. The roadcall related variable, TVM/RCAL, was ignored because the definitions for what makes a true roadcall were unreliable. The variables related to total expense (e.g., PASM/TEX, TREV/TEX) were removed because total expense is not truly comparable across systems; there are no set parameters for depreciating capital costs. Finally, REV/RVH was so highly correlated with OREV/RVH that it was eliminated, to counter redundancy in the data.

With each exploratory factor analytic pass through the data sets, the variables were checked against the factor structure to determine if remaining variables presented any of the structural problems mentioned above. The factor loading pattern resulting from each of the exploratory analyses was evaluated to identify that set of variables which best determined the emerging underlying structure of the data. With each pass through the data, the underlying structure became more clearly defined. The number of cases entering into the analysis had increased from 128 to 198 and the same general solution appeared across the four different sets of data.

The final set of thirty performance indicators that remained after the fourth pass through the data reflected a strong set of performance indicator variables. These portrayed such highly consistent factor loadings across all four data sets that it was evident that the most salient features of the performance concept model had been identified.

Table 2-2 lists the forty-eight performance indicator variables selected for analysis from the Section 15 data base. They are portrayed within the framework of the Fielding et al. conceptual model. Those variables eliminated prior to the final analysis are marked with an asterisk to offset them from the final set of thirty performance indicators used in subsequent analyses.

TABLE 2-2. FORTY-EIGHT PERFORMANCE INDICATOR VARIABLES USED IN ANALYSES

COST EFFICIENCY MEASURES

TVH/EMP	*RVM/FUEL
RVH/OEMP	*TVM/FUEL
TVM/EMP	TVM/MEXP
PVEH/ADM	TVM/MNT
PVEH/OP	*TVM/RCAL
PVEH/MNT	RVH/OEXP
*TVH/AVEH	TVM/OEXP
TVH/PVEH	RVH/TWG
*TVM/AVEH	RVH/OWAG
TVM/PVEH	RVH/VMWG
*RVM/TVM	RVH/ADWG

SERVICE EFFICIENCY MEASURES

TPAS/RVH	*RVH/LSUB
TPAS/RVM	*RVH/SSUB
TPAS/PVH	RVH/OSUB
*PASM/TPS	*RVH/TSUB
TVM/ACC	*TPAS/LOA
RVH/ACC	*TPAS/TSUB
REV/PVEH	*REV/TSUB
*REV/RVH	REV/OSUB
OREV/RVH	PAS/OSUB
REV/TPAS	

COST EFFECTIVENESS MEASURES

PAS/OEXP
*PASM/OEX
PAS/TWAG
*PAS/FUEL
*PASM/TEX
OREV/OEXP
*TREV/TEX

*Variable omitted prior to final analysis

FINAL FACTOR ANALYSIS ON THIRTY PERFORMANCE INDICATOR RATIO VARIABLES

The final factor analysis was carried out on the cleaned set of thirty performance indicator ratio variables. After all the data cleaning and verifying strategies, after all the exploratory passes through the data and after all the considerations for data quality, these thirty variables were chosen to represent the best possible information on performance currently available in the Section 15 data base.

Principal component factor analysis with varimax orthogonal rotation was carried out on the four different sets of thirty performance indicator variables. Two different computer routines were used--SPSS-PA1[6] and BMDP-P4M[7]. The latter was used to compare, as closely as possible, the current analyses with the previous work.

The patterns of factor loadings were so similar between the reported data, weighted data and the two sets of logs that it appeared very convincing that the underlying structure in the data set had, indeed, been found. Appendix E contains the factor pattern matrices for the final factor analyses.

The set of the "best" thirty performance indicators was analyzed across the four different data sets. With the exception of the raw, unweighted data exactly the same number of factors, in the same order, and portraying the same factor loading pattern, emerged.

Seven factors, accounting for approximately 85% of the variance emerged from the analysis. Table 2-3 portrays the pattern of factor loadings for the final weighted data set. Factors One, Two and Three represent output per dollar cost, utilization of service and revenue generation per expense, respectively. These first three factors directly relate to the three major categories of the performance concept model--cost efficiency, service effectiveness and cost-effectiveness outlined by Fielding, et al.

Factors Four, Five and Six represent labor efficiency, vehicle efficiency and maintenance efficiency respectively. Finally, Factor Seven is clearly related to safety. Only the raw data set portrayed an eighth factor. It seemed to be weakly

---

[6]N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner and D. H. Bent, <u>SPSS, statistical package for the social sciences</u>. (New York: McGraw-Hill, 1975.)

[7]W. J. Dixon, Ed., <u>BMDP statistical software 1981</u>. (Los Angeles: University of California Press, 1981.)

FIGURE 2-3. ROTATED ORTHOGONAL FACTOR LOADINGS MATRIX

FACTOR 1

OUTPUT PER DOLLAR COST

TVM/OEXP	.90**
RVM/TWG	.87
RVH/OEXP	.87*
RVH/OWAG	.83
TVM/MEXP	.71
RVH/OSUB	.61
RVH/VMWG	.58

Percent of Variance Explained: 25.7

FACTOR 2

UTILIZATION OF SERVICE

PAS/OEXP	.93
PAS/TWAG	.86
TPAS/RVH	.86*
TPAS/PVH	.84
TPAS/RVM	.83**
PAS/OSUB	.67

Percent of Variance Explained: 16.6

FACTOR 3

REVENUE GENERATION/ EXPENSE

REV/OSUB	.92**
OREV/OEXP	.91*
OREV/RVH	.84
REV/PVEH	.77
REV/TPAS	.70
RVH/OSUB	.66
PAS/OSUB	.61

Percent of Variance Explained: 12.6

FACTOR 4

LABOR EFFICIENCY

TVH/EMP	.91*
RVH/OEMP	.88**
TVH/PVEH	.53
TVM/EMP	.51

Percent of Variance Explained: 9.1

FACTOR 5

VEHICLE EFFICIENCY

TVM/PVEH	.92*
PVEH/OP	−.77
TVH/PVEH	.77**

Percent of Variance Explained: 7.2

FACTOR 6

MAINTENANCE EFFICIENCY

TVM/MNT	.94*
PVEH/MNT	.90**

Percent of Variance Explained: 6.6

FACTOR 7

SAFETY

TVM/ACC	.93*
RVH/ACC	.93**

Percent of Variance Explained: 5.6

TOTAL AMOUNT OF VARIANCE EXPLAINED: 83%

NOTE: * = First marker variable
** = Second marker variable
A cut-off value of .5 used throughout.

related to pubic assistance. However, the raw data set of variables had been based on reported information alone, without disaggregating the multi-mode information on revenues and subsidies. Thus, the weakly defined eighth factor appeared to be only an artifact of the aggregated data.

VERIFYING THE FINAL 1980 FACTOR ANALYSIS

The adequacy and strength of the final solution were determined by Thurstone's five criteria for detecting simple structure solutions in factor analysis results[8]. His criteria are as follows:

1. There should be at least one zero in each row of the factor loading matrix.
2. If *m* common factors appear in the structure, each column of the factor loading matrix should have at least *m* zeros.
3. For every pair of identified factors of the factor loading matrix:
    a. there should be several variables that load highly on one of the factors and minimally on the other.
    b. a large proportion of the variables should load minimally on both factors (when there are four or more factors).
    c. there should be only a small number of high loading variables on both factors.

The rotated factor loading structure was compared against Thurstone's criteria for evaluating structure for its "simpleness" and met each of the qualifying conditions. This was convincing evidence that a clear, underlying structure in the data had been found.

In interpreting and portraying the factor loading pattern, an arbitrary cut-off of .5 had been used as a factor load value. The high-loading, i.e., representative variables for any factor were identified with a .5 factor load, but .5 is strictly an arbitrary choice. Factor loadings of .3 and above are commonly listed among those high enough to provide some interpretative value. However, values of .45 or less, generally do not provide a very good basis for factor interpretation.[9] It was felt

---

[8] H. H. Harmon, Properties of different types of factor solutions. Modern factor analysis. (Chicago: University of Chicago Press, 1967), pp. 97-99.

[9] A. L. Comrey, A first course in factor analysis. (New York: Academic Press, 1973).

that a high cut-off value would make for easier and clearer interpretation of the factors.

The next question was: How much of the variance of the final factor solution was not being accounted for by the identified "high-loading" variables. The data were tested by regressing the high loading variables against the full set of variables representing each factor. For each factor, approximately 95% of the information was still being represented. Overall, 86% of the total variance of the original factor structure was represented in the subset of high-loading variables.

Reliability

A third question regarding the set of high-loading variables that defined the factor structure centered on the reliability--in a statistical sense--of the grouped variables. Cronbach's Alpha was calculated for each group of variables gathered together on a particular factor.

Cronbach's Alpha can be used to evaluate the internal consistency of a group of variables to see if they essentially target the same underlying information.[10] Alpha values range from zero to one with a value equal to one representing perfect reliability, or internal consistency in this case. An alpha value of .8 is considered very reliable.

Standardized Item Alpha was calculated for each group of high-loading variables on each factor, and for each of the four sets of slightly different performance indicators. The alpha values hovered around the .8 criterion on the weighted data set and were all well above .8 on the log set of the weighted data. This was true on all factors except Factor 5 which produced an uninterpretable alpha value. Factor 5 measures the positive and negative poles of the vehicle efficiency concept as shown in the negative and positive factor loadings. Thus, it confounds the calculation of standardized item alpha.

Factor Structure Stability

Two final questions were raised regarding the 1980 final factor analysis. They both focused on a single concern--how "globally" relevant was the final factor structure? Would the underlying structure of the data remain stable over different theoretical assumptions or an increase in data cases?

---

[10] E. G. Carmines and R. A. Zeller, Reliability and validity assessment. (Beverly Hills, Calif.: Sage Publications, 1979.)

Classical inferential factor analyses were carried out on the four performance indicator variable sets. As noted previously, this type of analysis assumes that the data comes from a random sample of cases from a larger population. All solutions and reported statistics are mathematically adjusted to predict values as they would exist in a larger population. Thus, it is conceivable that if a factor structure is somewhat weakly defined, a different structure could emerge from an inferential solution than from a principal components analysis. However, results from both the inferential and principal components analyses were consistent across the four data sets.

To test whether the final structure in the analyses would remain stable over an increase in data cases, an estimation procedure for missing data was used. The BMDP statistical computing package includes a program whereby missing data values can be estimated. Multiple regression on the variables with data is used to predict a "most likely estimate" for any case missing data on some subset of the variables in the analysis. When no prediction can be made from other available data, the mean of the variable of interest is used to replace the missing value. When any case is missing too much of its data, it is not used in the estimation procedure.

A final set of factor analyses was carried out on the four sets of performance indicators where missing values had been replaced with estimates. The number of cases then being analyzed increased from 194 to 280. It was plausible that an increase in the number of cases being analyzed could shift a weak or unstable factor solution to a different factor structure. The final set of factor analytic solutions carried out from the data sets which included estimated values were entirely consistent with the earlier results.

Thus, after rigorous testing of the final 1980 factor analysis, it was found that: 1) the same general underlying structure had consistently appeared across all checking routines; 2) not only the same factors appeared, but they also appeared in the same order and 3) with minor fluctuations, the factor loading patterns were generally the same. Therefore, it was concluded that a stable, consistent and reliable simple structure had been detected out of the larger group of performance indicators.

COMPARISON OF 1980 FINAL FACTOR ANALYSIS TO 1979 ANALYSIS

One of the motivations in analyzing this data in this way was to provide a comparison with the previous attempt to use Section 15 data for performance evaluation.

The earlier attempt was carried out on the first year (FY 1979) data. As might be expected, there were many more problems with the first year of collected data than with the second year of data. The former data set was fraught with missing data problems, imprecisely reported data, and less careful checking procedures before and after analysis.

For the final factor analysis on the 1979 data, one set of raw reported data consisting of thirty-six performance indicator variables was analyzed. A total of ninety-eight cases (out of 311) were in the analysis; the rest dropped out due to the snowball effects of missing data. Only a superficial grooming of the data was done. Thus, many erratic values and questionable zeros remained in the data.

For the final factor analysis on the 1980 data, four sets of similar data consisting of thirty performance indicator variables were analyzed. The data was carefully groomed for accidental or inconsistent values and strategies were developed to differentiate valid zeros from "missing data zeros." All in all, there was much greater confidence in the 1980 data set by the time the current set of factor analyses was begun than was possible for the 1979 data set.

Comparison of the two final factor structures--from the FY 1979 data analysis and from the FY 1980 data analysis--shows that the same first two factors emerge in the same order in both years. Output per dollar cost and utilization of service are Factors One and Two respectively for both factor analyses. Since the first few factors usually account for a large amount of the total variance in the data set, it was clear that the first two key features of performance had been identified in both years. Appendix E also includes the factor loading matrix resulting from the FY 1979 data analysis.

From that point on, the factor structures diverged across years. The remaining seven factors from the total of nine factors in the earlier analyses were as follows: vehicle efficiency, fuel efficiency, public assistance, social effectiveness, maintenance efficiency, revenue per expense and safety. Because the set of performance indicators used in the analyses had differed across years, it was difficult to compare the two any further.

Fuel efficiency and social effectiveness related variables had been dropped in the current analysis. The former did not lend themselves to valid cross system comparisons and the latter were not valid when based on other than service area population. Thus, the two data sets differed somewhat in the variables used for the analyses.

In the 1979 data, weighting strategies had not been used to disentangle the aggregated revenue and subsidy information. Thus, variables relevant to those areas were clearly contaminated and invalid for cross-system single mode analyses. The pattern of variation in such variables would have clearly been different from the other variables in the analysis, and the identification of a public assistance factor in the earlier analysis attests to that fact.

The 1979 analysis, when compared with the current set of analyses, shows that the underlying structures are not so different, but that the two data sets from which the analyses began were clearly different. In the current research there was a great deal more confidence concerning the variables chosen and especially regarding the quality of the data itself. It was strongly felt that the 1980 data analyses had, in fact, detected the key underlying concepts of performance for this data. The increase in number of cases analyzed, the many analyses on the four parallel sets of data, and finally, the rigorous verifying and validating procedures provided a great deal of confidence in the final results.

Further, the fact that both years of data had detected many of the same concepts, despite the poorer quality of the 1979 data, provided stronger validation for the conceptual model of transit performance. However, the final structures detected with the FY 1979 and FY 1980 data were different. The order in which factors emerged from the data was not the same. This was partly due to the use of somewhat different sets of performance indicator variables and partly the result of using the much cleaner and more complete set of FY 1980 data. Since the 1980 data had been so carefully cleaned and verified, it was evident that in the current analyses not only the underlying concepts had been detected, but their relative importance to each other and across the larger set of available data had also been determined.

SELECTING REPRESENTATIVE MARKER VARIABLES

A result of this research was the establishment of a small, unique subset of performance indicators that are particularly useful for performance evaluation by individual transit managers for systems of all sizes. The goal was to identify the minimum amount of data necessary to convey the maximum amount of performance information.

To accomplish this, the factor loading data in the rotated factor structure solutions on the final variable sets were used. High factor loadings represent a high correlation of a particular variable with a particular factor. When a variable has a high factor loading on only one factor, it can be said to "represent" that factor both statistically and conceptually.

To select a small subset of easily accessible performance indicators from the final factor structure five criteria were used: 1) Representativeness of a variable vis-a-vis a factor was reflected in a high factor loading on only one factor. 2) The distribution of values in the variable had to be as close to normal-like as possible. 3) Ease of collection of the variable was assessed via the percentage of data missing. 4) The variable had to have been well captured by the factor structure in general (high communality). 5) The variable selected had to be easily understood by transit managers.

Seven representative or "marker" variables were selected from the final factor structure--one variable representing each factor. Seven "alternate markers" were also identified. These alternates could be used equally well for assessing performance. The seven representative "marker" variables and their alternates are listed in Table 2-4 and 2-5 respectively.

TABLE 2-4. "MARKER" VARIABLES BEST REPRESENTING THE UNDERLYING PERFORMANCE CONCEPTS

FACTOR	PERFORMANCE CONCEPT	BEST "MARKER" FOR PERFORMANCE INDICATOR CONCEPT
1	Output per $ Cost	(RVH/OEXP) Revenue Vehicle Hour per Operating Expense
2	Utilization of Service	(TPAS/RVH) Unlinked Passenger Trips per Revenue Vehicle Hour
3	Revenue Generation per Expense	(OREV/OEXP) Operating Revenue per Operating Expense
4	Labor Efficiency	(TVH/EMP) Total Vehicle Hours per Total Employees
5	Vehicle Efficiency	(TVM/PVEH) Total Vehicle Miles per Peak Vehicle
6	Maintenance Efficiency	(TVM/MNT) Total Vehicle Miles per Maintenance Employee
7	Safety	(TVM/ACC) Total Vehicle Miles per Accident

TABLE 2-5. BEST SET OF "MARKER" VARIABLES AND THEIR ALTERNATES

FACTOR INDICATOR	BEST "MARKER"	GOOD ALTERNATE PERFORMANCE
1	RVH/OEXP	(TVM/OEXP) Total Vehicle Miles per Operating Expense
2	TPAS/RVH	(TPAS/RVM) Unlinked Passenger Trips per Revenue Vehicle Mile
3	OREV/OEXP	(REV/OSUB) Operating Revenue per Operating Subsidy
4	TVH/EMP	(RVH/OEMP) Revenue Vehicle Hours per Operating Employee
5	TVM/PVEH	(TVH/PVEH) Total Vehicle Hours per Peak Vehicle
6	TVM/MNT	(PVEH/MNT) Peak Vehicle per Maintenance Employee
7	TVM/ACC	(RVH/ACC) Revenue Vehicle Hours per Accident

The first three factors account for about 55% of the variance in the data. This demonstrates that for a quick performance evaluation, the first three "markers" would suffice. This small subset of statistics also provides information for each dimension of the performance concept model discussed previously. Thus, by using only three key statistics, a transit analyst could target the most salient performance concepts for individual and cross-sectional transit agency analysis.

The markers and the alternate set of markers are highly reliable (alpha range is from .802 to .937). Thus, with a maximum of seven variables from a much larger data set, the performance of a transit system can be evaluated. To assess the three major categories represented in the Fielding, et al., conceptual model, the first three "marker" variables would be sufficient. Further, any one of the seven factor concepts identified could be assessed by means of the relevant "marker" variable.

WHO IS NOT WELL REPRESENTED IN THE FACTOR ANALYSIS?

The FY 1980 Section 15 data is somewhat biased toward the larger systems. Although one-third of the systems reporting have twenty-five and under vehicles, it is this group which is consistently missing the largest percentage of its data.

Approximately 16% of this group's vehicle miles or vehicle hours data, 39% of its passenger data and 9% of its maintenance expense data is missing. In the final set of thirty performance indicator variables used in the factor analysis the small system group was missing from 7 % to 37% of its data. Thus, the small systems group was not well represented in the factor analysis.

This could have introduced a bias in the final solution. However, when the estimation of missing values procedure was used on the data, the factor structure that emerged was consistent with other results. Therefore, it was concluded that the final factor structure would remain stable even with increased representation from the smaller systems.

CONCLUSION

The FY 1980 Section 15 data has been used to identify and test the most easily accessible and parsimonious set of performance indicators for fixed route transit. The research had two objectives: first to find the minimum amount of data necessary to provide solid and stable performance evaluation capability, and second to test the validity of results obtained from the previous analysis of FY 1979 data.

The use of factor analysis on a large set of performance indicator ratios gleaned from the data the structure of the key underlying performance concepts. From the factor structure, a small subset of seven variables was identified and tested against the larger data structure. These seven variables are the most salient performance indicators currently available in the Section 15 data base. They can be used together or individually to assess fixed route transit performance.

There is a great deal of confidence in the data used and in the final results. Rigorous cleaning, verifying and grooming procedures carried out before analysis insured that the input data was as complete as possible. Careful decisions regarding which variables to keep and/or drop from the analysis provided the best possible set of performance indicators available for cross-sectional analysis in the Section 15 data. The use of four parallel data sets and several exploratory factor analyses detected the simple underlying structure of the data. Finally, the rigorous testing and validation of that underlying factor structure was convincing that the most salient performance indicator concepts had been found. The strongly consistent and stable structure in the data led to identification of the key variables for evaluation. These too measured up to testing and verifying procedures. Given the quality of the Section 15 data at hand it is felt that the most salient features for performance evaluation have been determined.

A globally-relevant set of performance indicators has been detected. These variables can be used for peer group comparisons because variables that were problematic for such comparisons were detected then dropped from the analysis (e.g., fuel efficiency and social effectiveness variables, are not given to cross-system analysis).

The strength of this research lies in both the quality of the data used and the rigor with which the results were tested. A relevant set of performance concepts has been identified and linked to easily accessible "marker" variables which can be used for cross-system assessment within peer groups defined by characteristics of transit operations in the next chapter.

CHAPTER 3
PEER GROUP FORMATION AND USE

INTRODUCTION

Using performance indicators for comparison requires the clustering of similar systems into groups, otherwise comparisons are misleading. There are many different ways of clustering transit systems--by size, by mode, by state, etc. This chapter describes a method in which transit systems are clustered by operating characteristics--by size, peak to base requirements and speed. Twelve peer groups are established and compared with peer groups established using the seven performance variables defined in Chapter 2. Peer groups defined by operating variables were found to be superior for performance analysis. Transit agencies clustered into these twelve groups can be reliably compared using the seven performance indicators defined in Chapter 2. These peer groups have stability over time.

Comparison of performance of transit systems and discussion of changes in the transit industry across years is facilitated by comparison of systems which are similar in their operating characteristics. Analysts and policy makers can be misled by comparing performance of systems which are essentially unlike one another. Construction of peer groups of transit systems allows individual systems to be compared to others which are similar, rather than with systems which differ in their operating environments. In addition, the relationship between operating characteristics and performance can be examined by focusing on differences in performance across peer groups with different operating characteristics. Finally, transit industry changes across years can be viewed in relation to operating characteristics of systems by comparing peer groups.

TYPOLOGY FOR TRANSIT

Separating transit systems into peer groups which share similar operating characteristics is analogous to separating any set of objects into a small number of groups in which members of the same group are more similar to each other than to objects in other groups, and the groups differ from one another. Problems of this sort are common in the social and biological sciences and in applied settings such as marketing research. One example of the application of such analysis in marketing research is the clustering of neighborhoods based on demographic characteristics from census data as the basis for targeting market segments. In biological sciences researchers often use cluster analysis as an aid to classifying plants or animals into clusters based on their anatomical similarity. The results of such analyses are the assignment of each object to one and only one of the groups or clusters.

The initial question in the formation of peer groups of transit agencies based on operations is how operating environment is to be measured. Ideally one would use demographic variables such as service area population density to determine the operating environment of each system. However, since these are not available in a form compatible with the level of reporting in the Section 15 data, four differentiating variables were chosen to measure inherent differences in operations. These variables are: total vehicle miles, number of peak vehicles, speed and peak to base ratio. Each reflects some aspect of the operating environment within which a transit system operates. Total vehicle miles and number of peak vehicles measure the overall size of the system. Total vehicle miles relates to maintenance and capital needs of the transit system because it measures the actual usage of vehicles. Peak vehicles reflect the daily maximum capacity of the system and the resultant labor needs in terms of drivers and management. Differences in speed capture the difference between urban and suburban systems. Peak to base ratio indicates the degree to which a system is oriented to peak service. In the absence of demographic data directly measuring service area characteristics like population density, household income and trip patterns, these variables, which are available in the Section 15 data, tap important variations in operating characteristics of transit systems.

Formation of peer groups requires grouping together agencies which have similar profiles across these four operating variables. For example, two agencies which are both large, slow and have high peak to base ratios should be placed in the same peer group, whereas systems which are small and fast should be assigned to a different peer group. The goal is to construct peer groups so that agencies within a group are similar to each other, and different from agencies in other peer groups. On the average agencies in one peer group will have profiles of operations which are distinct from those of other groups.

Cluster Analysis

Several data analysis methods exist for such analysis, including cluster analysis, multidimensional scaling, and Q factor analysis. In this research cluster analysis was chosen as the analytic tool for constructing peer groups because in contrast to multidimensional scaling and Q factor analysis, cluster analysis provides a grouping of the objects into a number of distinct groups, and cluster analysis routines are available which handle a large number of cases, such as are present in the Section 15 data. Cluster analysis is a technique ideally suited for forming peer groups because it provides an objective means for defining how similar objects are and an objective means for forming peer groups based on these similarities.

Cluster analysis is a general term referring to a large number of procedures which have in common the goal of constructing groups of items (either data cases or variables) based on their similarity across a profile of observations. The result of a cluster analysis is the formation of a number of groups of items and the assignment of each item to one of these groups. A summary of many of the techniques for doing cluster analysis can be found in Everitt.[1]

Cluster analysis, and similar techniques which construct groupings of the data, differ from methods such as discriminant analysis which attempt to classify objects into known groups. The latter type of analyses are different from cluster analysis in that they require that the groups be known in advance, whereas cluster analysis constructs the groups.

The most common and frequently used clustering methods are "hierarchical" clustering methods. Such procedures form clusters in a series of steps. The most common of these methods begins with each object belonging in a cluster by itself, and each step joins two clusters from the previous step into one more inclusive cluster. The procedure continues, joining clusters at each step until at the final step all objects are joined into one, all inclusive, cluster. At each step in the process, cases which are relatively more similar to each other will be in the same group. Since a hierarchical clustering solution provides a series of groupings from one in which each case is an individual cluster, to one in which all cases are joined into the same cluster, the researcher must choose a level in the hierarchichal series of clusters which provides a useful and meaningful number of clusters of the data.

The decision as to the number of clusters present in the data is an important issue in any cluster analysis. Some clustering methods provide a single partition of the data into a pre-specified number of groups. The K-means procedure discussed below is an example of such a procedure. However, most hierarchical procedures provide a series of clusters, from least to most inclusive. The researcher must

---

[1]B. Everitt, Cluster Analysis. (London: Heinemann, 1980.)

decide which set of clusters provides the most meaningful and useful grouping of the data.

DESCRIPTION OF CLUSTERING TECHNIQUES

In forming the peer groups of transit systems, three hierarchical clustering techniques were used in order to insure that the final results were not simply a function of the particular technique which was chosen. In this section these three techniques are explained.

The most important feature which distinguishes among clustering techniques is the rule by which items are included in a cluster, and by which clusters are joined together. Many rules exist for doing this. The three different methods used in this research were: single link, centroid and K-means clustering. These differ in the rule used for forming clusters. Descriptions provided here are fairly basic, and the reader who would like more detailed descriptions is referred to Dixon.[2]

Single Link Clustering

Single link clustering is a hierarchical clustering technique. The procedure starts with information about the similarity (or dissimilarity) among all pairs of items to be clustered. In the current analysis, the input was the dissimilarity between pairs of transit agencies based on their operating characteristics. The single link method starts initially with each case, here a transit agency, as a distinct cluster. The analysis proceeds through a series of steps, at each step combining two clusters (or individual cases) to form a larger cluster. The criterion used to join clusters is that the two clusters are joined which have the smallest difference linking any single member of one cluster with any single member of the other cluster. In other words, the two individual cases in different clusters which are most similar cause their respective clusters to be joined. The process continues until all cases are joined into a single, all inclusive cluster.

Centroid Clustering

The centroid method is similar to the single link clustering method in that it proceeds by joining clusters (or cases) in a series of steps, however it differs in the rule it uses to join the clusters. The centroid method assigns cases to clusters, or

---

[2] W.J. Dixon, Ed., BMDP Statistical Software 1981. (Los Angeles: University of California Press, 1981.)

joins clusters together, on the basis of the distance between a case and the center of a cluster, or the distance between the centers of two clusters. At the initial stage of the centroid clustering procedure each case is a single cluster. At each pass of the clustering process the two clusters which are closest together are joined to form a new cluster. This process continues until at the final step all cases are joined into a single cluster. The closeness between clusters, which is used as the basis for joining clusters, is the Euclidean distance between the locations of the clusters.[3] The location of a cluster is based on its values on the original variables in the analysis (combined across the members of the cluster). When two cases or clusters are identical in their values on the variables, the distance between them will be zero. When two cases or clusters are quite different in their values, the distance between them will be quite large.

When cases are combined into a new cluster the location of this new cluster on the variables is computed by taking the average of the values on each of the variables, weighted by the number of cases in the cluster. This location is called the centroid of the cluster, and is used for computing the distance from that cluster to other clusters.

<u>K-means Clustering</u>

K-means clustering is similar to centroid clustering in many respects, however, rather than joining small clusters to form larger ones, this method divides large clusters into smaller ones. This is often referred to as divisive rather than agglomerative clustering. In addition, this procedure only reports solutions for previously specified number of clusters. The K-means program begins with all cases in one cluster and then at each step in the clustering procedure divides a cluster into two smaller clusters. Clusters are divided on the basis of the distance between their centers. (See the discussion of centroid clustering for a definition of a cluster's center.) The division of large clusters into smaller ones proceeds until a prespecified number of clusters is produced. The final step in the K-means procedure is the

---

[3]The Euclidean distance between two single cases (i and j) defined across the variables (k) is:

$$D_{ij} = [\sum_k (X_{ik} - X_{jk})^2]^{1/2}$$

Where X_{ik} is the value for case i on variable k and X_{jk} is the value for case j on variable k.

reevaluation of the cluster assignment of each case, and the reassignment of cases to new clusters, if another cluster is closer than the original.

One of the drawbacks of the K-means procedure is the need to specify the number of clusters the program is to report. This is a problem for exploratory data analysis since the number of clusters present in the data usually is not known.

General Issues in Cluster Analysis

Two issues are common to all of these clustering methods; first, the decision as to the number of clusters in the data, and second, the problem of how to combine the information about cases into a measure of the similarity or dissimilarity among cases.

There is no fixed rule for deciding the number of clusters present in the data. In the most extreme case each item could be assigned to its own individual cluster, or, on the other hand, all cases could be combined into one all inclusive cluster. The issue is to choose a point in the series of clusters, from least to most inclusive, which provides a useful and meaningful grouping of the data.

Choice as to the number of clusters is made in view of the substantive research problem and the group structure of the items being clustered. In the current context the problem is to choose a number of peer groups of transit agencies so that there are enough groups to capture the major differences among agencies, but so that there are not so many groups that the fine grained distinctions among them are not useful. In addition, it is important to have groups which are neither too small, so that a given agency has few peers, nor too large, so that members of a group differ greatly from each other. Complex statistical procedures exist for making this decision, but were not used in this research. In single link and centroid clustering solutions, the decision about the number of clusters is made after one views the results of the analysis. However, in the K-means clustering the number of clusters must be specified prior to the analysis.

The second issue which is common to clustering analyses is how to measure the similarity among the items to be clustered. In the current analysis 274 transit agencies had complete data on each of the four operating characteristics: total vehicle miles, number of peak vehicles, speed and peak to base ratio. The question is, how should the information on these variables be combined to measure how similar transit agencies are to each other in their operations. Two issues need to be considered in arriving at the measure.

First, these four variables are measured on quite different scales. For example, total vehicle miles (in 10,000's) range from 1 to 10869, while peak to base ratios range from .6 to 5.0. If one were to combine these into a single measure, the differences on variables with large values (for example total vehicle miles) would swamp the differences on variables with small values (for example peak to base ratios), giving total vehicle miles extra weight in the calculation. In order to overcome this problem it is necessary to express all variables on the same scale. This is done by standardizing all variables by transforming them to Z scores. The mean of a variable is subtracted from each value on the variable, and then the value is divided by the standard deviation of the variable. The resulting standardized variables all have means of zero and standard deviations of one.

The second question is how to combine four measures on operations for each agency into one measure of dissimilarity between each pair of agencies. In this research the dissimilarity between agencies was measured by taking the Euclidean distance between cases across the four standardized operating characteristics. The formula for Euclidean distance is given in footnote 3, above.

PEER GROUPS BASED ON OPERATING CHARACTERISTICS

This section describes the formation of peer groups of transit agencies based on their operating characteristics. The final peer groups were formed using centroid method, hierarchical clustering. Single link and K-means clustering were also used to group the agencies into clusters; however, these groups were judged to be less satisfactory than those produced by the centroid method.

Formation of Peer Groups using Centroid Clustering

The twelve peer groups of fixed route, motor bus systems were defined on the basis of four operating variables: total vehicle miles, number of peak vehicles, speed and peak to base ratio. Centroid, hierarchical clustering, as implemented in the BMDP package of statistical analysis programs, was used to form the clusters. The centroid cluster analysis included 274 of the 304 transit agencies in the FY 1980 Section 15 data. The remaining 30 agencies were missing data on one or more of the four operating variables and were excluded from the analysis. All variables were standardized to Z scores (as described above) prior to analysis. The closeness of clusters was measured using the Euclidean distance between their locations.

The analysis produced a hierarchical series of clusters. Inspection of the final solution indicated that there were twelve clear clusters of transit agencies, and that two agencies could not be assigned to any cluster. Peer group assignments for the 274 transit agencies are in Appendix I.

The results of this method of clustering were superior to any of the other clustering methods since it produced distinct clusters of moderate size and assigned all but two of the agencies to clusters. The final solution provided twelve peer groups ranging in size from 2 to 78 members. This was judged to be a more satisfactory division of transit agencies than that resulting from either single link or K-means clustering.

Single Link Clustering

Two clustering analyses were done using single link clustering. One was done with all four operating variables, (total vehicle miles, number of peak vehicles, speed and peak to base ratio) and a second analysis was done using three of these variables (excluding total vehicle miles). The single link analysis using the four variables included 274 of the 304 systems. The remaining thirty systems had data missing on one or more of the four operating measures and were excluded from the analysis. The single link analysis with three variables included 275 of the agencies.

Neither the single link clustering with four operating variables, nor the one with three variables produced a useful or meaningful grouping of the transit agencies. At an intermediate level in the clustering analysis with four variables there were 34 separate clusters. One cluster contained 78 agencies, a second cluster contained 41 agencies and the remaining 32 clusters were quite small with between 2 and 10 members. There was no other point in the set of clusters which provided a better grouping of the data.

The single link analysis with three variables also failed to provide a useful set of clusters. The results were similar to the analysis with four variables. At an intermediate level there were 30 clusters, one containing 78 cases, one with 20 and the others with between 2 and 10 cases each.

Both of these clustering solutions exhibit a problem which is common in single link clustering, called "chaining". In such a result cases are added one after another to a single cluster, rather than being placed in a number of distinct clusters. This does not provide a useful grouping of the data.

K-means Clustering

The K-means clustering procedure was used employing all four operating variables. Two solutions were produced, one with ten and one with twelve clusters.

These values were chosen because they were in the range of the number of clusters produced by the centroid method.

The ten cluster solution produced three groups with only one or two members in each and an additional three groups with more than 60 members each. Seventy-two percent of the agencies fell into one of these three large groups. The remaining four clusters had between 14 and 24 cases each. Clusters which are quite large or quite small are not useful for practical purposes, therefore this analysis was rejected as the basis for defining the peer groups.

The solution with twelve clusters slightly reduced the sizes of the three large groups from the ten cluster analysis by forming new groups or placing cases from these large groups into other groups. However, each of these three large groups still had more than 56 members each. In addition, there were four groups with only one or two members each. Several factors lead us to reject this as our final solution for definition of the peer groups. Most importantly, the need for prior specification of the number of clusters present in the data presumes that one knows in advance the number of peer groups present in the sample of transit agencies. Second, the groups which are produced using this method seem to be quite sensitive to a few cases which have extreme or unusual values on a single variable. That is, clusters with very few members are formed to accomodate cases which have extreme values on one variable thus forcing the other cases to be lumped into a few large clusters.

Comparison of the twelve cluster solution using the K-means procedure with the twelve peer groups defined on the basis of the centroid method provides a means for checking the peer group solution. If two different methods reveal similar clusters of transit agencies, confidence in the groups is increased. We can be more confident that the results are not due to a peculiarity of the particular analytic technique. Comparison of the peer groups from the centroid clustering method with those from the K-means technique provides such a test. Comparing the twelve peer groups produced by these two methods reveals that for 80% of the cases peer group assignment is the same based on the two different methods. Four of the twelve peer groups from the centroid clustering solution were kept entirely intact in the K-means analysis, though in three groups from the centroid analysis other agencies were added in the K-means solution. In seven other peer groups from the centroid analysis more than half of the cases remained together, and one small peer group with 8 cases was divided among four K-means groups. This indicates a high correspondence between the two methods, and adds support to the centroid clustering solution.

COMPARISON OF 1979 AND 1980 CLUSTERS

There are substantive differences between the cluster structures discovered in the 1979[4] and 1980 data. In 1980 there are twelve clusters, whereas there were only eight in 1979. In part this is a consequence of the fact that fewer cases entered into the 1979 data analysis because of missing data problems. In addition the 1980 data analysis was based on more valid and reliable data. Despite these differences, there is a basic underlying similarity between the two analyses.

One hundred and eighty seven transit systems entered into the analyses in both years. Five of the 1979 cluster groups are essentially the same in 1980. The three peer groups that changed between the two years were those of smaller systems. Many more of these smaller systems entered into the analysis in 1980 and thus the peer group structure is finer grained for this size range in 1980. Many of the new groups in 1980 are composed of these smaller systems. Overall, the 1980 analysis can be considered more accurate and detailed for the smaller systems.

Even the peer groups that are essentially the same had minor differences. Each transit system that did not stay with its peer group between the two years was examined to see if it had changed in any way. About half of these had substantial differences in their basic operating characteristics and thus a change in peer group is an accurate reflection of a change in the transit system. These kinds of changes are expected. The other systems which changed peer groups were at the boundaries of their peer group, i.e. they were somewhat extreme in some characteristic relative to their peer group. These changes are also minor and logical given the increased detail of the 1980 analysis and the more careful determination of borders in that analyis.

VALIDATION OF PEER GROUPS

This section examines whether agencies in the twelve peer groups formed using the centroid method of clustering in fact differ in their operating characteristics, or whether, on the other hand, the groupings of systems fail to capture differences in the operating characteristics of their members.

---

[4]Shirley C. Anderson and Gordon J. Fielding, Comparative analysis of transit performance. Final Report No. UMTA-CA-11-0020-1. (Irvine, Calif.: University of California, Institute of Transportation Studies, January, 1982.) (NTIS No. PB82-196478).

Several analyses were done in order to demonstrate the validity and robustness of the twelve peer groups. The analyses reported in this chapter focus on the internal validity of the peer groups; that is, whether they reflect differences on the original variables which were used to form them. Chapter 4 reports on the predictive capabilities of the peer groups; that is, their relationship to other factors, such as performance, which were not used in their formation.

Description of Peer Groups

One of the most straightforward demonstrations that the peer groups differ in their operating characteristics is to examine the average characteristics of the transit agencies in each peer group. Statistics describing the total vehicle miles, number of peak vehicles, speed and peak-to-base ratios of the peer groups are presented in Table 3-1. Inspection of these values indicates that although there is variation within each peer group, the peer groups do in fact differ from each other in their operating characteristics. To the extent that we can describe the differences in these groups, and make predictions about peer group membership based on operating characteristics, our confidence in the validity of the groups is increased.

The two private bus companies in Peer Group 1 stand out because of their extremely high average speed. They are the smallest in size and the lowest in peak to base ratios relative to other peer groups.

Peer Group 2 consists of transit providers primarily located in small urban areas or suburban areas across the United States with populations under 500,000. They are small (1 to 46 peak vehicles), fast (17 to 22 miles per hour) and have average peak to base ratios.

Although Peer Group 3 is a cross-national group, Southwestern systems are disproportionately represented. While a few systems are in the suburban fringes of major urban areas, most are in small cities or towns. These systems are small (2 to 74 peak vehicles) with low peak to base ratios (1.0 to 1.15) and above average speeds.

Peer Group 4 draws from all parts of the country despite its small size. These systems serve small cities with suburban characteristics. Systems in Peer Group 4 have a high average speed (15.9 to 16.8 miles per revenue vehicle hour). They tend to be small (fewer than 50 peak vehicles) with low peak to base ratios. Their speed is consistant with their suburban locations.

TABLE 3-1. DESCRIPTIVE STATISTICS FOR PEER GROUPS

Peer Group (n)		Peak Vehicles	Vehicle Miles (10,000)	Speed	Peak to Base Ratio
1	mean	13	193.8	27.88	1.02
	std. dev.	17	269.7	2.99	0.03
	minimum	1	3.0	25.77	1.00
	maximum	25	384.5	30.00	1.04
2	mean	14	86.7	19.35	1.24
	std. dev.	12	98.8	1.42	0.39
	minimum	1	6.6	17.21	1.00
	maximum	46	405.6	21.34	2.30
3	mean	20	101.6	14.51	1.10
	std. dev.	18	95.4	0.65	0.16
(44)	minimun	2	4.4	13.54	0.80
	maximum	74	435.1	15.65	1.50
4	mean	22	108.0	16.23	1.10
	std. dev.	15	89.8	0.38	0.05
(7)	minimum	10	39.2	15.88	1.00
	maximum	47	257.6	16.76	1.15
5	mean	26	83.7	8.91	1.32
	std. dev.	30	93.9	0.91	0.39
(15)	minimum	1	1.4	7.50	0.57
	maximum	107	318.5	10.86	2.10
6	mean	28	126.7	12.19	1.11
	std. dev.	36	168.0	0.63	0.12
(45)	minimum	2	7.3	10.79	1.00
	maximum	192	850.8	13.49	1.39
7	mean	57	203.7	12.80	1.83
	std. dev.	50	180.1	1.50	0.27
(78)	minimum	4	14.5	9.63	1.37
	maximum	223	817.7	16.26	2.47
8	mean	138	453.7	12.69	2.88
	std. dev.	104	366.0	2.03	0.32
(33)	minimum	3	4.7	8.33	2.31
	maximum	387	1349.4	18.14	3.61

TABLE 3-1. DESCRIPTIVE STATISTICS FOR PEER GROUPS (con.)

Peer Group (n)		Peak Vehicles	Vehicle Miles (10,000)	Speed	Peak to Base Ratio
9	mean	230	1259.3	15.72	1.40
	std. dev.	72	316.2	1.03	0.28
(8)	minimum	96	769.0	14.56	1.11
	maximum	329	1635.4	17.32	1.86
10	mean	393	1723.0	11.10	1.76
	std. dev.	94	451.1	1.78	0.33
(8)	minimum	260	1058.6	8.18	1.10
	maximum	506	2385.3	13.65	2.07
11	mean	889	3465.7	13.53	2.48
	std. dev.	251	1055.0	2.12	0.42
(13)	minimum	666	2405.8	10.17	1.66
	maximum	1573	5688.0	18.40	3.14
12	mean	2477	9850.2	10.58	1.74
	std. dev.	789	1331.6	3.62	0.22
(3)	minimum	1914	3843.4	6.45	1.60
	maximum	3378	10868.7	13.23	2.00
Total	mean	125	519.9	13.40	1.68
	std. dev.	316	1270.3	2.89	0.94
	minimum	1	1.4	4.81	0.57
	maximum	3378	10868.7	30.00	13.00
	number	297	279	277	297

Peer Group 5 is unusual in that nearly half of its members are private bus companies in the urban New York City area, while most of the rest are small mid-western city agencies. The systems in this group are distinguished by their very low speeds. They are slightly below average in size, and average in peak to base ratios.

Peer Group 6 draws systems from most regions of the United States but with a particular emphasis on the Midwest and South central regions. While a few medium sized cities are included in this group, many of the systems serve small towns or somewhat rural areas; three-quarters of these systems are in areas with populations under 250,000. Systems in this peer group range in size, but are generally below average in number of peak vehicles. They have low peak to base ratios.

Members of the largest peer group, Peer Group 7, are found in all parts of the United States. They primarily serve small cities and large towns (77,000 to 500,000), although a number are in towns in metropolitan New York. Systems in this peer group are average in size and speed, but above average in peak to base ratios.

Peer Group 8 has primarily Midwestern and Eastern small to medium-sized cities, although a few of its members are from the outer suburban sections of New York and Chicago. It differs from other peer groups in its high average peak to base ratio (all above 2.3). Systems in this peer group range widely in speed and size, though there are no systems over 400 peak vehicles in this group.

Systems in Peer Group 9 are all from the Southwestern areas of the United States. They predominate in suburban, low density areas with populations between .5 and 1.5 million. Systems in this peer group are above average in size and speed, and about average in their peak to base ratios.

Transit systems in Peer Group 10 are all public agencies in large urban areas (1 to 3 million), in most areas of the United States except the Northeast. These systems have an above average number of peak vehicles (260 to 506) and usually below average speeds, with a wide range of peak to base ratios. Peer Group 10 is similar to Peer Group 11, though the systems are smaller on average and have slightly lower peak to base ratios.

Peer Group 11 includes public transit agencies in major urban areas (1.4 to 16 million) in all regions of the United States. They have a high number of peak vehicles (666 to 1573) and are second in size only to Peer Group 12. These systems are above average in peak to base ratio and are average in speed.

The transit agencies in Peer Group 12 are the major public transit providers in the three largest urban areas of the United States. All three have over 1900 peak vehicles. They are one of the two slowest groups of systems, and they have slightly above average peak to base ratios.

Relationship between Operating Characteristics and Peer Groups

The description of individual peer groups illustrates the relationship between peer groups and operating characteristics. However, summary measures of the strength of the overall relationships between peer group membership and each of the operating characteristics is useful. The eta coefficient provides a summary of the degree of association between a number of groups (such as the peer groups) and another variable (for example, an operating measure).[5] The eta coefficient squared is interpreted as the proportion of variance in an operating characteristic which can be accounted for by peer group membership. Table 3-2 presents four eta coefficients, each describing the relationship between one of the four operating characteristics and the twelve peer groups. These results show that the peer groups capture a large portion of the variability among the agencies on all four of the operating variables. However, the groups seem to be most strongly related to differences in the size of the systems.

TABLE 3-2. RELATIONSHIP BETWEEN OPERATING CHARACTERISTICS AND PEER GROUPS

Operating Characteristic	Eta	Eta^2
Total Vehicle Miles	.968	.938
Number of Peak Vehicles	.952	.907
Speed	.874	.764
Peak to Base Ratio	.915	.837

PREDICTING PEER GROUP MEMBERSHIP

Another approach to looking at the relationship of operating characteristics to peer group membership is to ask whether a system's peer group membership can be predicted from its operating characteristics. Two methods, discriminant analysis and construction of a decision typology, were used to predict peer group membership from operating characteristics.

---

[5]N.H. Nie, C.H. Hull, J.G. Jenkins, K. Steinbrenner and D.H. Bent, SPSS: Statistical Package for the Social Sciences. (New York: McGraw Hill, 1975.)

Discriminant Analysis

Discriminant analysis is a statistical technique for combining information on a number of variables to make a prediction about the group membership of a case or a number of cases. The logic of this technique is the reverse of cluster analysis. Whereas cluster analysis attempts to construct groups of objects from their characteristics on a number of variables, discriminant analysis takes the groups as given, and attempts to find the best combination of the variables to predict membership in these groups. It may be used either to test the assignment of cases to groups, or to make group assignments for new cases with unknown group membership. Discriminant analysis here is used in a descriptive manner since use of cluster analysis to form the peer groups on operating characteristics makes subsequent statistical tests of the differences among groups on operating variables illegitimate.

Four operating variables (total vehicle miles, number of peak vehicles, speed and peak-to-base ratio) were used to predict the most likely peer group assignment for each transit agency. Of the 271 cases in the analysis, the group membership for 246 (91%) was predicted correctly. The discriminant analysis also reported a second most likely peer group assignment for each agency. Of the 25 cases whose group membership was incorrectly predicted on the first pass, 19 were correctly predicted on the second pass. This is a rate of 98% correct on either the first or the second prediction.

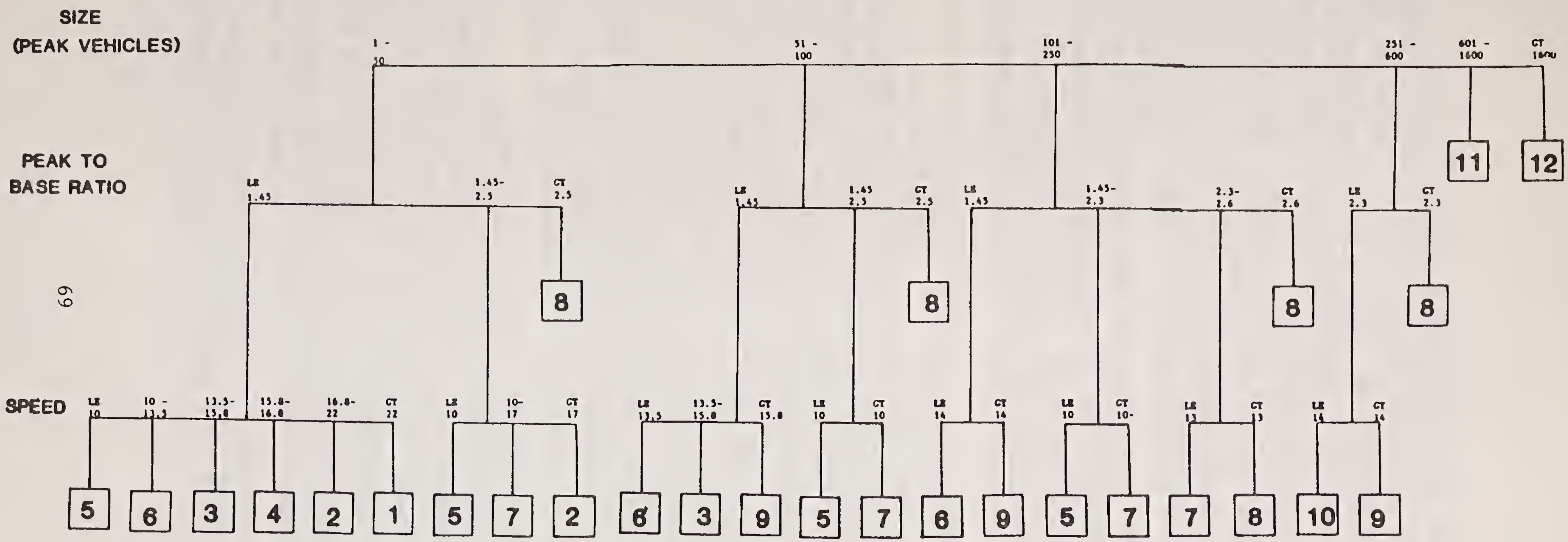
Results of this analysis indicate that the membership in peer groups can be predicted quite accurately from information on operating characteristics. The fact that discriminant analysis uses a different mathematical model to combine information on operating characteristics than does cluster analysis, lends additional confidence to the conclusion that peer groups do capture differences among transit agencies on operating characteristics.

Typology

Another way to demonstrate the validity of the differences among the peer groups on the operating characteristics is to construct a set of decision rules for assigning agencies to peer groups based on their operating characteristics. This also has great practical significance since although cluster analysis constructs a set of groups from data on characteristics of the agencies, it does not handle the problem of the assignment of new cases.

Figure 3-1 presents a decision tree which makes a prediction of peer group membership for each transit system based on its number of peak vehicles, peak to

FIGURE 3-1. PEER GROUP TYPOLOGY



NOTE: LE = LESS THAN OR EQUAL TO GT = GREATER THAN base ratio and speed. Since total vehicle miles and number of peak vehicles are highly correlated for the sample of cases, only the number of peak vehicles was necessary in the decision tree to distinguish among the peer groups.

By starting at the top of the decision tree, and following the path corresponding to the operating characteristics of a system, a transit agency can be assigned to its appropriate peer group. A test of this typology on the FY 1980 data correctly predicted peer group membership of 97% of the cases. This typology could also be used to predict the peer group membership for agencies which did not report data in FY 1980, or to construct peer groups from data reported for other years.

The success of this method of predicting peer group membership on the basis of operating characteristics lends further support to the validity of the peer groups in terms of capturing differences among agencies on operating characteristics.

CLUSTER ANALYSIS ON PERFORMANCE VARIABLES

An analysis was done to develop peer groups on the basis of performance for several reasons. Since peer groups based on operating characteristics capture significant differences on performance, it is of some interest to determine if the inverse is true--those cases most similar in performance will also be similar in their operating characteristics. Peer groups based on performance could also be used as a research tool for exploring possible causes of higher performance. Although a set of four operating variables were used to make peer groups, there may be many other features of a transit system--such as its management form, allocation of expenses to various functions or geographical location--which contribute to performance. Performance peer groups could be used to generate hypotheses about which other facets of operations lead to specific patterns of performance.

The first set of marker variables identified in the previous chapter were used as the measures of performance. One cluster solution was based on all seven markers, another on the first three. The first three were used for an alternate solution because they are the most important in terms of the variance in performance they capture. They also represent the three major aspects of the performance model.

The same methods were used for clustering the cases on performance as were used for the operating variables. All values were standardized prior to the analysis. Euclidean distance was used as the measure of dissimilarity and the centroid method of clustering was used.

The solution found with seven performance variables provided neither a complete nor useful set of peer groups. Only 205 out of 304 cases had enough data to enter the analysis. Of these only 140 (68% of 205) entered into a distinct peer group. The others were left as outliers or grouped into tiny clusters of 2 or 3 cases; too small for practical application. Thus, overall, only 46% of the 304 transit systems became members of useful peer groups. It was concluded that with so many independent measures of performance, there simply were not coherent patterns of performance across all seven variables. Many cases were statistically too different to cluster with other cases.

The peer group solution with the first three performance indicators--RVH/OEXP, TPAS/RVH and OREV/OEXP--gave a more useful division. Two hundred thirty transit agencies had enough data to enter the analysis and all but seven of these became members of a peer group. Thus 76% of the 304 cases were clustered on the basis of performance. Seven performance peer groups were formed with between 13 and 66 members each.

Table 3-3 shows how much association there is between peer group membership based on performance and the seven performance marker variables. The proportion of variance accounted for by performance peer group membership is given by the eta^2 values. As would be expected, the three performance indicators that were used to form these peer groups have much of their variability accounted for, as shown by eta^2 coefficients of .622 to .689. However, comparison to Table 3-2 reveals that discrimination on performance is much less than that on operating characteristics. The lowest eta^2 on an operating characteristic is .764. Thus peer groups based on performance are not as distinct from each other as are those based on operating characteristics.

The peer groups based upon three performance indicators also discriminate on the four other marker variables, although to a much lower degree. For vehicle efficiency, less than 4% of the variation can be accounted for by performance peer group membership. This suggests that the performance groups do not capture much of the differences on vehicle efficiency. Labor efficiency is well accounted for with an eta^2 of .202. Maintenance efficiency and safety fall in between labor and vehicle efficiency.

Table 3-4 relates the performance peer groups to the operating characteristics. Size, as measured by the number of peak vehicles, is well differentiated between

TABLE 3-3. RELATIONSHIP BETWEEN PERFORMANCE CHARACTERISTICS AND PERFORMANCE PEER GROUPS

Performance Characteristic	Eta	Eta^2
Cost Efficiency (RVH/OEXP)	.788	.622
Service Utilization (TPAS/RVH)	.807	.651
Revenue Generation (OREV/OEXP)	.830	.689
Labor Efficiency (TVH/EMP)	.450	.202
Vehicle Efficiency (TVM/PVEH)	.191	.037
Maintenance Efficiency (TVM/MNT)	.281	.079
Safety (TVM/ACC)	.324	.105

TABLE 3-4. RELATIONSHIP BETWEEN OPERATING CHARACTERISTICS AND PERFORMANCE PEER GROUPS

Performance Characteristic	Eta	Eta^2
Peak Vehicles	.525	.276
Peak to Base Ratio	.404	.163
Speed	.222	.050

performance peer groups. Peak to base ratio has a lesser but still important amount of its variance accounted for. Speed is poorly differentiated between the performance peer groups.

The descriptive statistics for each performance peer group on each marker variable and operating variable are given in Appendix F.

The peer groups based on three performance indicators are quite similar to those produced by the cluster analysis based on seven performance variables. Some of the peer groups in the seven variable analysis are all members of the same peer group in the three variable analysis. Some others are divided between two of the peer groups in the three variable analysis. Overall, the three variable peer group analysis is better because it encompasses more cases, forms a more distinctive pattern of clustering and maintains much of the structure found in the seven variable solution.

However, the solution found for the cluster analyses on operating variables and the one on performance variables have little in common. No more than 43% of any of the seven peer groups found with performance indicators fell into the same peer group based on operating characteristics. In fact, on the average, only 13.2% of the cases from the performance clusters moved together into the same cluster based on operating variables. Or put another way, each peer group based on performance has members from about seven of the peer groups based on operating characteristics.

Although the peer groups based on performance capture some of the differences in operating characteristics, each of these peer groups covers a larger range of size and peak to base ratios than any of the peer groups based on operating characteristics. Thus they are useful in demonstrating that managers are able to affect patterns of efficiency and effectiveness despite constraints determined by operating characteristics. Good performance is not just the province of transit systems with the optimal size or peak to base ratio.

USES OF PEER GROUPS BASED ON OPERATING AND PERFORMANCE VARIABLES

Viable peer groups were found through analysis on both operating characteristics and performance indicators. Since these groupings are quite different from each other, it is necessary to consider which is better for managers to use when evaluating their transit system. Most other authors[6,7,8] who have worked on developing methods for evaluating transit performance have limited their analyses to transit systems which are similar in size, type of service area and characteristics of the population to be served (e.g. % of elderly, wage levels). The principle behind these approaches is that transit managers are forced to work within a given set of parameters.

It is unfair and uninformative to compare transit systems which function under totally different sets of circumstances which are not directly under the transit system's control. This principle has guided this research as well. But, unlike many of the previous efforts, the object of this research was to develop a method of comparative evaluation which is nation-wide in scope. Many states do not have enough transit systems to make adequate norms for comparison and performance audits select systems in ways that can be misleading. Even those states with many transit systems, such as New York and Michigan, have systems which operate under quite different circumstances, such as the differences between the New York metropolitan area and upstate New York.

The peer groups defined in this research are based upon operating characteristics which are measured by information obtainable from Section 15 data. Comparisons wthin these groups are more valid than comparing systems with similar performance because they are based upon inherent characteristics of transit operations reflecting the demands of the service area and management's response. Further, it was found that more complete information was available for basic operating variables than for performance indicators. Only two hundred and thirty cases could be clustered on the basis of performance, while 274 could be clustered on the basis of operating characteristics. Operating characteristics are also more

---

[6]Dennis F. McCrossen, Choosing performance indicators for small transit systems. Transportation Engineering, 48(3), March 1978, 26-30.

[7]Kumares C. Sinha, David P. Jukins and Oreste M. Bevilacqua, Stratification approach to evaluation of urban transit performance. Transportation Research Record No. 761. (Washington, D.C.: Transportation Research Board, 1980), pp. 20-27.

[8]James M. Holec, Dianne S. Schwager and Angel Fandialan, Use of federal Section 15 data in transit performance: Michigan program. Transportation Research Record No. 746. (Washington, D.C.: Transportation Research Board, 1980), pp. 36-38.

likely to be stable from year to year, thus allowing the same basic set of peer groups to be used repeatedly.

Peer groups based on operating characteristics provide more discriminatory power across the entire set of performance indicators. The analysis based on performance indicators found that not all performance indicators varied significantly between peer groups, and only the first three performance indicators were clearly different for each peer group. Also, when clustering is based upon performance, there is not much to be learned by comparing a transit system to those with similar performance; they will already be too similar to make small distinctions meaningful. Also, a case at the lower end of performance in its performance peer group, would be similar to the best performer in an adjacent peer group. It doesn't encourage improvement when management realizes that they are the worst performer in the set of best performing transit systems. They just argue over misplacement.

Therefore, use of peer groups based upon operating characteristics is advocated. Comparisons of performance are more indicative of how well management is performing; more systems can be included in such an analysis; the peer groups are more stable, and such comparison has more discriminatory power for revealing particularly weak or strong systems operating under similar circumstances.

CHAPTER 4
THE PERFORMANCE OF PEER GROUPS

INTRODUCTION

The previous three chapters have presented the technical aspects of how this project identified the major dimensions of bus transit performance and established peer groups of bus companies that share similar operating characteristics. The main goal of this chapter is to demonstrate how the combined use of peer groups and key performance indicators creates a powerful tool for understanding differences in transit performance. The first section describes the performance of each peer group on each performance indicator compared to national norms. The second answers the question: Does overall performance on each indicator vary significantly between peer groups? The third section examines the issue of whether the operating characteristics which were used to create the peer groups have a structured relationship to each of the performance indicators. And the chapter concludes with a discussion of further uses for this type of performance evaluation.

The performance indicators and peer groups were identified through statistical means, taking into account the reliability of different parts of the Section 15 data base, patterns of correlations between variables and other mathematical properties inherent to the data. These statistical analyses strongly support the choices of performance indicators and the structuring of peer groups. However, a statistically significant analysis does not guarantee that the results of the analysis will have important implications in practical contexts. While more statistical analysis will be presented in this chapter to further substantiate the validity of the results, emphasis will be placed on demonstrating that the peer groups and performance indicators capture major, distinctive patterns of performance. Graphs and verbal descriptions of the data will be used in this chapter in preference to statistical tables, so that the patterns of relationships may be shown without recourse to statistics or technical interpretations.

Another major emphasis will be placed on showing how the inherent operating chracteristics of transit systems relate to performance. Operating characteristics of a transit system such as size, speed and peak to base ratio have a structured, although complex relationship to transit performance, in both direct and indirect ways. Oram has described how labor costs are greatly increased by high peak to base ratios.[1] Speed is also significantly related to cost per vehicle mile.[2]

Operating characteristics also act indirectly on transit performance by serving as proxies for environmental and demographic characteristics of the service area. A number of studies have found significant relationships between service area characteristics and transit performance. Giuliano found that for California transit systems, the size of the service area, population density and the size of the urban area are related to cost efficiency and labor efficiency.[3] In cross-sectional national studies, Nelson[4] and Miller[5] found, respectively, that demographic variables (low income households, relatively young or old populations, percentage of auto-less families) and environmental factors (city age, city size) significantly influence bus transit costs. In their discussions of their results, these authors link the environmental and demographic variables to the operating characteristics of transit systems used in this study.

Speed is related to the population density of a service area, the traffic congestion on major roads and the kinds of routes operated by a transit system (e.g., express vs local). Peak to base ratios indicate whether a transit system is oriented to work bound commuters (high peak to base ratio) or to transit dependent populations such as the elderly, the low income and students. To some degree the peak to base ratio will also reflect environmental factors such as the lack of parking and insufficient highway capacity which influence service utilization in older Eastern cities. The size of a transit system (measured by both the number of peak

---

[1]Richard L. Oram, Peak period supplements: The contemporary economics of urban bus transport in the U.K. and U.S.A. Progress in Planning, 1979, 12(2), 81-154.

[2]James H. Miller and John C. Rea, comparison of cost models for urban transit. Highway Research Record No. 435 (Washington, D.C.: Transportation Research Board, 1973), pp. 11-19.

[3]Genevieve Giuliano, The effect of environmental factors on the efficiency of public transit service. Transportation Research Record No. 797 (Washington, D.C.: Transportation Research Board, 1981), pp. 11-16.

[4]Gary R. Nelson, An econometric model of urban bus transit operations. (Unpublished Ph.D. Dissertation. Rice University, 1972.) Available from University Microfilms International as No. 72-26457.

[5]David R. Miller, Differences among cities, differences among firms, and costs of urban bus transport. Journal of Industrial Economics, 1970, 19(1), 22-32.

vehicles and total vehicle miles) is related to other constraints on operations: organized labor units are more influential in larger agencies, efficient route scheduling is more difficult, and these cause diseconomies of scale reducing the advantages gained through service integration.

Since there is more evidence for the environmental and demographic influences on transit performance than for the direct influence of operating characteristics, it is necessary to assess the actual impact of the four selected operating characteristics on the performance indicators used in this study. Further, all of the previously cited studies used relatively small samples of transit systems (less than 35 in each case) which do not purport to represent the national transit industry. The data reported here are more comprehensive than that used in the earlier research. In addition, the other studies did not assess the relationship of operating and environmental factors to a set of performance measures that represent distinct major dimensions of transit performance. Thus the relationship of operating characteristics to the major dimensions of performance will be examined in detail.

PERFORMANCE PROFILES OF PEER GROUPS

Each peer group can be characterized by its relative strengths and weaknesses across the seven performance indicators. While two peer groups may look quite similar on any given performance measure, no two peer groups were identical across all seven of them. In addition, no single peer group can be credited with the best overall performance. There are apparently tradeoffs between measures. For instance, all peer groups with high ridership also have relatively expensive service.

The performance profile of each peer group was created by comparing its average (mean) score on each performance indicator to the average for the entire nation as represented in Section 15 data. On the graphs that follow, the national average is indicated by a zero on the vertical axis. Scores above the zero indicate above average performance, ranging up to one standard deviation above the national mean. Scores below the zero indicate below average performance, ranging down to one standard deviation below the national mean. Numerical data used to construct these graphs are given in Appendix G.

<u>Although each peer group is compared to the national average for each performance indicator, this is a descriptive device and not intended to act as an absolute standard of performance</u>. Standards need to be developed relative to each peer group because each is operating under different constraints. In this sample of transit systems, about half are below the national average and half are above the

national average on each performance indicator. Although no peer group averages more than about one standard deviation above or below the national mean, approximately 30% of the individual transit systems will be more than one standard deviation from the mean.

For ease of comprehension, the discussion for each peer group refers to the general concepts measured by the performance indicators. However, the graphs represent performance on the seven marker variables identified in Chapter 2 and are shown in Table 2-4. Table 4-1 summarizes which marker varibles represent which general concept.

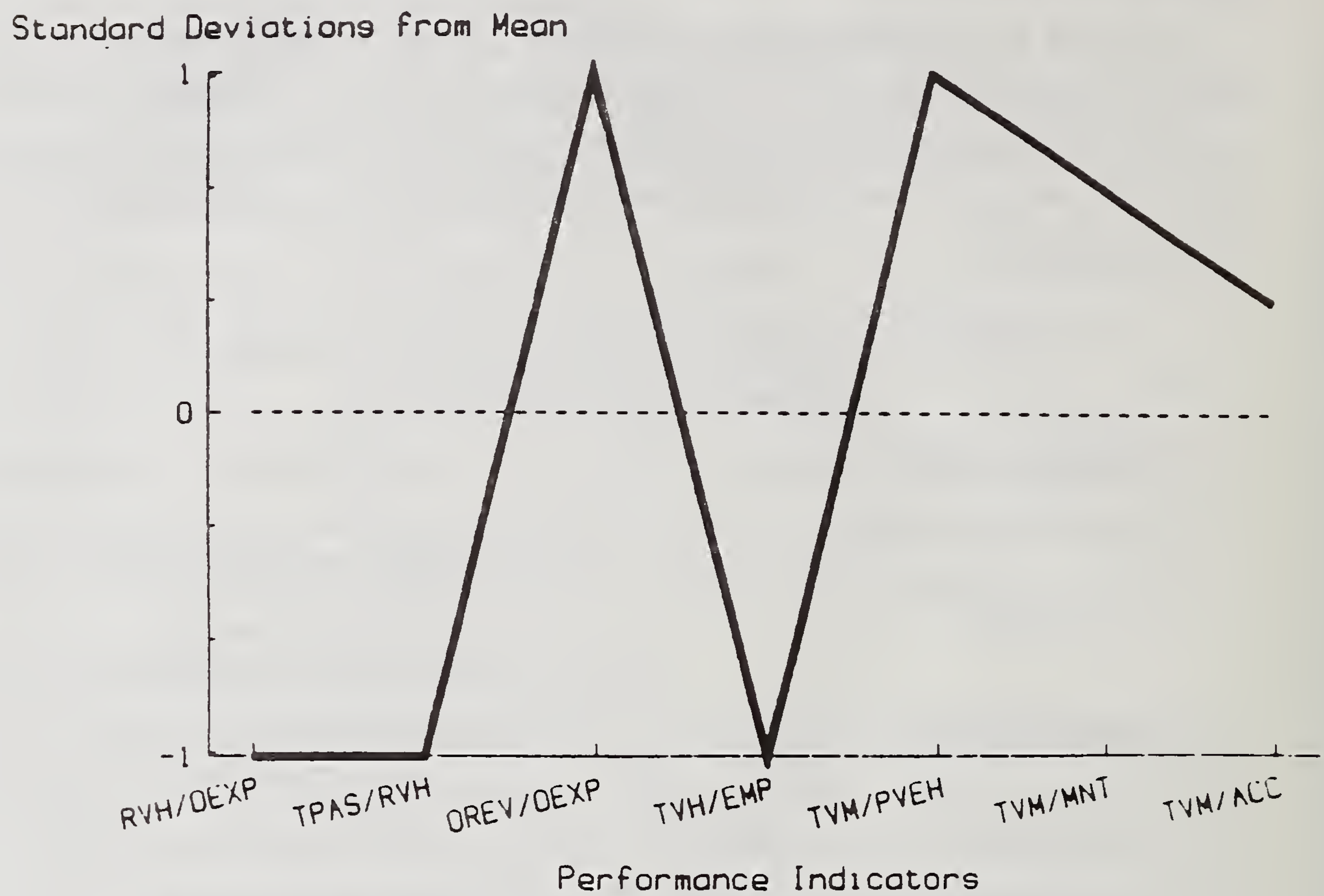
TABLE 4-1. THE RELATION OF MARKER VARIABLES TO PERFORMANCE CONCEPTS

Marker Variable	Concepts
RVH/OEXP	Cost Efficiency Output per Dollar Cost
TPAS/RVH	Service Utilization Service Effectiveness
OREV/OEXP	Revenue Generation Cost Effectiveness
TVH/EMP	Labor Efficiency
TVM/PVEH	Vehicle Efficiency
TVM/MNT	Maintenance Efficiency
TVM/ACC	Safety

Peer Group 1

The two private bus companies in Peer Group 1 stand out because of their high average speed and long passenger trips. Although both have exceptionally high revenue generation performance, they are below the national average on the measures of cost efficiency, service utilization and labor efficiency. The high score on vehicle efficiency is a result of one company's unusually high mileage per peak vehicle. However, because of their high speed, both companies generate many vehicle miles and do well in both safety and maintenance efficiency.

FIGURE 4-1. PERFORMANCE PROFILE FOR PEER GROUP 1



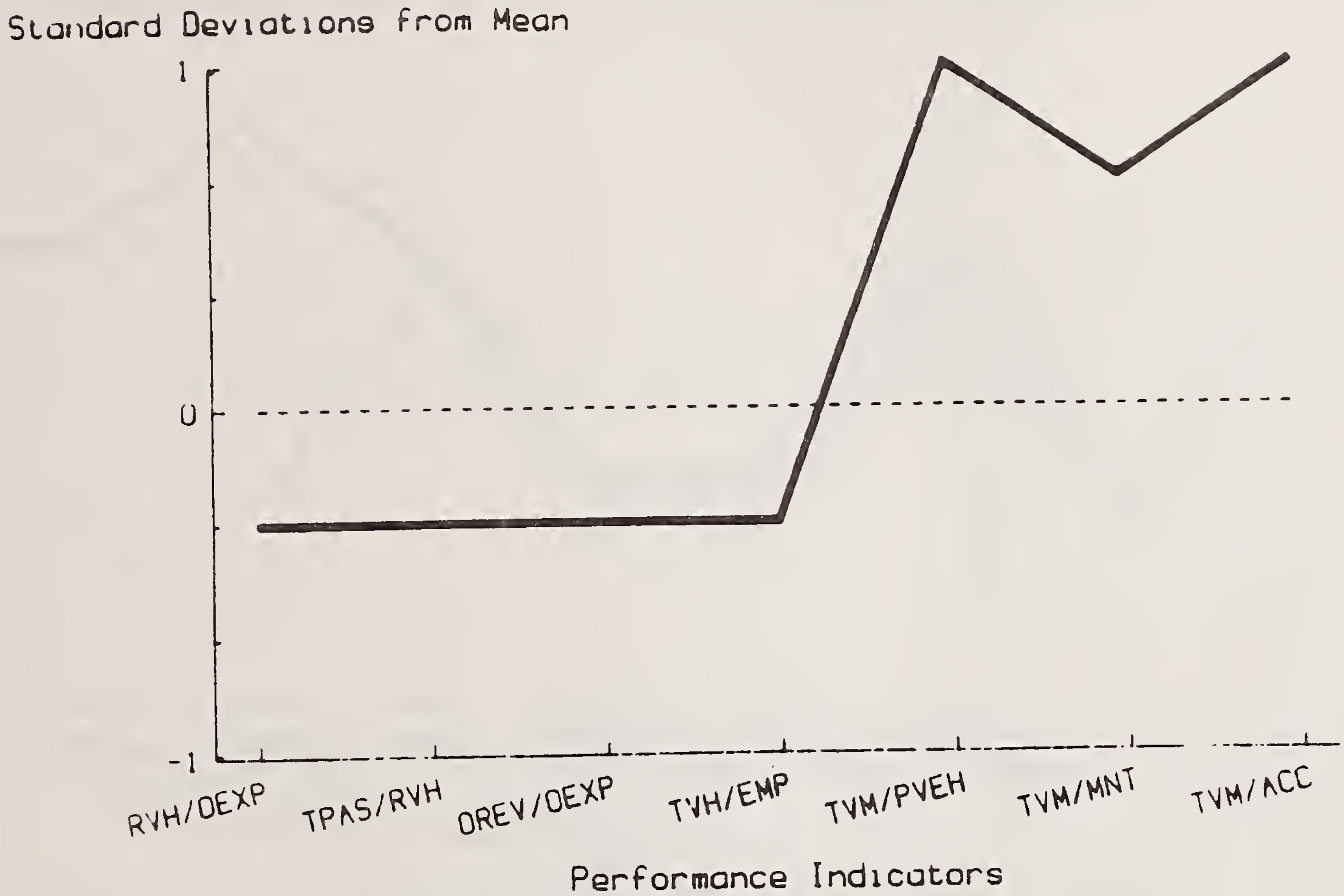
Zero on vertical axis equals national mean.
N = 2

Although there were only two agencies represented in the FY 1980 data for this peer group, it is anticipated that many more of this type will report in the future. Since FY 1982, agencies operating under contract to public agenices were encouraged to report, because a metropolitan area's share of federal transit assistance is determined, in part, by the revenue vehicle service miles operated by all companies in the region.

Peer Group 2

The 16 systems in Peer Group 2 are small, fast bus companies in small urban or suburban areas across the United States. They excel in vehicle efficiency, safety

FIGURE 4-2. PERFORMANCE PROFILE FOR PEER GROUP 2



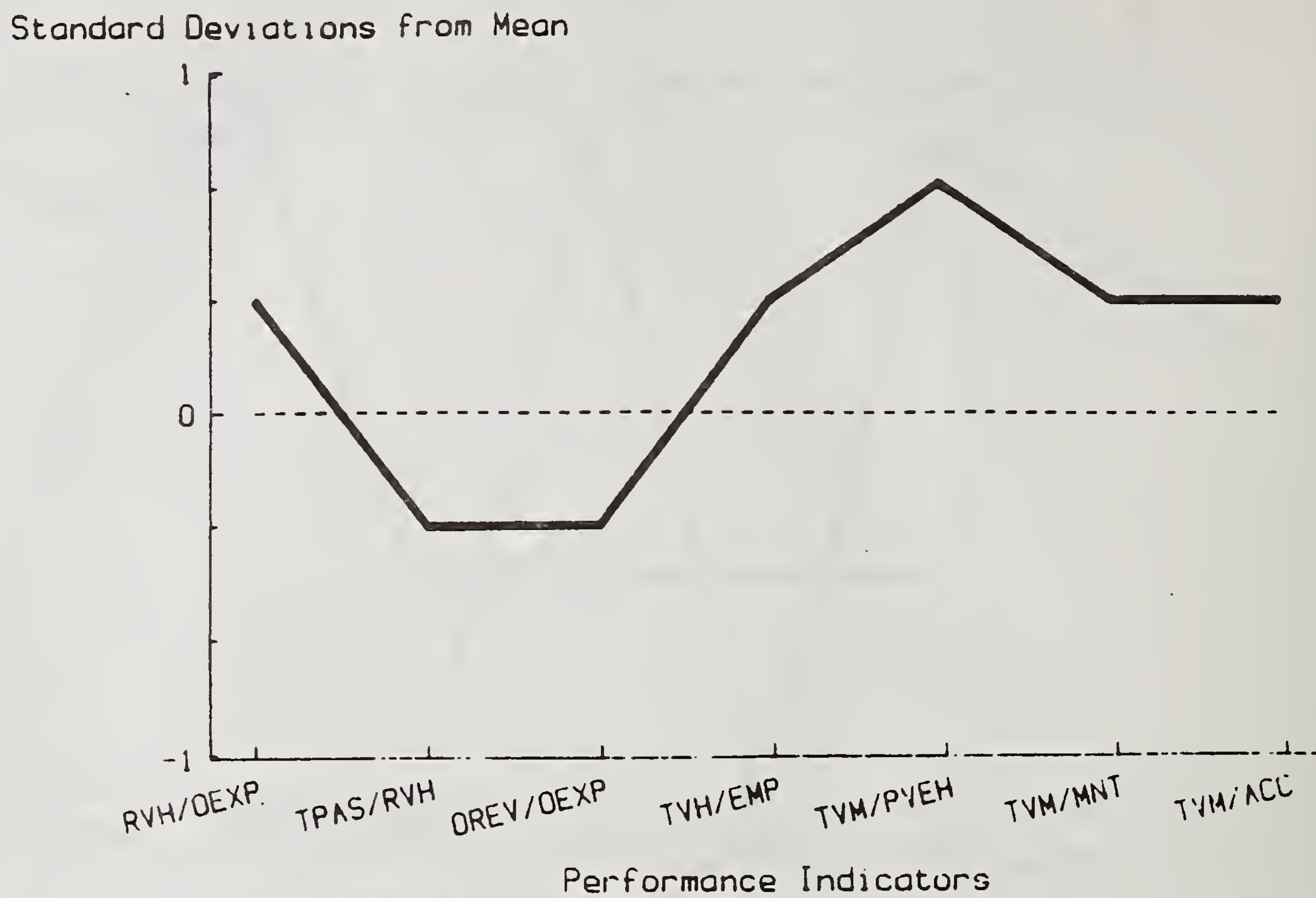
Zero on vertical axis equals national mean
N = 16

and maintenance efficiency--having high mileage per peak vehicle, per accident and per maintenance employee. As a group, these systems have a below average performance on all other measures although there is great variation in cost efficiency, service effectiveness and ratio of operating revenue to operating expense.

Peer Group 3

Peer Group 3 is a cross-national group but draws disproportionately from the Southwest. These 44 systems are small and about average in both speed and peak to base ratio. Peer Group 3's performance profile is quite similar to Peer Group 6's

FIGURE 4-3. PERFORMANCE PROFILE FOR PEER GROUP 3



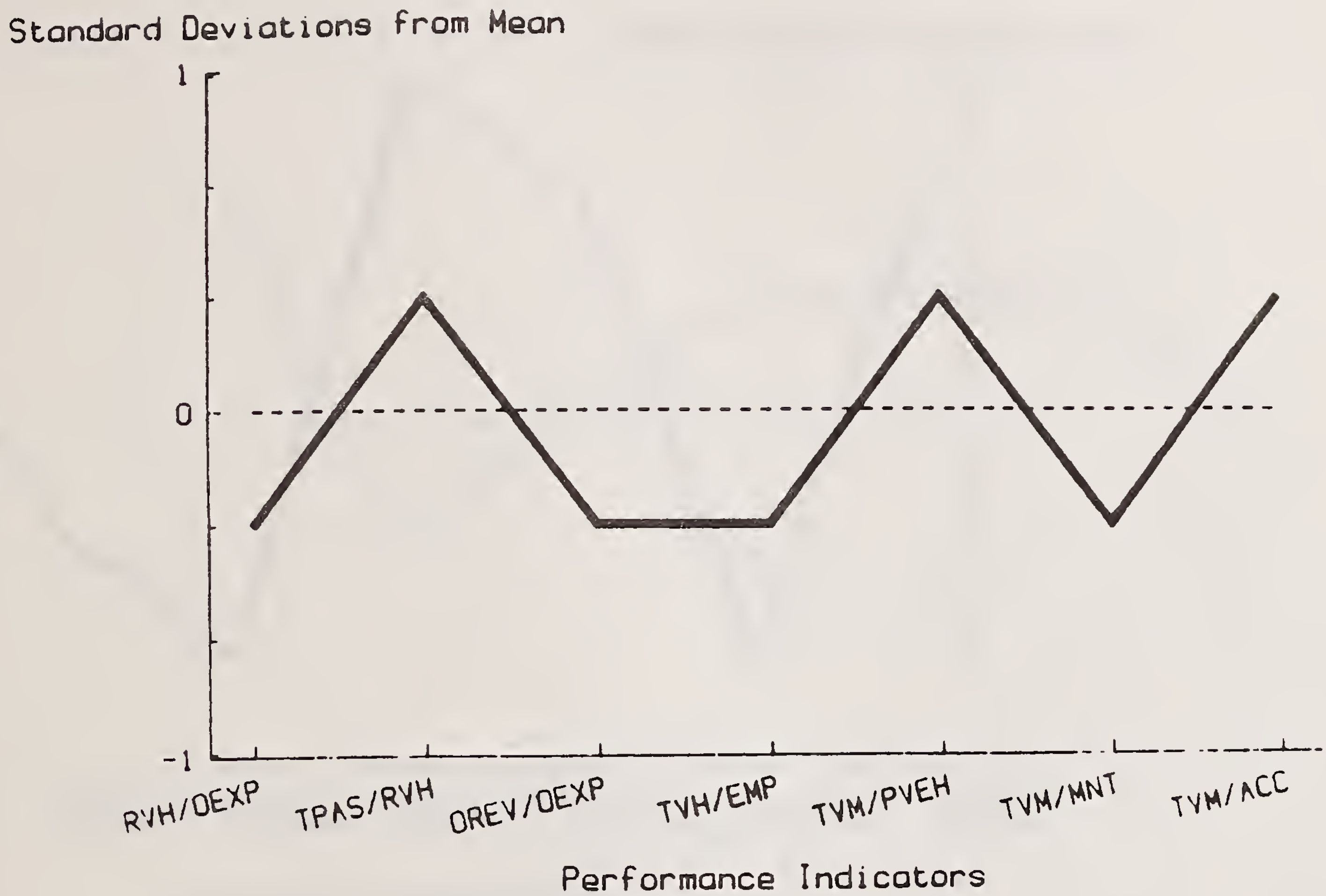
Zero on vertical axis equals national mean
N = 44

except that a slightly lower level of cost efficiency is traded off for a higher level of vehicle efficiency, reflecting the somewhat higher average speeds of this group. Although Peer Group 3 has a just below average number of passengers per hour, its revenue generation is the second lowest among the peer groups. This possibly reflects the state operating assistance available to many of these systems and the local desire to retain low fares.

Peer Group 4

Although Peer Group 4 draws from all parts of the country, it contains only seven systems. These systems are small and serve small cities with suburban characteristics, as indicated by their high average speed (15.9 to 16.8 mph). The

FIGURE 4-4. PERFORMANCE PROFILE FOR PEER GROUP 4



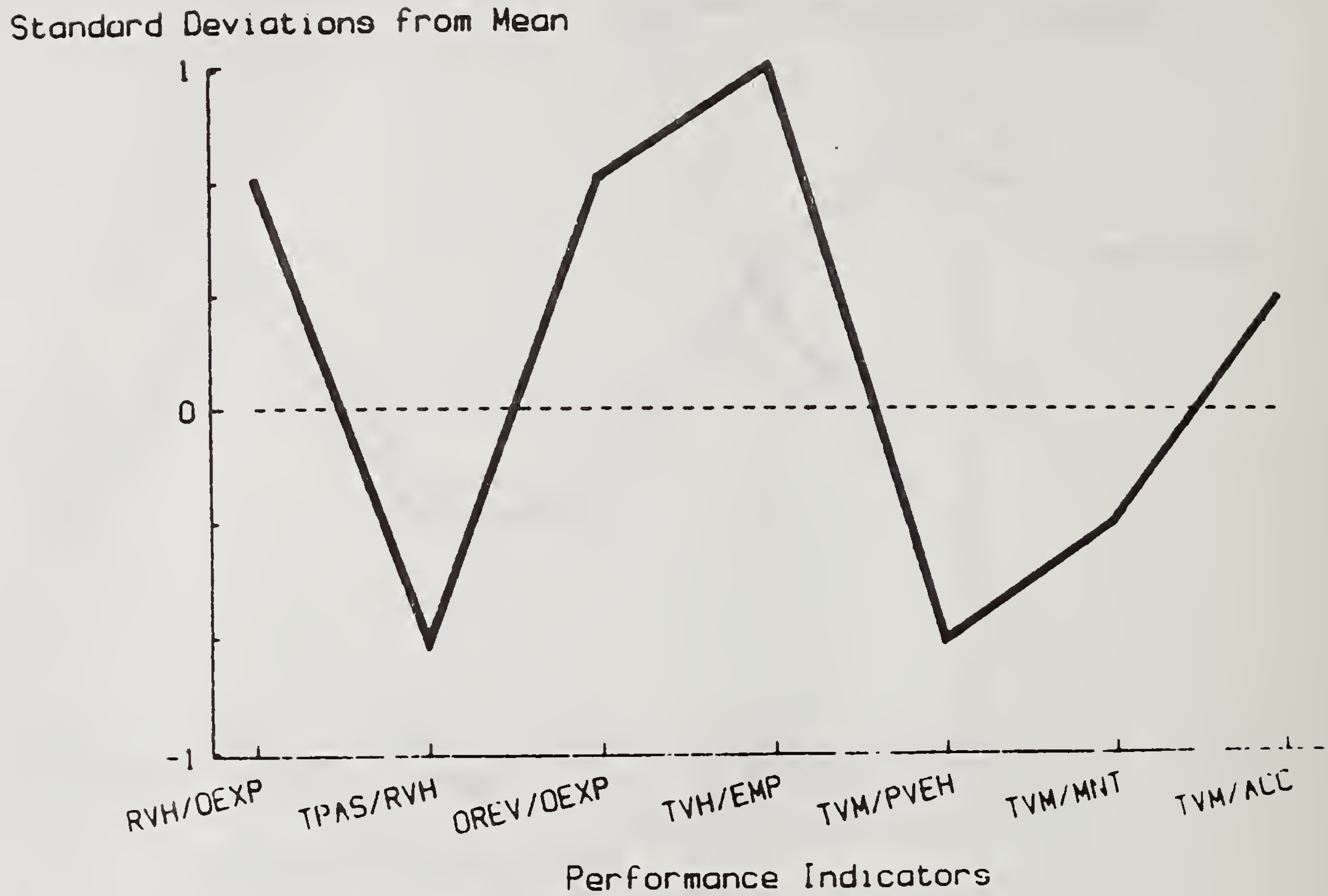
Zero on vertical axis equals national mean
N = 7

group's performance profile hovers just around average on all measures. It is just above average on service effectiveness, vehicle efficiency and safety. All other measures are slightly below average.

Peer Group 5

Peer Group 5 is unusual in that nearly half of its fifteen members are private bus companies in the New York, New Jersey metropolitan area, while most of the rest are in smaller Midwestern cities. Peer Group 5's performance profile is the inverse of Group 9--with well above average performance in cost efficiency, revenue generation and employee efficiency. Below average performance showings are indicated for service effectiveness and vehicle efficiency although this is in

FIGURE 4-5. PERFORMANCE PROFILE FOR PEER GROUP 5



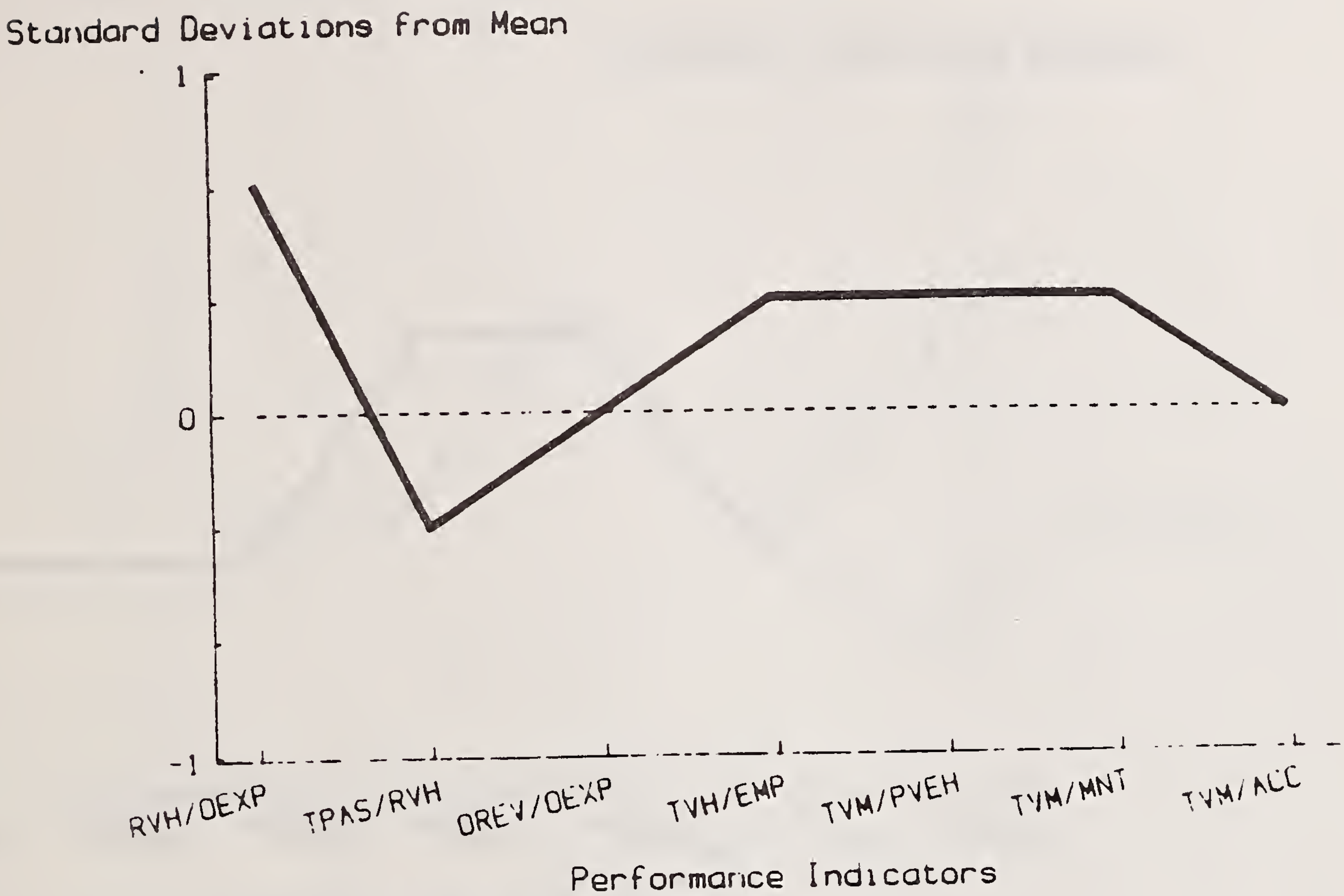
Zero on vertical axis equals national mean
N = 15

large part a result of long passenger trip lengths and slow speeds. Because this group has a mixture of private and public companies, revenue generation is a very heterogeneous variable, and this peer group has both the highest and lowest levels of subsidization to be found among all transit systems.

<u>Peer Group 6</u>

Peer Group 6 draws from most regions of the United States but with a particular emphasis on the Midwest and deep South central regions. The forty-five members of this group are small to medium systems with a low peak to base ratio. Peer Group 6 does well in cost efficiency, labor efficiency, vehicle efficiency and

FIGURE 4-6. PERFORMANCE PROFILE FOR PEER GROUP 6



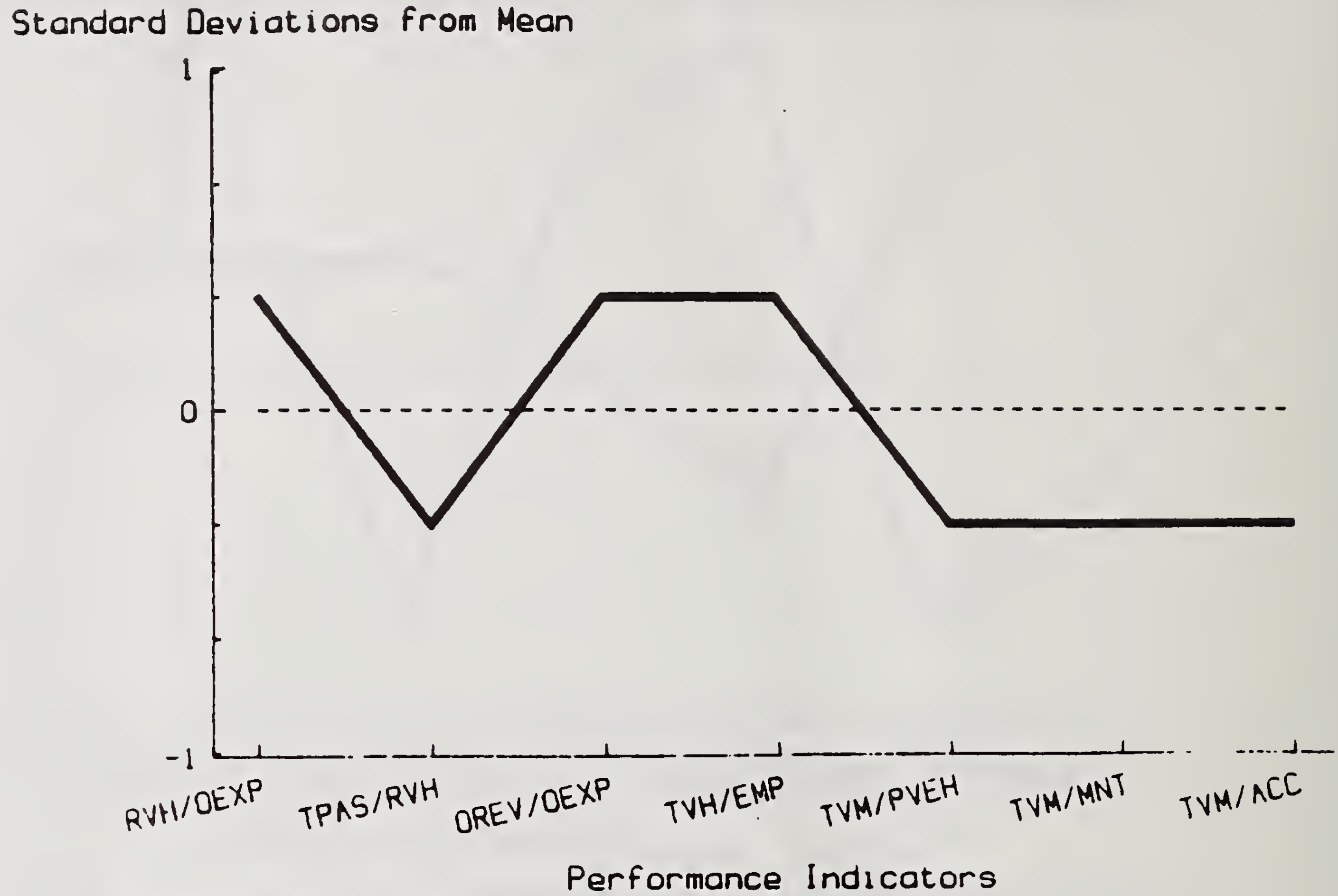
Zero on vertical axis equals national mean
N = 45

maintenance efficiency. Their below average service effectiveness as a group is actually a function of a few systems with very low numbers of passengers per service hour. They are average in safety and revenue generation. Overall, this group has a very high level of performance, rivalled only by Peer Group 3.

<u>Peer Group 7</u>

As members of the largest peer group, the seventy-eight Peer Group 7 systems are found in all parts of the United States. They range in size from small to medium, have average speed and slightly above average peak to base ratios. They are slightly above average in cost efficiency, labor efficiency and cost-

FIGURE 4-7. PERFORMANCE PROFILE FOR PEER GROUP 7



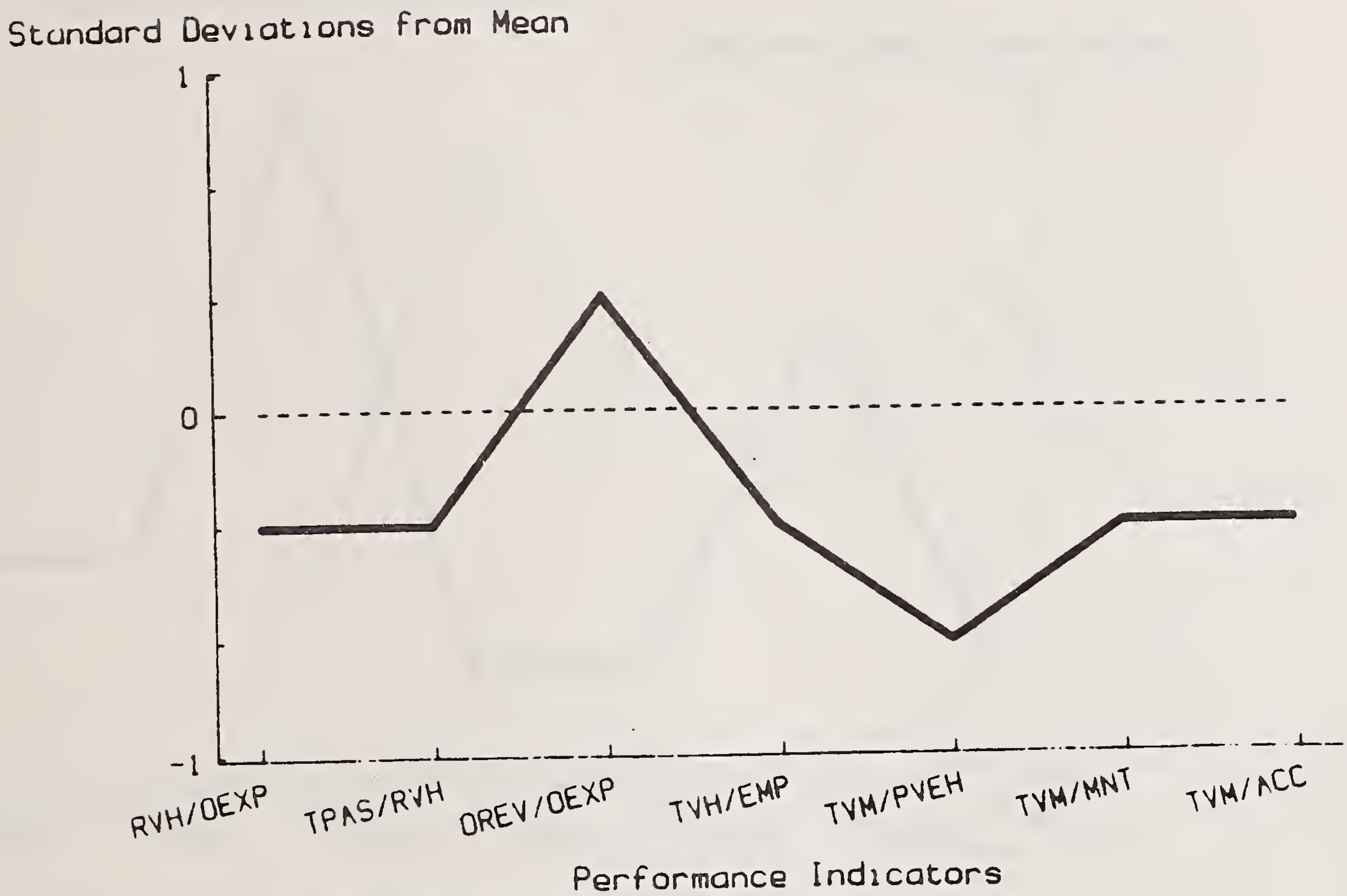
Zero on vertical axis equals national mean.
N = 78

effectiveness. They are slightly below average on all other measures. Although the group is large, the systems are relatively homogeneous on most measures although there are exceptions and outliers.

<u>Peer Group 8</u>

Peer Group 8 has thirty-three small to medium transit systems which share in having a higher than average peak to base ratio. They primarily serve cities. The major strength of Peer Group 8 is that it has a slightly above average proportion of operating revenue relative to operating expense, i.e., they recover a relatively high level of their expenses from the fare box. However, they have the lowest vehicle

FIGURE 4-8. PERFORMANCE PROFILE FOR PEER GROUP 8



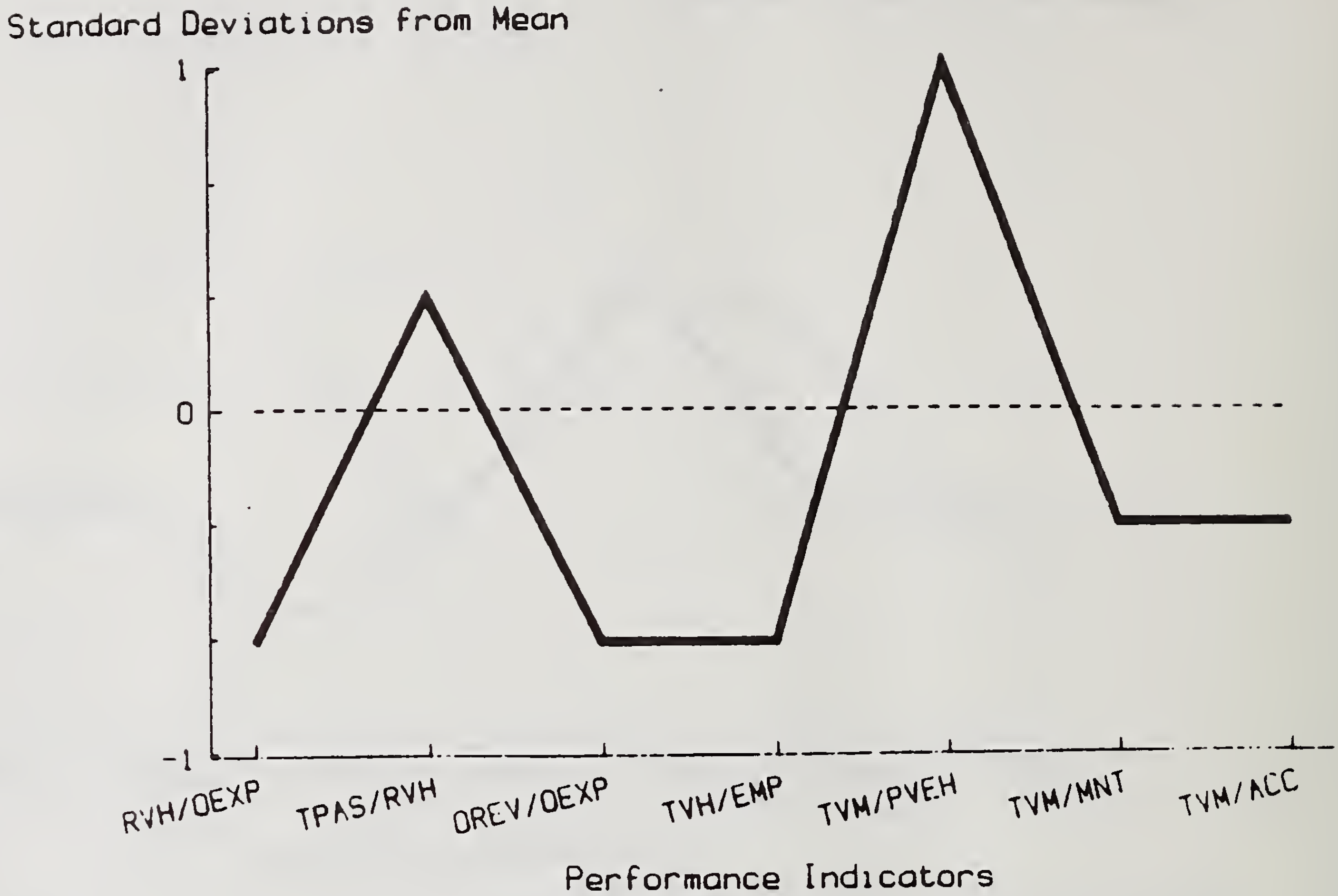
Zero on vertical axis equals national mean.
N = 33

efficiency, probably because of their strong peak orientation. They are slightly below average on the other measures of performance with little variation among them.

Peer Group 9

Peer Group 9 systems are all from the Southwestern region of the United States in suburban areas. These eight systems are above average in size and speed. In this peer group very high vehicle efficiency and above average numbers of passengers are weighed against quite low cost-efficiency, revenue generation, and labor efficiency. Safety and vehicle maintenance efficiency are just below average.

FIGURE 4-9. PERFORMANCE PROFILE FOR PEER GROUP 9



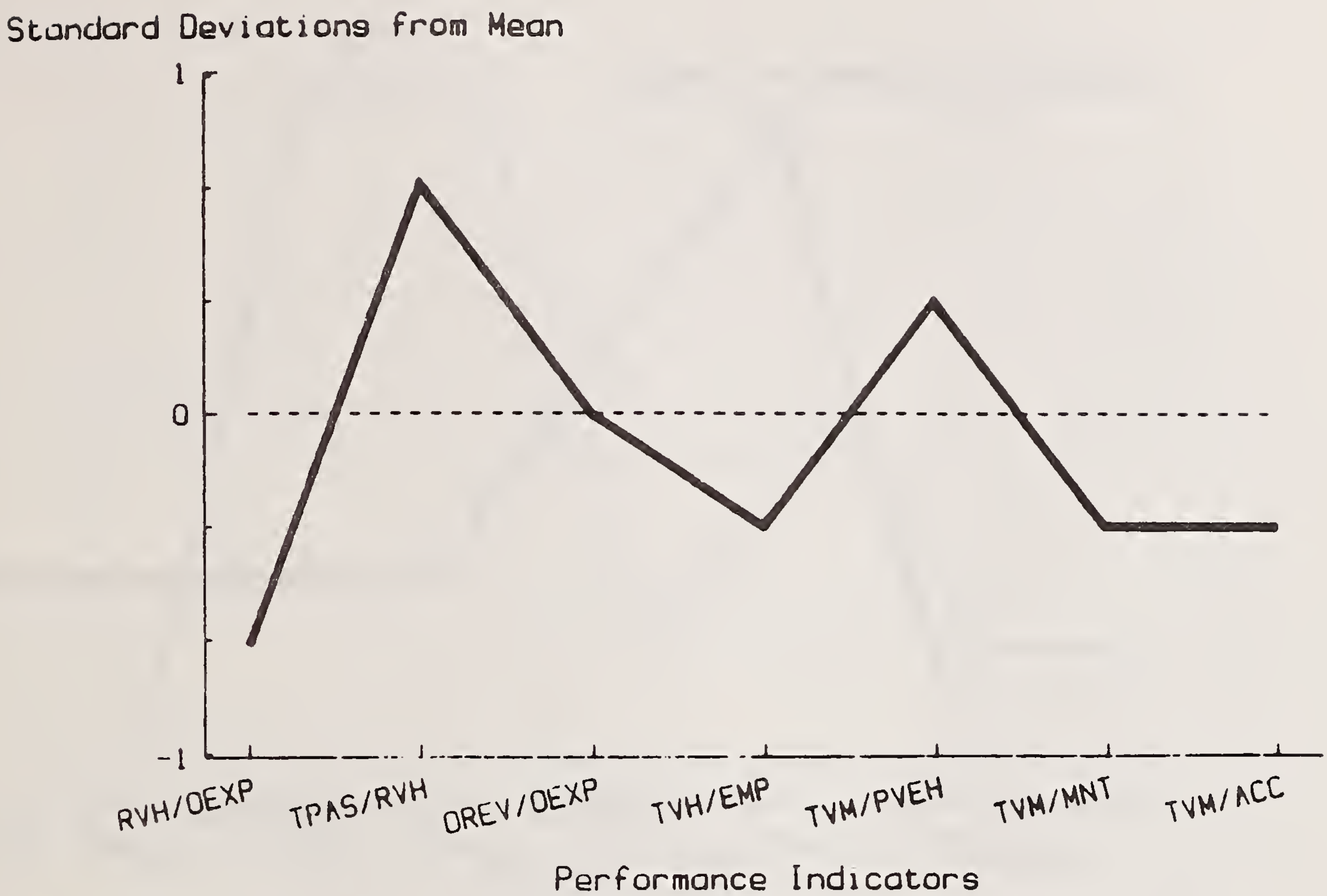
Zero on vertical axis equals national mean.
N = 8

However, the low average on revenue generation is a result of two systems with unusually high levels of local assistance.

Peer Group 10

Peer Group 10's eight transit systems are above average in size, below average in speed, and they serve major urban areas in most parts of the United States, except the Northeast. Peer Group 10 is well above average in the number of passengers per hours that are carried by systems in this group. It is also slightly above average in vehicle efficiency. Its low point is cost efficiency, with revenue

FIGURE 4-10. PERFORMANCE PROFILE FOR PEER GROUP 10



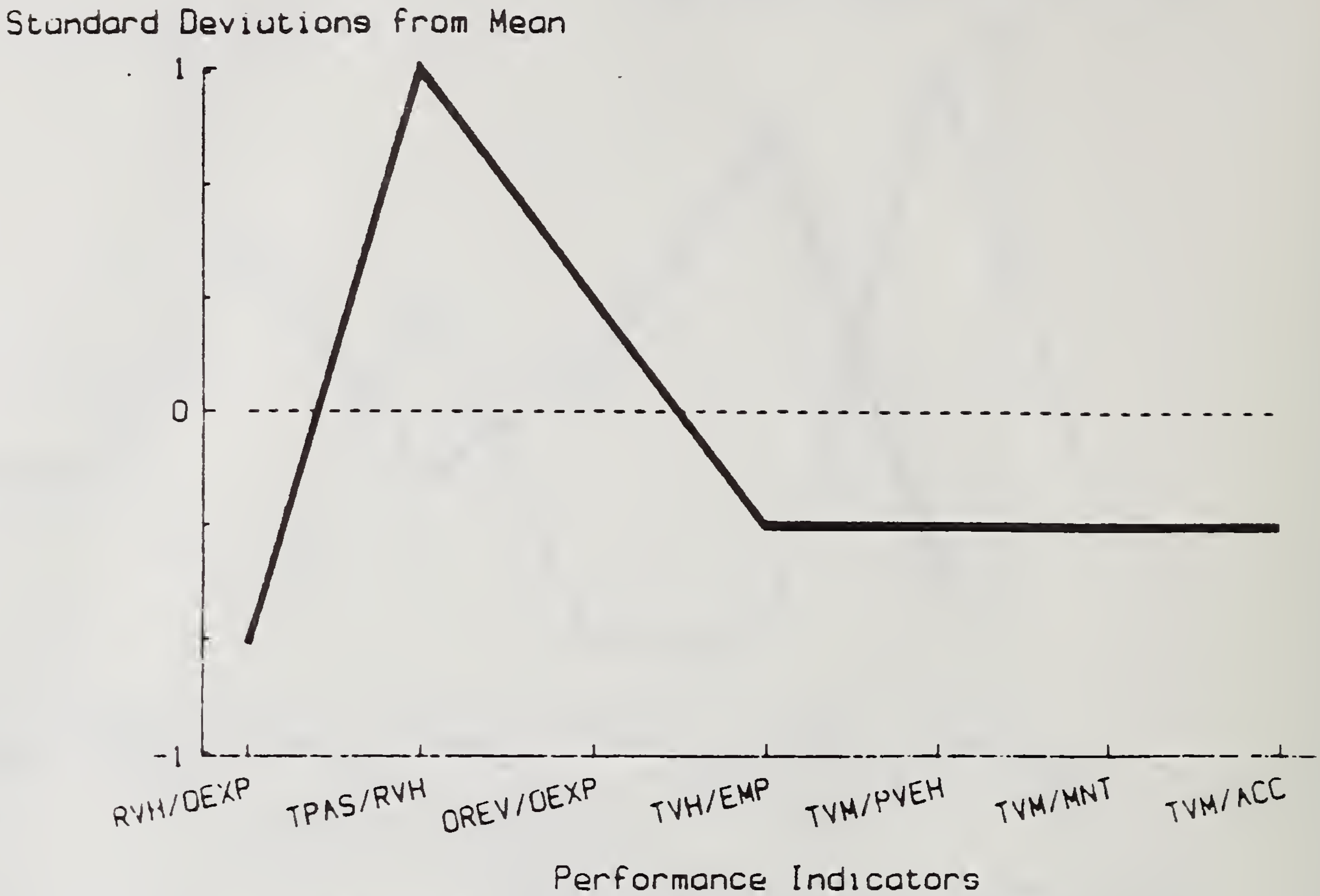
Zero on vertical axis equals national mean.
N = 8

vehicle hours per operating expenses being well below average. On the other measures of performance, Peer Group 10 is just slightly below the mean for all transit systems.

Peer Group 11

Peer Group 11 systems are primarily very large, public transit agencies in major urban areas. These thirteen systems are above average on peak to base ratio. Peer Group 11's major strength is the number of passengers per hour carried. Only Peer Group 12 carries more passengers. Peer Group 11 also generates an above average amount of revenue. However, the service provided by the systems in Peer Group 11

FIGURE 4-11. PERFORMANCE PROFILE FOR PEER GROUP 11



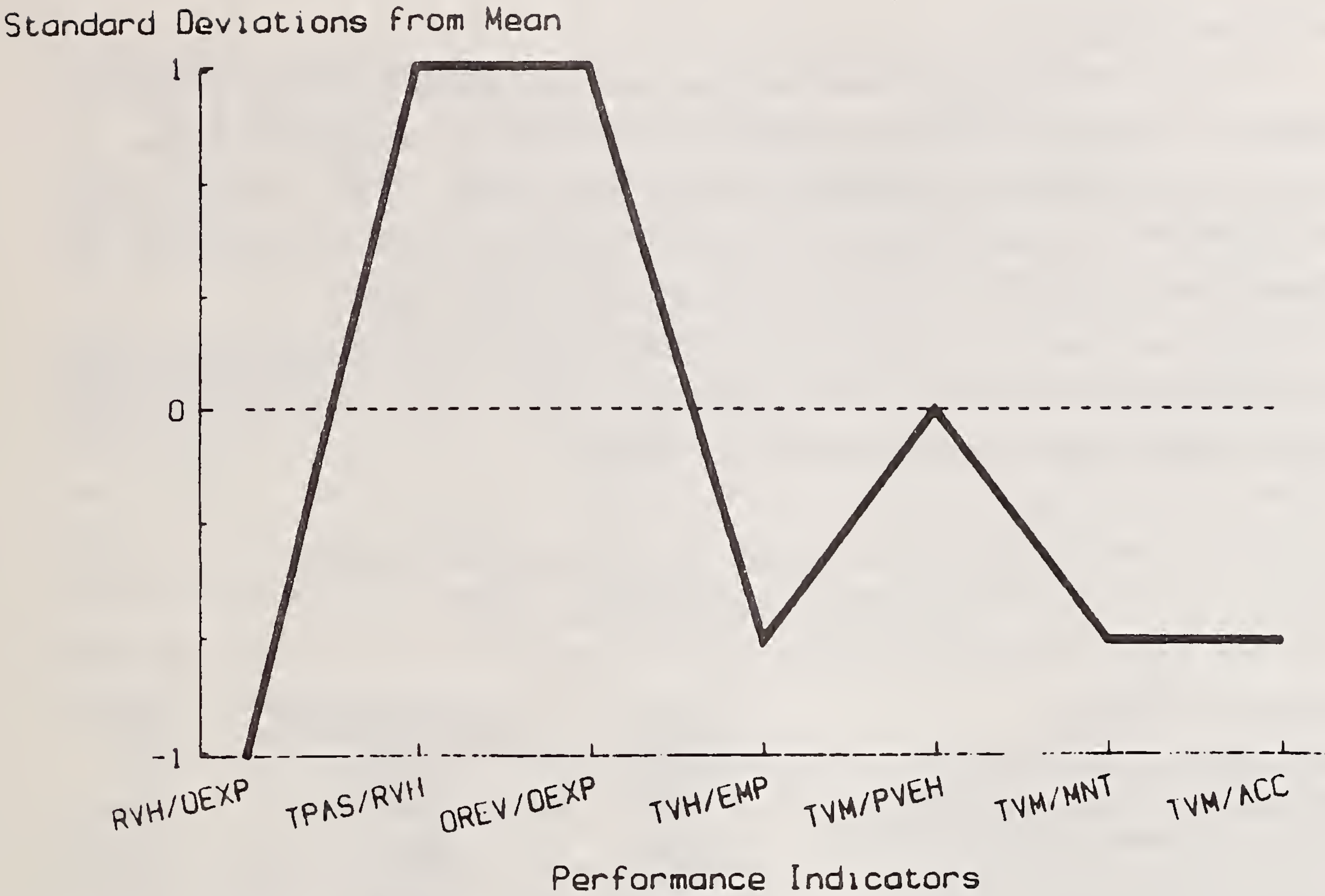
Zero on vertical axis equals national mean.
N = 13

tends to be more expensive than average. On safety, labor efficiency, vehicle efficiency and maintenance efficiency, Peer Group 11 is slightly below average in performance.

Peer Group 12

The transit agencies in Peer Group 12 are the major public transit providers in the three largest urban areas of the United States. All three have well above average ridership per service hour and a resultingly high level of operating revenue. However, they have uniformly high expenses and low labor efficiency. Vehicle

FIGURE 4-12. PERFORMANCE PROFILE FOR PEER GROUP 12



Zero on vertical axis equals national mean.
N = 3

efficiency is average because two systems are below average but one is well above average. They also have well below average safety and maintenance records as a group.

COMPARISON OF PERFORMANCE INDICATORS ACROSS PEER GROUPS

While each peer group has its own distinctive performance profile, this does not prove that each performance indicator by itself captures important differences between peer groups. Two methods were used to explore this issue. A series of statistical tests were used to see if the peer groups were significantly different on each performance indicator. Then, each performance indicator was displayed graphically to show the important, structured ways that the performance indicators vary across peer groups.

Peer Group Differences

Analysis of variance tests whether there are statistically significant differences between the means of a set of groups. It does this by measuring whether the variation between groups is greater than that within groups. Thus, if the groups are more different from each other than are the individual transit systems within the groups, there is a significant difference between the groups overall.

A one-way analysis of variance was done for each performance indicator using the SPSS program Breakdown.[6] Table 4-2 shows the F score, eta coefficient and level of significance for each performance indicator.

TABLE 4-2. RELATIONSHIP BETWEEN PERFORMANCE VARIABLES AND PEER GROUPS

Performance Measure	F score	Significance Level	Eta
Cost Efficiency (RVH/OEXP)	9.839	.0000	.55
Service Utilization (TPAS/RVH)	6.319	.0000	.49
Revenue Generation (OREV/OEXP)	3.466	.0002	.36
Labor Efficiency (TVH/EMP)	5.533	.0000	.44
Vehicle Efficiency (TVM/PVEH)	24.300	.0000	.71
Maintenance Efficiency (TVM/MNT)	4.441	.0000	.41
Safety (TVM/ACC)	5.854	.0000	.46

For each performance indicator, there are highly significant differences between groups. Since the standard cutoff point for significance is a probability level not exceeding .05, it can be seen on the table where significance levels are listed that all performance indicators far exceed this standard. The measures of revenue generation and maintenance efficiency are slightly less differentiated between peer groups than the others as shown by their lower F scores and slightly lower eta coefficients. Vehicle efficiency shows the greatest differentiation between groups as shown by its larger F score (24.3000) and large eta (.71).

However, this analysis does not indicate which peer groups are most different or whether the differences are important. It is possible that one or two peer groups

---

[6]N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner and D. H. Bent, SPSS: Statistical package for the social sciences. (New York: McGraw-Hill, 1975.)

are radically different from the others, and the others are indistinguishable from each other. Thus representation of the actual data will be used to show the structure of peer group differences.

Comparison of Peer Groups

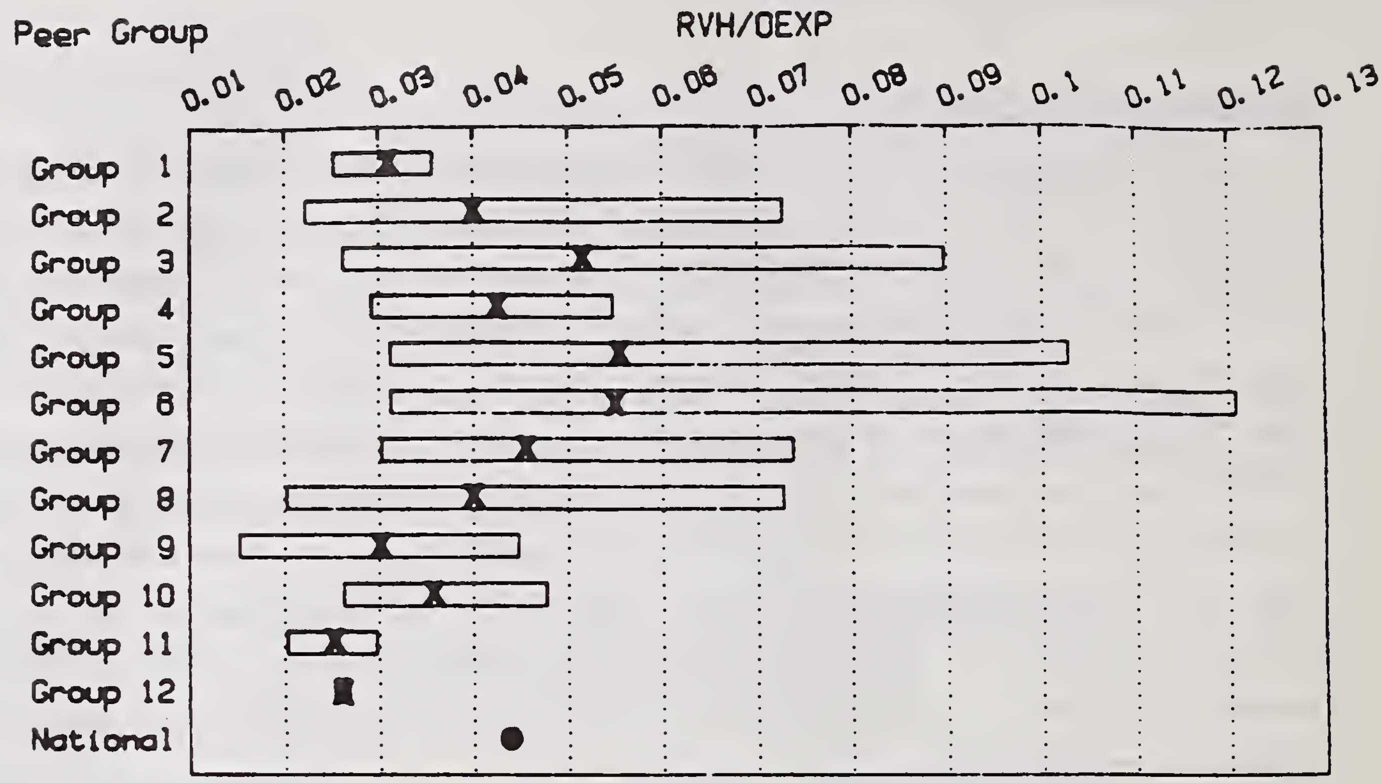
The following set of graphs compares the peer groups on each of the seven performance indicators. Each peer group is represented by a bar with an X on it. The length of the bar shows the range of values that the transit systems in that group achieved on that performance measure. The left end of the bar represents the minimum value and the right end represents the maximum value. The X on each bar shows the average (mean) for that peer group. Since a higher value indicates better performance, the bars with X's farther to the right represent better average performance. The longer bars represent more variation on a performance measure. When an X is not centered on a bar, it shows that the peer group is not evenly distributed around the mean. Therefore, the side of the bar which is longest typically contains one or two values which are extreme compared to the rest of the systems in the peer group.

Each performance indicator will be described in terms of how the peer groups compare on average performance, how much the peer groups overlap in performance and what patterns exist across the peer groups.

Cost Efficiency

The pattern of values for cost efficiency shows wide differences between peer groups. Two peer groups, 11 and 12, not only have a significantly lower average for hours per dollar of operating expense, their best performing members do not do as well as the poorest performing members of groups 4, 5, 6 and 7. Groups 11 and 12 are the very largest systems in this sample. Groups 1 and 2, which are peer groups of small systems, also show lower cost efficiency. This suggests that size reduces cost efficiency at the extremes. The two peer groups which have the next largest members, 9 and 10, also have less cost efficiency on the average but they overlap the performance of the peer groups with smaller systems. The peer groups with small to medium systems such as 3, 5 and 6 do the best on this performance indicator but they also show the most variation. This suggests that small-to medium size expedites performance but does not necessarily guarantee it. Overall, this performance measure clearly differentiates between groups in expected ways.

FIGURE 4-13. THE RANGE FOR COST EFFICIENCY BY PEER GROUP



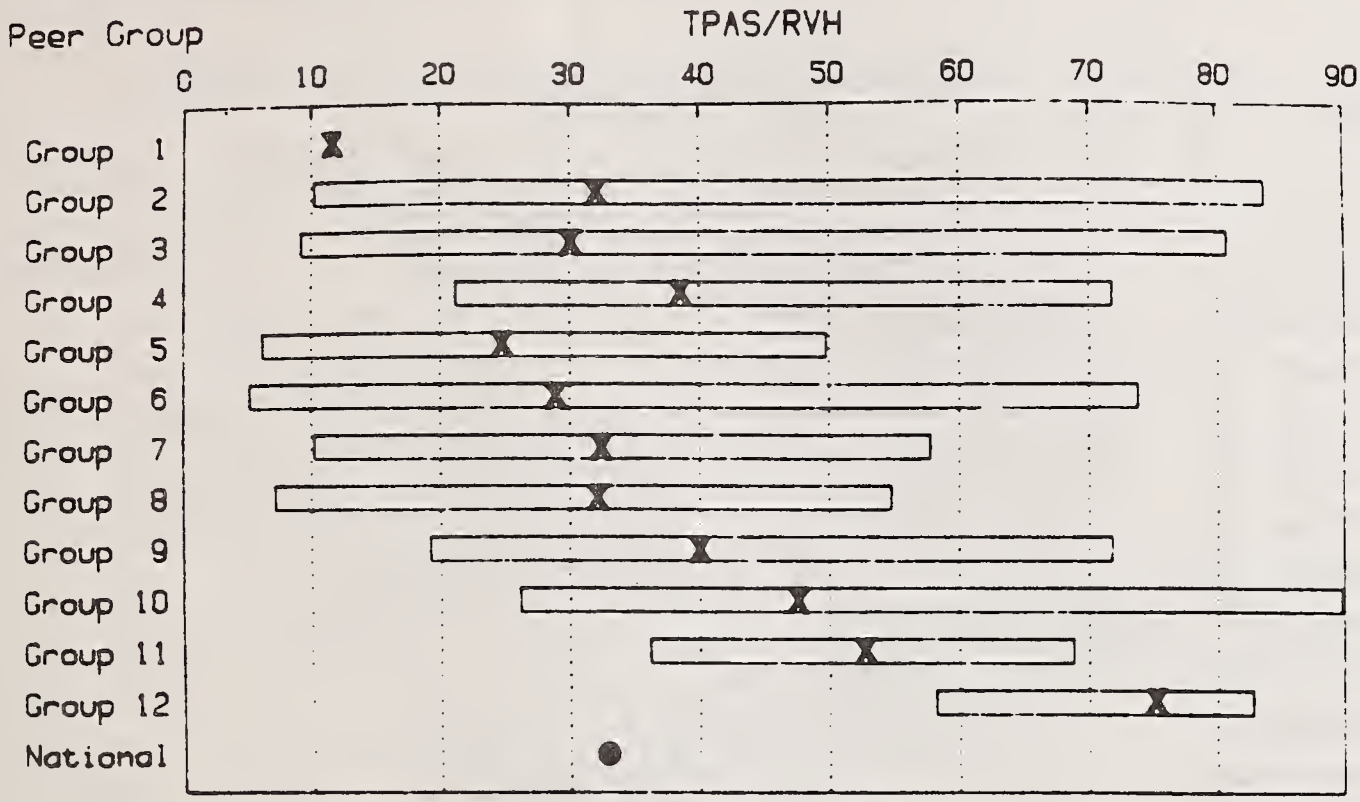
Cost efficiency = Revenue Vehicle Hours per Dollar of Operating Expense.

X on bar shows peer group mean.
X in circle shows national mean.

<u>Service utilization</u>

The pattern for service utilization is the inverse of that for cost efficiency with poor performers on cost efficiency doing best on service utilization. The peer groups of large systems, 11 and 12, carry the most passengers per service hour. Some of the peer groups with smaller systems, such as 5, 7 and 8 do not even overlap group 12. However, some peer groups with fairly large members (i.e., Group 9) do slightly worse than the smaller systems of Group 4. Unlike many of the other performance indicators, practically every peer group shows wide variation in performance. However, despite this variation, the pattern for service utilization clearly varies by peer groups. Since this performance indicator and the one measuring cost efficiency give radically different versions of which transit systems

FIGURE 4-14. THE RANGE FOR SERVICE UTILIZATION BY PEER GROUP



Service utilization = Unlinked Passenger Trips per Revenue Vehicle Hour.

X on bar shows peer group mean.
X in circle shows national mean

are performing best, the need to use a multi-faceted approach to measuring performance is demonstrated. Researchers who advocate a single measure of performance, even if it is a mixed measure of efficiency and effectiveness statistics like cost per passenger, mask important differences in performance.[7]

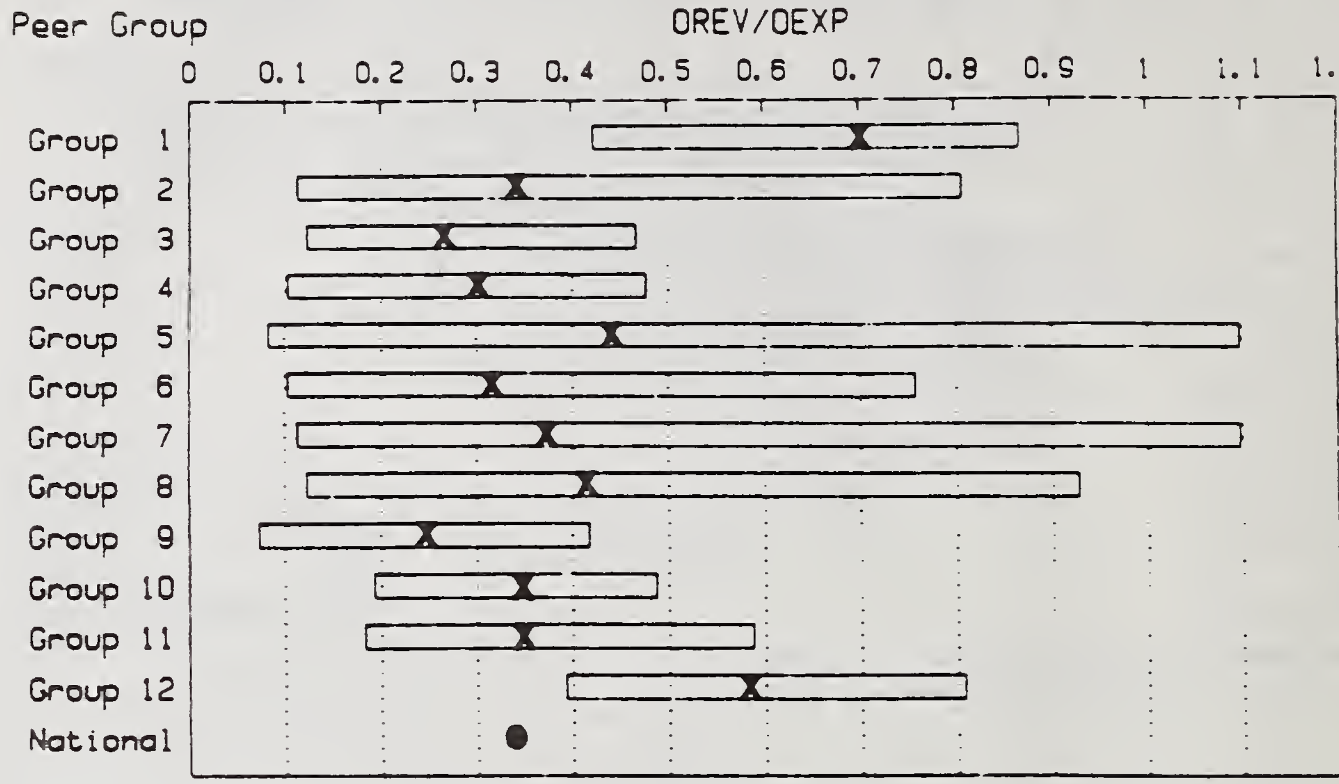
Revenue Generation

The amount of revenue generated by passenger fares and auxiliary sources of earned income (such as advertising on buses) reveals a pattern distinctive from that

---

[7]Timothy A. Patton, Transit performance indicators. Transportation Systems Center Staff Study #SS-67-U.3-01. (Cambridge, Mass.: U.S. Department of Transportation, Transportation Systems Center, 1983.)

FIGURE 4-15. THE RANGE FOR REVENUE GENERATION BY PEER GROUP



Revenue generation = Dollar of Operating Revenue per Dollar of Operating Expense.

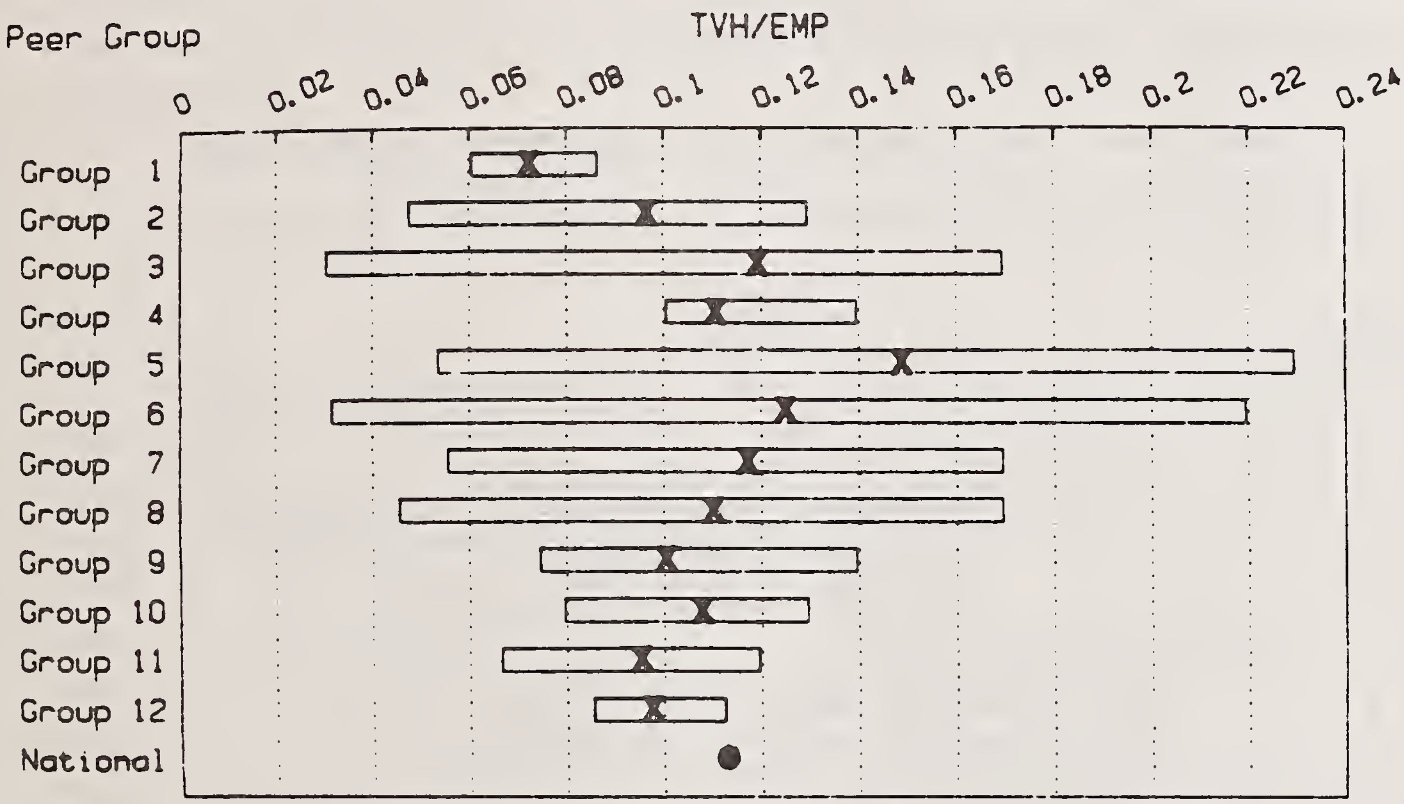
X on bar shows peer group mean.
X in circle shows national mean.

of the previous two performance indicators. All the peer groups cover a relatively wide range of values and the pattern shown here of high/low performers is not reflected on any other measure. Peer Groups 1 and 12 both excel, but probably for quite different reasons. Group 12 systems have extremely high passenger loads and thus are able to generate passenger revenue by volume of passengers. Group 1 systems on the other hand are express commuter service providers which charge high fares and have a low level of subsidization. Group 5 is somewhat unusual in that it combines commuter and inter-city service providers, and resultingly it has the most variation on this measure.

The lower performing groups (3, 4, 6 and 9) tend to have low peak to base ratios but some other groups with low peak to base ratios do average or better (e.g., 2).

For most peer groups, the measure of revenue generation has a distinctive pattern but groups 2, 5, 7, and 8 are not easily differentiated from each other because they are each so varied.

FIGURE 4-16. THE RANGE FOR LABOR EFFICIENCY BY PEER GROUP



Labor Efficiency = Total Vehicle Hours (10,000's) per Employee.

X on bar shows peer group mean.
X in circle shows national mean.

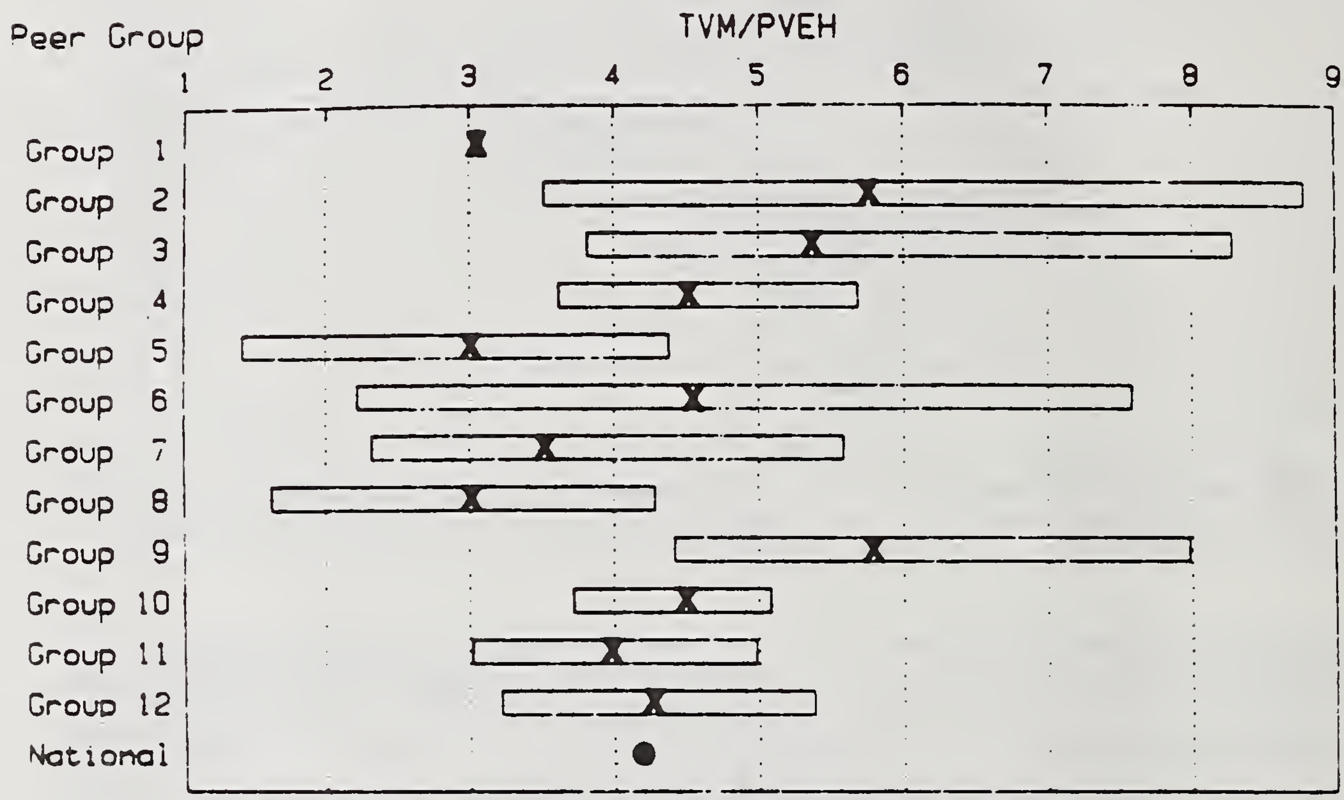
Labor Efficiency

The pattern for labor efficiency is quite similar to that for cost efficiency; the very small systems (1 and 2) and the very large systems (11 and 12) show the lowest labor efficiency. Small to medium size systems, particularly in groups 3, 5 and 6, are the most efficient but each group has members that are quite inefficient. In fact, the poor performing members of those groups are less labor efficient than any of the members of the peer groups whose average efficiency is the lowest.

Unexpectedly, a high peak to base ratio does not result in lower labor efficiency. Group 8 has the highest peak to base ratio by far but exactly average labor efficiency. Group 5 also has a relatively high peak to base ratio but is the most labor efficient group.

Although there is some overlap among the peer groups on this measure, the pattern of average scores and distributions of values within peer groups demonstrate the usefulness of this measure.

FIGURE 4-17. THE RANGE FOR VEHICLE EFFICIENCY BY PEER GROUP



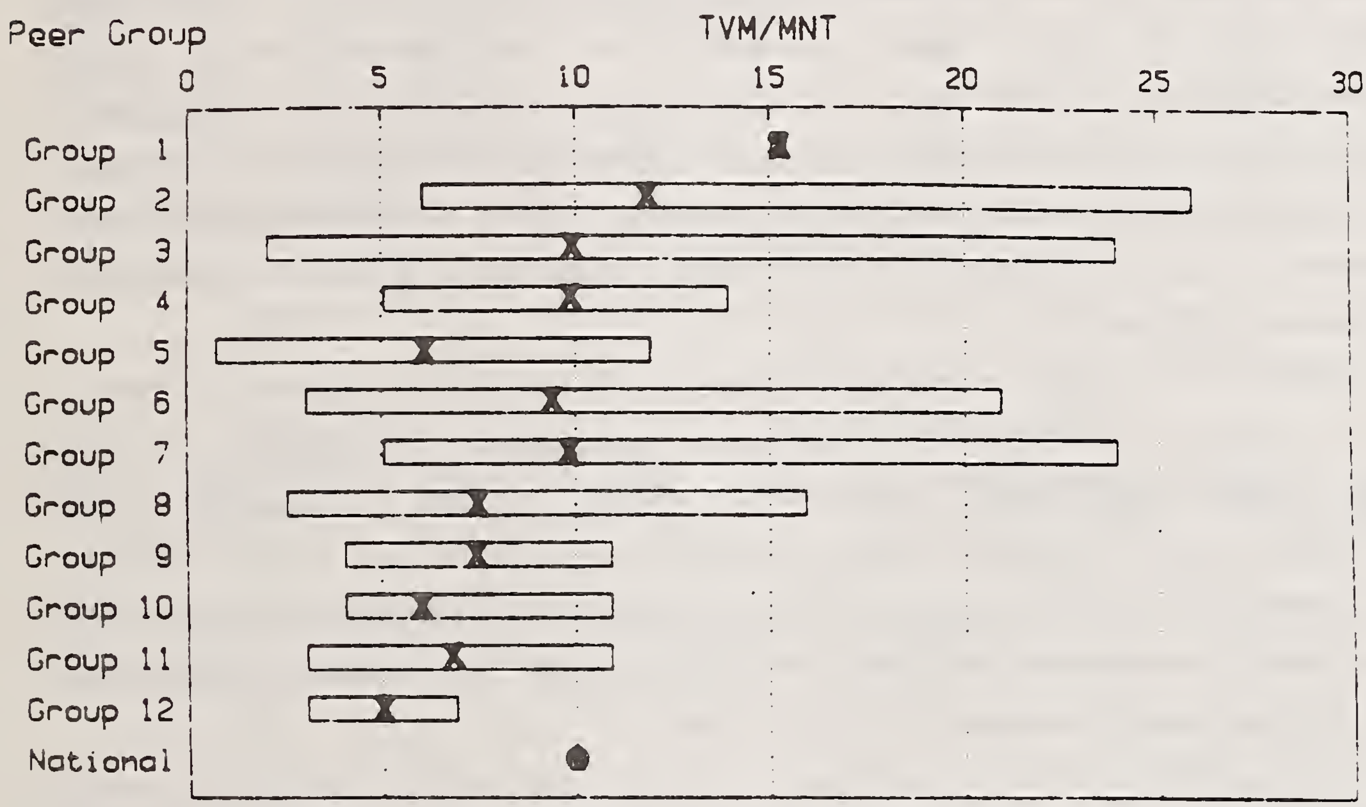
Vehicle Efficiency = Total Vehicle Miles (10,000's) per Peak Vehicle.

X on bar shows peer group mean.
X in circle shows national mean.

Vehicle Efficiency

Different peer groups display quite different levels of vehicle efficiency. Peer Group 8, which is characterized by a high peak to base ratio, averages very low on vehicle efficiency. This is not surprising considering that with an average peak to base ratio of nearly 3, only a third of the buses employed by these systems are in use for more than a few hours a day. Other peer groups with relatively high peak to base ratios (11 and 7) also display relatively low vehicle efficiency. Peer Group 5 also has low vehicle efficiency but as a consequence of its very slow speed. In this peer group, most buses are in use throughout the day but they generate relatively few miles each.

FIGURE 4-18. THE RANGE FOR MAINTENANCE EFFICIENCY BY PEER GROUP



Maintenance Efficiency = Total Vehicle Miles (10,000's) per Maintenance Employee.

X on bar shows peer group mean.
X in circle shows national mean.

The peer groups with good vehicle efficiency get nearly twice as many miles per peak vehicle as the low performers. Groups 2, 7 and 9 stand out as having good performance. Not only do they average high vehicle efficiency, their poorest performing members tend to do better than the average members of all other peer groups.

This measure captures major distinctions between peer groups since peer groups vary markedly on both average performance and range.

Maintenance Efficiency

Maintenance efficiency is notable for its great range. The values range from 10,000 total vehicle miles to 250,000 total vehicle miles per maintenance employee. The peer groups with smaller transit systems (2,3,6,7) are most likely to encompass a wide range of values. Since maintenance can be done in-house by transit employees or as an outside service, caution must be used when comparing systems on this measure. The peer groups of smaller systems show the greatest diversity on this measure because small systems are less likely to keep a full maintenance staff. However, for larger systems this performance indicator is a reliable measure of maintenance efficiency.

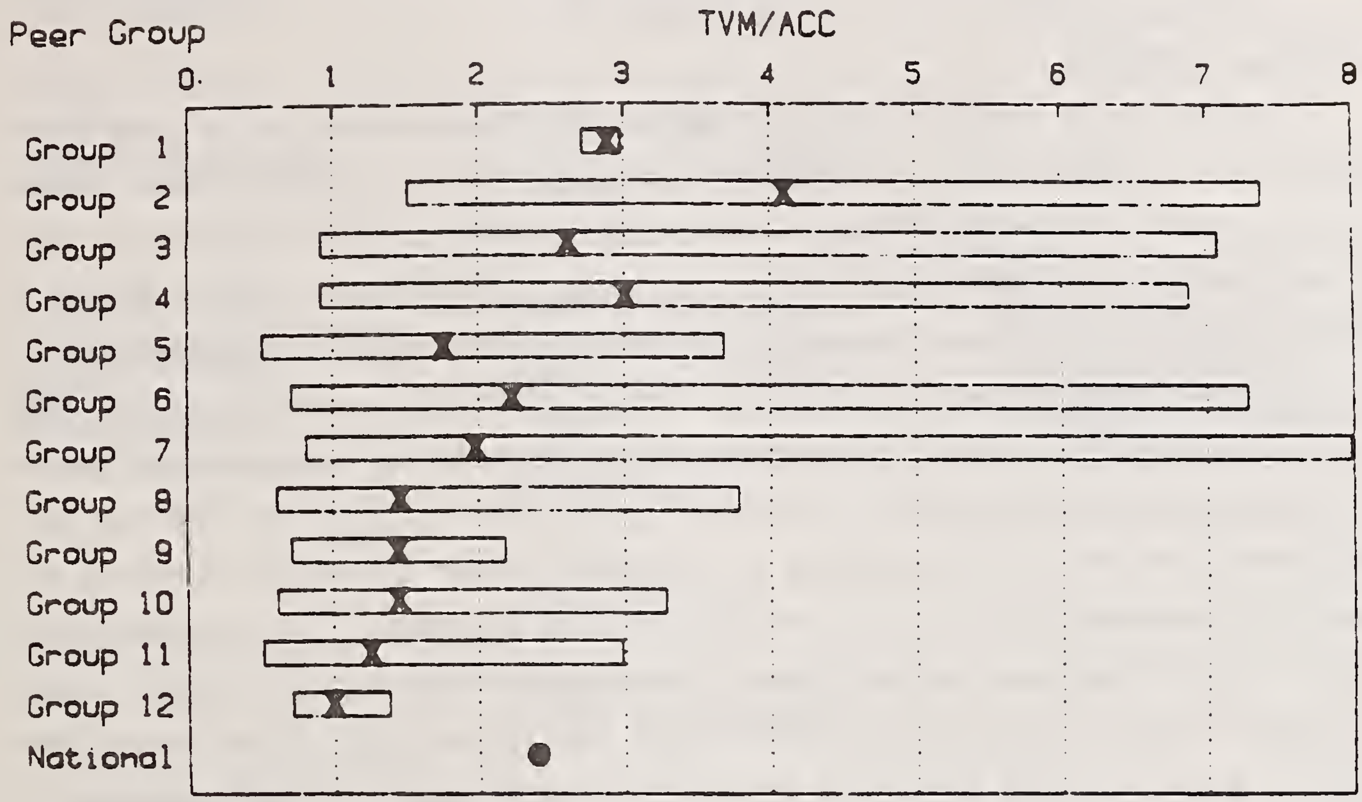
Although the larger system peer groups (8-12) have lower average efficiency, this is in part a consequence of different maintenance arrangements as noted above. Several peer groups with smaller systems (3 and 5) have the lowest performing systems suggesting that in-house maintenance can be less efficient under some circumstances. Peer Group 5, which is the slowest peer group, also shows lower maintenance efficiency--in part because the vehicles travel fewer miles relative to service hours.

Safety

Like maintenance efficiency, safety shows great variation in some peer groups and great consistency in others. The peer groups of smaller companies, 2-7, tend to have greater safety, even when extreme outlying cases are eliminated. Peer Group 2 is the safest--probably because it consists of small companies operating in relatively uncongested areas as shown by their high average operating speed. Group 1, the other group of fast systems, also shows superior safety.

Peer groups of larger systems (9-11) operate in denser urban areas and tend to have a relatively high proportion of their vehicles operating during the congested

FIGURE 4-19. THE RANGE FOR SAFETY BY PEER GROUP



Safety = Total Vehicle Miles (10,000's) per Accident.

X on bar shows peer group mean.
X in circle shows national mean.

peak hours, and thus have the lowest safety. Slow systems (5) and those operating most during congested peak hours (8) also show lower performance.

This safety measure must be used with some caution, particularly for peer groups with extreme values, because it is likely that not all systems used the same definitions of accidents. Despite this caution, safety measures do seem to reflect patterned differences between peer groups.

THE RELATION OF OPERATING AND PERFORMANCE VARIABLES

The preceding sections have implicitly linked certain operating characteristics to specific aspects of performance. For instance, it was noted that peer groups of large systems have the lowest cost efficiency. This section of the chapter will examine the relations between operating and performance variables in a more systematic way. First, each performance indicator will be examined to discover which operating characteristics can affect performance on that measure. Then a brief analysis will be done to show that operating characteristics help shape the performance profile of each peer group across the seven performance measures.

Correlation Analysis

Although it is easy to link one operating characteristic at a time to a performance indicator, this is an overly simplistic view of how the peer groups actually relate to the operating characteristics. Most of the peer groups were formed through the interaction of the operating characteristics. Thus, except in extreme cases, several characteristics of a transit system must be considered before it is clear which peer group it belongs to. Thus, it is not totally accurate to say that it is the largeness of transit systems that influence performance in such and such a way, because the performance of the peer groups also varies in accordance with their speed and peak to base ratios as well. To further complicate matters, the operating characteristics are not independent of one another. Large systems also tend to have higher peak to base ratios. So it may look like size is a crucial factor in a specific kind of performance when in fact the crucial factor is the peak to base ratio, or even that both large size and a high peak to base ratio are necessary to produce a particular kind of performance.

A series of simple and multiple regression correlation analyses were performed to disentangle the effects of the four operating characteristics on the seven performance indicators (see technical details in Appendix H). Table 4-3 shows an overview of the correlation results. Under each performance indicator, the operating variables that significantly and independently correlate with it are listed. They are given in the order of their relative importance.

Further work needs to be done on the complex interrelations shown in Table 4-3, but some tentative conclusions can be drawn. The general results of one multiple regression analysis are shown on Table 4-3 and discussed below. Correlation coefficients and beta weights are not shown since these vary between analyses.

As shown in the table, each operating characteristic contributes to differences in performance in significant ways. Each appears to be important for several different aspects of performance as well. The size of a transit system appears to be its most important characteristic. The number of peak vehicles not only affects performance on six of the performance indicators, it is the most important for four of them. Total vehicle miles, another measure of size, also gives an independent

TABLE 4-3. OPERATING VARIABLES THAT CORRELATE WITH PERFORMANCE INDICATORS

Cost Efficiency (RVH/OEXP)	Service Utilization (TPAS/RVH)	Revenue Generation (OREV/OEXP)
Peak Vehicles	Peak Vehicles	None
Speed	Peak to Base Ratio	
Total Vehicle Miles	Total Vehicle Miles	
	Speed	

Labor Efficiency (TVH/EMP)	Vehicle Efficiency (TVM/PVEH)	Maintenance Efficiency (TVM/MNT)
Speed	Peak to Base Ratio	Peak Vehicles
Peak Vehicles	Speed	Total Vehicle Miles
Total Vehicle Miles	Total Vehicle Miles	
	Peak Vehicles	

Safety (TVM/ACC)
Peak Vehicles
Total Vehicle Miles

contribution to six performance indicators. However, its contribution is relatively small in each case.

Speed also has major importance. It contributes to performance on four of the seven measures and is most important for labor efficiency. Peak to base ratio appears to have the least importance. It appears for only two of the performance indicators. However, it is the most important operating characteristic in relation to vehicle efficiency. The results for speed and peak to base ratio may be underestimated in these results, because in other analyses they are much more important. However, across all analyses they were found to have some importance and cannot be overlooked when analyzing how operating conditions affect performance.

The type of correlational analysis done here cannot capture other apparent features of the relationship between operating characteristics and performance. In some instances the size of a transit system has a negative effect on performance at both large and small extremes of system size. These results do not adequately assess that relationship. Also, there appear to be threshold effects. Very high speed, as exhibited in Peer Group 1, seems to affect performance, but below this level it is much less important. The correlation analysis cannot capture this type of situation either. However, earlier discussions of patterns in the data point out these relationships in the data and they appear to be valid.

Notably, revenue generation is not significantly predicted by any of the operating characteristics. This is in accord with the patterns shown in Figure 4-15. The averages for peer groups are scattered across the figure. In addition, each peer grop covers a broad range of values. For instance, in Peer Group 5, the highest system generates 10 times as much proportional revenue from the fare box than the lowest member of the group. This result is not surprising since fare levels are often set by policy makers outside the transit system. A special local sales tax will mandate lower fares while most private bus companies do not have equivalent access to many subsidies and must generate most of their money from operations.

<u>Profile Analysis</u>

Taken one at a time, the performance indicators do show significant relationships to operating characteristics. In addition, the patterns across all seven performance indicators further confirm the individual variable findings. To demonstrate this, the performance profiles of selected peer groups were compared to show how operating statistics relate to overall patterns of performance. Figure 4-20A compares the performance profiles of the three peer groups of largest systems (10, 11. and 12). Figure 4-20B compares the performance profiles of three peer groups of smaller systems (2,5,6). The profiles for the large systems are similar--the strengths and weaknesses of each system are the same although the magnitudes vary. The same holds true for the small groups. In addition, the patterns for large and small systems are the inverse of each other. While large groups are strong in service utilization (TPAS/RVH) and Revenue Generation (OREV/OEXP), the small systems are weak. The small systems are above average in Vehicle Efficiency (TVM/PVEH) and Maintenance Efficiency (TVM/MNT) and the large systems are relatively lower in these areas.

The next set of graphs, Figures 4-21A and 4-21B, portray a comparison of relatively slow and fast groups. The overall profiles are much less similar than the size comparison. However, it can be seen that the fast groups are most similar in their high vehicle efficiency (TVM/PVEH) and safety (TVM/ACC) and the slow groups are relatively lower on those attributes.

The last comparison, Figures 4-22A and 4-22B, shows the peer groups with high and low peak to base ratios. As with speed, these profiles are similar only in

FIGURE 4-20 A. PERFORMANCE PROFILES OF LARGE TRANSIT SYSTEMS

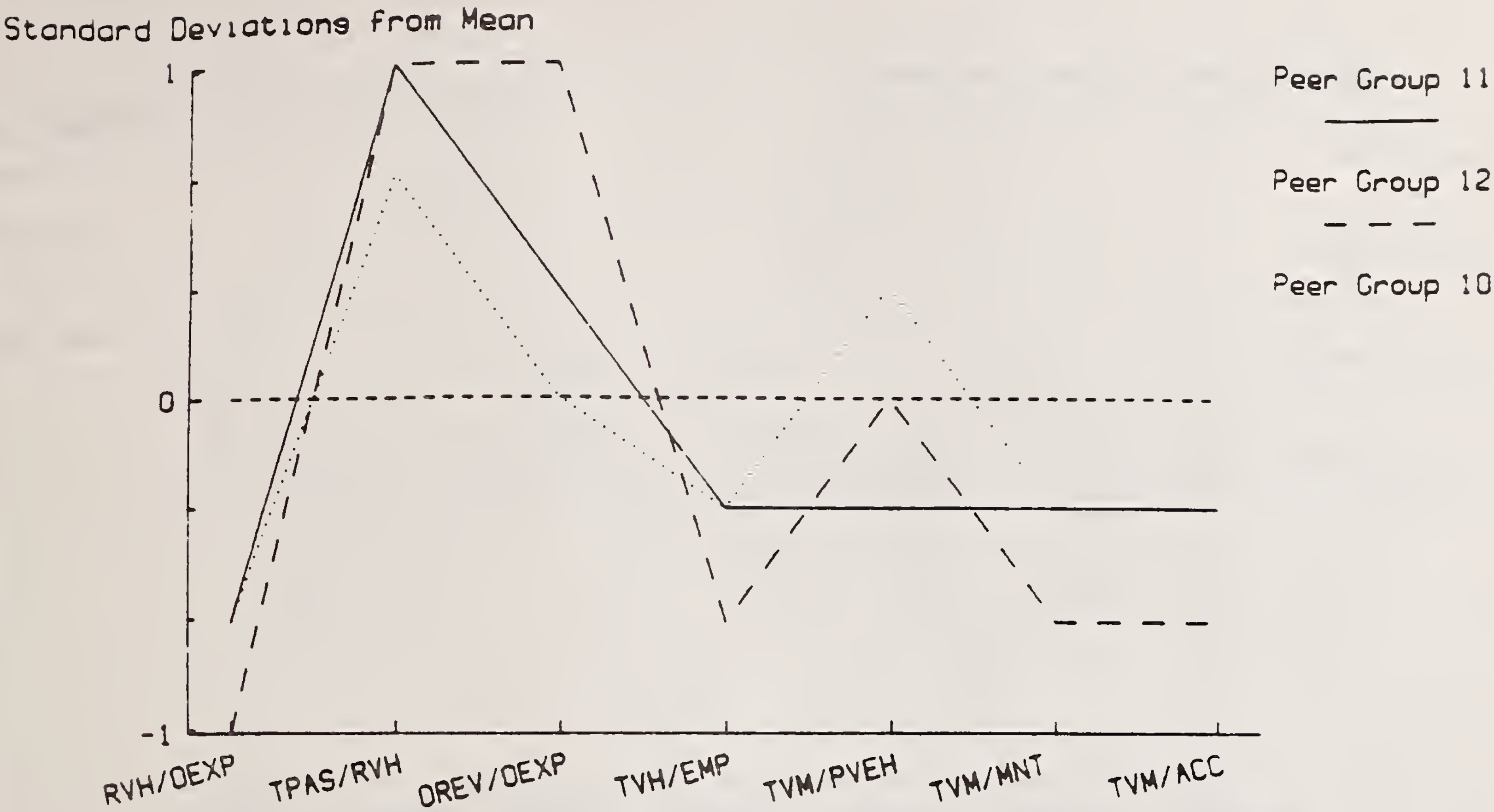
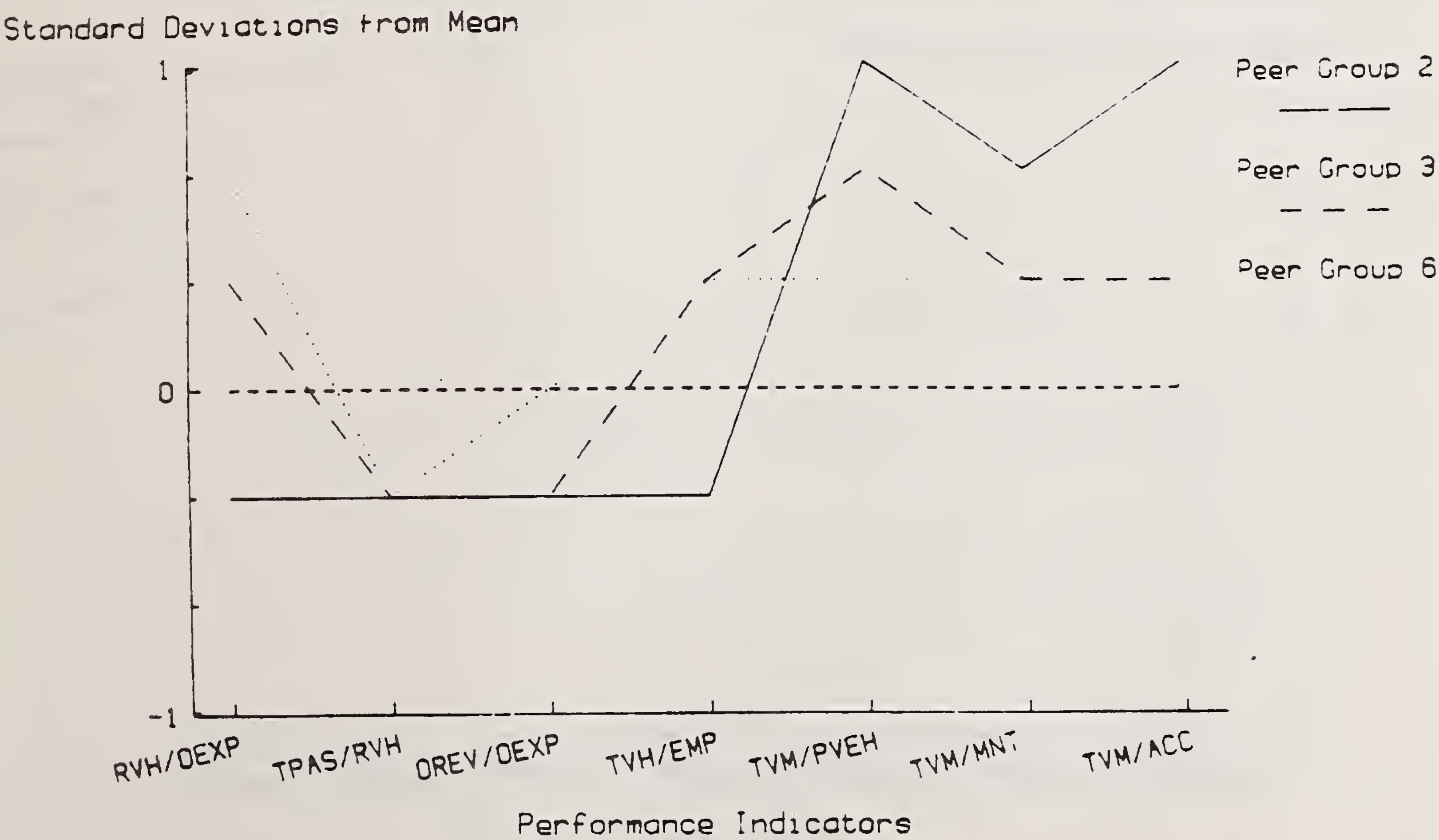


FIGURE 4-20 B. PERFORMANCE PROFILES OF SMALL TRANSIT SYSTEMS



Zero on vertical axis equals national mean.

FIGURE 4-21A. PERFORMANCE PROFILES OF FAST TRANSIT SYSTEMS

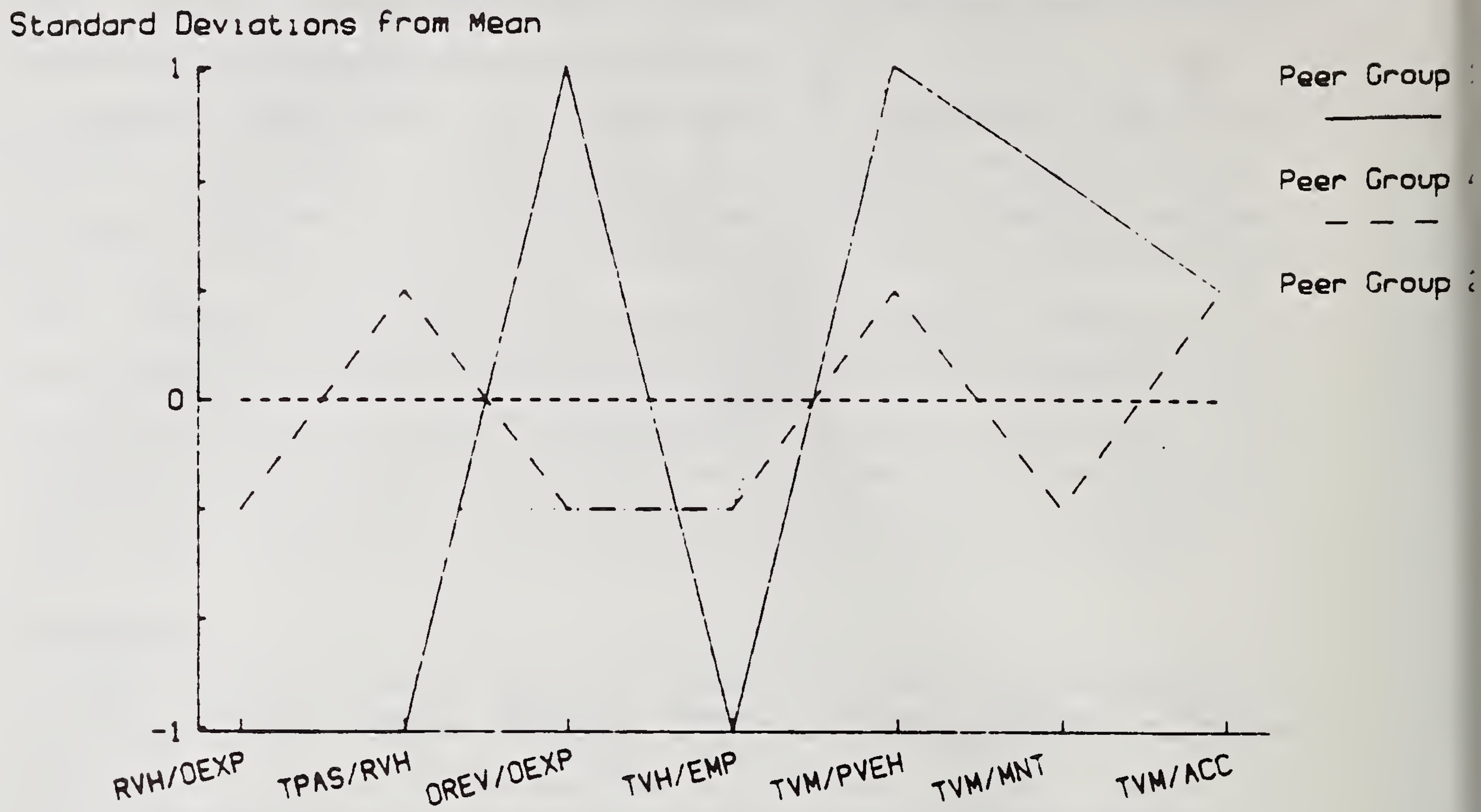
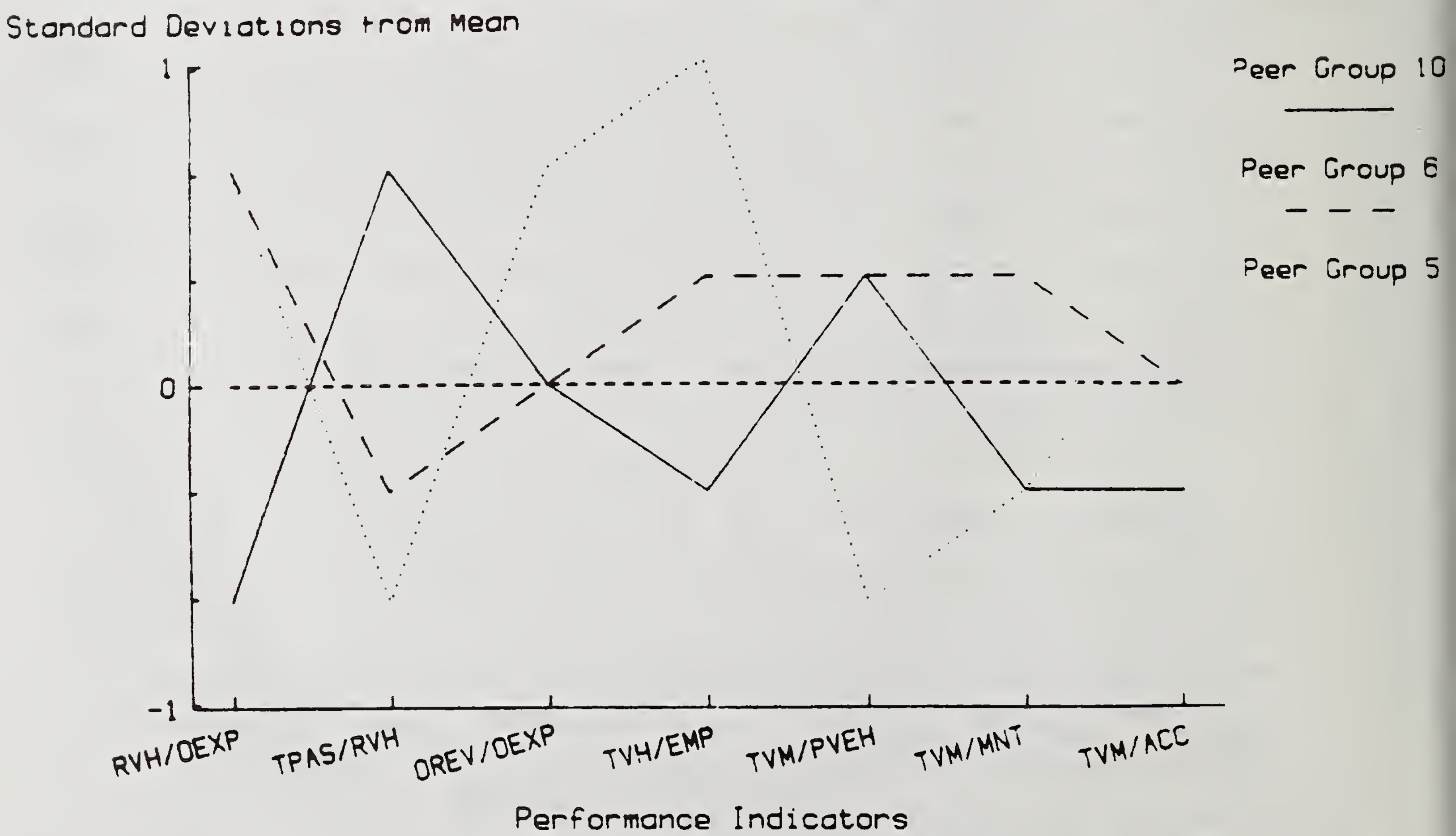


FIGURE 4-21 B. PERFORMANCE PROFILES OF SLOW TRANSIT SYSTEMS



Zero on vertical axis equals national mean.

FIGURE 4-22A. PERFORMANCE PROFILES OF HIGH PEAK TO BASE TRANSIT SYSTEMS

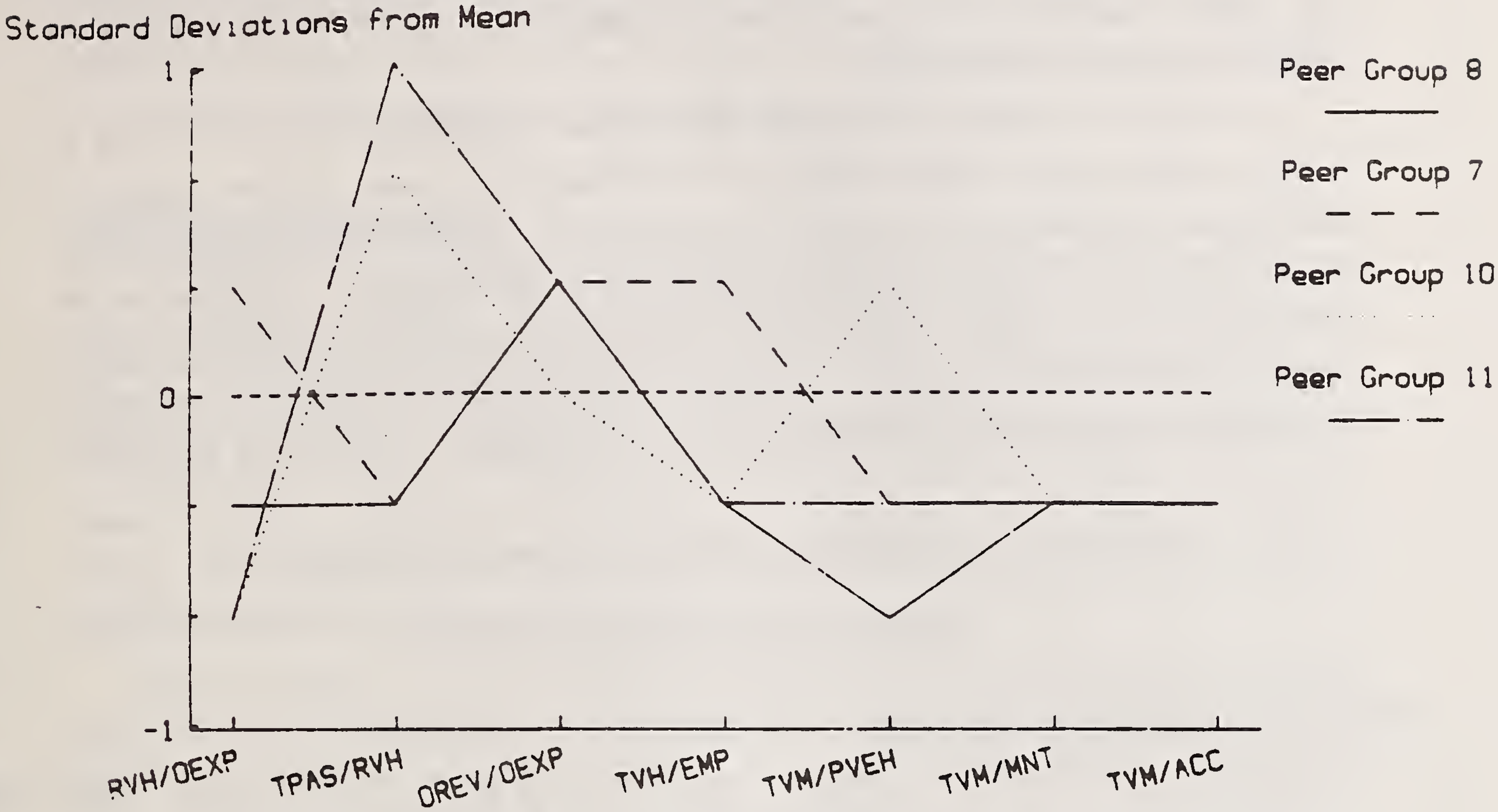
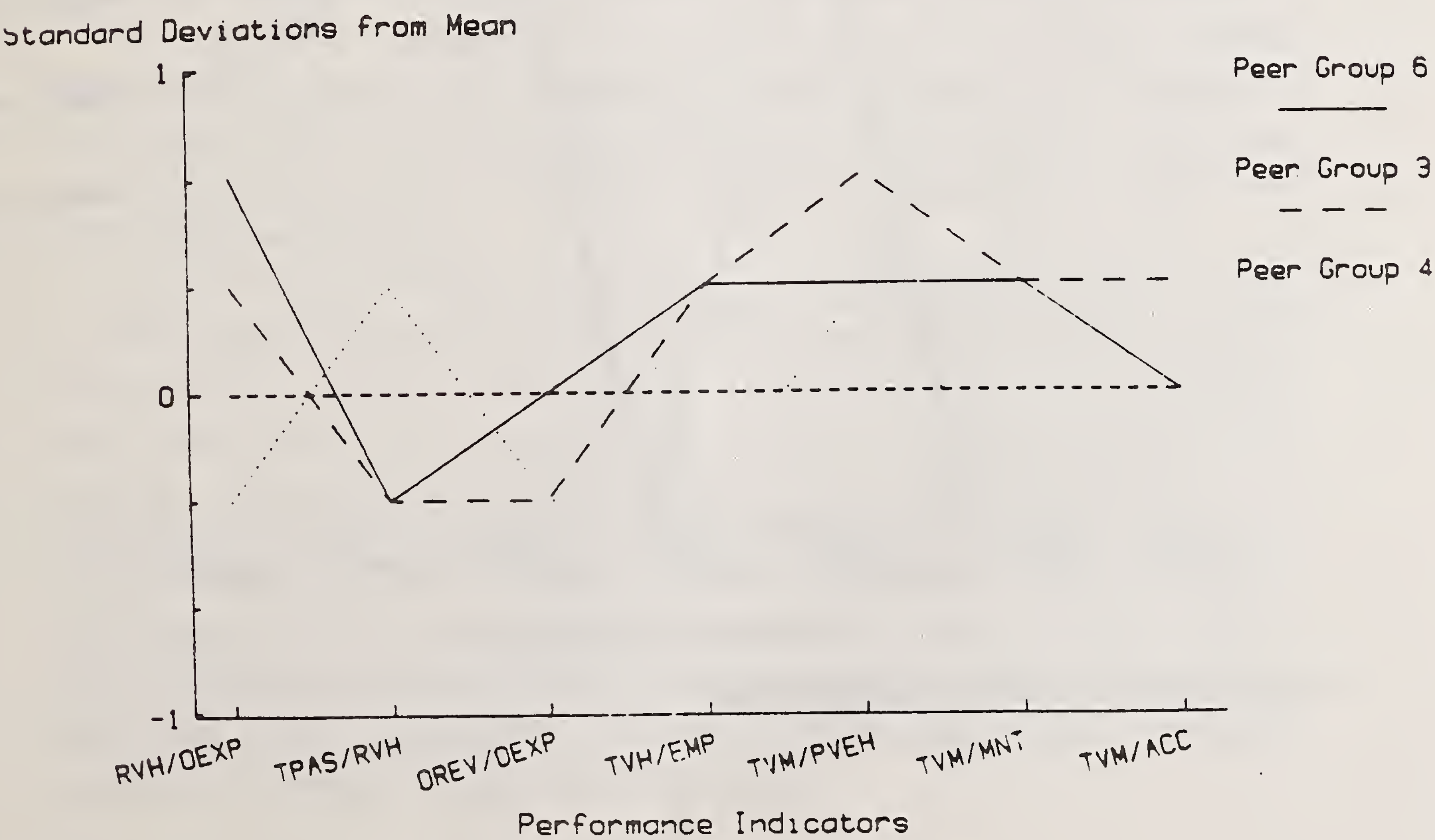


FIGURE 4-22B. PERFORMANCE PROFILES OF LOW PEAK TO BASE TRANSIT SYSTEMS



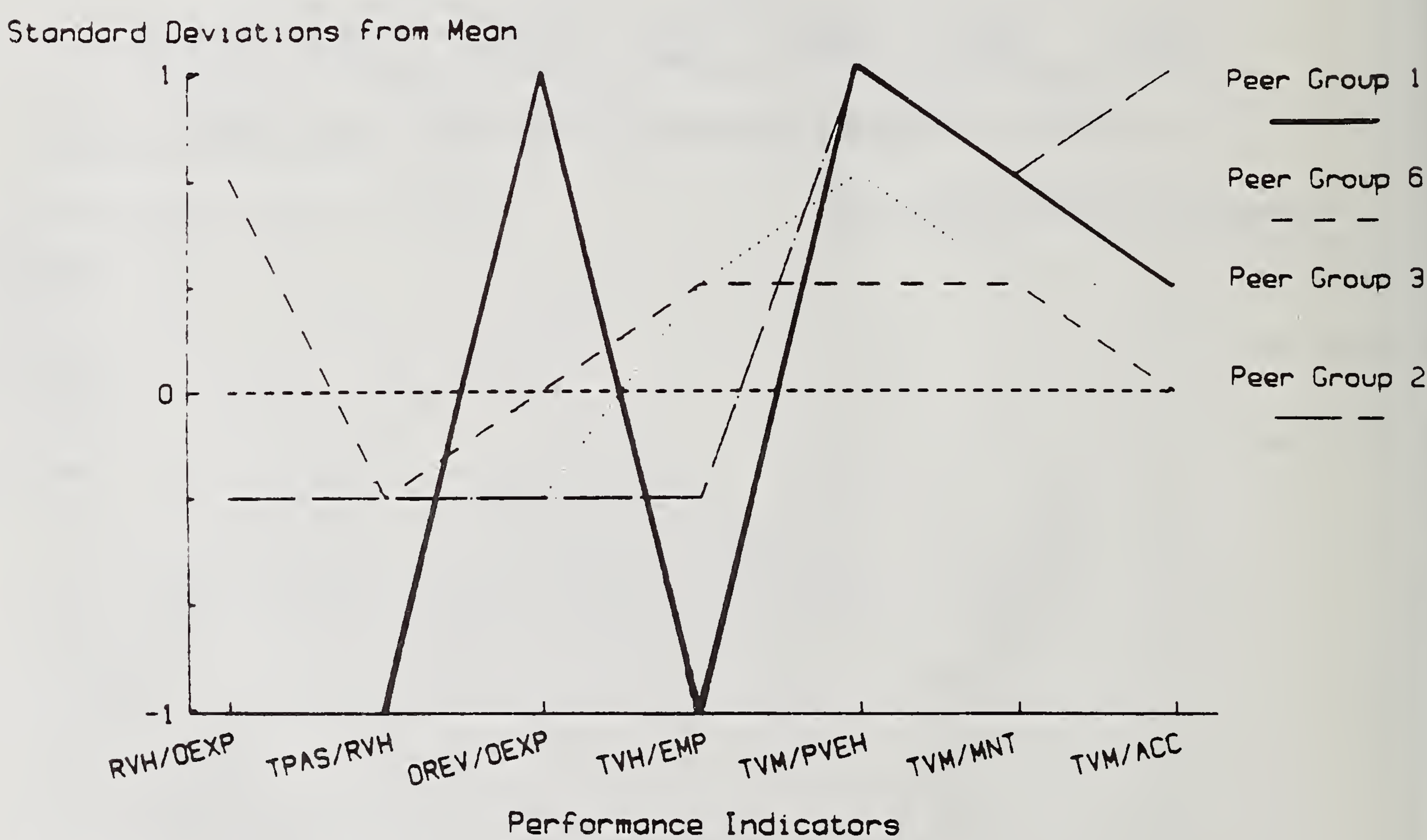
Zero on vertical axis equals national mean.

specific ways. The high peak to base ratio groups have above average revenue generation (OREV/OEXP) and generally low vehicle efficiency (TVM/PVEH). The low peak to base groups are low on revenue generation (OREV/OEXP) and high on vehicle efficiency (TVM/PVEH).

The results of these comparisons confirm the correlation results. The size of transit systems has a broad impact on performance, and peak to base ratios and speed make less general, more specific contributions to performance. For clarity's sake only some of the peer groups have been shown on each chart. The peer groups with average size, peak to base ratio or speed fall somewhere between the clear patterns shown for these figures. Certain exceptions also show the more important influence of speed or peak to base ratios at the extremes. Figure 4-23 shows Peer

FIGURE 4-23. PERFORMANCE PROFILES TO COMPARE PEER GROUP 1

TO OTHER SMALL SYSTEM PEER GROUPS



Zero on vertical axis equals national mean.

Group 1, the peer group with the smallest average size in comparison with the other small peer groups from Figure 4-20B. Peer Group 1's profile of performance differs radically from the other small groups on the first four performance indicators. This confirms the earlier finding that performance is a consequence of the joint impact of the operating characteristics of bus companies.

<u>Aggregation of Peer Groups</u>

The similarity of performance profiles, particularly the similarities based on size comparisons as shown in Figures 4-20A and 4-20B, superficially suggest that some peer groups may be so similar that they could form one peer group. However they are similar only in comparison to the rest of the nation. A close examination of Figures 4-13 through 4-19 reveals that no two peer groups share both the same mean and range of values on any performance indicator. Thus within its original peer group a transit system could be performing well above the mean, but in an aggregated peer group it could be operating below the mean.

Also, as stated in the final section of Chapter 3, it is more meaningful to assess how well transit management is doing relative to peers who begin with the same operating characteristics, not relative to systems which are similar only in performance. Aggregating peer groups on the basis of a similar performance profile would undermine the use of this system of performance evaluation as a tool for managers.

However, some aggregation of peer groups can be done through a slightly different interpretation of the cluster analysis based upon operating characteristics. Peer groups 3 and 4 are most similar and could form one peer group. These peer groups differ only in terms of their average speed, 16.2 mph for Peer Group 4 and 14.5 mph for Peer Group 3.

An even larger aggregated peer group can be formed from Peer Groups 3, 4 and 6. Once again, these peer groups differ mostly on speed, ranging from 16.2 mph for Peer Group 4 to 12.2 mph for Peer Group 6. At this point, however, the aggregated peer group would have 96 members--over a third of the transit systems in the analysis. This aggregated peer group would also begin to reduce the importance of speed as an operating characteristic because this large peer group would cover almost the entire range of speed. The performance profiles of Peer Groups 3 and 4 also differ substantially because of their differences in speed. As major performance differences would be lost through this aggregation, it is not recommended that these peer groups be aggregated except for specific purposes where differences caused by variations in average speed are irrelevant.

Based upon the cluster analysis, there are no other peer groups that can be legitimately joined together without violating important differences in their operating characteristics.

USES OF PERFORMANCE EVALUATION

Managerial Uses

The performance indicators and peer groups have primarily been established for the use of transit managers. In conjunction, they form a diagnostic tool with which managers can pinpoint problem areas within their own transit operation.

Peer groups create sets of transit systems with similar operating conditions that constrain performance. To date most peer comparisons have been limited to within state comparisons.[8] When national sets have been created, they have often resulted in comparisons between systems with dissimilar operating conditions.[9] Now a consistent set of peer groups has been created and methods outlined for placing systems with each group.

For each peer group, a set of norms has been established on each performance indicator. These norms are the means and standard deviations listed in Appendix G. Using these statistics, a transit manager can review performance of his/her transit system, and determine whether it has been functioning above or below the mean. If the system is more than one standard deviation above the mean, then that system is in the top (roughly) 15% of its peer group.

Of course, knowing a system's problem areas within the seven dimensions of performance does not automatically reveal the underlying causes. For instance, a low labor efficiency score can mean many things. It could mean that drivers are inefficiently scheduled to cover the transition from the peak to base period. Or it could mean that there are too many maintenance personnel relative to the size of the fleet.

---

[8] James M. Holec, Dianne S. Schwager, and Angel Fandalian, Use of federal Section 15 data in transit performance: Michigan program. Transportation Research Record No. 765. (Washington, D.C.: Transportation Research Board, 1980), pp. 36-38.

[9] ATE Management Co., Inc., Phase I report to the Los Angeles County Transportation Commission on issues related to performance audits of major Los Angeles County bus operators. (Los Angeles: Los Angeles County Transportation Commission, 1978), p. 79.

In addition, a system's performance should be judged in relation to its own objectives and policies. A transit system which wishes to optimize the reliability and comfort of bus service might choose to be less maintenance efficient in pursuit of these alternate goals.

The peer groups and performance indicators also give transit managers a way of evaluating their performance over time. Secular trends such as inflation or a nationwide increase in ridership because of a gas crisis, make it hard to directly compare performance from year to year. However, by comparing a system's standing in relation to the other members of a peer group over time corrects for many of the external changes.

Other research has shown that different kinds of transit systems change in different ways over time.[10] Peer groups allow comparison of transit systems that are expected to change in similar ways.

Research Uses

This system of performance evaluation provides a useful research tool in several ways. The key performance indicators provide a non-arbitrary set of performance measures whose validity and reliability have been established.

The peer groups are useful in research because they control for basic differences between types of transit systems. For instance, a study comparing the relative benefits of hiring private management firms would find invalid results if private management was used only by small and comparatively efficient systems. Making such comparisons within peer groups provides a clearer picture of the special benefits or problems with private management.

The peer groups are also useful for research comparison over time. For instance, it is useful to look at whether different types of transit systems respond to government assistance in different ways. The results from recent studies on the effects of subsidies in transit would have been more convincing had the longitudinal studies been selected from the same peer group.[11]

---

[10] Leland C. Barbour and Robert J. Zerillo, Transit Performance in New York State. (Albany: New York Department of Transportation, 1981.)

[11] John Pucher, Anders Markstedt and Ira Hirschman, Impacts of subsidies on the costs of urban public transit. Journal of Transportation Economics and Policy, 1983, 17(2), 155-176.

Future Applications

This system of evaluation for transit performance is being taught to transit managers in a series of workshops, as noted in the introduction to this report. An instructional manual is being prepared which will explain the performance indicators and peer groups in terms of their use in performance evaluation using simple statistical and microcomputer applications. Case studies present examples of what can be learned through performance evaluation.

CONCLUSION

This chapter has integrated the results of Chapters One through Three in a number of ways. Each peer group was shown to have a distinctive performance profile across the seven performance indicators. When each performance indicator is examined in detail, the peer groups are shown to differ significantly not only in their average performance but also in the distribution of values on each indicator variable.

The operating characteristics relate to performance in complex ways. The size of a transit system is the most important of the operating characteristics, being of major importance in explaining variation on four performance indicators. Speed and peak to base ratio are also of importance for differences on the three other performance indicators. And each operating characteristic contributes in small, but important ways, to differences in performance on indicators where they are not the most important influence.

BIBLIOGRAPHY

American Public Transit Association. Operating statistics report 1980: Transit system operating statistics for calendar/fiscal year 1979. Washington, D.C.: American Public Transit Association, October 1980.

American Public Transit Association. Operating statistics report 1981: Transit system operating statistics for calendar/fiscal year 1980. Washington, D.C.: American Public Transit Association, October 1981.

Anderson, Shirley C. and Fielding, Gordon J. Comparative analysis of transit performance. Final report No. UMTA-CA-11-0020-1. Irvine, California: University of California, Institute of Transportation Studies, January 1982. (NTIS No. PB 82-196478.)

ATE Management Co., Inc. Phase I report to the Los Angeles County transportation commission on issues related to performance audits of major Los Angeles County bus operators. Los Angeles, California: Los Angeles County Transportation Commission, 1978.

Barbour, L. C. and Zerrillo, R. J. Transit Performance in New York State. Albany, NY: New York State Department of Transportation, 1981.

California. Business and Transportation Agency. Transportation Development Act: Statutes as amended and related sections of the California Administrative Code as adopted by the Secretary of the Business and Transportation Agency. Report No. DMT-032. Sacramento, California: California Department of Transportation, Division of Mass Transportation, February 1978.

California. Office of the Controller. Financial transactions concerning transit operators and non-transit claimants under Transportation Development Act: Annual report for fiscal year 1980-1981. Sacramento, California: State of California, Office of the Controller, 1982.

Carmines, E. G. and Zeller, R. A. Reliability and validity assessment. Beverly Hills, California: Sage Publications, 1979.

Comrey, A. L. A first course in factor analysis. New York: Academic Press, 1973.

Dixon, W. J., ed. BMDP statistical software 1981. Los Angeles: University of California Press, 1981.

Everitt, B. E. Cluster analysis. London: Heinemann, 1980.

Fielding, Gordon J., Glauthier, Roy E. and Lave, Charles A. Development of performance indicators for transit. Final report No. UMTA-CA-11-0014-78-1. Irvine, California: University of California, Institute of Transportation Studies, December 1977. (NITS No. PB 278-678.)

Fielding, Gordon J., Glauthier, Roy E. and Lave, Charles A. Performance indicators for transit management. Transportation, 1978, 7(4), pp. 365-379.

Fielding, Gordon J. Changing objectives for American transit. Part I 1950-1980. Transport Reviews, 1983, 3(3), pp. 287-299.

Fielding, Gordon J. Changing objectives for American transit. Part 2. Management's reponse to hard times. Transport Reviews, 1983, 3(4), pp. 341-362.

Fielding, Gordon J. and Faust, Katherine. Dimensions of bus performance for peer groups of transit agencies in fiscal years 1980 and 1981 using Section 15 data. (Working Paper No. 83-5.) Irvine, California: University of California, Institute of Transportation Studies, December 1983.

Fielding, Gordon J., Brenner, Mary E. and de la Rocha, Olivia. Using Section 15 data for transit performance analysis. Interim report No. UMTA-CA-11-0026-1. Irvine, California: University of California, Institute of Transportation Studies, January 1983.

Fielding, Gordon J., Brenner, Mary E. and de la Rocha, Olivia. Using Section 15 data: Adapting and evaluating the magnetic tape version for statistical analysis. (Working Paper No. 83-6.) Irvine, California: University of California, Institute of Transportation Studies, Jaunary 1983.

Giuliano, Genevieve. The effect of environmental factors on the efficiency of public transit service. Transportation Research Record No. 797. Washington, D.C.: Transportation Research Board, 1981, pp. 11-16.

Harmon, H. H. Properties of different types of factor solutions. Modern factor analysis. Chicago: University of Chicago Press, 1967, pp. 93-109.

Holec, James M., Schwager, Dianne S. and Fandalian, Angel. Use of federal Section 15 data in transit performance: Michigan program. Transportation Research Record No. 746. Washington, D.C.: Transportation Research Board, 1980, pp. 36-38.

Iowa Department of Transportation. Uniform data management system: System development and testing. Report No. DOT-I-81-2. Ames, Iowa: Iowa Department of Transportation, October 1980.

Kruskal, J. B. Transformations of data. International encyclopedia of the social sciences, Vol. 15. David Sills, Ed. New York: MacMillan Co., 1968, pp. 182-193.

McCrosson, Dennis F. Choosing performance indicators for small transit systems. Transportation Engineering, March 1978, 48(3), pp. 26-30.

Miller, David R. Differences among cities, differences among firms, and costs of urban bus transport. Journal of Industrial Economics, November 1970, 19(1), pp. 22-32.

Miller, James H. The use of performance-based methodologies for the allocation of transit operating funds. Traffic Quarterly, October 1980, 34(4), pp. 555-574.

Miller, James H. and Rea, John C. Comparison of cost models for urban transit. Highway Research Record, No. 435, Washington, D.C.: Transportation Research Board, 1973, pp. 11-19.

Nelson, Gary R. An econometric model of urban bus transit operations. Unpublished Ph.D. dissertation. Rice University, 1972. Available from University Microfilms International as No. 72-26457.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. and Bent, D. H. SPSS: statistical package for the social sciences. New York: McGraw-Hill, 1975.

Oram, Richard L. Peak period supplements: The contemporary economics of urban bus transport in the U.K. and U.S.A. Progress in Planning, 1979, 12(2), pp. 81-154.

Ostrom, Elinor et al. Community organization and the provision of police services. Beverly Hills, California: Sage Professional Papers in Administrative and Policy Studies, 1973.

Patton, Timothy A. Transit performance indicators. Transportation Systems Center Staff Study #SS-67-U.3-01. Cambridge, Mass.: U.S. Department of Transportation, Transportation Systems Center, 1983.

Post, Buckley, Schuh and Jernigan, Inc. Florida transit system performance measures and standards. Tallahasse, FL: Florida Department of Transportation, Public Transportation Operations Division, 1979.

Pucher, John, Markstedt, Anders and Hirschman, Ira. Impacts of subsidies on the costs of urban public transit. Journal of Transportation Economics and Policy, 1983, 17(2), pp. 155-176.

Rivlin, Alice M. Systematic thinking for social action. Washington, D.C.: Brookings Institution, 1971.

Sinha, Kuares C., Jukins, David P. and Bevilacqua, Oreste E. Stratification approach to evaluation of urban transit performance. Transportation Research Record No. 761. Washington, D.C.: Transportation Research Board, 1980, pp. 20-27.

U.S. Department of Transportation, Transportation Systems Center. National urban mass transportation statistics: 1981 Section 15 report. Report No. UMTA-MA-06-0107-83-1. Springfield, Virginia: National Technical Information Service, November 1982.

U.S. Department of Transportation, Transportation Systems Center. National urban mass transportation statistics: First annual report Section 15 reporting system: Transit financial and operating data reported for fiscal years ending between July 1, 1978 and June 30, 1979. Report No. UMTA-MA-06-0107-81-1. Washington, D.C.: U.S. Government Printing Office, May 1981.

U.S. Department of Transportation, Urban Mass Transportation Administration. National urban mass transportation statistics: Second annual report, Section 15 reporting system. Report No. UMTA-MA-06-0107-82-1. Washington, D.C.: U.S. Department of Transportation, July 1982.

Vaziri, M. and Deacon, J. A. Application of Section 15 and census data-to transit decision making. Final report No. UMTA-KY-11-0002-83. Springfield, Virginia: National Technical Information Service, 1983.

APPENDIX A

SUMMARY OF TSC DATA FILES AND VARIABLES REORGANIZED

TSC FILE NAME	VARIABLES
NRSDWK*	Number of Vehicles in Operation
	Total Vehicle Miles
	Total Vehicle Hours
	Total Revenue Miles
	Total Revenue Hours
	Revenue Capacity Miles
	Unlinked Passenger Trips
	Unlinked Passenger Miles
	Average Time per Unlinked Trip
	Vehicle Operators Full-time
	Vehicle Operators Part-time
	Total Service Persons
MOPRTN	Number of Revenue Vehicles
EMPSCH	Total Employees Operating Labor
	Total Employees Capital Labor
	Transportation Executive, Professional and Supervisory Personnel Operating Labor
	Transportation Executive, Professional and Supervisory Personnel Capital Labor
	Transportation Support Personnel Operating Labor
	Transportation Support Personnel Capital Labor
	Revenue Vehicle Operators Operating Labor
	Revenue Vehicle Operators Capital Labor
	Maintenance Executive, Professional and Supervisory Personnel Operating Labor
	Maintenance Executive, Professional and Supervisory Personnel Capital Labor
	Maintenance Support Personnel Operating Labor
	Maintenance Support Personnel Capital Labor
	Revenue Vehicle Maintenance Mechanics Operating Labor
	Revenue Vehilce Maintenance Mechanics Capital Labor
	Other Maintenance Mechanics Operating Labor
	Other Maintenance Mechanics Capital Labor
	Vehicle Servicing Personnel Operating Labor
	Vehicle Servicing Personnel Capital Labor
	General Administration Executive, Professional and Supervisory Personnel Operating Labor
	General Administrative Executive, Professional and Supervisory Personnel Capital Labor
	General Administration Support Personnel Operating Labor
	General Administration Support Personnel Capital Labor

*All variables originating in file NRSDWK had three versions, one each for a typical weekday, Saturday, and Sunday.

TSC FILE NAME	VARIABLES
OWSS	Total Operating Time Platform Time Line Service Total Nonoperating Paid Work Time
NRSTDY	Number of Vehicles in Operation AM Peak Number of Vehicles in Operation Midday Number of Vehicles in Operation PM Peak Total Vehicle Miles AM Peak Total Vehicle Miles Midday Total Vehicle Miles PM Peak Total Vehilce Revenue Miles AM Peak Total Vehicle Revenue Miles Midday Total Vehicle Revenue Miles PM Peak Total Vehicle Revenue Hours AM Peak Total Vehicle Revenue Hours Midday Total Vehicle Revenue Hours PM Peak Vehicle Operators Full Time AM Peak Vehicle Operators Full Time Midday Vehicle Operators Full Time PM Peak Vehicle Operators Full Time AM Peak Vehicle Operators Part Time PM Peak Vehicle Operators Part Time Midday Vehicle Operators Part Time PM Peak
BUSWAY	Directional Miles on Exclusive Right of Way Directional Miles on Controlled Access Right of Way Directional Miles on Mixed Traffic Right of Way
MNPENC	Total Roadcalls Kilowatt Hours of Propulsion Power Gallons of Diesel Fuel Gallons of Gasoline Gallons of LPG or LNG Gallons of Bunker Fuel
ACCDNT	Total Accidents (Computed)
UAREA	Urban Area Population
WESPSC	Total Hours of Operation Saturday Total Hours of Operation Sunday
REVSCH	Passenger Fares for Transit Service Special Transit Fares School Bus Service Revenues Freight Tariffs Charter Service Revenues Auxiliary Transportation Revenues Non-Transportation Revenues Taxes Levied by Transit System Local Cash Grants and Reimbursements

TSC FILE NAME	VARIABLES
REVSCH, continued	Local Special Fare Assistance Federal Cash Grants and Reimbursements Subsidy from other Sections of Operations Total Revenue
FGCA	Total Federal Assistance for Capital Revenue
NFGCA	State General Revenues State Dedicated Revenues State Total Assistance Local General Revenues Local Dedicated Revenues Local Total Assistance
FGRA	Total Federal Assistance for Operating Revenue
NFGRA	State General Revenues State Dedicated Revenues State Total Assistance Local General Revenues Local Dedicated Revenues Local Total Assistance
XDMOF/XTFO	Operators Sal and Wgs Veh Opr Other Sal and Wgs Veh Opr Fringe Benefits Veh Opr Services Veh Opr Operators Sal and Wgs Veh Maint Other Sal and Wgs Veh Maint Frings Benefits Veh Maint Services Veh Maint Operators Sal and Wgs Nonveh Maint Other Sal and Wgs Nonveh Maint Frings Benefits Non Veh Maint Services Non Veh Maint Operators Sal and Wgs Genl Admin Other Sal and Wgs Genl Admin Frings Benefits Genl Admin Services Genl Admin
XMFT/XF	Total Veh Operation Expense Total Veh Maintenance Expense Total Nonveh Maintenance Expense Total Genl Admin Expense
XO	Total Expenses for Published Reports
TRSYS	Total System Operating Expense from Form 301 Transit System ID Number Transit System name Single or Multimode

APPENDIX B

VARIABLES AND THEIR RESPECTIVE WEIGHTING FACTOR TO DISAGGREGATE MOTOR BUS STATISTICS

Variable	Weighting Factor
Passenger Revenue	Motor bus passengers/total passengers
Special Transit Fare	Motor bus passengers/total passengers
School Bus Revenue	All designated as motorbus
Freight Tariffs	All designated as motor bus
Charter Service	All designated as motor bus
Auxiliary Revenue	Number of motor bus vehicles/total vehicles (excluding demand responsive vehicles)
Non-transportation revenue	Motor bus operating expense/total operating expense
Taxes Levied by transit system	Motor bus operating expense/total operating expense
All cash grants (state, local, Fed.)	Motor bus operating expense/total operating expense
Total Employee Wages	Motor bus employees/total employees
Revenue Vehicle Operator Wages	Motor bus drivers/total drivers
Revenue Vehicle Maintenance Wages	Motor bus maintenance employees/total maintenance employees
Non-revenue Vehicle Maintenance Wages	Motor bus maintenance employees/total maintenance employees
Fringe Benefits	Motor Bus Employees/Total Employees

APPENDIX C

DESCRIPTIVE STATISTICS FOR FOUR DATA SETS OF PERFORMANCE INDICATOR VARIABLES

TABLE C-1. DESCRIPTIVE STATISTICS FOR 48 PERFORMANCE INDICATOR VARIABLES FROM RAW REPORTED DATA

Variable	Number of Cases	Mean	Standard Deviation	Variance	Skewness	Kurtosis
TVH/EMP	275	.116	.034	.001	3.242	26.160
RVH/OEMP	274	.167	.043	.002	1.693	10.335
TVM/EMP	274	1.532	.446	.119	2.307	16.321
PVEH/ADM	287	3.146	1.744	3.043	2.219	9.139
PVEH/OP	291	.579	.169	.028	1.367	3.703
PVEH/MNT	278	2.390	1.640	2.690	4.352	27.081
TVH/AVEH	280	.245	1.07	.011	2.868	13.845
TVH/PVEH	278	.317	.090	.008	1.387	6.977
TVM/AVEH	279	3.296	1.916	3.670	6.598	68.464
TVM/PVEH	277	4.241	1.463	2.140	2.105	12.108
RVM/TVM	279	.941	.075	.006	−2.783	11.343
RVM/FUEL	275	4.809	9.554	91.273	16.085	263.908
TVM/FUEL	273	5.143	10.581	111.955	16.098	263.567
TVM/MEXP	267	3.404	1.669	2.786	2.059	7.254
TVM/MNT	264	9.928	8.481	71.929	5.594	39.196
TVM/RCAL	272	.848	2.764	7.642	12.178	174.989
RVH/OEXP	274	.045	.015	.0002	1.204	3.063
TVM/OEXP	274	.631	.228	.052	1.414	3.374
RVH/TWG	274	.102	.295	.087	10.830	125.118
RVH/OWAG	268	.101	.060	.004	6.572	72.432
RVH/VMWG	250	.962	8.258	68.191	15.746	248.599
RVH/ADWG	270	.917	.761	.579	3.890	23.561
TPAS/RVH	238	32.848	16.452	270.668	1.012	1.395
TPAS/RVM	239	2.560	1.510	2.280	2.277	10.643
TPAS/PVH	248	9.354	5.236	27.411	1.522	4.350
PASM/TPS	231	4.556	4.756	22.619	6.309	52.941
TVM/ACC	264	2.359	2.344	5.495	3.904	19.440
RVH/ACC	264	.171	.197	.039	6.154	52.574
REV/PVEH	296	2.567	2.068	4.276	3.937	26.238
REV/RVH	279	9.124	7.551	57.019	3.034	12.405
OREV/RVH	279	9.258	7.579	57.438	3.004	12.213
REV/TPAS	249	.327	.313	.098	4.416	25.515
RVH/LSUB	250	3.190	37.707	1,421.806	15.508	243.274
RVH/SSUB	224	15.175	144.965	21,014.951	13.075	180.880
RVH/OSUB	279	.087	.099	.010	7.458	68.602
RVH/TSUB	279	.072	.100	.010	7.591	70.471
TPAS/LOA	227	159.545	2001.678	4,006,716.275	14.968	224.953
TPAS/TSUB	249	2.237	5.622	31.601	15.069	231.922
REV/TSUB	302	.747	1.755	3.081	7.132	59.108
REV/OSUB	302	.862	1.745	3.045	7.082	58.705

TABLE C-1 (continued)

Variable	Number of Cases	Mean	Standard Deviation	Variance	Skewness	Kurtosis
PAS/OSUB	249	2.723	5.628	31.669	13.530	201.382
PAS/OEXP	246	1.332	.605	.366	.847	1.463
PASM/OEX	228	5.251	3.151	9.931	1.782	4.950
PAS/TWAG	246	2.846	8.285	68.636	12.059	156.202
PAS/FUEL	243	11.412	24.536	601.996	14.657	223.480
PASM/TEX	230	5.049	3.194	10.202	1.884	5.330
REV/OEXP	297	.415	1.006	1.012	16.309	275.601
TREV/TEX	282	1.016	.216	.047	4.090	45.771

TABLE C-2: DESCRIPTIVE STATISTICS FOR 48 PERFORMANCE INDICATOR VARIABLES FROM WEIGHTED REPORTED DATA

Variable	Number of Cases	Mean	Standard Deviation	Variance	Skewness	Kurtosis
TVH/EMP	275	.116	.034	.001	3.242	26.160
RVH/OEMP	274	.167	.043	.002	1.693	10.335
TVM/EMP	274	1.532	.446	.119	2.307	16.321
PVEH/ADM	287	3.146	1.744	3.043	2.219	9.139
PVEH/OP	291	.579	.169	.028	1.367	3.703
PVEH/MNT	278	2.390	1.640	2.690	4.352	27.081
TVH/AVEH	280	.245	1.07	.011	2.868	13.845
TVH/PVEH	278	.317	.090	.008	1.387	6.977
TVM/AVEH	279	3.296	1.916	3.670	6.598	68.464
TVM/PVEH	277	4.241	1.463	2.140	2.105	12.108
RVM/TVM	279	.941	.075	.006	-2.783	11.343
RVM/FUEL	275	4.809	9.554	91.273	16.085	263.908
TVM/FUEL	273	5.143	10.581	111.955	16.098	263.567
TVM/MEXP	267	3.404	1.669	2.786	2.059	7.254
TVM/MNT	264	9.928	8.481	71.929	5.594	39.196
TVM/RCAL	272	.848	2.764	7.642	12.178	174.989
RVH/OEXP	274	.045	.015	.0002	1.204	3.063
TVM/OEXP	274	.631	.228	.052	1.414	3.374
RVH/TWG	274	.102	.295	.087	10.830	125.118
RVH/OWAG	268	.101	.060	.004	6.572	72.432
RVH/VMWG	250	.962	8.258	68.191	15.746	248.599
RVH/ADWG	270	.917	.761	.579	3.890	23.561
TPAS/RVH	238	32.848	16.452	270.668	1.012	1.395
TPAS/RVM	239	2.560	1.510	2.280	2.277	10.643
TPAS/PVH	248	9.354	5.236	27.411	1.522	4.350
PASM/TPS	231	4.556	4.756	22.619	6.309	52.941
TVM/ACC	264	2.359	2.344	5.495	3.904	19.440
RVH/ACC	264	.171	.197	.039	6.154	52.574
REV/PVEH	292	2.450	1.702	2.897	2.611	10.475
REV/RVH	276	8.720	6.611	43.708	3.111	14.918
OREV/RVH	276	8.854	6.637	44.051	3.063	14.527
REV/TPAS	249	.315	.307	.094	4.701	28.346
RVH/LSUB	246	3.256	38.012	1444.885	15.382	239.348
RVH/SSUB	220	16.018	148.788	2,2137.919	12.531	166.949
RVH/OSUB	273	.079	.072	0.005	7.195	64.990
RVH/TSUB	274	.096	.383	.147	14.994	237.376
TPAS/LOA	224	167.343	2,016.147	4,064,849.818	14.835	221.305
TPAS/TSUB	246	4.424	33.277	1,107.326	15.069	231.922
REV/TSUB	293	1.414	10.846	117.634	16.430	276.681
REV/OSUB	293	.916	2.290	5.242	7.794	69.911
PAS/OSUB	244	2.349	1.683	2.831	2.194	6.665
PAS/OEXP	246	1.332	.605	.366	.847	1.463
PASM/OEX	228	5.251	3.151	9.931	1.782	4.950
PAS/TWAG	246	2.846	8.285	68.636	12.059	156.202
PAS/FUEL	243	11.412	24.536	601.996	14.657	223.480
PASM/TEX	230	5.049	3.194	10.202	1.884	5.330
REV/OEXP	293	.402	.970	.941	16.300	274.266
TREV/TEX	282	1.016	.216	.047	4.090	45.771

TABLE C-3: DESCRIPTIVE STATISTICS FOR LOGARITHMS (BASE 10) OF 48 PERFORMANCE INDICATOR VARIABLES FROM RAW REPORTED DATA

Variable	Number of Cases	Mean	Standard Deviation	Variance	Skewness	Kurtosis
TVH/EMP	275	-0.951	.117	.014	-0.663	6.772
RVH/OEMP	274	-0.792	.114	.013	-1.130	7.221
TVM/EMP	274	.168	.127	.016	-1.058	6.799
PVEH/ADM	287	.440	.228	.052	-0.331	1.174
PVEH/OP	291	-0.254	.122	.015	-0.153	2.064
PVEH/MNT	278	.319	.216	.047	.228	3.694
TVH/AVEH	280	-0.642	.161	.026	.154	2.546
TVH/PVEH	278	-0.515	.123	.015	-0.427	2.075
TVM/AVEH	279	.479	.171	.029	.657	4.105
TVM/PVEH	277	.604	.142	.020	-0.236	2.033
RVM/TVM	279	-0.028	.041	.002	-3.741	20.206
RVM/FUEL	275	.613	.164	.027	3.192	32.415
TVM/FUEL	273	.642	.156	.024	4.106	40.678
TVM/MEXP	267	.488	.197	.039	-0.131	1.045
TVM/MNT	264	.927	.225	.051	.471	5.239
TVM/RCAL	272	-0.463	.485	.236	.942	1.199
RVH/OEXP	274	-1.370	.144	.021	-0.054	.300
TVM/OEXP	274	-0.226	.149	.022	.051	.619
RVH/TWG	274	-1.163	.252	.064	2.779	16.081
RVH/OWAG	268	-1.040	.187	.035	.324	2.731
RVH/VMWG	250	-0.430	.284	.081	3.518	26.557
RVH/ADWG	270	-0.133	.278	.077	.246	.555
TPAS/RVH	238	1.458	.240	.058	-0.752	.940
TPAS/RVM	239	.339	.257	.066	-0.582	.990
TPAS/PVH	248	.900	.266	.071	-0.818	1.223
PASM/TPS	231	.573	.231	.053	1.617	4.332
TVM/ACC	264	.260	.286	.082	.801	.868
RVH/ACC	264	-0.883	.282	.080	.923	1.710
REV/PVEH	296	.317	.281	.079	-0.127	1.636
REV/RVH	279	.864	.278	.077	.352	.732
OREV/RVH	279	.864	.278	.077	.352	.732
REV/TPAS	249	-0.582	.263	.069	.659	2.772
RVH/LSUB	250	-0.745	.567	.322	2.106	8.629
RVH/SSUB	223	-0.373	.751	.563	1.614	4.226
RVH/OSUB	279	-1.191	.252	.063	.000	3.893
RVH/TSUB	279	-1.282	.320	.102	.235	2.173
TPAS/LOA	227	.746	.598	.357	2.171	8.863
TPAS/TSUB	249	.150	.369	.136	.268	1.713
REV/TSUB	302	-0.432	.447	.200	.525	2.001
REV/OSUB	302	-0.331	.382	.146	.722	2.857
PAS/OSUB	249	.245	.300	.090	-0.184	0.347
PAS/OEXP	246	.074	.225	.051	-0.959	1.560
PASM/OEX	228	.647	.263	.069	-0.530	.953
PAS/TWAG	246	.280	.281	.079	1.403	9.783
PAS/FUEL	243	.947	.250	.063	.591	7.600
PASM/TEX	230	.617	.314	.099	-2.295	15.407
REV/OEXP	297	-0.496	.242	.059	1.126	8.330
TREV/TEX	282	-0.008	.172	.030	-11.675	172.255

TABLE C-4: DESCRIPTIVE STATISTICS FOR LOGARITHMS (BASE 10) OF 48 PERFORMANCE INDICATOR VARIABLES FROM WEIGHTED REPORTED DATA

Variable	Number of Cases	Mean	Standard Deviation	Variance	Skewness	Kurtosis
TVH/EMP	275	-0.951	.117	.014	-0.663	6.772
RVH/OEMP	274	-0.792	.114	.013	-1.130	7.221
TVM/EMP	274	.168	.127	.016	-1.058	6.799
PVEH/ADM	287	.440	.228	.052	-0.331	1.174
PVEH/OP	291	-0.254	.122	.015	-0.153	2.064
PVEH/MNT	278	.319	.216	.047	.228	3.694
TVH/AVEH	280	-0.642	.161	.026	.154	2.546
TVH/PVEH	278	-0.515	.123	.015	-0.427	2.075
TVM/AVEH	279	.479	.171	.029	.657	4.105
TVM/PVEH	277	.604	.142	.020	-0.236	2.033
RVM/TVM	279	-0.028	.041	.002	-3.741	20.206
RVM/FUEL	275	.613	.164	.027	3.192	32.415
TVM/FUEL	273	.642	.156	.024	4.106	40.678
TVM/MEXP	267	.488	.197	.039	-0.131	1.045
TVM/MNT	264	.927	.225	.051	.471	5.239
TVM/RCAL	272	-0.463	.485	.236	.942	1.199
RVH/OEXP	274	-1.370	.144	.021	-0.054	.300
TVM/OEXP	274	-0.226	.149	.022	.051	.619
RVH/TWG	274	-1.163	.252	.064	2.779	16.081
RVH/OWAG	268	-1.040	.187	.035	.324	2.731
RVH/VMWG	250	-0.430	.284	.081	3.518	26.557
RVH/ADWG	270	-0.133	.278	.077	.246	.555
TPAS/RVH	238	1.458	.240	.058	-0.752	.940
TPAS/RVM	239	.339	.257	.066	-0.582	.990
TPAS/PVH	248	.900	.266	.071	-0.818	1.223
PASM/TPS	231	.573	.231	.053	1.617	4.332
TVM/ACC	264	.260	.286	.082	.801	.868
RVH/ACC	264	-0.883	.282	.080	.923	1.710
REV/PVEH	292	.307	.272	.074	-0.320	1.496
REV/RVH	275	.851	.261	.068	.084	.461
OREV/RVH	275	.859	.260	.068	.065	.468
REV/TPAS	247	-0.605	.243	.059	.486	3.340
RVH/LSUB	245	-0.722	.562	.316	2.187	8.862
RVH/SSUB	218	-0.343	.743	.552	1.708	4.404
RVH/OSUB	273	-1.128	.242	.058	1.919	8.125
RVH/TSUB	273	-1.255	.312	.097	.867	3.927
TPAS/LOA	222	.788	.600	.360	2.223	8.640
TPAS/TSUB	244	.191	.379	.144	.876	4.356
REV/TSUB	293	-0.406	.465	.217	.909	2.851
REV/OSUB	293	-0.308	.388	.150	1.118	3.007
PAS/OSUB	244	.279	.286	.082	-0.146	.352
PAS/OEXP	246	.074	.225	.051	-0.959	1.560
PASM/OEX	228	.647	.263	.069	-0.530	.953
PAS/TWAG	246	.280	.281	.079	1.403	9.783
PAS/FUEL	243	.947	.250	.063	.591	7.600
PASM/TEX	230	.617	.314	.099	-2.295	15.407
REV/OEXP	293	-0.496	.242	.059	1.126	8.330
TREV/TEX	282	-0.008	.172	.030	-11.675	172.255

APPENDIX D

CORRELATION MATRIX FOR 48 PERFORMANCE INDICATORS FROM THE WEIGHTED RAW DATA SET

	TVHIEMP	RVHIOEMP	TVMIEMP	PVEHIADM	PVEHIOP	PVEHIMNT	TVHIAVEH	TVHIPVEH	TVMIAVEH	TVMIPVEH
TVHIEMP	1.00000	0.23086	0.84460	0.19008	0.14944	0.05302	0.10288	0.10655	-0.15796	-0.19578
RVHIOEMP	0.23086	1.00000	0.35856	0.21165	0.33886	0.01334	0.38354	0.53430	0.09144	0.17410
TVMIEMP	0.84460	0.35856	1.00000	0.18064	0.12426	0.11589	0.15055	0.20263	0.12207	0.17058
PVEHIADM	0.19008	0.21165	0.18064	1.00000	0.41992	0.16834	-0.05611	-0.15977	-0.20317	-0.32614
PVEHIOP	0.14944	0.33886	0.12426	0.41992	1.00000	0.21707	-0.14264	-0.46105	-0.32244	-0.63382
PVEHIMNT	0.05302	0.01334	0.11589	0.16834	0.21707	1.00000	-0.09940	-0.14463	-0.11192	-0.16080
TVHIAVEH	0.10288	0.38354	0.15055	-0.05611	-0.14264	-0.09940	1.00000	0.58647	0.78347	0.33613
TVHIPVEH	0.10655	0.53430	0.20263	-0.15977	-0.46105	-0.14463	0.58647	1.00000	0.36962	0.68348
TVMIAVEH	-0.15796	0.09144	0.12207	-0.20317	-0.32244	-0.11192	0.78347	0.36962	1.00000	0.62522
TVMIPVEH	-0.19578	0.17410	0.17058	-0.32614	-0.63382	-0.16080	0.33613	0.68348	0.62522	1.00000
RVHITVM	0.13069	0.42777	0.12292	0.03217	-0.17312	-0.05151	0.13194	0.34401	0.01778	0.18083
RVMIFUEL	-0.02173	-0.00275	-0.01993	-0.05739	0.06514	0.04912	0.01160	-0.05549	0.00884	-0.04546
TVMIFUEL	-0.02327	-0.01247	-0.02177	-0.05762	0.06985	0.05056	0.00878	-0.06369	0.00789	-0.05090
TVMIMEXP	-0.06240	-0.01049	0.08562	0.11881	-0.00281	0.23703	-0.01339	0.01428	0.09003	0.15906
TVMIMNT	-0.02175	0.08235	0.17857	0.02144	-0.06656	0.90450	0.03692	0.13341	0.12960	0.22599
TVMIRCAL	-0.05793	-0.14145	-0.08353	0.04681	0.09113	0.05649	-0.10597	-0.17645	-0.03193	-0.10331
RVHIOEXP	0.13831	0.49553	0.22645	0.19655	0.10044	0.06576	0.17750	0.25593	0.04295	0.11423
TVMIOEXP	-0.11920	0.19229	0.15208	0.00943	-0.02804	0.05867	0.08052	0.14389	0.24579	0.36957
RVHITNG	0.10148	0.51198	0.20942	0.16750	0.03616	0.11739	0.12505	0.30899	0.01746	0.18721
RVHIOWAG	0.04894	0.39070	0.12033	0.01398	-0.01169	0.02075	0.05186	0.23644	-0.01007	0.18290
RVHIVMWG	0.04760	0.11794	0.11672	0.28860	-0.09682	0.30769	0.03004	0.14954	0.02353	0.13972
RVHIADWG	0.07195	0.36940	0.10703	0.33093	0.11580	0.06192	0.03463	0.16175	-0.08907	-0.01622
TPASIRVH	-0.20998	-0.35237	-0.28853	-0.23328	-0.24494	-0.18313	-0.05039	-0.13035	0.03384	-0.03859
TPASIRVM	-0.05571	-0.21719	-0.25705	-0.16592	-0.19088	-0.18334	-0.00480	-0.06820	-0.10696	-0.19656
TPASIPVH	-0.14962	-0.04456	-0.18083	-0.29148	-0.42975	-0.23587	0.16069	0.25930	0.16776	0.24227
PASMITPS	-0.07509	-0.31189	0.02619	-0.07357	-0.03266	0.11056	-0.11448	-0.21888	0.16705	0.10074
TVHIACC	-0.00362	0.21672	0.00978	0.03345	0.06222	0.06102	0.06375	0.07321	0.05445	0.07915
RVHIACC	0.09524	0.39171	0.02357	0.15226	0.18183	0.05065	0.12277	0.13599	-0.04762	-0.05207
WRVIPVEH	-0.17031	-0.18563	-0.13998	-0.06925	-0.17885	-0.13918	0.12063	-0.01199	0.22022	0.08003
WREVIRVH	-0.16310	-0.41617	-0.15608	0.02705	0.08082	0.00040	-0.06901	-0.34989	0.08880	-0.18004
WDRVIRVH	-0.16469	-0.42079	-0.15863	0.02533	0.08646	-0.00164	-0.07078	-0.35638	0.08819	-0.18957
WRVIIPAS	-0.06699	-0.17033	-0.05703	0.17346	0.28664	0.11307	-0.05483	-0.31062	0.01086	-0.22577
RVHIWLSB	-0.00917	-0.03649	0.00796	0.02292	-0.03250	0.01651	0.00934	-0.00257	0.02150	0.01267
RVHIWSSB	-0.01301	0.03169	0.03238	-0.06846	0.04917	0.00323	-0.01095	-0.02938	0.03062	0.01827
RVHIWOSB	-0.09154	-0.26407	-0.19932	-0.04094	-0.19095	-0.06414	-0.08727	-0.13097	-0.08566	-0.12782
RVHIWTSB	-0.07447	-0.26488	-0.18876	-0.02846	-0.15316	-0.05629	-0.08547	-0.16465	-0.08905	-0.16537
TPASIWLO	-0.01301	-0.04600	0.00053	0.01842	-0.03333	0.00072	0.00375	-0.01162	0.01900	0.00783
PASIWTSB	-0.10942	-0.32941	-0.23839	-0.11621	-0.23643	-0.08014	-0.11286	-0.15908	-0.09858	-0.14113
WRVIWTSB	-0.09493	-0.36110	-0.12327	0.06200	0.08135	0.05320	-0.12602	-0.30411	-0.04415	-0.20897
WRVIWOSB	-0.10723	-0.37148	-0.13513	0.06137	0.06481	0.04419	-0.12914	-0.29837	-0.04095	-0.19711
PASIWOSB	-0.11976	-0.33722	-0.25035	-0.12731	-0.25307	-0.08871	-0.11728	-0.14923	-0.10094	-0.13090
PASIOEXP	-0.14462	-0.08598	-0.17643	-0.10539	-0.18801	-0.11828	-0.02699	-0.01613	-0.01703	-0.00658
PASMIDEX	-0.15688	-0.30313	-0.09430	-0.16482	-0.17106	0.02358	-0.12004	-0.19728	0.10131	0.06905
PASITRAC	-0.14615	-0.08516	-0.14969	-0.12048	-0.24535	-0.07004	-0.06645	0.01852	-0.01823	0.07203
PASIFUEL	-0.03306	-0.04048	-0.06314	-0.08159	0.03960	0.02162	0.00653	-0.07223	-0.01069	-0.07994
PASMITEX	-0.14727	-0.30211	-0.08255	-0.12507	-0.18359	0.02632	-0.13075	-0.17780	0.08003	0.07532
WDRVIOXP	-0.15135	-0.17410	-0.10542	0.14089	0.15993	0.00505	-0.05033	-0.29013	0.03655	-0.18373
TREVITEX	-0.01338	0.04066	0.02694	-0.03535	-0.10153	0.01048	-0.07887	0.16889	-0.04708	0.19099

	RVM1TVM	RVM1FUEL	TVM1FUEL	TVM1MEXP	TVM1MNT	TVM1RCAL	RVH1OEXP	TVM1OEXP	RVH1TWG	RVH1OWAG
RVH1EMP	0.13069	-0.02173	-0.02327	-0.06240	-0.02175	-0.05793	0.13831	-0.11928	0.10148	[illegible]
RVH1OEMP	0.42777	-0.00275	-0.01247	-0.01049	0.08235	-0.14105	0.49553	0.19229	0.51198	0.39070
TVM1EMP	0.12292	-0.01993	-0.02177	0.08562	0.17857	-0.08353	0.22645	0.15208	0.20942	0.12033
PVEH1ADM	0.03217	-0.05739	-0.05762	0.11881	0.02144	0.04681	0.19655	0.00943	0.16750	0.01398
PVEH1OP	-0.17312	0.06514	0.06985	-0.00281	-0.06656	0.09113	0.10044	-0.02804	0.03616	-0.01169
PVEH1MNT	-0.05151	0.04912	0.05056	0.23703	0.90450	0.05649	0.06576	0.05867	0.11739	0.02075
TVH1AVEH	0.13194	0.01160	0.00878	-0.01339	0.03692	-0.10597	0.17750	0.08052	0.12505	0.05186
TVH1PVEH	0.34401	-0.05549	-0.06369	0.01428	0.13341	-0.17645	0.25593	0.14389	0.30899	0.23644
RVH1AVEH	0.01778	0.00884	0.00789	0.09003	0.12960	-0.03193	0.04295	0.24579	0.01746	-0.01007
TVM1PVEH	0.18883	-0.04546	-0.05090	0.15906	0.22599	-0.10331	0.11423	0.36957	0.18721	0.18290
RVM1TVM	1.00000	-0.00839	-0.03286	0.00401	0.04796	-0.14472	0.35923	0.13791	0.37985	0.34253
RVM1FUEL	-0.00839	1.00000	0.99968	0.03632	0.02335	-0.02672	0.03579	0.03950	0.04318	0.04306
TVM1FUEL	-0.03286	0.99968	1.00000	0.03589	0.02185	-0.02421	0.02662	0.03502	0.03276	0.03317
TVM1MEXP	0.00401	0.03632	0.03589	1.00000	0.26971	0.15550	0.50450	0.61075	0.48413	0.43778
TVM1MNT	0.04796	0.02335	0.02185	0.26971	1.00000	-0.00270	0.10588	0.18971	0.19249	0.08940
TVM1RCAL	-0.14472	-0.02672	-0.02421	0.15550	-0.00270	1.00000	-0.01588	0.04066	0.08230	0.12712
RVH1OEXP	0.35923	0.03579	0.02662	0.50450	0.10588	-0.01588	1.00000	0.83755	0.80422	0.71663
TVM1OEXP	0.13791	0.03950	0.03502	0.61075	0.18971	0.04066	0.83755	1.00000	0.64737	0.61407
RVH1TWG	0.37985	0.04318	0.03276	0.48413	0.19249	0.08230	0.80422	0.64737	1.00000	0.94700
RVH1OWAG	0.34253	0.04306	0.03317	0.43778	0.08940	0.12712	0.71663	0.61407	0.94700	1.00000
RVH1VMWG	0.19669	0.02360	0.01815	0.49035	0.41144	0.06102	0.43434	0.40219	0.55741	0.45155
RVH1ADWG	0.28362	-0.02781	-0.03487	0.02339	0.04990	-0.07895	0.36842	0.17904	0.39102	0.24962
TPAS1RVH	-0.06827	-0.04318	-0.04170	-0.19537	-0.16318	-0.08902	-0.42174	-0.31687	-0.45789	-0.36936
TPAS1RVM	-0.04567	-0.04474	-0.04298	-0.26109	-0.21885	-0.11945	-0.35122	-0.41506	-0.41843	-0.36621
TPAS1PVH	0.17279	-0.06109	-0.06499	-0.21731	-0.10316	-0.20942	-0.26675	-0.25724	-0.28453	-0.28059
PASM1IPS	-0.33616	-0.01450	-0.00647	0.03324	0.17824	0.36658	-0.21870	0.03795	-0.19383	-0.15138
TVM1ACC	0.12216	-0.01289	-0.01651	0.11451	0.09728	0.28100	0.18241	0.15314	0.37668	0.45783
RVH1ACC	0.16479	-0.01786	-0.02177	0.03025	0.03645	0.25892	0.23915	0.05844	0.37612	0.40272
[illegible]	-0.12254	-0.07897	-0.07462	-0.32401	-0.09218	0.11173	-0.42902	-0.34377	-0.47767	-0.50291
WREV1RVH	-0.45029	-0.05188	-0.04013	-0.18524	-0.08067	0.29853	-0.47339	-0.29321	-0.50166	-0.50507
WORV1RVH	-0.45710	-0.05001	-0.03812	-0.18944	-0.08441	0.29505	-0.47660	-0.29347	-0.50874	-0.51154
WRV1TPAS	-0.36618	-0.03363	-0.02386	-0.09366	-0.01585	0.32307	-0.21578	-0.13376	-0.16010	-0.19025
[illegible]	-0.05921	-0.00633	-0.00499	0.01852	0.03568	-0.01779	-0.02891	-0.00686	-0.05081	-0.06387
RVH1WSSB	0.03672	-0.00820	-0.00888	0.12050	0.01388	-0.03366	0.11958	0.17230	0.13059	0.11467
RVH1WOSB	-0.09721	-0.02105	-0.01867	0.12385	-0.09728	0.11289	0.15251	0.14532	0.09492	0.15636
RVH1WTSB	-0.11725	-0.01219	-0.00942	0.10490	-0.10284	0.15272	0.14096	0.13192	0.09356	0.15299
[illegible]	-0.06284	-0.00719	-0.00567	0.01679	0.01367	-0.02455	-0.03816	-0.01391	-0.06451	-0.07401
PAS1WTSB	-0.09879	-0.01713	-0.01476	0.04649	-0.10804	0.10015	0.01895	0.02706	0.00821	0.09216
WRV1WTSB	-0.46675	-0.03808	-0.02653	0.02112	-0.04923	0.37827	-0.20050	-0.06969	-0.21367	-0.20613
WRV1WOSB	-0.46018	-0.04398	-0.03249	0.01667	-0.05352	0.35810	-0.21342	-0.07850	-0.23262	-0.22239
[illegible]	-0.09398	-0.02229	-0.02002	0.04842	-0.11166	0.07604	0.00946	0.01968	-0.00825	0.07925
PAS1OEXP	0.19632	-0.00930	-0.01459	0.18029	-0.08806	-0.12568	0.26428	0.26970	0.08420	0.12862
PASM1OEX	-0.08220	-0.00966	-0.00851	0.12558	0.06168	0.12532	0.04885	0.25138	-0.06078	0.00861
PAS1IWAG	0.22086	0.00287	-0.00370	0.21811	-0.00874	-0.09332	0.16610	0.19309	0.27278	0.35494
[illegible]	-0.01322	0.98678	0.98670	-0.00071	-0.01175	-0.05929	-0.02161	-0.02646	-0.02561	[illegible]
PASM1TEX	-0.09130	-0.00055	0.00058	0.14604	0.06228	0.14491	-0.01580	0.16819	-0.01582	0.04475
WORV1OXP	-0.19775	-0.03026	-0.02431	-0.02856	-0.06307	0.20155	0.00529	0.10288	-0.19547	-0.25222
TREV1TEX	-0.00085	0.00931	0.00884	0.00890	0.06797	-0.01988	-0.04819	-0.03564	0.08173	0.08063

	RVH1VMWG	RVH1ADWG	TPAS1RVH	TPAS1RVM	TPAS1PVH	PASM1TPS	TVM1ACC	RVH1ACC	WRV1PVEH	WREV1RVH
TVM1EMP	0.04760	0.07195	-0.20998	-0.05571	-0.14962	-0.07509	-0.00362	0.09524	-0.17031	-0.16310
RVH1OEMP	0.11794	0.36940	-0.35237	-0.21719	-0.04456	-0.31189	0.21672	0.39171	-0.18563	-0.41617
TVM1EMP	0.11672	0.10703	-0.28855	-0.25705	-0.18083	0.02619	0.00978	0.02357	-0.13998	-0.15008
PVEH1ADM	0.28860	0.33093	-0.23328	-0.16592	-0.29148	-0.07357	0.03345	0.15226	-0.06925	0.02705
PVEH1OP	-0.09682	0.11580	-0.24494	-0.19088	-0.42925	-0.03266	0.06222	0.18183	-0.17885	0.08082
PVEH1MNT	0.30769	0.06192	-0.18315	-0.18334	-0.23587	0.11056	0.06102	0.05065	-0.13918	0.00040
TVH1AVEH	0.03004	0.03463	-0.05039	-0.00480	0.16069	-0.11448	0.06375	0.12277	0.12063	-0.06901
TVH1PVEH	0.14954	0.16175	-0.13035	-0.06820	0.25930	-0.21888	0.07321	0.13599	-0.01199	-0.34989
TVM1AVEH	0.02353	-0.08907	0.03384	-0.10696	0.16776	0.16705	0.05445	-0.04762	0.22022	0.08880
TVM1PVEH	0.13972	-0.01622	-0.03859	-0.19656	0.24227	0.10074	0.07915	-0.05207	0.08003	-0.18004
RVH1TVM	0.19669	0.28362	-0.06827	-0.04567	0.17279	-0.33616	0.12216	0.16479	-0.12254	-0.45029
RVH1FUEL	0.02360	-0.02781	-0.04318	-0.04474	-0.06109	-0.01450	-0.01289	-0.01786	-0.07897	-0.05188
TVM1FUEL	0.01815	-0.03487	-0.04170	-0.04298	-0.06499	-0.00647	-0.01651	-0.02177	-0.07462	-0.04013
TVM1MEXP	0.49035	0.02339	-0.19537	-0.26109	-0.21731	0.03324	0.11451	0.03025	-0.32401	-0.18524
TVM1MNT	0.41144	0.04990	-0.16318	-0.21885	-0.10316	0.12824	0.09728	0.03645	-0.09218	-0.08067
TVM1RCAL	0.06102	-0.07895	-0.08902	-0.11945	-0.20942	0.36658	0.28100	0.25892	0.11173	0.29853
RVH1OEXP	0.43434	0.36842	-0.42174	-0.35122	-0.26675	-0.21870	0.18241	0.23915	-0.42902	-0.47339
TVM1OEXP	0.40219	0.17904	-0.31687	-0.41506	-0.25724	0.03795	0.15314	0.05844	-0.34377	-0.29321
RVH1TRC	0.55741	0.39102	-0.45789	-0.41843	-0.28453	-0.19383	0.37668	0.37612	-0.47767	-0.50166
RVH1OWAG	0.45155	0.24962	-0.36936	-0.36621	-0.24059	-0.15138	0.45783	0.40272	-0.50297	-0.50507
RVH1VMWG	1.00000	0.08054	-0.21249	-0.20647	-0.12966	-0.00765	0.14402	0.11773	-0.22659	-0.22679
RVH1ADWG	0.08054	1.00000	-0.23121	-0.15591	-0.11034	-0.18268	0.02336	0.08923	-0.09556	-0.16199
TPAS1RVH	-0.21249	-0.23121	1.00000	0.91355	0.87424	-0.07503	-0.08593	-0.17895	0.48544	0.34235
TPAS1RVM	-0.20647	-0.15591	0.91355	1.00000	0.86435	-0.19378	-0.18060	-0.17421	0.44692	0.25431
TPAS1PVH	-0.12966	-0.11034	0.87424	0.86435	1.00000	-0.20440	-0.08239	-0.13589	0.47857	0.13512
PASM1TPS	-0.00765	-0.18268	-0.07503	-0.19378	-0.20440	1.00000	0.08338	0.01434	0.34877	0.56422
TVM1ACC	0.14402	0.02336	-0.08593	-0.18060	-0.08239	0.08338	1.00000	0.90689	-0.18969	-0.13087
RVH1ACC	0.11773	0.08923	-0.17895	-0.17421	-0.13589	0.01434	0.90689	1.00000	-0.18507	-0.16830
WRV1PVEH	-0.22659	-0.09556	0.48544	0.44692	0.47857	0.34877	-0.18969	-0.18507	1.00000	0.79562
WREV1RVH	-0.22679	-0.16199	0.34235	0.25431	0.13512	0.56422	-0.13087	-0.16830	0.79562	1.00000
WORV1RVH	-0.23188	-0.16545	0.34202	0.25360	0.13159	0.56402	-0.13526	-0.17214	0.79263	0.99885
WRV1TPAS	-0.10909	-0.00615	-0.27064	-0.29812	-0.38618	0.53431	0.03889	0.04570	0.40499	0.68172
RVH1WLSB	0.01988	-0.02932	0.07470	0.06052	0.05684	0.01333	-0.03749	-0.02906	-0.00720	-0.00560
RVH1WSSB	0.05418	0.00491	-0.04355	-0.06009	-0.04084	-0.04919	-0.03992	-0.03677	-0.04790	-0.03970
TVH1WOSB	0.08309	-0.03972	0.15357	0.17295	0.02258	0.13046	-0.02468	-0.01322	0.04555	0.12599
RVH1WTSB	0.06482	-0.02843	0.16706	0.19030	0.02137	0.13925	0.00116	0.01084	0.05648	0.15807
TPAS1WLO	-0.00429	-0.03100	0.08184	0.06761	0.06125	0.01130	-0.04722	-0.04024	-0.00158	-0.00115
PAS1WTSB	-0.00835	-0.07914	0.30469	0.31290	0.15013	0.06728	-0.02010	-0.02837	0.05126	0.10591
WRV1WTSB	-0.05517	-0.09117	0.15564	0.11451	-0.04442	0.52514	-0.04535	-0.07457	0.49159	0.82963
WRV1WOSB	-0.05079	-0.10454	0.16391	0.11799	-0.03499	0.52660	-0.06547	-0.09319	0.50604	0.83310
PAS1WOSB	-0.00648	-0.09566	0.33251	0.33635	0.17797	0.05610	-0.03678	-0.04375	0.05426	0.09961
PAS1OEXP	0.10423	0.01142	0.71820	0.68694	0.68559	-0.24188	-0.02324	-0.09837	0.14562	-0.01382
PASM1OEX	0.10389	-0.12452	0.44705	0.30756	0.30100	0.60599	0.01049	-0.08901	0.36453	0.37091
PAS1TWAG	0.22339	0.03537	0.66076	0.58065	0.63247	-0.20783	0.19550	0.02894	0.05477	-0.06781
PAS1FUEL	-0.01076	-0.05235	0.10677	0.11205	0.07757	-0.04876	-0.03613	-0.04254	-0.01321	-0.01329
PASM1TEX	0.12865	-0.10185	0.40348	0.26452	0.25704	0.61223	-0.00448	-0.08729	0.34141	0.35682
[illegible]	-0.00874	0.03133	0.23447	0.20159	0.10442	0.40674	-0.14369	-0.14060	0.73633	0.70177
TREV1TEX	0.06901	0.04444	-0.21945	-0.24943	-0.16313	0.12798	-0.09920	-0.05994	-0.03623	-0.03121

	WORV1RVH	WRV1TPAS	RVH1WLSB	RVH1WSSB	RVH1WOSB	RVH1WTSB	TPAS1WLO	PAS1WTSB	WRV1WTSB	WRV1WOSB
[illegible]	-0.16469	-0.06699	-0.00917	-0.01301	-0.09154	-0.07447	-0.01301	-0.10942	-0.09493	[illegible]
RVH1OEMP	-0.42079	-0.17033	-0.03649	0.03169	-0.26407	-0.26488	-0.04600	-0.32941	-0.36110	-0.37148
TVM1EMP	-0.15863	-0.05703	0.00796	0.03238	-0.19932	-0.18876	0.00053	-0.23839	-0.12327	-0.13513
PVEH1ADM	0.07533	0.17346	0.02292	-0.06846	-0.04094	-0.02846	0.01842	-0.11621	0.06200	0.06137
[illegible]	0.08646	0.28664	-0.03250	0.04917	-0.19095	-0.15316	-0.03333	-0.23643	0.08135	[illegible]
PVEH1MNT	-0.00164	0.11307	0.01651	0.00323	-0.06414	-0.05629	0.00072	-0.08014	0.05320	0.04419
TVH1AVEH	-0.07078	-0.05483	0.00934	-0.01095	-0.08727	-0.08547	0.00375	-0.11286	-0.12602	-0.12914
TVH1PVEH	-0.35638	-0.31062	-0.00257	-0.02938	-0.13097	-0.16465	-0.01162	-0.15908	-0.30411	-0.29837
[illegible]	0.08819	0.01086	0.02150	0.03062	-0.08566	-0.08905	0.01900	-0.09858	-0.04415	[illegible]
[illegible]										
TVM1PVEH	-0.18457	-0.22577	0.01267	0.01827	-0.12782	-0.16537	0.00783	-0.14113	-0.20849	-0.19711
RVM1TVM	-0.45710	-0.36618	-0.05921	0.03672	-0.09721	-0.11725	-0.06284	-0.09879	-0.46675	-0.46016
RVM1FUEL	-0.05001	-0.03363	-0.00633	-0.00820	-0.02105	-0.01219	-0.00719	-0.01713	-0.03808	-0.04398
TVM1FUEL	-0.03812	-0.02386	-0.00499	-0.00888	-0.01867	-0.00942	-0.00567	-0.01476	-0.02653	-0.03249
TVM1MEXP	-0.18944	-0.09366	0.01852	0.12050	0.12385	0.10490	0.01699	0.04649	0.02112	0.01667
TVM1MNT	-0.08441	-0.01585	0.03568	0.01388	-0.09728	-0.10284	0.01367	-0.10804	-0.04923	-0.05352
TVM1RCAL	0.29505	0.32307	-0.01779	-0.03366	0.11289	0.15272	-0.02455	0.10015	0.37827	0.35810
RVH1OEXP	-0.47660	-0.21578	-0.02891	0.11958	0.15251	0.14096	-0.03816	0.01895	-0.20050	-0.21382
TVM1OEXP	-0.29347	-0.13376	-0.00686	0.17230	0.14532	0.13192	-0.01391	0.02706	-0.06969	-0.07850
RVH1TWG	-0.50874	-0.16010	-0.05081	0.13059	0.09492	0.09356	-0.06451	0.00821	-0.21367	-0.23262
RVH1OWAG	-0.51154	-0.19025	-0.06387	0.11467	0.15636	0.15299	-0.07401	0.09216	-0.20613	-0.22239
RVH1VMWG	-0.23188	-0.10909	0.01988	0.05418	0.08309	0.06482	-0.00429	-0.00835	-0.05517	-0.05079
RVH1ADWG	-0.16545	-0.00615	-0.02932	0.00491	-0.03972	-0.02843	-0.03100	-0.07914	-0.09117	-0.10454
TPAS1RVH	0.34202	-0.27064	0.07470	-0.04355	0.15357	0.16706	0.08184	0.30469	0.15564	0.16391
TPAS1RVM	0.25360	-0.29812	0.06052	-0.06009	0.17295	0.19030	0.06761	0.31290	0.11451	0.11799
TPAS1PVH	0.13159	-0.38618	0.05684	-0.04084	0.02258	0.02137	0.06125	0.15013	-0.04442	-0.03499
PASM1TPS	0.56402	0.53431	0.01333	-0.04919	0.13046	0.13925	0.01130	0.06728	0.52514	0.52660
TVM1ACC	-0.13576	0.03889	-0.03749	-0.03992	-0.02468	0.00116	-0.04722	-0.02010	-0.04535	-0.06547
RVH1ACC	-0.17214	0.04570	-0.02906	-0.03677	-0.01322	0.01084	-0.04024	-0.02837	-0.07457	-0.09319
WRV1PVEH	0.79263	0.40499	-0.00720	-0.04790	0.04555	0.05648	-0.00158	0.05126	0.49159	0.50604
WREV1RVH	0.99885	0.68172	-0.00560	-0.03970	0.12899	0.15807	-0.00115	0.10591	0.82963	0.83310
WORV1RVH	1.00000	0.67995	-0.00654	-0.03910	0.12678	0.15508	-0.00198	0.10342	0.82620	0.83020
WRV1TPAS	0.67995	1.00000	-0.04677	-0.02859	0.08096	0.09271	-0.04574	-0.02489	0.64597	0.65113
RVH1WLSB	-0.00654	-0.04677	1.00000	-0.01264	-0.03076	-0.01633	0.99868	-0.00354	-0.01635	-0.02386
RVH1WSSB	-0.03910	-0.02859	-0.01264	1.00000	0.02034	0.03911	-0.01185	0.00392	-0.00730	-0.01679
RVH1WOSB	0.12678	0.08096	-0.03076	0.02034	1.00000	0.96329	-0.02959	0.93822	0.46873	0.49392
RVH1WTSB	0.15508	0.09271	-0.01633	0.03911	0.96329	1.00000	-0.01603	0.96507	0.50863	0.49318
TPAS1WLO	-0.00198	-0.04574	0.99868	-0.01185	-0.02959	-0.01603	1.00000	-0.00176	-0.01506	-0.02198
PAS1WTSB	0.10342	-0.02489	-0.00354	0.00392	0.93822	0.96507	-0.00176	1.00000	0.39818	0.38804
WRV1WTSB	0.82620	0.64597	-0.01635	-0.00730	0.46873	0.50863	-0.01506	0.39818	1.00000	0.98739
WRV1WOSB	0.83020	0.65113	-0.02386	-0.01679	0.49392	0.49318	-0.02198	0.38804	0.98739	1.00000
PAS1WOSB	0.09764	-0.04519	-0.00971	-0.00481	0.95466	0.94727	-0.00745	0.99016	0.38064	0.39097
PAS1OEXP	-0.01596	-0.45061	0.06220	0.03373	0.33545	0.33791	0.06673	0.39840	0.03617	0.03536
PASM1OEX	0.36887	0.02290	0.07582	-0.01898	0.36134	0.37294	0.07690	0.36785	0.37525	0.37219
PAS1TWAG	-0.07283	-0.42994	0.02685	0.04670	0.30607	0.31608	0.02778	0.39466	0.02460	0.01778
TPAS1FUEL	-0.01153	-0.08349	0.00510	-0.01670	0.00756	0.01864	0.00587	0.03388	-0.01937	-0.02449
PASM1TEX	0.35312	0.02694	0.08669	-0.02983	0.35616	0.37146	0.08712	0.37221	0.38086	0.37427
WORV1OXP	0.78477	0.55430	-0.01418	0.01433	0.20792	0.22582	-0.00923	0.11870	0.67756	0.68198
TREV1TEX	-0.03483	0.04840	0.01573	-0.02519	0.09971	-0.00023	0.00797	-0.01115	0.05802	0.11777

	PAS1WOSB	PAS1OEXP	PASM1OEX	PAS1TWAG	PAS1FUEL	PASM1TEX	WORV1OXP	TREV1TEX
TVM1EMP	-0.11976	-0.14462	-0.15888	-0.14615	-0.03306	-0.14727	-0.15135	-0.01338
RVH1OEMP	-0.33722	-0.08598	-0.30313	-0.08516	-0.04048	-0.30211	-0.17410	0.04066
TVH1EMP	-0.25035	-0.17643	-0.09430	-0.14969	-0.06314	-0.08255	-0.10542	0.02694
PVEH1ADM	-0.12731	-0.10539	-0.16482	-0.12048	-0.08159	-0.12507	0.14089	-0.03535
PVEH1OP	-0.25307	-0.18801	-0.17106	-0.24535	0.03960	-0.18359	0.15993	-0.10153
PVEH1MNT	-0.08871	-0.11828	0.02358	-0.07004	0.02162	0.02632	0.00505	0.01048
TVH1AVEH	-0.11728	-0.02699	-0.12004	-0.06645	0.00653	-0.13075	-0.05033	-0.07887
TVH1PVEH	-0.14923	-0.01613	-0.19728	0.01852	-0.07223	-0.17780	-0.29013	0.16889
TVM1AVEH	-0.10094	-0.01703	0.10131	-0.01823	-0.01069	0.08003	0.03655	-0.04708
TVM1PVEH	-0.13090	-0.00658	0.06905	0.07203	-0.07994	0.07532	-0.18373	0.19099
RVM1TVM	-0.09398	0.19632	-0.08220	0.22086	-0.01322	-0.09130	-0.19775	-0.00085
RVM1FUEL	-0.02229	-0.00930	-0.00966	0.00287	0.98678	-0.00055	-0.03026	0.00931
TVM1FUEL	-0.02002	-0.01459	-0.00851	-0.00370	0.98670	0.00058	-0.02431	0.00884
TVM1MEXP	0.04842	0.18029	0.12558	0.21811	-0.00071	0.14604	-0.02856	0.00890
TVM1MNT	-0.11166	-0.08806	0.06168	-0.00874	-0.01175	0.06228	-0.06307	0.06797
TVM1RCAL	0.07604	-0.12568	0.12532	-0.09332	-0.05929	0.14491	0.20155	-0.01988
RVH1OEXP	0.00946	0.26428	0.04885	0.16610	-0.02161	-0.01580	0.00529	-0.04819
TVH1OEXP	0.01968	0.26970	0.25138	0.19309	-0.02646	0.16819	0.10288	-0.03561
RVH1TWG	-0.00825	0.08420	-0.06078	0.27278	-0.02561	-0.01582	-0.19547	0.08173
RVH1OWAG	0.07925	0.12862	0.00861	0.35494	-0.01568	0.04475	-0.25222	0.08063
RVH1VMWG	-0.00648	0.10423	0.10389	0.22339	-0.01076	0.12865	-0.00874	0.06901
RVH1ADWG	-0.09566	0.01142	-0.12452	0.03537	-0.05235	-0.10185	0.03133	0.04444
TPAS1RVH	0.33251	0.71820	0.44705	0.66076	0.10677	0.40348	0.23447	-0.21945
TPAS1RVM	0.33635	0.68694	0.30756	0.58065	0.11205	0.26452	0.20159	-0.24943
TPAS1PVH	0.17797	0.68559	0.30100	0.63247	0.07757	0.25704	0.10442	-0.16313
PASM1TPS	0.05610	-0.24188	0.60599	-0.20783	-0.04876	0.61223	0.40674	0.12798
TVH1ACC	-0.03678	-0.02324	0.01049	0.19550	-0.03613	-0.00448	-0.14369	-0.09920
RVH1ACC	-0.04375	-0.09837	-0.08901	0.02894	-0.04254	-0.08729	-0.14060	-0.05994
WRV1PVEH	0.05426	0.14562	0.36453	0.05477	-0.01321	0.34141	0.73633	-0.03623
WREV1RVH	0.09961	-0.01382	0.37041	-0.06781	-0.01329	0.35682	0.78197	-0.03121
WORV1RVH	0.09764	-0.01596	0.36887	-0.07283	-0.01153	0.35312	0.78477	-0.03483
WRV1TPAS	-0.04519	-0.45061	0.02290	-0.42994	-0.08349	0.02694	0.55430	0.04840
RVH1WLSB	-0.00971	0.06220	0.07582	0.02685	0.00510	0.08669	-0.01418	0.01573
RVH1WSSB	-0.00481	0.03393	-0.01898	0.04670	-0.01670	-0.02983	0.01433	-0.02519
RVH1WOSB	0.95466	0.33545	0.36134	0.30607	0.00756	0.35616	0.20792	0.09971
RVH1WTSB	0.94727	0.33791	0.37294	0.31608	0.01864	0.37146	0.22582	-0.00023
TPAS1WLO	-0.00745	0.06623	0.07690	0.02778	0.00587	0.08712	-0.00923	0.00797
PAS1WTSB	0.99016	0.39840	0.36785	0.39460	0.03388	0.37221	0.11870	-0.01115
WRV1WTSB	0.38064	0.03617	0.37525	0.02460	-0.01937	0.38086	0.67756	0.05802
WRV1WOSB	0.39097	0.03536	0.37219	0.01778	-0.02449	0.37427	0.68198	0.11777
PAS1WOSB	1.00000	0.41859	0.37018	0.40802	0.03288	0.37057	0.11172	0.01955
PAS1OEXP	0.41859	1.00000	0.52479	0.87185	0.10945	0.44142	0.25735	-0.21493
PASM1OEX	0.37018	0.52479	1.00000	0.47076	0.04322	0.96012	0.51351	-0.01398
PAS1TWAG	0.40802	0.87185	0.47076	1.00000	0.10757	0.46032	0.08704	-0.11929
PAS1FUEL	0.03288	0.10945	0.04322	0.10757	1.00000	0.04621	0.00082	-0.02547
PASM1TEX	0.37057	0.44142	0.96012	0.46032	0.04621	1.00000	0.43676	0.11510
WORV1OXP	0.11172	0.25735	0.51351	0.08704	0.00082	0.43676	1.00000	-0.11108
TREV1TEX	0.01955	-0.21493	-0.01398	-0.11929	-0.02547	0.11510	-0.11108	1.00000

APPENDIX E

FACTOR LOADING MATRICES FOR FOUR 1980 DATA SETS AND ONE 1979 DATA SET

TABLE E-1. FACTOR LOADING MATRIX FOR FINAL FACTOR ANALYSIS ON RAW DATA PERFORMANCE INDICATORS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor
TVH1EMP	0.13607	-0.06445	-0.09637	0.87215	0.06008	0.18614	0.12524	0.18701
RVH1OEMP	0.32279	-0.07904	-0.03455	0.86420	-0.06850	-0.01118	0.13836	-0.04053
TVM1EMP	0.28510	-0.25829	-0.03470	0.44470	0.41043	0.33880	-0.04102	0.23531
PVEH1ADM	-0.06574	-0.05391	-0.04944	0.29784	-0.27706	0.08850	0.02543	0.71185
PVEH1OP	0.15065	-0.24197	0.64931	0.30440	-0.78396	0.07215	0.03674	0.09794
PVEH1MNT	0.02747	-0.08193	-0.05308	0.05412	-0.25570	0.92393	0.01515	0.00290
TVH1PVEH	0.06797	0.09797	-0.02548	0.52355	0.73445	-0.05423	0.07792	-0.19530
TVM1PVEH	0.17096	-0.10785	0.00045	0.05079	0.92785	0.05899	0.01326	-0.09721
TVM1MEXP	0.67693	0.01871	-0.15914	-0.06319	0.14693	0.34002	-0.02830	0.09368
TVM1MNT	0.08557	-0.07386	-0.05546	0.07153	0.16718	0.94776	0.01826	-0.03385
RVH1OEXP	0.87465	-0.01690	-0.12441	0.36232	-0.04711	-0.00424	0.09380	0.02387
TVM1OEXP	0.89551	-0.12596	-0.09479	0.05379	0.16294	0.08541	0.02009	0.02179
RVH1TWG	0.88751	-0.02621	-0.14354	0.30052	-0.02360	0.03173	0.21067	-0.01769
RVH1OWAG	0.87019	-0.01081	-0.12276	0.17579	-0.09048	-0.06074	0.26876	-0.14484
RVH1VMWG	0.53366	0.07541	-0.07792	0.01047	0.11522	0.45674	0.07996	0.20684
TVM1ADWG	0.33971	-0.02659	-0.08490	0.34336	-0.01008	-0.10012	-0.05855	0.18827
TPAS1RVH	-0.36462	0.05963	0.14740	-0.18685	0.02280	-0.04756	-0.06202	-0.00002
TPAS1RVM	-0.38518	0.83021	0.24096	0.03513	-0.10214	-0.10693	-0.07939	-0.10513
TPAS1PVH	-0.29734	0.85103	0.15408	0.07047	0.28858	-0.06388	-0.06149	-0.14223
TVM1ACC	0.20381	-0.06719	-0.02267	-0.00044	0.07116	0.05377	0.94825	-0.00978
RVH1ACC	0.12356	-0.05808	-0.03608	0.20501	-0.05235	-0.00399	0.94358	0.01060
REV1PVEH	-0.33094	0.30253	0.00704	0.06694	0.17661	-0.08188	-0.04520	-0.1761[illegible]
OREV1RVH	-0.37950	0.19195	0.84624	-0.14277	-0.04942	-0.05980	-0.06122	-0.02156
REV1TPAS	-0.08674	-0.45150	0.76428	-0.12799	0.00488	0.00697	0.13547	0.00475
RVH1QSUB	0.62572	0.04989	0.41290	0.10441	-0.07865	-0.04935	0.01609	0.56144
REV1QSUB	0.03397	0.03192	0.75598	-0.12361	-0.04336	-0.04488	-0.01193	0.51497
PAS1QSUB	0.22995	0.59337	0.36164	-0.09788	-0.08118	-0.04218	-0.08383	0.57480
PAS1QEXP	0.24290	0.93164	0.00740	-0.05011	-0.01635	-0.02699	-0.06097	0.09391
PAS1TWAG	0.34936	0.86503	-0.06394	-0.10439	0.02320	0.01654	0.08760	0.07996
ORV1OEXP	-0.01203	0.27900	0.89129	0.01545	-0.10341	-0.05404	-0.09526	0.05086

TABLE E-2. FACTOR LOADING MATRIX FOR FINAL FACTOR ANALYSIS ON LOGARITHMS (BASE 10) OF RAW DATA PERFORMANCE INDICATORS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
TVHLEMP	0.10973	-0.00711	-0.04217	0.91172	-0.03833	0.25868	0.05356
RVHLOEMP	0.22936	-0.07883	-0.01106	0.87974	-0.00603	-0.00379	0.11423
TVMLEMP	0.19935	-0.16876	0.08549	0.51474	0.22502	0.64315	0.00329
PVEHLADM	-0.03767	-0.01963	0.08719	0.46327	-0.54610	0.23542	-0.10097
PVEHLOP	0.09258	-0.18181	0.07845	0.34883	-0.77099	0.07031	0.02752
PVEHLMNT	0.09501	-0.03922	-0.07538	0.07032	-0.49140	0.77225	0.07225
TVHLPVEH	0.03828	0.10735	-0.11512	0.51307	0.77465	-0.05240	0.05704
TVMLPVEH	0.10876	-0.04553	0.01038	0.08577	0.91042	0.28107	0.00534
TVHLMEXP	0.67923	0.01436	-0.09921	-0.11118	0.09450	0.49531	-0.02911
TVMLMNT	0.16287	-0.06764	-0.06566	0.12401	0.12606	0.92739	0.07299
RVHLOEXP	0.87996	0.02666	-0.11454	0.34238	-0.03897	-0.01109	0.13247
TVMLOEXP	0.86493	-0.09781	-0.03313	0.00934	0.14695	0.29502	0.06930
RVHLTWG	0.88881	-0.00713	-0.16201	0.29169	-0.00791	0.00440	0.20104
RVHLOWAG	0.87245	-0.03709	-0.16072	0.17016	-0.01072	-0.11871	0.23491
RVHLVMWG	0.70374	0.05178	-0.20060	0.08813	0.02843	0.39787	0.11755
RVHLADWG	0.40188	0.06936	-0.00814	0.61216	-0.05324	-0.00574	-0.02448
TPASLRVH	-0.39777	0.85443	0.15933	-0.20848	0.05672	-0.02223	-0.13896
TPASLRVM	-0.40363	0.86222	0.10424	-0.02394	-0.07034	-0.20038	-0.10633
TPASLPVH	-0.30521	0.83881	0.10466	0.04834	0.41620	-0.05571	-0.08857
TVMLACC	0.23016	-0.16341	-0.05074	-0.02832	0.07159	0.14637	0.92298
RVHLACC	0.23219	-0.09715	-0.09191	0.14327	-0.02538	-0.01381	0.94451
REVLPVEH	-0.44038	0.21943	0.78476	0.03704	0.31178	-0.08656	-0.06145
OREVLRVH	-0.49068	0.17168	0.81315	-0.18013	-0.01955	-0.05382	-0.09643
REVLTPAS	-0.17798	-0.59787	0.75628	-0.00769	-0.07406	-0.03927	0.02113
RVHLQSUB	0.71194	0.08264	0.43081	0.24800	-0.27518	0.09112	0.03625
REVLQSUB	0.09596	0.19332	0.94259	0.02609	-0.16959	0.01944	-0.05205
PASLQSUB	0.26873	0.74764	0.48084	0.03799	-0.13996	0.05721	-0.08075
PASLOEXP	0.17983	0.95707	0.09471	0.01098	0.03498	-0.03219	-0.05977
PASLTWAG	0.31176	0.91955	0.03721	0.01789	0.05483	-0.02040	0.01745
ORVLOEXP	-0.02656	0.22204	0.92748	-0.00253	-0.05776	-0.07503	-0.04408

TABLE E-3. FACTOR LOADING MATRIX FOR FINAL FACTOR ANALYSIS ON WEIGHTED DATA PERFORMANCE INDICATORS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
TVMIEMP	0.14527	-0.05323	-0.05192	0.89556	-0.01306	0.18744	0.12304
RVHIOEMP	0.28021	-0.07682	-0.05514	0.84451	-0.05068	-0.00674	0.19029
[illegible]	0.31125	-0.26012	0.08900	0.50656	0.31793	0.37069	-0.06009
PVEHIADA	0.03696	-0.02462	0.14654	0.38904	-0.49540	0.15002	-0.08640
PVEHIOP	0.11462	-0.23389	0.07030	0.27323	-0.76991	0.02749	0.09144
PVEHIMNT	0.01573	-0.07922	-0.05954	0.04725	-0.26230	0.90120	0.03939
[illegible]	0.05055	0.08922	-0.09615	0.52603	0.75336	-0.05005	0.08671
TVMIPVEH	0.17674	-0.12570	0.02348	0.09172	0.91912	0.09588	-0.00715
TVMIMEXP	0.71157	0.01965	-0.09165	-0.03421	0.09891	0.35024	-0.05003
TVMIMNT	0.08159	-0.07742	-0.05642	0.08329	0.15171	0.94298	0.02809
[illegible]	0.87323	-0.01344	-0.08706	0.36180	-0.05542	-0.02565	0.11457
TVMIOEXP	0.90598	-0.13058	-0.02355	0.07077	0.14829	0.08229	0.02408
RVHITWG	0.87443	-0.02766	-0.11562	0.29671	-0.01478	0.01176	0.23495
RVHIOWAG	0.83410	-0.01812	-0.12971	0.14901	-0.03232	-0.09970	0.32198
[illegible]	0.57575	0.08501	-0.01053	0.05090	0.03257	0.47465	0.04317
RVHIADWG	0.36727	-0.01904	-0.02208	0.37009	-0.07292	-0.09126	-0.08527
TPASIRVH	-0.37062	0.85578	0.13472	-0.19481	0.04849	-0.04151	-0.06541
TPASIRVM	-0.41556	0.82987	0.14765	-0.00409	-0.04532	-0.13055	-0.05615
[illegible]	-0.32181	0.84273	0.08420	0.05091	0.30046	-0.06856	-0.04719
TVMIACC	0.20617	-0.06922	-0.01633	-0.00195	0.06955	0.06870	0.93293
RVHIACC	0.12259	-0.05434	-0.04083	0.10593	-0.05904	0.00237	0.93332
WRVIPVEH	-0.39692	0.23732	0.77492	0.06566	0.30170	-0.07970	-0.02432
[illegible]	[illegible]0.41352	0.12914	0.84345	-0.15062	-0.02683	-0.03681	-0.06190
WRVITPAS	-0.11697	-0.49906	0.70423	-0.09569	0.04236	0.01298	0.17202
RVHIWQSB	0.61532	0.09626	0.65955	0.16806	-0.19307	-0.02367	0.00306
WRVIWQSB	-0.00474	0.05019	0.92540	-0.05530	-0.11946	-0.01697	-0.01933
[illegible]	0.19007	0.67303	0.61486	-0.02889	-0.16879	-0.01777	-0.09631
PASIOEXP	0.25149	0.93042	0.07165	-0.04266	-0.01496	-0.02369	-0.06099
PASITWAG	0.35735	0.86048	0.01280	-0.09339	0.02804	0.02569	0.08833
WDRVIOXP	0.03411	[illegible]22097	0.90634	0.05084	-0.10380	-0.03990	-0.08535

TABLE E-4. FACTOR LOADING MATRIX FOR FINAL FACTOR ANALYSIS ON LOGARITHMS (BASE 10) OF WEIGHTED DATA PERFORMANCE INDICATORS

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
[redacted]	0.10867	-0.01353	-0.01638	0.91249	-0.03391	0.25879	0.05234
RVHLOEMP	0.22795	-0.08535	0.00757	0.88071	-0.00432	-0.00282	0.10966
TVHLEMP	0.17857	-0.18789	0.11969	0.51197	0.24226	0.63490	0.01545
PVEHLADM	-0.05300	-0.02920	0.09786	0.46583	-0.53795	0.23462	-0.09251
[redacted]	0.08214	-0.19087	0.10167	0.35145	-0.76686	0.07425	0.02769
PVEHLMNT	0.10135	-0.04712	-0.03962	0.07518	-0.48899	0.77920	0.06594
TVHLPVEH	0.04946	0.11393	-0.12227	0.51064	0.77149	-0.05513	0.05236
TVMLPVEH	0.10153	-0.05960	0.01079	0.08054	0.41864	0.27142	0.01284
[redacted]	0.69187	0.01076	-0.06794	-0.10572	0.08695	0.49834	-0.03614
TVMLMNT	0.16420	-0.07856	-0.03100	0.12525	0.13379	0.92772	0.07190
RVHLOEXP	0.89283	0.02145	-0.06785	0.34961	-0.04795	-0.00645	0.12410
TVMLOEXP	0.86410	-0.11121	0.01772	0.01329	0.14722	0.29356	0.07126
[redacted]	0.89505	-0.01707	-0.11696	0.30007	-0.01183	0.01142	0.19415
RVHLOWAG	0.87802	-0.04608	-0.12020	0.17819	-0.01509	-0.11121	0.22778
RVHLVMWG	0.72631	0.05289	-0.17267	0.09652	0.01605	0.40667	0.10194
RVHLADWG	0.59434	0.0580[illegible]	0.02294	0.61603	-0.04862	-0.00555	-0.02010
[redacted]	-0.39404	0.65657	0.15716	-0.20860	0.05581	-0.01828	-0.13957
TPASLRVM	-0.39225	0.86943	0.09875	-0.02255	-0.07652	-0.19341	-0.11245
TPASLPVH	-0.29746	0.84301	0.10025	0.04742	0.41419	-0.05349	-0.09132
TVMLACC	0.22759	-0.16164	-0.04989	-0.02880	0.07454	0.14353	0.92746
[redacted]	0.23670	-0.09412	-0.09337	0.14447	-0.02723	-0.01347	0.94361
WRVLPVEH	-0.43048	0.21065	0.78755	0.07135	0.38026	-0.04180	-0.07609
WORVLRVM	-0.49039	0.15112	0.82160	-0.16701	-0.03207	-0.00723	-0.12186
WRVLTPAS	-0.14638	-0.64845	0.72613	0.02621	-0.07970	0.01119	0.01413
[redacted]	0.65041	0.09895	0.60672	0.20467	-0.20710	-0.00322	0.09214
WRVLWOSB	0.03204	0.17486	0.96646	0.00433	-0.14264	-0.00622	-0.02636
PASLWOSB	0.16319	0.76345	0.57513	-0.01684	-0.10811	-0.01711	-0.04438
PASLOEXP	0.19295	0.95575	0.12508	0.01591	0.02768	-0.02460	-0.06631
[redacted]	0.32101	0.91356	0.07253	0.02476	0.05057	-0.01026	0.01103
WORVLUXP	0.02722	0.19890	0.95250	0.04110	-0.07255	-0.01531	-0.06156

TABLE E-5. FACTOR LOADING MATRIX FOR FY 1979 FACTOR ANALYSIS

CONCEPT	VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	FACTOR 9
V	RVH/OWAG	.935	.000	.000	.000	.000	.000	.000	.000	.000
V	RVH/OEXP	.927	.000	.000	.000	.000	.000	.000	.000	.000
V	RVH/TWG	.924	.000	.000	.000	.000	.000	.000	.000	.000
IX	TREV/RVH	-.807	.000	.000	.000	.000	.000	.000	.000	.000
IX	REV/RVH	-.705	.000	.000	.000	.000	.000	.000	.000	.000
VI	TPAS/RVM	.000	.889	.000	.000	.000	.000	.000	.000	.000
XI	PAS/OEXP	.000	.887	.000	.000	.000	.000	.000	.000	.000
VI	TPAS/PVH	.000	.877	.000	.000	.000	.000	.000	.000	.000
XI	PAS/TWAG	.000	.866	.000	.000	.000	.000	.000	.000	.000
VI	TPAS/RVH	.000	.859	.000	.000	.000	.000	.000	.000	.000
IX	REV/PVEH	-.489	.502	.000	.000	.000	.000	.000	.000	.000
II	TVM/PVEH	.000	.000	.885	.000	.000	.000	.000	.000	.000
II	TVH/PVEH	.000	.000	.877	.000	.000	.000	.000	.000	.000
I	PVEH/OP	.000	.000	-.802	.000	.000	.000	.000	.000	.000
II	TVH/AVEH	.000	.000	.633	.000	.000	.000	.000	.000	.000
III	TVM/FUEL	.000	.000	.000	.987	.000	.000	.000	.000	.000
III	RVM/FUEL	.000	.000	.000	.986	.000	.000	.000	.000	.000
III	PAS/FUEL	.000	.000	.000	.958	.000	.000	.000	.000	.000
X	REV/OSUB	.000	.000	.000	.000	.989	.000	.000	.000	.000
X	PAS/OSUB	.000	.000	.000	.000	.978	.000	.000	.000	.000
X	RVH/OSUB	.000	.000	.000	.000	.977	.000	.000	.000	.000
VII	RVH/POP	.000	.000	.000	.000	.000	.926	.000	.000	.000
VII	TPAS/POP	.000	.000	.000	.000	.000	.865	.000	.000	.000
VII	TPAS/ELD	.000	.000	.000	.000	.000	.851	.000	.000	.000
IV	TVM/MNT	.000	.000	.000	.000	.000	.000	.934	.000	.000
I	PVEH/MNT	.000	.000	-.483	.000	.000	.000	.784	.000	.000
IV	TVEH/MEXP	.540	.000	.000	.000	.000	.000	.643	.000	.000
I	TVM/EMP	.000	.000	.497	.000	.000	.000	.643	.000	.000
XII	REV/OEXP	.000	.000	.000	.000	.000	.000	.000	.960	.000
XII	REV/TEX	.000	.000	.000	.000	.000	.000	.000	.933	.000
VIII	RVH/ACC	.000	.000	.000	.000	.000	.000	.000	.000	.944
VIII	TVM/ACC	.000	.000	.000	.000	.000	.000	.000	.000	.924
	VP	5.217	4.667	3.606	3.052	3.011	2.677	2.623	2.036	1.948

The above factor loading matrix has been rearranged so that the columns appear in decreasing order of variance explained by factors. The rows have been rearranged so that for each successive factor, loadings greater than .5000 appear first. Loadings less than .4500 have been replaced by zero.

APPENDIX F

DESCRIPTIVE STATISTICS FOR PERFORMANCE PEER GROUPS

Peer Group (n)		RVH/OEXP	TPAS/RVH	OREV/OEXP
Group 1	mean	0.029	61.758	0.330
(24)	s.d.	0.008	14.341	0.122
Group 2	mean	0.072	13.780	0.215
(13)	s.d.	0.010	5.030	0.044
Group 3	mean	0.047	17.278	0.145
(17)	s.d.	0.007	8.043	0.034
Group 4	mean	0.036	32.997	0.237
(42)	s.d.	0.005	5.813	0.048
Group 5	mean	0.040	41.248	0.420
(36)	s.d.	0.007	8.216	0.053
Group 6	mean	0.053	24.582	0.348
(66)	s.d.	0.008	6.882	0.090
Group 7	mean	0.033	39.516	0.716
(15)	s.d.	0.009	12.926	0.154
unclustered	mean	0.071	43.475	0.639
(7)	s.d.	0.036	28.140	0.284

		TVM/ACC	PVEH	PKTOBS	SPEED
Group 1	mean	1.972	625.500	1.974	13.374
(24)	s.d.	3.029	592.302	0.655	2.410
Group 2	mean	4.742	28.769	1.526	13.126
(13)	s.d.	4.748	45.710	0.751	3.148
Group 3	mean	2.628	25.294	1.605	11.727
(17)	s.d.	1.929	25.455	0.823	2.537
Group 4	mean	1.848	102.140	1.594	13.923
(42)	s.d.	0.840	118.674	0.530	1.946
Group 5	mean	1.588	119.343	1.930	13.068
(36)	s.d.	0.957	155.199	0.650	1.568
Group 6	mean	2.319	33.015	1.473	13.441
(66)	s.d.	1.852	40.147	0.572	2.410
Group 7	mean	1.902	188.533	2.446	13.270
(15)	s.d.	1.504	284.301	0.928	4.271
unclustered	mean	2.609	498.429	1.326	12.233
(7)	s.d.	1.770	1269.923	0.630	4.362

APPENDIX F (continued)

		TVH/EMP	TVM/PVEH	TVM/MNT
Group 1	mean	0.094	4.142	6.616
(24)	s.d.	0.019	0.978	2.394
Group 2	mean	0.154	4.431	16.1992
(13)	s.d.	0.031	1.566	19.5102
Group 3	mean	0.120	4.412	11.7037
(17)	s.d.	0.024	2.411	11.4433
Group 4	mean	0.114	4.283	10.0824
(42)	s.d.	0.029	1.003	8.9546
Group 5	mean	0.114	3.858	8.2342
(36)	s.d.	0.022	0.8979	2.4065
Group 6	mean	0.118	4.388	9.9930
(66)	s.d.	0.020	1.236	4.6703
Group 7	mean	0.111	3.735	8.0744
(15)	s.d.	0.024	1.692	3.0784
unclustered	mean	0.136	3.588	7.2172
(7)	s.d.	0.062	1.600	3.7552

APPENDIX G

DESCRIPTIVE STATISTICS FOR PERFORMANCE INDICATORS

BY PEER GROUP

Peer Group (n)		RVH/OEXP	TPAS/RVH	OREV/OEXP	TVH/EMP
1 (2)	mean	.031	11.5	.64	.073
	s.d.	.008	---	.32	.020
	minimum	.025	11.5	.42	.060
	maximum	.036	11.5	.87	.087
2 (16)	mean	.040	32.1	.34	.095
	s.d.	.014	25.6	.21	.024
	minimum	.022	9.7	.11	.047
	maximum	.073	84.5	.81	.133
3 (44)	mean	.051	30.1	.26	.118
	s.d.	.013	16.0	.09	.025
	minimum	.026	9.2	.11	.028
	maximum	.090	81.3	.47	.170
4 (7)	mean	.042	37.9	.30	.112
	s.d.	.010	18.2	.14	.014
	minimum	.029	21.0	.09	.100
	maximum	.055	71.5	.48	.140
5 (15)	mean	.056	24.6	.42	.145
	s.d.	.019	11.2	.29	.049
	minimum	.031	6.3	.08	.053
	maximum	.103	49.9	1.10	.229
6 (45)	mean	.055	28.8	.31	.124
	s.d.	.017	13.2	.14	.031
	minimum	.031	5.4	.09	.031
	maximum	.121	73.5	.76	.220
7 (78)	mean	.045	31.8	.36	.117
	s.d.	.010	10.6	.17	.017
	minimum	.030	5.0	.11	.055
	maximum	.074	58.0	1.11	.166
8 (33)	mean	.040	32.1	.34	.095
	s.d.	.012	14.2	.19	.023
	minimum	.020	7.1	.12	.045
	maximum	.073	54.8	.93	.170
9 (8)	mean	.030	40.1	.24	.099
	s.d.	.009	15.4	.13	.023
	minimum	.015	19.0	.07	.074
	maximum	.045	72.2	.42	.14
10 (8)	mean	.035	46.9	.34	.106
	s.d.	.008	25.2	.11	.014
	minimum	.026	26.1	.19	.080
	maximum	.048	89.8	.49	.128

Peer Group		RVH/OEXP	TPAS/RVH	OREV/OEXP	TVH/EMP
11 (13)	mean	.025	52.9	.348	.095
	s.d.	.003	10.1	.120	.012
	minimum	.020	36.0	.178	.066
	maximum	.030	69.0	.587	.116
12 (3)	mean	.026	74.7	.581	.098
	s.d.	.001	14.1	.210	.014
	minimum	.025	58.5	.386	.085
	maximum	.027	83.4	.807	.113
Total	mean	.045	32.8	.337	.113
	s.d.	.015	16.5	.290	.030
	minimum	.015	5.4	.070	.028
	maximum	.121	89.8	1.100	.229

Peer Group		TVM/PVEH	TVM/MNT	TVM/ACC
1	mean	3.0	15.3	2.9
	s.d.	---	.1	.2
	minimum	3.0	15.2	2.7
	maximum	3.0	15.4	3.0
2	mean	5.8	11.9	4.1
	s.d.	1.4	5.0	2.2
	minimum	3.5	5.8	1.5
	maximum	8.8	25.7	7.4
3	mean	5.3	9.7	2.6
	s.d.	.9	3.9	1.5
	minimum	3.7	1.8	.8
	maximum	8.2	23.5	7.1
4	mean	4.5	9.6	3.0
	s.d.	.8	3.6	2.0
	minimum	3.6	5.3	.9
	maximum	5.6	14.0	6.9
5	mean	3.0	6.1	1.7
	s.d.	.7	2.8	.9
	minimum	1.4	.6	.4
	maximum	4.4	12.3	3.7
6	mean	4.5	9.2	2.3
	s.d.	.9	3.7	1.6
	minimum	2.2	3.0	.6
	maximum	7.6	21.4	7.3
7	mean	3.6	9.2	1.9
	s.d.	.7	3.4	1.2
	minimum	2.3	4.6	.7
	maximum	6.1	24.3	8.0
8	mean	3.0	7.5	1.4
	s.d.	.6	2.7	.7
	minimum	1.6	2.5	.5
	maximum	4.3	15.6	3.8
9	mean	5.7	7.5	1.5
	s.d.	1.1	2.2	.5
	minimum	4.4	3.7	.6
	maximum	8.0	11.0	2.2
10	mean	4.4	6.2	1.5
	s.d.	.5	2.8	.9
	minimum	3.7	3.7	.5
	maximum	5.1	11.5	3.3

Peer Group		TVM/PVEH	TVM/MNT	TVM/ACC
11	mean	3.9	6.7	1.23
	s.d.	.6	2.0	.72
	minimum	3.1	2.6	.50
	maximum	5.0	11.0	3.00
12	mean	4.2	5.4	1.00
	s.d.	1.1	2.1	.34
	minimum	3.2	3.2	.74
	maximum	5.4	7.4	1.40
Total	mean	4.2	9.90	2.40
	s.d.	1.5	8.40	2.30
	minimum	1.1	.68	.48
	maximum	8.0	25.70	8.00

Note: The figures for the total are for the entire set of transit systems reporting Section 15 data for FY1980. Thus some of the transit systems included in the total are not in a peer group because they were missing data and could not be assigned to a peer group.

APPENDIX H

TECHNICAL NOTES ON THE CORRELATIONAL ANALYSIS OF THE RELATIONSHIP BETWEEN OPERATING CHARACTERISTICS AND PERFORMANCE

The relationship between operating characteristics and performance was explored in two major ways. In the first, the peer groups as a whole were considered the unit of analysis. With this approach, the important distinctions in size, speed and peak to base ratio as made by the cluster analysis were preserved. In the second way, the individual transit systems were the units of analysis. This allowed for more complex multiple regression analyses because there were more cases and the data was true interval level data.

The peer group data was analyzed with a series of Spearman rank order correlations between each operating characteristic and each of the seven performance indicators. The mean of each variable for each peer group was used as the value in the correlation.

Size as measured by peak vehicles was the most important variable, being significantly correlated to four performance indicators each. Of the performance indicators, only Revenue Generation was not significantly correlated with any of the four operating characteristics. Labor Efficiency was only correlated with speed. Each other performance indicator was correlated with several operating characteristics.

The rank order correlation is not an entirely satisfactory version of the relations between operating and performance variables. It does not correct for correlations between the operating variables, some of which are substantially correlated. For instance, peak vehicles and total vehicle miles have a Pearson's correlation of .98. It is not clear whether each of these variables makes an independent contribution to explaining variance in performance, or if they are essentially redundant.

The second phase of analysis involved a series of multiple regressions with the four operating characteristics as independent variables and each performance indicator as the criterion. Since the operating variables have extremely non-normal distributions, as noted in Chapter 1, the regression analysis was done with the raw data and with log 10 transformations of the independent variables. A stepwise procedure was used in which the independent variables entered the equation in order of their relative importance in accounting for variance in the criterion.

The results with the transformed variables are given in Chapter 4 in Table 4-3. In summary, size as measured by number of peak vehicles was the most important independent variable, although each operating variable contributed significant explanatory power for several performance indicators. Only Revenue Generation was not significantly correlated with operating characteristics. In general these results are quite similar to those for the rank order correlations.

The results with the untransformed data differ in several ways from the other correlation results. Peak to base ratio is the most important independent variable, being most important in three equations and entering into all but one of the equations. The correlation coefficients are lower with the untransformed data, except for Revenue Generation which reaches significance in this analysis.

However, all the analyses are in accord with the conclusion that each of the operating characteristics makes a significant contribution to explaining differences in performance; and that performance in Revenue Generation is least related to operating characteristics.

Before a definitive statement can be made on the relation between opearting characteristics and performance, several analytical issues need to be explored further:

- Were the optimal transformations done on the data for each variable?
- How can non-linear relationships be better described in this context?
- Is there a problem with multi-collinearity between the number of peak vehicles and total vehicle miles?
- How can the numerous suppression effects be better understood? While the most common suppression was between peak vehicles and total vehicle miles (which is not unexpected in light of their high correlation) there was evidence of suppression between other variables as well.

APPENDIX I

FY 1980 DATA FOR TRANSIT AGENCIES BY PEER GROUP

KEY

COLUMN	VARIABLE	
1	ID	: Transit System ID Number
2	PEER GROUP	: Peer Group ID Number
3	URBAN AREA	: Urban Area Code
4	PVEH	: Peak Vehicles
5	TVM	: Total Vehicle Miles
6	SPEED	: Miles per Hour
7	PKTOBS	: Peak to Base Ratio
8	RVH/OEXP	: Revenue Vehicle Hours per Operating Expense
9	TPAS/RVH	: Passenger Trips per Revenue Vehicle Hour
10	OREV/OEXP	: Weighted Operating Revenue per Operating Expense
11	TVH/EMP	: Vehicle Hours per Employee
12	TVM/PVEH	: Total Vehicle Miles per Peak Vehicle
13	TVM/MNT	: Vehicle Miles per Maintenance Employee
14	TVM/ACC	:Total Vehicle Miles per Accident

Notes:
1. A '-9' indicates missing data.
2. Total Vehicle Miles and Total Vehicle Hours are in units of 10,000.

If your copy of this Appendix
is illegible, request ITS
Working Paper 83-5 from:
Institute of Trans. Studies
Univ. of Calif., Irvine
Irvine, CA 92717

PEER_GROUP_1

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PK HRS	RVH/DEXP	TPAS/RVH	WKRV/DXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
10	1	205	1	3	30.00	1.00	.025	11.50	.423	.060	3.04	15.18	3.0
6005	1	15	25	384	25.77	1.04	.036	-9.00	.870	.087	15.38	15.38	2.7

PEER_GROUP_2

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PK HRS	RVH/DEXP	TPAS/RVH	WKRV/DXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
12	2	170	28	168	17.21	1.40	.022	25.11	.130	.120	6.00	9.95	3.2
2010	2	177	9	52	17.21	1.29	.056	11.35	.269	-9.000	5.80	-9.00	-9.0
2058	2	1	46	406	20.49	2.30	.027	9.70	.750	.102	8.82	10.56	6.4
2065	2	61	2	14	21.34	1.00	.044	17.10	.192	.047	7.24	14.48	3.6
2069	2	63	2	7	19.71	1.00	.023	-9.00	.258	.108	3.62	-9.00	7.2
2074	2	1	16	74	19.19	1.00	.024	84.51	.805	.073	4.64	10.45	2.2
3021	2	4	3	18	17.79	1.00	.068	-9.00	.511	.110	5.85	5.85	3.5
4014	2	160	12	42	19.33	1.00	.046	20.49	.529	.082	3.50	8.09	4.2
4038	2	121	16	93	18.05	1.00	.049	17.32	.395	.092	5.83	9.32	1.8
5018	2	197	11	48	18.47	1.38	.035	-9.00	.210	.106	4.38	16.08	6.
5023	2	35	1	7	21.25	1.00	.048	-9.00	.305	.070	6.58	10.96	6.5
5029	2	215	15	79	20.75	1.36	.037	82.23	.111	.061	5.25	9.26	-9.
5084	2	3	24	124	20.49	2.00	.031	24.86	.197	.092	5.16	9.53	2.3
7007	2	78	1	7	21.01	1.00	.060	17.97	.177	.133	7.44	-9.00	7.4
9010	2	2	19	129	18.95	1.06	.030	30.82	.272	.119	6.77	25.74	1.4
9018	2	193	20	119	18.33	1.11	.038	44.32	.361	.106	5.93	14.81	1.6

PEER GROUP 3

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PKTOBS	RVH/DEXP	PAS/RVH	WORV/OXP	IVM/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
7	3	144	52	335	13.76	1.18	.040	26.49	.214	.117	6.44	10.03	2.81
1007	3	257	9	52	15.39	1.00	.052	29.37	.472	.140	5.77	-9.00	5.19
1052	3	152	2	10	15.38	1.00	.090	12.38	.172	.115	5.17	6.90	5.17
2015	3	112	4	4	13.66	.80	-9.000	79.06	-9.000	.028	-9.00	1.76	1.46
2049	3	1	7	45	15.19	1.00	.054	22.27	.448	.129	6.47	8.23	3.77
2060	3	76	72	270	15.64	1.01	.026	39.67	.354	.083	3.75	6.32	1.16
2072	3	-9	11	69	15.46	1.00	-9.000	-9.00	-9.000	-9.000	6.27	-9.00	-9.00
3003	3	251	5	25	15.31	1.00	.058	-9.00	-9.000	.134	5.04	8.39	5.04
3008	3	236	19	110	13.81	1.00	.054	30.16	.388	.116	5.81	7.88	.89
3012	3	185	26	118	13.99	1.24	.043	81.26	.291	.102	4.55	6.97	2.89
3020	3	4	4	18	13.73	1.00	.056	-9.00	.361	.100	4.53	6.05	2.59
3026	3	254	11	64	14.42	1.10	.048	28.42	.254	.125	5.83	9.16	1.23
3027	3	159	15	70	15.57	1.50	.034	19.04	.166	.108	4.68	7.98	2.51
4009	3	123	18	81	13.60	1.38	.061	19.75	.331	.107	4.52	6.26	2.81
4012	3	141	44	184	14.64	1.42	.042	-9.00	.297	.119	4.17	10.80	1.61
4021	3	222	6	30	14.70	1.20	.045	55.52	.337	.080	5.06	6.19	2.33
4028	3	243	12	74	14.98	1.00	.054	14.55	.279	.115	6.19	14.02	3.91
4030	3	242	30	123	15.36	1.00	.063	25.41	.373	.109	4.09	11.14	1.33
4032	3	168	24	114	13.78	1.00	.062	30.78	.320	.119	4.76	9.44	3.68
4045	3	202	4	19	14.87	1.00	.064	27.37	.240	.084	4.77	4.77	1.59
5002	3	155	17	109	14.57	1.13	.057	16.16	.321	.143	6.40	12.10	2.42
5034	3	213	11	58	15.00	1.00	.049	36.46	.212	.109	5.25	10.51	-9.00
5041	3	211	13	53	13.74	1.00	.044	-9.00	.117	.085	4.11	4.86	1.30
5043	3	142	16	85	14.44	1.00	.068	24.96	.401	.109	5.30	12.29	-9.00
5047	3	241	12	62	13.94	1.00	.048	-9.00	.170	.123	5.16	11.25	1.12
5057	3	78	26	120	14.08	1.44	.035	25.12	.129	.142	4.62	-9.00	1.37
5068	3	3	19	79	13.71	1.36	.045	28.80	.265	.158	4.18	-9.00	1.12
5074	3	3	4	20	14.77	1.33	.051	20.73	.285	.122	5.10	10.21	1.70
6001	3	158	17	82	15.13	1.31	.063	13.33	.172	.164	4.85	-9.00	2.23
6010	3	135	26	109	13.92	1.24	.066	9.16	.167	.126	4.19	10.57	-9.00
6019	3	71	73	296	14.37	1.24	.041	35.75	.261	.113	4.06	8.23	2.11
6025	3	219	12	64	14.20	1.00	.059	22.29	.331	.128	5.30	9.79	2.54
6035	3	184	4	28	14.26	1.00	.069	15.47	.380	.121	7.07	11.79	7.07
6037	3	253	5	27	14.57	1.00	.073	17.70	.196	.170	5.41	-9.00	4.51
9004	3	115	21	105	13.66	1.05	.044	38.60	.238	.134	5.00	19.43	2.62
9005	3	258	6	30	13.54	1.00	.034	24.50	.191	.091	4.97	7.64	1.24
9006	3	227	43	353	14.64	1.00	.045	26.72	.159	.117	8.21	33.54	2.94
9007	3	173	12	87	15.65	1.00	.050	22.22	.190	.119	7.27	12.46	5.81
9017	3	225	9	47	14.30	1.00	.046	27.89	.173	.148	5.26	-9.00	2.15
9022	3	2	16	101	15.28	1.00	.040	-9.00	.111	.140	6.29	12.57	1.70
9027	3	80	74	435	14.01	1.35	.040	34.02	.192	.116	5.88	7.34	1.18
9028	3	6	15	72	14.26	1.00	.044	32.09	.227	.137	4.81	11.09	2.49
9035	3	86	22	118	15.08	1.10	.033	41.19	.265	.105	5.36	12.28	-9.00
9041	3	2	21	113	14.16	1.05	.031	57.26	.303	.115	5.37	9.99	1.66

PEER_GROUP_4

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PK/BUS	RVH/DEXP	UPAS/RVH	WHRV/OXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
4	4	28	10	42	16.75	1.11	.029	21.00	.097	.105	4.15	13.85	6.9
1043	4	128	24	89	15.99	1.14	.042	51.24	.440	.100	3.70	8.08	2.9
2064	4	1	11	39	16.26	1.00	.054	21.42	.355	.121	3.57	6.54	3.5
4035	4	69	38	214	15.88	1.15	.055	28.44	.482	.141	5.63	13.99	1.2
5030	4	217	14	65	15.95	1.08	.037	38.95	.181	.111	4.62	5.30	2.3
8007	4	176	11	50	16.03	1.10	.041	33.06	.235	.107	4.55	7.36	.9
9020	4	153	47	258	16.76	1.09	.034	71.50	.340	.102	5.48	12.26	2.8

PEER_GROUP_5

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PK/BUS	RVH/DEXP	UPAS/RVH	WHRV/OXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
2	5	93	71	240	8.20	1.58	.056	27.38	.328	.195	3.38	8.58	2.6
2029	5	1	29	87	9.08	1.45	.043	49.91	.829	.122	3.00	7.01	2.3
2041	5	1	34	78	9.44	1.79	.033	24.45	1.073	-9.000	2.28	-9.00	1.3
2046	5	1	107	318	7.96	2.10	.031	-9.00	.803	.104	2.98	4.49	1.1
2061	5	1	6	18	8.15	1.00	.060	16.12	.440	.157	2.94	5.39	2.5
2163	5	1	8	23	9.61	1.33	.055	18.21	.308	.125	2.86	4.58	1.2
2070	5	1	20	53	7.93	1.18	.070	6.36	.245	.228	2.66	6.64	-9.0
3017	5	95	1	1	7.50	1.00	.103	30.00	.084	.229	1.44	-9.00	1.4
3024	5	117	41	150	9.38	1.37	.050	34.56	.438	.148	3.67	6.63	2.9
5013	5	21	1	4	9.32	1.00	.040	21.83	.080	.147	3.70	12.33	3.7
5021	5	203	6	11	9.22	1.50	.076	-9.00	.333	.094	1.90	3.80	1.0
6009	5	239	20	88	9.57	1.25	.072	-9.00	.585	.172	4.38	5.47	.8
6030	5	26	1	3	8.00	1.00	.076	-9.00	.410	.053	3.39	.68	.4
6033	5	99	38	142	9.31	1.73	.055	19.36	.283	.119	3.77	6.99	1.4
8004	5	234	4	38	10.86	.57	.055	22.23	.165	.141	-9.00	-9.00	1.1

PEER GROUP 6

ID	PEER GROUP	URBAN AREA	#VEH	TVM	SPEED	PT HRS	RVH/DEXP	PAS/RVH	WHRV/DXP	TVM/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
1004	6	138	36	148	11.55	1.00	.031	33.84	.219	.093	4.10	5.20	2.5
1015	6	250	15	32	12.21	1.00	.077	26.23	.574	.092	3.16	3.60	4.6
1047	6	152	8	35	12.73	1.33	.051	-9.00	.390	.122	4.35	10.84	5.7
1050	6	57	34	140	11.42	1.26	.046	-9.00	.370	.101	4.10	9.97	1.6
1057	6	171	17	79	12.06	1.06	.051	24.87	.338	.121	4.66	21.41	1.5
2009	6	177	5	24	13.05	1.00	.047	42.30	.476	.167	4.77	-9.00	1.4
2051	6	1	13	64	12.14	1.00	.088	14.81	.268	.183	4.95	12.87	5.3
2055	6	1	9	37	11.45	1.13	.059	21.53	.531	.135	4.15	9.33	1.9
3002	6	118	19	76	12.48	1.00	.046	22.07	.285	.111	3.99	6.32	2.1
3007	6	129	29	133	11.98	1.12	.056	23.82	.336	.126	4.58	10.21	1.0
3009	6	179	10	37	11.95	1.00	.073	33.36	.763	.129	3.70	7.40	1.2
3025	6	105	30	130	12.69	1.25	.040	37.45	.313	.114	4.33	7.22	3.2
4006	6	270	11	52	12.54	1.22	.055	35.41	.324	.126	4.75	5.80	-9.0
4023	6	137	21	97	12.07	1.31	.053	-9.00	.309	.112	4.60	6.66	1.5
4025	6	122	41	188	12.01	1.28	.058	30.44	.566	.121	4.60	6.73	1.7
4026	6	119	11	63	12.04	1.00	.082	21.08	.233	.131	5.71	7.86	3.4
4027	6	48	35	128	11.38	1.00	.032	44.34	.232	.075	3.66	9.79	1.2
4029	6	40	105	600	13.00	1.07	.045	-9.00	.226	.118	5.72	9.82	1.0
4036	6	212	19	35	11.64	1.19	.066	27.06	.328	.109	4.45	7.04	3.8
4037	6	74	49	288	12.69	1.00	.072	13.19	.177	.151	5.87	11.51	.8
5006	6	164	21	103	11.99	1.00	.058	20.39	.352	.160	4.89	11.81	1.7
5019	6	35	8	36	12.18	1.00	.079	-9.00	.274	.157	4.47	14.31	1.0
5035	6	134	37	152	11.45	1.16	.052	19.37	.171	.160	4.10	9.85	1.0
5037	6	174	12	92	12.89	1.00	.051	6.73	.145	.129	7.63	-9.00	3.9
5051	6	212	12	58	13.26	1.09	.051	26.91	.201	.102	4.80	9.45	1.3
5054	6	200	11	54	11.68	1.00	.047	-9.00	.166	.115	4.95	6.80	1.5
5060	6	180	34	161	13.13	1.26	.039	25.38	.244	.112	4.75	8.87	1.9
5061	6	181	13	54	12.85	1.08	.038	31.26	.209	.103	4.17	6.33	1.3
5063	6	92	17	79	12.41	1.00	.051	28.37	.242	.129	4.65	9.88	1.5
5065	6	85	2	7	11.61	1.00	.039	5.41	.098	.096	3.63	4.84	7.2
5090	6	220	6	23	11.01	1.00	.052	8.67	.100	.109	3.83	11.48	2.8
5093	6	238	8	35	12.54	1.00	.072	-9.00	.233	.106	4.32	17.27	1.2
5097	6	-9	24	63	11.69	1.26	.048	73.50	.219	.031	2.61	2.98	2.[illegible]
6015	6	259	11	52	12.49	1.10	.055	47.95	.474	.138	4.72	6.93	1.6
6016	6	167	13	56	12.36	1.18	.063	29.70	.360	.122	4.28	6.95	.7
6023	6	201	9	27	11.51	1.29	.072	31.89	.389	.098	3.05	8.31	1.[illegible]
6026	6	199	13	64	12.27	1.08	.054	28.16	.266	.134	4.94	6.43	1.[illegible]
6034	6	261	3	51	13.07	1.00	.121	17.01	.512	.216	6.72	12.64	5.[illegible]
6038	6	214	12	46	11.82	1.20	.062	26.71	.273	.130	3.82	9.16	5.[illegible]
7009	6	78	18	71	10.79	1.13	.037	-9.00	.128	.096	3.96	3.17	1.4
9001	6	182	5	34	12.46	1.00	.039	29.51	.459	.118	6.72	11.2[illegible]	.[illegible]
9008	6	2	92	359	11.43	1.23	.042	55.87	.523	.158	3.91	14.[illegible]	.[illegible]
9009	6	6	192	851	13.49	1.10	.033	31.93	.243	.170	4.43	13.29	.[illegible]
9012	6	124	75	333	12.60	1.30	.041	26.96	.230	.123	3.80	-9.[illegible]	[illegible]
9023	6	2	129	605	12.64	1.39	.026	41.07	.248	.135	4.6[illegible]	13.15	[illegible]

PEER_GROUP_7

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PT/BUS	RVH/OEXP	PAS/RVH	WKRV/OXP	IVH/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
3	7	67	92	320	13.44	2.24	.038	31.64	.331	.108	3.48	8.99	1.6
5	7	17	15	70	14.12	1.67	.040	5.01	.123	.135	4.69	24.23	-9.0
6	7	252	9	34	15.22	1.80	.030	-9.00	.184	.102	3.73	-9.00	1.5
9	7	194	30	110	12.07	1.67	.042	32.74	.123	.127	3.66	11.43	1.8
11	7	205	20	77	13.72	1.67	.053	-9.00	.209	.124	3.85	13.05	1.0
1001	7	31	204	819	13.23	1.59	.037	50.79	.342	.107	4.01	8.61	.9
1005	7	111	21	71	10.40	1.62	.045	34.11	.284	.135	3.36	11.76	2.8
1006	7	149	73	190	10.32	2.09	.046	32.26	.168	.115	2.61	6.56	1.2
1008	7	47	156	505	12.48	1.86	.037	33.62	.227	.096	3.24	8.70	1.9
1014	7	84	65	216	11.68	1.48	.033	37.31	.365	.110	3.33	7.89	1.46
1016	7	172	47	135	12.10	1.88	.047	30.91	.501	.105	2.88	5.41	.8
1041	7	57	13	42	11.73	1.86	.044	22.50	.579	.126	3.25	16.92	4.2
1055	7	65	110	397	12.55	1.96	.037	57.76	.454	.108	3.60	7.36	.9
1056	7	110	25	93	10.77	1.47	.048	22.63	.389	.129	3.73	10.71	.9
2002	7	51	186	636	12.55	2.09	.039	34.51	.390	.150	3.42	7.07	4.9
2003	7	120	29	100	13.61	1.53	.045	42.97	.395	.105	3.45	9.53	.9
2006	7	1	6	23	10.83	2.00	.039	-9.00	.354	.098	3.80	4.56	1.9
2017	7	59	17	49	15.71	1.70	.053	35.58	.572	.055	2.89	8.13	1.1
2018	7	59	137	422	11.84	2.17	.039	42.84	.355	.122	3.08	7.85	1.9
2025	7	1	56	238	12.29	1.37	.039	40.84	.747	.141	4.25	11.33	2.5
2034	7	1	93	255	11.02	2.21	.033	54.07	.680	.115	2.75	7.40	2.0
2040	7	1	140	357	-9.00	1.79	-9.000	-9.00	1.080	.094	2.55	12.22	1.9
2044	7	1	89	232	11.35	1.75	.030	26.08	.948	.110	2.61	3.75	1.1
2050	7	1	14	50	14.28	1.56	.037	29.75	.479	.106	3.58	10.02	3.3
2052	7	1	7	43	14.33	1.75	.056	-9.00	.429	.115	6.15	10.75	3.0
2067	7	1	24	80	10.95	2.00	.033	34.68	.554	.122	3.35	6.18	.3
3001	7	126	57	293	16.26	2.19	.042	22.12	.385	.104	5.15	7.97	1.5
3006	7	56	168	543	10.69	2.47	.040	52.54	.608	.116	3.23	6.96	.8
3010	7	63	52	206	13.21	1.73	.037	36.22	.336	.117	3.95	5.50	4.4
3013	7	116	61	186	11.88	2.35	.040	10.31	.339	.122	3.06	10.36	1.3
3014	7	89	61	168	11.45	2.26	.036	52.31	.387	.110	2.75	7.81	-9.0
3015	7	98	50	180	12.14	1.79	-9.000	37.83	-9.000	.127	3.60	9.63	2.6
3031	7	61	90	232	12.64	2.05	.030	43.72	.447	.109	2.58	3.44	1.8
4002	7	108	64	222	13.89	1.73	.046	27.31	.331	.119	3.46	10.55	1.1
4005	7	321	26	71	10.98	2.17	.058	25.75	.391	.137	2.71	7.84	2.7
4007	7	172	37	141	11.29	1.95	.060	37.62	.362	.137	3.80	12.61	3.2
4008	7	75	88	360	13.70	2.00	.041	40.42	.375	.116	4.32	9.76	1.3
4011	7	192	13	46	14.29	1.86	.047	29.23	.307	.100	3.51	22.73	2.5
4015	7	109	28	84	15.16	2.00	.044	33.81	.269	.105	2.99	3.75	2.7
4017	7	125	77	179	10.62	1.37	.044	15.94	.282	.129	3.75	6.16	2.7
4041	7	210	62	298	14.49	1.79	.094	22.10	.454	.130	4.79	10.66	1.3
4043	7	81	33	137	13.04	1.45	.050	21.46	.432	.119	4.14	7.3	1.3
4044	7	146	27	97	12.95	1.39	.057	44.06	.552	.117	3.58	7.49	1.1

PEER GROUP 7 (continued)

ID	PEER GROUP	URBAN AREA	PVEH	PVM	SPEED	PTRIPS	RVH/HEXP	PMS/RVH	WHRV/HXP	PVH/LMP	PVM/PVLH	PVM/MNI	PVM/MIC
5001	7	154	35	72	13.62	2.27	.051	35.48	.337	.125	2.88	8.00	1.8
5009	7	273	16	36	13.55	2.29	.043	33.80	.308	.115	2.26	7.70	-9.0
5016	7	32	223	754	12.89	2.19	-9.000	-9.00	-9.000	.094	3.38	7.37	.9
5020	7	191	10	31	12.72	2.00	.051	13.12	.176	.088	3.12	6.24	4.4
5028	7	277	15	65	13.08	1.50	.057	24.85	.110	.166	4.34	-9.00	2.8
5032	7	68	76	151	13.[illegible]	1.44	.076	11.74	.183	.097	4.18	5.58	.8
5033	7	64	69	244	13.49	1.77	.042	27.51	.236	.115	3.53	7.16	1.6
5039	7	139	14	76	12.21	1.75	.072	21.69	.256	.130	2.61	7.93	-9.0
5044	7	96	59	214	11.26	1.59	.050	21.75	.264	.132	3.63	11.91	1.2
5045	7	3	24	273	12.90	1.68	.032	27.83	.349	.118	3.25	7.00	-9.0
5052	7	73	40	172	12.39	1.48	.076	-9.00	.208	.107	4.30	7.97	1.6
5055	7	102	4	14	13.94	2.00	.055	-9.00	.197	.104	3.62	4.83	2.4
5056	7	25	39	153	14.19	1.63	.035	22.94	.217	.099	4.06	6.88	.9
5058	7	102	38	117	11.27	1.81	.035	33.26	.188	-9.000	3.07	7.15	.9
5059	7	162	34	96	11.50	1.48	.032	45.60	.213	.107	2.83	6.01	1.5
5082	7	3	80	413	13.49	1.70	.037	34.95	.423	.128	5.16	7.24	1.4
6002	7	79	59	361	11.43	1.55	.055	34.83	.244	-9.000	4.42	-9.00	1.3
6007	7	36	92	232	12.70	2.14	.045	29.36	.339	.110	3.06	9.88	1.1
6012	7	163	12	43	12.56	1.50	.054	38.52	.233	.110	3.61	10.83	2.0
6018	7	60	80	312	15.84	2.05	.042	20.86	.238	.108	3.90	8.21	.7
6021	7	36	35	179	15.93	1.92	.054	44.58	.691	.157	5.55	15.95	1.2
6022	7	83	33	123	11.06	1.43	.053	-9.00	.389	.111	3.74	6.67	1.1
6024	7	91	43	135	12.76	1.48	.055	-9.00	.536	.133	4.30	8.41	-9.0
6029	7	26	19	88	14.40	2.38	.037	-9.00	.455	.097	4.64	6.30	1.5
7001	7	171	38	153	10.60	1.46	.074	-9.00	-9.000	.155	4.02	10.39	1.0
7002	7	49	159	603	13.35	2.11	.043	31.90	.317	.118	3.81	9.31	1.0
7008	7	151	31	97	12.68	1.92	.048	27.24	.313	.114	3.13	9.60	1.6
7011	7	349	21	59	12.25	1.75	.043	35.27	.394	.108	2.81	8.19	1.0
7012	7	187	24	62	9.63	2.18	.046	-9.00	.418	.139	2.57	10.29	1.9
7013	7	169	20	72	13.36	2.00	.053	-9.00	.303	.138	3.60	12.01	8.0
7014	7	150	22	80	13.27	1.57	.049	24.81	.377	.107	3.62	9.37	1.1
8005	7	104	75	197	15.89	1.75	.063	22.72	.301	.142	5.64	15.17	1.[illegible]
9032	7	27	163	686	14.05	1.63	.036	30.04	.275	.104	4.21	7.60	1.[illegible]
9033	7	72	96	524	14.74	1.55	.040	27.55	.250	.123	5.45	10.93	2.4
9042	7	2	27	94	12.03	1.93	.038	31.28	.250	.140	3.48	9.49	2.9

PEER_GROUP_8

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PK/OBS	RVH/UEXP	IPAS/RVH	WORKV/UXP	IVH/EMP	IVM/PVEH	IVM/MNI	IVM/AUC
1048	8	53	240	702	12.11	2.61	.039	37.87	.453	.111	2.92	8.24	1.2
2004	8	23	357	1040	10.96	2.93	.037	39.80	.467	.097	2.91	4.19	1.1
2013	8	41	212	700	11.83	2.68	.032	46.78	.404	.107	3.30	6.53	.7
2039	8	1	85	209	9.56	3.40	.025	-9.00	.701	.095	2.46	3.73	1.3
2043	8	1	197	448	8.33	3.03	.032	52.99	.717	.120	2.27	4.43	1.0
2045	8	1	105	267	9.46	2.92	.029	53.50	.662	.115	2.54	4.45	.7
2059	8	1	191	615	18.14	3.35	.020	36.50	.821	.072	3.22	6.02	2.0
3004	8	77	83	242	14.37	3.61	.042	16.15	.367	.088	2.92	6.73	1.3
3011	8	208	26	60	11.34	2.39	.036	12.14	.161	.096	2.32	7.54	2.3
4001	8	97	61	222	15.25	2.65	.070	29.77	.927	.100	3.64	6.00	1.5
4003	8	38	240	1038	13.69	2.40	.041	33.20	.453	.111	4.33	9.09	1.0
4004	8	54	128	492	14.25	2.67	.041	54.77	.436	.118	3.85	8.49	1.6
4018	8	34	215	760	13.18	2.39	.033	40.22	.235	.095	3.54	6.50	.8
4019	8	21	79	264	13.35	2.82	.035	33.00	.422	.110	3.34	5.89	-9.0
4040	8	46	166	625	13.87	2.31	.043	38.69	.481	.115	3.77	7.62	3.0
4042	8	44	158	487	14.32	3.51	.041	-9.00	.494	.127	3.08	11.59	3.7
5003	8	206	27	67	10.34	2.70	.058	8.61	.208	.165	2.48	12.17	1.3
5005	8	103	143	424	12.55	2.65	.041	10.79	.388	.125	2.97	8.84	1.5
5011	8	87	74	183	9.22	2.96	.040	14.46	.124	.104	2.47	5.39	.6
5013	8	21	351	1207	13.01	2.90	.034	36.31	.305	.099	3.44	5.41	-9.0
5017	8	35	33	107	12.72	2.75	.030	37.30	.351	.045	3.25	2.54	.8
5022	8	50	175	669	12.81	2.87	.062	15.74	.222	.170	3.82	15.56	1.6
5025	8	147	86	243	13.03	3.19	.033	35.68	.371	.107	2.82	7.83	.7
5050	8	30	166	598	13.05	2.59	.037	-9.00	.475	.097	3.60	7.71	.5
5064	8	3	3	5	13.38	3.00	.053	9.38	.146	.097	1.57	7.84	.7
5077	8	3	13	25	15.61	3.25	.024	35.89	.250	-9.000	1.92	-9.00	-9.0
5080	8	3	113	421	13.28	2.63	.035	30.10	.365	.124	3.73	9.36	1.4
5091	8	0	23	47	13.41	3.29	.061	44.07	.242	.110	2.05	9.41	1.2
6004	8	15	387	1349	14.16	2.97	.036	45.43	.549	.101	3.49	8.82	.7
6017	8	43	62	199	11.36	2.82	.078	18.14	.287	.103	3.21	6.63	1.3
7005	8	22	260	982	12.89	2.57	.035	34.01	.299	.109	3.78	6.59	1.0
7010	8	82	80	239	13.22	2.96	.035	36.04	.426	.122	2.99	7.73	2.2
8003	8	224	14	34	10.49	2.80	.073	7.13	.306	.130	2.44	11.77	2.4

PEER GROUP 9

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PKTOBS	RVH/DEXP	TPAS/RVH	WORV/OXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
6011	9	33	329	1460	14.67	1.86	.040	41.01	.244	.117	4.44	8.11	1.57
8001	9	52	229	1133	16.42	1.73	.031	32.46	.095	.074	4.95	5.42	2.19
9002	9	55	303	1635	14.56	1.58	.029	72.22	.415	.095	5.40	6.73	1.21
9013	9	25	244	1256	14.93	1.28	.015	39.36	.070	.077	5.15	3.74	1.66
9019	9	39	172	887	16.51	1.32	.025	36.83	.258	.084	5.16	8.14	-9.00
9026	9	19	225	1332	16.27	1.11	.024	47.25	.408	.109	5.92	7.93	.66
9030	9	19	96	769	17.32	1.17	.045	18.98	.235	.142	8.01	10.99	1.99
9036	9	2	241	1602	15.10	1.16	.028	32.31	.208	.093	6.65	9.15	1.52

PEER GROUP 10

ID	PEER GROUP	URBAN AREA	PVEH	TVM	SPEED	PKTOBS	RVH/DEXP	TPAS/RVH	WORV/OXP	TVH/EMP	TVM/PVEH	TVM/MNT	TVM/ACC
8	10	28	473	2131	12.88	2.00	.028	32.50	.306	.107	4.50	11.46	1.7
2001	10	29	273	1403	8.18	1.11	.048	29.01	.303	.109	5.14	4.45	3.3
2007	10	1	260	1059	11.15	1.95	.033	-9.00	.453	.117	4.07	6.01	2.0
5008	10	16	475	2089	11.27	1.80	.046	46.77	.489	.128	4.40	9.27	1.0
6008	10	13	355	1752	13.65	2.03	.026	29.85	.271	.079	4.94	3.68	.5
6032	10	26	368	1371	10.33	2.07	.041	73.98	.461	.101	3.73	4.53	.9
8006	10	24	506	2385	11.94	1.66	.032	26.10	.191	.102	4.71	6.24	1.2
9015	10	6	434	1595	9.43	1.49	.028	89.78	.273	.106	3.67	4.22	.7

PEER GROUP 11

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PK/OBS	RVH/DEXP	IPAS/RVH	WORV/DXP	TVH/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
1	11	17	666	3010	18.40	2.82	.023	49.58	.298	.088	4.52	8.43	1.28
1003	11	7	811	2496	12.70	2.56	.020	52.76	.250	.066	3.08	2.63	2.34
2068	11	76	1149	5688	16.00	1.66	.026	36.45	.587	.101	4.95	7.09	1.42
3019	11	4	1076	3761	10.17	2.21	.030	67.08	.509	.099	3.50	4.24	.96
3022	11	11	785	3377	13.10	2.18	.030	38.88	.406	.101	4.30	6.98	.69
3030	11	8	1573	5374	11.05	2.97	.024	52.53	.440	.097	3.42	5.36	1.11
3034	11	14	786	2684	12.12	3.14	.027	57.99	.448	.103	3.42	7.42	.84
4022	11	20	738	2967	13.05	2.68	.030	53.97	.251	.082	4.02	6.22	3.01
5015	11	9	689	2406	14.89	2.06	.021	68.62	.254	.116	3.49	7.69	.50
5027	11	12	829	2885	13.92	2.91	.026	61.90	.320	.096	3.48	7.55	.59
5031	11	5	948	4349	13.50	2.47	.022	44.37	.255	.099	4.59	6.86	1.06
7006	11	10	783	2870	12.56	2.30	.024	43.87	.178	.094	3.67	5.42	.81
9014	11	6	718	3188	14.40	2.21	.025	59.55	.334	.099	4.44	10.96	1.38

PEER GROUP 12

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PK/OBS	RVH/DEXP	IPAS/RVH	WORV/DXP	TVH/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
2008	12	1	3378	10869	6.45	1.63	.027	83.45	.807	.113	3.22	3.21	1.42
5066	12	3	2178	8343	12.07	2.00	.025	82.20	.549	.085	3.90	5.52	.9[illegible]
9021	12	2	1914	10339	13.23	1.60	.026	58.51	.386	.097	5.40	7.35	.7[illegible]

NOT ASSIGNED TO A PEER GROUP

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PK/OBS	RVH/DEXP	IPAS/RVH	WORV/DXP	TVH/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
5024	99	58	67	81	4.81	2.68	.047	14.68	.107	.113	1.21	2.91	.65
9016	99	6	235	872	21.72	5.11	.017	31.03	.522	.065	3.71	10.07	3.7[illegible]

NOT_INCLUDED_IN_PEER_GROUP_ANALYSIS

ID	PEER GROUP	URBAN AREA	PVEH	IVM	SPEED	PKIUBS	RVH/UEXP	IPAS/RVH	WRKV/OXP	IVH/EMP	IVM/PVEH	IVM/MNT	IVM/ACC
1002	-9	189	32	70	-9.00	1.78	-9.000	-9.00	.327	.064	2.18	4.10	2.5
1013	-9	106	11	-9	-9.00	1.00	-9.000	-9.00	-9.000	-9.000	-9.00	-9.00	-9.0
1040	-9	-9	15	-9	-9.00	1.00	-9.000	-9.00	.212	-9.000	-9.00	-9.00	-9.0
1045	-9	152	5	-9	-9.00	1.67	-9.000	-9.00	.482	-9.000	-9.00	-9.00	-9.0
1046	-9	-9	-9	-9	-9.00	-9.00	-9.000	-9.00	.303	-9.000	-9.00	-9.00	-9.0
1053	-9	7	2	-9	-9.00	1.00	-9.000	-9.00	.122	-9.000	-9.00	-9.00	-9.0
1064	-9	-9	5	-9	20.00	1.25	.048	-9.00	-9.000	-9.000	-9.00	-9.00	-9.0
2012	-9	177	52	-9	-9.00	1.00	-9.000	-9.00	.599	-9.000	-9.00	-9.00	-9.0
2016	-9	59	-9	-9	-9.00	-9.00	-9.000	-9.00	-9.000	-9.000	-9.00	-9.00	-9.0
2021	-9	112	-9	87	13.97	-9.00	.042	-9.00	.409	.104	-9.00	6.44	2.1
2056	-9	1	4	-9	-9.00	1.33	-9.000	-9.00	.382	.094	-9.00	-9.00	-9.0
2066	-9	1	55	334	-9.00	1.67	-9.000	-9.00	.619	-9.000	6.07	83.47	3.7
3016	-9	98	1	-9	-9.00	1.00	-9.000	-9.00	.159	-9.000	-9.00	-9.00	-9.0
3018	-9	165	-9	-9	-9.00	-9.00	-9.000	-9.00	.343	-9.000	-9.00	-9.00	-9.0
3028	-9	11	13	93	12.60	13.00	-9.000	8.82	-9.000	.422	7.13	30.92	-9.(
4020	-9	275	5	-9	-9.00	1.00	-9.000	-9.00	.170	-9.000	-9.00	-9.00	-9.0
4033	-9	113	5	-9	-9.00	1.00	-9.000	-9.00	.201	-9.000	-9.00	-9.00	-9.0
4034	-9	18	425	-9	-9.00	1.38	-9.000	-9.00	.464	-9.000	-9.00	-9.00	-9.0
4039	-9	48	-9	232	12.59	-9.00	.051	40.58	.454	.112	-9.00	12.68	.9
5004	-9	255	9	-9	-9.00	1.00	-9.000	-9.00	.269	-9.000	-9.00	-9.00	-9.0
5010	-9	45	84	-9	-9.00	1.79	-9.000	-9.00	.243	-9.000	-9.00	-9.00	-9.0
5053	-9	210	-9	-9	-9.00	-9.00	-9.000	-9.00	.238	-9.000	-9.00	-9.00	-9.0
5086	=9	3	86	-9	-9.00	1.04	-9.000	-9.00	.311	-9.000	-9.00	-9.00	-9.0
5087	-9	3	13	-9	-9.00	2.60	.041	50.51	.438	.100	-9.00	-9.00	-9.0
6014	-9	276	11	-9	-9.00	1.00	-9.000	-9.00	.415	-9.000	-9.00	-9.00	-9.0
7015	-9	70	-9	-9	-9.00	-9.00	-9.000	-9.00	.321	-9.000	-9.00	-9.00	-9.0
7016	-9	22	14	-9	-9.00	2.30	-9.000	-9.00	.165	-9.000	-9.00	-9.00	-9.0
8003	-9	204	8	-9	-9.00	1.00	-.196	-9.00	.293	-9.000	-9.00	-9.00	-9.0
9031	-9	-9	18	-9	-9.00	1.06	-9.000	-9.00	.143	-9.000	-9.00	-9.00	-9.0
9043	-9	2	5	-9	-9.00	1.25	-9.000	-9.00	-9.000	-9.000	-9.00	-9.00	-9.0