U.S. Department of
Transportation

Bus Transit Monitoring Manual

Volume 1: Data Collection Program Design

August 1981

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTA-IT-09-9008-81-1		

4. Title and Subtitle	5. Report Date
Bus Transit Monitoring Manual Volume 1 Data Collection Design	August 1981
	6. Performing Organization Code
7. Author(s) John Attanucci, Imogene Burns, Nigel Wilson	8. Performing Organization Report No.
9. Performing Organization Name and Address Multisystems, Inc. 1050 Massachusetts Ave. Cambridge, MA 02138 ATE Management & Service Co. 617 Vine Street Cincinnati, OH 45202	10. Work Unit No. 11. Contract or Grant No. DOT-UT-90005
12. Sponsoring Agency Name and Address U.S. Department of Transportation Urban Mass Transportation Administration 400 Seventh Street, S.W. Washington, D.C. 20590	13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes

UMTA Project Director - Mr. Brian McCollom, UPM-13

16. Abstract

The objective of the Bus Transit Monitoring Study is to develop a method for the design of comprehensive, statistically based data collection programs for the bus transit industry to support the short range planning process. The two volumes of this manual document this method and provide transit operators with step-by-step procedures to develop their own individually-tailored programs. This volume, Volume 1, Data Collection Design, explains the various components of a comprehensive, route-level data collection program, beginning with the determination of data needs and finishing with interpretation of the data. Volume 2, Sample Size Tables, provides an extensive set of tables for determining sample sizes for systems and routes of varying size and operating characteristics. A two-stage approach is described with a baseline phase to produce detailed profiles for each bus route, and a monitoring phase to gather limited data on a periodic basis. The advantages and disadvantages of various data collection techniques are discussed. Both the desired accuracy and the inherent variability of the data items are incorporated in the selection of a sampling plan. Allowance is made for the use of simple linear relationships between data items to reduce the overall cost of the data collection program where feasible. The recommended data collection program is shown to meet UMTA Section 15 reporting requirements for passenger related data.

17. Key Words	18. Distribution Statement
Bus, fixed route; evaluation techniques; transit management; planning analysis, methodology; planning, tools; productivity ; surveys; transit performance; TSM	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 (8-69)

Bus Transit Monitoring Manual

Volume 1: Data Collection Program Design

Final Report
August 1981

Prepared by
Multisystems, Inc.
1050 Massachusetts Ave.
Cambridge, Mass. 02138

with
ATE Management & Service Co.
617 Vine Street
Cincinnati, Ohio 45202

Prepared for
Office of Planning Assistance
Urban Mass Transportation Administration
Washington, D.C. 20590

Distributed by
Technology Sharing Program
Office of the Secretary of Transportation

DOT-I-81-30

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

FOREWORD

Many transit operators have adopted sets of service performance measures and standards and have developed plans to use them in a systematic evaluation. In many cases, however, transit operators have not been able to implement the measures and standards because they have had difficulty in developing a cost-effective system to collect the needed information. To assist these operators, UMTA's Office of Planning Assistance, through its Special Studies Program, has sponsored a study in data collection.

This two-volume manual is the product of this study which documents a method to develop comprehensive statistically based data collection programs that will enable transit operators to collect passenger-related data in a cost-effective manner. We believe the step-by-step procedures provided in this manual will be of value to transit operators in their efforts to improve their data collection systems.

Additional copies of this report are available from the National Technical Information Service (NTIS), Springfield, Virginia 22161. Please reference UMTA-IT-09-9008-81-2 on the request.

CHGraves

Charles H. Graves, Director
Office of Planning Assistance (UPM-10)
Urban Mass Transportation Administration
U.S. Department of Transportation
Washington, D.C. 20590

A. Linhares

Alfonso B. Linhares, Director
Office of Technology Sharing (I-40)
Office of the Secretary
U.S. Department of Transportation
Washington, D.C. 20590

The preparation of this report has been financed through a contract from the U.S. Department of Transportation, under the Urban Mass Transportation Act of 1964, as amended. The contents of this report were prepared by Multisystems, Inc. and ATE Management and Service Co., Inc. and do not necessarily reflect the official views or policies of the U.S. Department of Transportation of the Urban Mass Transportation Administration.

ACKNOWLEDGEMENTS

This report was prepared by Multisystems, Inc., with the assistance of ATE Management and Service Co., Inc. The authors are John Attanucci, Imogene Burns and Nigel Wilson. Technical input was received from Martin Flusberg and JoAnn Kruger of Multisystems, Jim Curry and Chester Hashizume of ATE, and Joel Horowitz. John Kirsch provided editorial assistance and the draft and final documents were largely typed by Gail Pasquale.

The authors would like to thank the respective operations planning staffs of the Northeastern Illinois Regional Transportation Authority (RTA) and the Chicago Transit Authority (CTA), who provided the opportunity to field test the methods described in this manual. Paul Oppenheim, Ed Bailey, Mike Bolton and Steve Wojtkiewicz of the RTA and Harry Hirsch, Steve Legler, Bob Vance and Sue Walker of the CTA provided invaluable assistance before, during and after the Chicago data collection trial. In addition to the RTA and CTA staffs, the following people served on a continuing project review panel and the authors wish to express their appreciation for their insight and direction:

J. Vincent Farrell	Rhode Island Public Transit Authority Providence, Rhode Island
Jack Garrity	Regional Transit Service Rochester, New York
Ron Hartman	American Public Transit Association Washington, D.C.
Thomas Niskala and John Pappas	HouTran Houston, Texas
Jack Rielly	Capital District Transportation Authority Albany, N.Y.
Mark Turnquist	Cornell University Ithaca, N.Y.

Finally, we would like to thank Brian McCollom of UMTA for his valuable comments and assistance throughout the project and in preparing this report.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

HOW TO USE THIS MANUAL

This manual consists of two volumes: Data Collection Design (Volume 1) and Sample Size Tables (Volume 2). Together, the two volumes of this manual provide transit properties with the necessary information to design a comprehensive bus service monitoring program.

This volume, Data Collection Design, explains the various components of a comprehensive data collection program, beginning with the determination of data needs and finishing with the interpretation of the data. The first five chapters provide a basic framework for step-by-step program design procedures which are presented in an instruction/example format in Chapter 6. As such, it is important for the user of this manual to read the first five chapters before attempting to use the prcedures outlined in Chapter 6. Once familiar with the basic concepts and practical considerations which are discussed in detail in Chapters 1-5, the user can proceed to use the design procedures in Chapter 6 where the underlying framework and assumptions are largely unstated. The user will also need to refer to the detailed statistical discussions and work sheets contained in Appendix A of this volume, in order to fully carry out the program design procedures outlined in Chapter 6.

Volume 2, Sample Size Tables, contains an extensive set of tables for determining sample sizes for systems and routes of varying size and operating characteristics. Volume 2 cannot be used alone, but is simply a reference document for users of Volume 1. Instructions for use of the sample size tables are included in Chapters 4 and 6 of Volume 1 as well as in Volume 2.

Chapter 1

INTRODUCTION

In recent years, there has been a growing awareness of the need to use public transportation resources more efficiently. It has become more important to carefully evaluate (or re-evaluate) all services, both current and planned. Recent research has considerably advanced the state-of-the-art of transit evaluation methods. A number of transit properties, large and small, have adopted sets of service performance measures and standards, and have developed on-going systematic evaluation programs for using them.

In many cases, however, improved evaluation procedures have not been supported by comprehensive data collection programs. Cost-effective programs are needed to provide the passenger-related performance data required by individual properties.

1.1 Previous Transit Data Collection Research

The last detailed study of U.S. transit data collection practices was conducted by the American Transit Association (ATA) more than thirty years ago. Between 1946 and 1949, the ATA published several reports describing techniques for traffic checking and schedule preparation. In 1946, the Manual of Traffic and Transit Studies was released describing detailed procedures for conducting twenty different data collection "studies." In 1947, the ATA began a four-part study into techniques for traffic checking and schedule development.[1] The first part consisted of an in-depth description of "sample" procedures based on methods used by the New Orleans Public Service Inc. In the second part, a survey of scheduling

---

[1] Further information and copies of this study are available from the American Public Transit Association, 1225 Connecticut Ave., Washington, D.C.

practices was carried out with responses reported from over seventy transit systems in North America. The third part of the study was a symposium of industry practices which provided commentary on the results of the first two study parts. In the last part of the study, selected areas for improved techniques were investigated.

For more than three decades, these ATA reports have constituted the only comprehensive reference source on techniques for data collection and analysis. While the reports have been extremely valuable to transit properties, they have significant limitations. First, the reports do not take into account service changes of recent years, such as multiple fare structures and transit passes. More importantly, the ATA manual does not explore issues such as the amount of data to be collected and the frequency of data collection. Many properties have very different practices with respect to sample size and frequency of collection, and it is likely that some collect too little data, while others collect too much.

1.2 Objective of Bus Transit Monitoring Study

The objective of the present study is to develop a comprehensive, statistically-based data collection program that will enable transit operators to collect in a cost-effective manner the passenger-related operations data that they need. Procedures have been developed which will allow properties to conduct the following tasks:

1) select the appropriate data collection techniques;

2) determine the proper sampling plans for different types of data; and

3) estimate the cost of collecting the data required for their own system.

These procedures have been summarized in a step-by-step approach which can be used to determine data needs and design data collection programs in individual transit properties.

A panel of experts in transit operations has assisted in this study. The panel, consisting primarily of managers and planners of both small and large transit properties, reviewed all findings and assisted in planning the general direction of the study. In addition, the review panel included a representative of the American Public Transit Association (APTA) and a statistical expert experienced in transit operations.

The initial phase of the study focused on defining the data needed by the transit industry for operations planning and management decision-making, and on the techniques currently used to collect these data. This information was collected through:

1) a review of reports prepared by a number of transit properties;

2) a survey conducted by the Massachusetts Bay Transportation Authority and the Tidewater Transportation District Commission; and

3) interviews with forty-one transit properties.

The results of this phase are described in Interim Report #1, Bus Transit Monitoring Study: Data Needs and Data Collection Techniques (April 1979, NTIS PB80-161409).

Using the information obtained from this review, a preliminary design of a general data collection program was developed. The preliminary program was then field-tested in the Chicago metropolitan area, with the cooperation of the Northeastern Illinois Regional Transportation Authority (RTA) and the Chicago Transit Authority (CTA). The field tests consisted of both actual data collection and the canvassing of RTA and CTA staff reactions to the preliminary program. The information obtained from the Chicago field-tests was then used to revise the preliminary approach. The revised program is presented in this new data collection manual.

1.3 Transition from Current Practices to the Proposed Program

Data are collected by most transit properties for a variety of activities including scheduling, detailed route planning, marketing, deficit allocation or funding reimbursement, and external reporting requirements. Since these activities may be conducted by different departments, the data collection for these activities may not be well coordinated, nor may the data collected be maintained in one central location for common use. It is often difficult in many properties to determine if the resources allocated to transit data collection are being used most effectively.

The approach outlined in this manual provides properties with the opportunity to reassess their current data collection practices with an emphasis on more efficiently collecting, processing, and maintaining the required route and system data. This approach formalizes the efforts currently being made in the industry to monitor performance of bus systems. It reorganizes into a systematic structure many actions now performed by most transit managers. Individual properties can either directly follow this approach or modify it based on their data collection experience.

The manual is intended for use by those responsible for developing data collection plans (e.g., planners, schedule supervisors, revenue analysts) in a property of any size. While it would be helpful to have a basic understanding of statistics and sampling theory, no prior knowledge is required to use the procedures contained in the manual. Each concept is fully explained and then incorporated into a step-by-step procedure.

1.4 Two-Phase Data Collection: Baseline and Continuous Monitoring

The proposed approach includes two distinct data collection phases. In the first phase, or the baseline data collection phase, the "base conditions" are defined for each route in the system. Base conditions include all the data needed for effective operations planning including total boardings, loads at key points on the route, running and arrival times, revenues, and passenger characteristics. The baseline phase presents a snapshot of system performance within a relatively short time span. Complete route "profiles" are developed from these data which facilitate comparisons among routes in specific subareas, garage divisions, function types, or in the system as a whole. Since the baseline phase includes the collection of all data items needed for service evaluation, including origin-destination data from a passenger survey, it provides an excellent opportunity to analyze the potential for major route restructuring or reallocation of equipment.

The baseline phase is also used to identify relationships among data items which may be used to reduce the effort needed for monitoring performance. If strong relationships are found on individual routes, they would permit the subsequent use of less expensive data collection techniques on those routes. For example, if the number of boarding passengers can accurately be predicted from farebox revenue, then farebox revenue could be used with an "average fare factor" to estimate total boardings.

In the monitoring phase of the data collection program, each route is checked periodically to detect changes which have occurred. By checking passengers, revenue and schedule adherence, a planner establishes the new route performance (within a given accuracy range) and decides whether a change has occurred which requires follow-up action. If none of the monitored data items changes significantly, it is assumed that the other data collected during the baseline phase (e.g., passenger origins and destinations, fare categories) have also remained stable.

While the baseline and monitoring data collection phases differ in the number of data items which are collected, the sampling requirements are similar. Thus, the monitoring phase is the baseline phase minus certain collection techniques.

This approach to data collection provides a property with the performance data necessary for routine planning and scheduling functions, as well as for external reports on both a route and systemwide level.

The two data collection phases are designed in the same way. Four important inputs are required:

1) a list of data required by the property and how frequently it is to be obtained;

2) an estimate of the required accuracy for each data item of interest;

3) key property and route characteristics; and

4) existing data or data obtained from a special "pretest" from which sample sizes can be determined.

Guidelines for determining each of these inputs are provided in this manual along with all of the necessary steps to design a comprehensive monitoring program.

1.5 Cost of a Monitoring Program

Cost is an obvious concern (and probably a manager's first question) in the development of a comprehensive data collection program. While costs vary widely depending on specific property characteristics and ridership patterns, some guidelines can be used to estimate the cost of a monitoring program.

By far the most costly component is the use of on-board traffic checkers to monitor total boardings. This cost can be avoided if, as is often the case, a property can obtain reliable data from drivers. Other techniques can also be substituted for on-board data collection if a strong relationship exists between total boardings and farebox revenue or maximum load on a particular route. These factors dramatically impact the total resources required by a property to carry out a comprehensive monitoring program.

Based on information from Chicago and other properties studied in this project, the range of checker resources required for typical bus system sizes has been estimated using average values for data variability, desired accuracy and route characteristics. The (full-time) traffic checker staff requirements shown in Table 1.1 assume that every route in the system is monitored four times a year. (If less frequent monitoring is desired, these requirements can be reduced proportionally.) Generally, the low end of the range given in Table 1.1 represents the case where reliable operator data are available; the upper end of the range represents the case where drivers do not collect boarding data. The range also reflects differences among property and route characteristics which directly impact required sample sizes and, therefore, total checker requirements. To determine where in the given range a particular property falls, the detailed procedures outlined in Chapter 6 should be used on a route-by-route basis.

Staff requirements for the baseline data collection phase for most properties would fall near the upper end of the indicated ranges (for a period of about 3 months). In addition, the cost of an on-board passenger survey on all routes should be added to the staff requirements in Table 1.1 for the baseline phase. More information on these and other data collection cost components is provided in Chapter 5.

1.6 "Section 15" Data Requirements

The data collection program outlined in this manual will provide a property with a wealth of information concerning passenger utilization of the system, including the data required by UMTA for the Section 15 "Transit Service Consumed Schedule" (Form 655). Section 15 requires three data items: unlinked passenger trips, passenger miles, and average time per unlinked passenger trip. These items are required on a systemwide basis for specified time periods during an average weekday, Saturday, and Sunday. These data are included in the data collection design procedures detailed in this manual. The procedures will allow a property to sample on a route level rather than on a systemwide random trip basis. Section 4.7

Table 1.1

Typical Checker Staff Requirements for Bus Systems of Different Sizes

Peak Buses	Off-Peak Buses	Average Daily Service Hours	Number of Traffic Checkers Required
25	22	12	½ - 1
50	40	12	1 - 2
100	70	14	1½- 4
300	215	15	3 - 7
500	250	16	6 - 13
750	470	17	8 - 15
1000	600	18	10- 19
2000	1100	19	20- 38

explains how data collected at the route level can be compiled to satisfy the systemwide reporting requirements of UMTA Section 15.

1.7 Organization of Manual

The two volumes of this manual provide transit properties with the necessary information to design a comprehensive bus service monitoring program. This volume, Data Collection Design, explains the various components of a monitoring program and presents a step-by-step procedure for designing a program. Volume 2, Sample Size Tables, provides an extensive set of tables for determining sample sizes for systems and routes of varying size and operating characteristics.

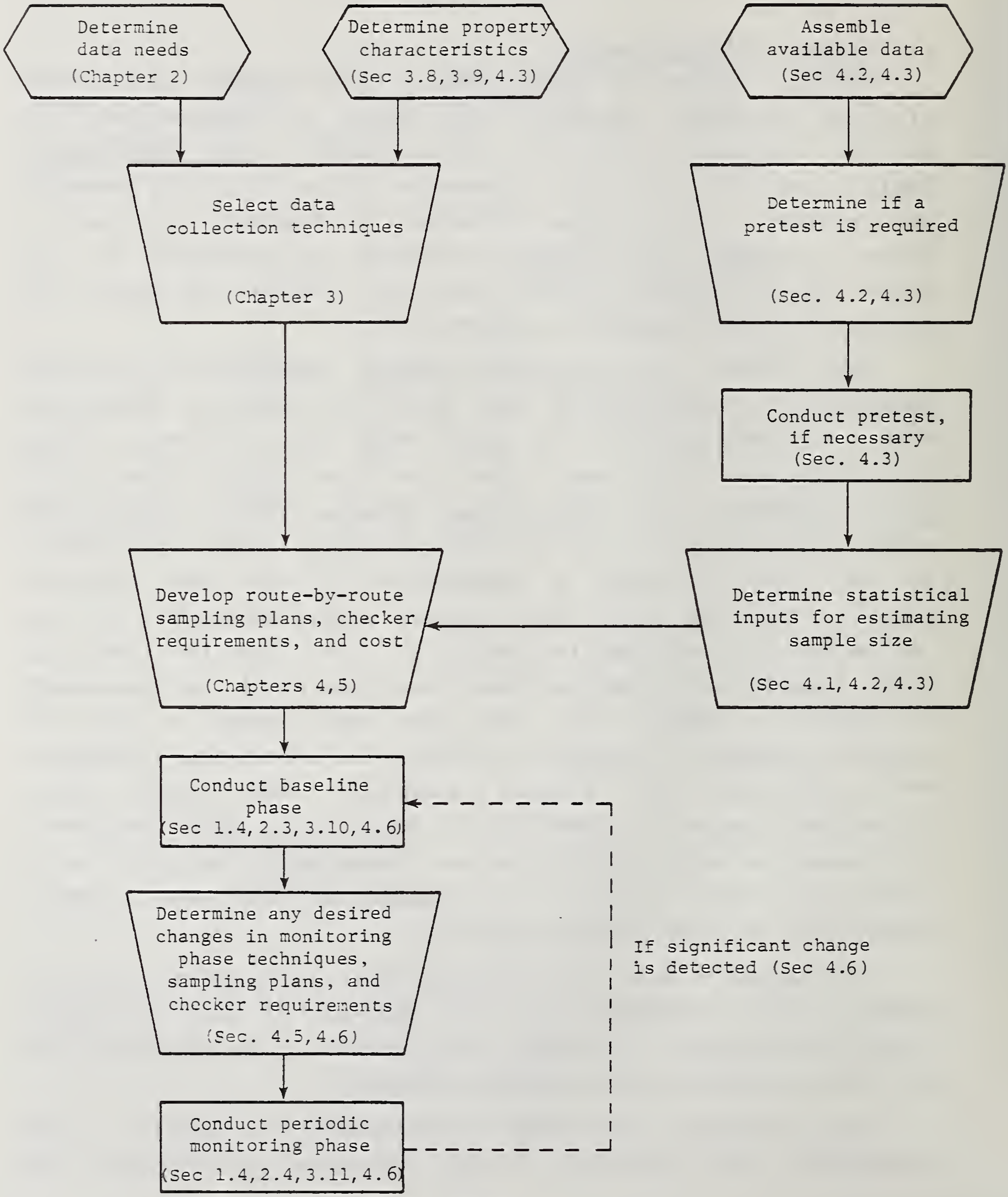
This volume, Data Collection Design, explains the various aspects of a comprehensive data collection program, beginning with the determination of data needs and finishing with interpretation of the data. The first five chapters provide a basic framework for step-by-step program design procedures which are presented in an instruction/example format in Chapter 6. As such, it would be informative for the user of this manual to read the first five chapters before attempting to use the procedures outlined in Chapter 6. Once familiar with the basic concepts and practical considerations which are discussed in detail in Chapters 2-5, the user can proceed to use the design procedures in Chapter 6 where the underlying framework and assumptions are largely unstated. The overall data collection program is summarized in Figure 1.1, which indicates the order in which activities are undertaken, as well as a reference to the section of the manual providing an in-depth description of each program activity.

In Chapter 2, the service-related data needs of the typical property are discussed in the context of the two-phase collection strategy. Guidance is provided on the determination of the requirements for a specific property.

Data collection techniques are described in Chapter 3. The advantages and limitations of each technique are outlined, and

Figure 1.1

Summary of Data Collection Program Design and Implementation



sample forms are provided for several techniques. The chapter includes recommendations for combining techniques during the baseline and monitoring collection phases for different property and route characteristics.

Chapter 4 describes the inputs and procedures needed to develop a sampling strategy, including appropriate sample sizes and guidelines on the timing of data collection efforts. Also introduced are the reference tables for determining required sample sizes, which are contained in Volume 2. Special sampling considerations to meet the UMTA Section 15 requirements are described in Section 4.7. The chapter concludes with a discussion of several procedures that a property could use to interpret samples.

Procedures for estimating the cost of a comprehensive monitoring program are provided in Chapter 5. The process of estimating checker and other resource requirements is explained, and some "rules-of-thumb" are provided for quick cost estimates.

In Chapter 6, the complete process for designing a comprehensive data collection program is detailed in sequential step-by-step procedures. The procedures incorporate the framework described in Chapters 2-5 and are to be followed during the actual design of a property's data collection program. An example is presented along with the discussion of each step to illustrate the procedures and calculations which would be performed by a transit property.

A technical discussion of sampling theory is presented in Appendix A, including detailed formulae and the statistical assumptions that underlie the discussion of sampling in Chapter 4. Appendix A also provides step-by-step instructions and work sheets for calculating some of the statistical inputs and data tests described in Chapter 4.

Finally, a discussion of various ways a property can classify its routes to simplify the sample size estimation procedure is presented in Appendix B.

Chapter 2

DATA NEEDS

The first step in the design of a comprehensive data collection program is to specify the data required by the operator. These needs depend on planning and other management activities and on external reporting requirements. Two key attributes of the data should be defined or estimated: how the data will be used, and how often they will be used.

2.1 Determining Data Needs

The data required by individual transit properties vary depending on the size and type of system operated and on specific management objectives. Those responsible for data collection should contact all appropriate management and supervisory personnel within the property to identify their data needs. The departments or staff to be contacted should include, but not necessarily be limited to:

- planning
- scheduling
- finance/revenue/budget
- transportation
- general manager

Each department (staff member) contacted should be asked to list the service-related data items used, how they are used, and how often they are used. Once a preliminary list of data needs has been compiled in this manner, it should be circulated to those originally contacted for review. The final list of data should also include those items required by outside agencies, such as a governing board, city council, state agency and the Urban Mass Transportation Administration (with special attention to UMTA Section 15 requirements).

2.2 Typical Data Needs of North American Transit Properties

The first task of this study included a review of the data needs reported by more than one hundred bus transit properties in North America. This included an analysis of the material collected from 71 transit properties by the Massachusetts Bay Transportation Authority (MBTA) in Boston and the Tidewater Transportation District Commission (TTDC) in Norfolk, Virginia.[1] These materials were supplemented by discussions with 41 other properties in order to focus directly on the data required by these properties and the data collection techniques currently employed.

These efforts resulted in the set of data items used by a large majority of the properties contacted. The set is shown in Table 2.1. Each of the data items listed was reported as being useful in one or more aspects of service management, including route planning, scheduling, marketing, funding reimbursement or deficit allocation, and external reporting. As discussed in the next section, all items listed in Table 2.1 are needed in the baseline data collection phase.

While this list is comprehensive, not all of the data items need to be maintained with the same currency. The data collection design procedures in this manual assume that collection of each item is performed systematically, but not with the same frequency. The following sections describe which data typically require frequent monitoring and which data will generally be collected during the initial baseline phase, but then less frequently.

2.3 Data Needs in the Baseline Phase

The nature of the data listed in Table 2.1 and the performance characteristics of the typical bus route suggest

---

[1] For further information on this effort, see Bus Service Evaluation Procedures: A Review, prepared by the Massachusetts Bay Transportation Authority and Tidewater Transportation District Commission, April 1979, NTIS Report No. PB79-296314.

Table 2.1

Data Needs in Baseline Phase

Route (or Stop) Specific	
Load (peak or other)*	
Bus arrival time	
Total boardings (i.e., passenger-trips)	
Revenue	
Boardings (or revenue) by fare category	
Passengers boarding and alighting by stop	
Transfer rates between routes	
Passenger characteristics and attitudes	
- age	- income
- handicap	- auto ownership
- sex	- auto availability
- job status	- home location
- attitudes toward level of service	
Passenger travel patterns	
- origin/destination	- work (school) trip mode
- work and/or school trip location	- non-work (school) travel patterns
- time of day of work (school) trip	- trip frequency
System-wide	
Unlinked passenger trips	
Passenger-miles	
Average unlinked passenger travel time	
Linked passenger trips	

* At specified points; not averaged throughout a trip.

that a two phase data collection program is appropriate for most transit properties. As described in Chapter 1, the purpose of the baseline phase includes:

- development of complete "baseline" route profiles
- provision of route performance data systemwide at the same point in time, providing the opportunity for a systematic analysis of route structure and vehicle allocation
- identification of relationships among individual data items which may allow less costly data collection techniques to be used in monitoring.

In order to develop comprehensive information on route performance in the baseline phase, all the items listed in Table 2.1 should be collected. The collection of these data will permit direct comparison among routes and analysis of alternative service plans, including route restructuring, reallocation of vehicles, and schedule modifications.

The data items in a sample comprehensive route profile are shown in Table 2.2. For each of items 1 to 5, an operator will generally be interested in the mean value and in the variation within each time period and from day to day. These five items will generally be used to derive measures of effectiveness for different routes (in terms of utilization and operating efficiency) as well as for operations planning and scheduling. Items 6 to 13 provide more specialized information which would be used for detailed route, sub-area or system planning (e.g., evaluation of through-routing, branching, short turning, limited or express services) as well as for studies of the property's fare structure and related policies.

Finally, items 14 and 15 provide information on the relationship between specific data items which are likely to exhibit particularly strong interrelationships. These relationships can be expressed in terms of "conversion factors" which may allow an operator to estimate one data item by directly measuring its related item, thus reducing the cost of

Table 2.2

Data Items in Sample Comprehensive Route Profile

General Effectiveness Data

1. Boardings per trip, per day
2. Revenue per trip, per day
3. Maximum load per trip
4. Running time by route segment
5. Difference between scheduled and actual arrival times

Data for Specialized Analyses

6. Distribution of boardings, revenue by fare category
7. Transfer rates per day
8. Passengers boarding and alighting by stop per trip
9. Average unlinked trip length per passenger
10. Average unlinked trip travel time per passenger
11. Passenger-miles per day
12. Passenger characteristics and attitudes
13. Passenger travel patterns

Data Collection Design Items

14. Relationship between boardings and revenue per trip
15. Relationship between boardings and maximum load per trip

monitoring a route. The data collected in the baseline phase allow a property to test these relationships for each route; if the statistical relationship is shown to be strong enough, then the conversion factor can be used during the monitoring phase, to estimate total boardings from observed revenue or peak load data. A more detailed explanation of tests for relationship and of conversion factors is included in Section 4.5.

2.4 Data Needs in the Monitoring Phase

Once a route profile is established during the baseline phase, an operator would want to regularly monitor each route for significant changes. In order to do this at reasonable cost, a subset of the data listed in Table 2.1 should be selected for periodic monitoring.

The three basic data items needed for tracking individual route performance in the monitoring phase are shown in Table 2.3.

Table 2.3

Data Needs in Monitoring Phase

Bus arrival time

Load at peak load point

One or more of the following:
- Total boardings
- Boardings by fare category
- Revenue

Bus arrival time must be collected periodically by all properties to ensure efficient scheduling and reliable service. Arrival times are usually collected in conjunction with either load or boarding counts. Load data are most often needed to determine appropriate service frequencies and are easily collected at the same time as bus arrival times (by using either a point or ride check as discussed in Chapter 3).

Total boardings, boardings by fare category, and revenue are alternative measures of the total utilization of the route. The choice of which one(s) to monitor will depend on the feasibility of different data collection techniques for the property, and on particular local needs. Certain data collection techniques yield two or more of these items at the same time, so that the property may be able to monitor directly a wider range of route performance measures.

This approach to monitoring assumes that if none of these three data items changes, no other data item listed in Table 2.1 has changed significantly since the baseline phase. Passenger on/off counts, characteristics, attitudes, origin-destination patterns, transfers, and some of the systemwide data required for Section 15 reports are all indirectly monitored through the collection of arrival time, load, passenger-trips, revenue or fare category data. It is highly unlikely that passenger travel characteristics will change without a corresponding change in the data items which measure schedule reliability, total passenger use, and revenue collected.

If significant changes are observed in an individual route during the monitoring period, another baseline phase needs to be conducted to update the route profile. Based on the routes analyzed in this study, it is recommended that the baseline phase be redone if total daily boardings change by 25 percent or more from the initial baseline phase.

The required level of detail for a given data item may vary from property to property. For example, one property may have a service standard (e.g., revenue per vehicle mile) which varies by time period. That property would therefore need mean revenue by time of day, while another property may only need total daily revenue. Typically, on routes which are scheduled based on observed demand, at least peak load would be required by selected time periods during the day (e.g., perhaps as short as the peak half-hour or 15-minute period). Routes which operate on policy headways may only require mean data summarized for one or two time periods during the day. In any case, the sampling procedures presented in the following chapters can be applied to any desired time period.

Chapter 3

DATA COLLECTION TECHNIQUES

A large number of data collection techniques are used by transit properties to obtain the data identified in Chapter 2. The seven principal data collection techniques are shown in Table 3.1. Each technique provides one or more of the data items listed in Table 2.1.

Some of these seven techniques are known by different names. For example, ride checks are also known as on-off checks and characteristic counts; point checks are often called standing checks, or load checks. For consistency, the names in Table 3.1 will be used throughout this manual.

Each of these techniques is described in the following sections (3.1-3.7) using examples to show their application and how the characteristics of a route (or property) can influence the types and extent of data obtained. Section 3.8 compares the seven techniques, with emphasis on the range of data items that can be obtained by each one. Section 3.9 discusses how to select appropriate combinations of techniques under various operating conditions. Finally, Sections 3.10 and 3.11 recommend specific techniques for use in the baseline and monitoring phases of data collection, respectively.

3.1 Ride Checks

In a ride check, a checker rides on-board the vehicle. Data collected typically include passengers on/off by stop, and arrival time at each stop or at a sub-set of stops. (See Figure 3.1 for a sample ride check form.) At some properties, boarding passengers may be counted by fare category. Experienced ride checkers on some systems also note whether the running speeds on route segments are appropriate. Finally, checkers performing ride checks may also record farebox readings at various points along the route.

Table 3.1

Seven Principal Data Collection Techniques

Technique (reference)	Description
Ride Check (Section 3.1)	Check taken on board vehicle, recording the number of passengers boarding and alighting at each stop and the bus arrival time at selected points.
Point Check (Section 3.2)	Check taken on street, estimating passengers on board vehicle and recording vehicle arrival time. Peak load count taken at peak load point. Multiple point checks include several points along a route.
Boarding Count (Secion 3.3)	On-board count of total number of passengers boarding, most often broken down by fare category.
Farebox Reading (Section 3.4)	Recording of farebox register reading at selected points. Requires registering fareboxes.
Revenue Count (Section 3.5)	Count of revenue in farebox vault, by bus.
Transfer Count (Section 3.6)	Count of transfer tickets collected on each bus which may involve specially-issued transfer tickets.
Survey (Section 3.7)	Variety of techniques in which passengers are asked to provide information.

Figure 3.1

Sample Form for Ride Checks

ROUTE ____________ BUS # ____________ DATE ____________

DAY	ROUTE NO.			DIR.		GAR/RUN NO.				TIME LV. TERM.				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					0									M

WEATHER ____________ TEMP ____________

CHECKER ____________

Bus Stop	Bus Stop Number (17 18 19 20)	Arriving Time: Scheduled (21 22 23 24 25)	Arriving Time: Actual (26 27 28 29 30)	Passengers: Boarding (31 32 33 34 35)	Passengers: Alighting (36 37 38 39 40)	
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			
		M	M			

Since the number of passengers boarding and alighting at each stop is recorded during a ride check, it is possible to determine the load as the bus leaves each stop. Thus, ride checks are an excellent method of monitoring passenger load at all points along the route.

Given the mileage between successive stops, ride checks can also be used to estimate passenger-miles. This Section 15 data item can be simply computed by multiplying the number of passengers on-board leaving each stop by the distance between that stop and the next stop.

3.2 Point Checks

For a point check, a checker stands at a bus stop and records selected data for passing buses. Data collected generally include estimates of passenger load and bus arrival time. (See Figure 3.2 for a sample point check form.) Passenger activity (i.e., boardings and alightings) at the stop where the check is being made can also be recorded by the on-street checker.

Most properties use point checks to observe the "peak load," on a route which is used as an input to scheduling decisions. Peak load is the load on the bus at the peak load point or that point on a route at which the majority of buses have the maximum number of passengers on-board. To measure peak load, one must know the location of the peak load point. Since this point can change, it is necessary to verify the peak load point periodically, generally through a ride check.

For long routes, or routes which serve a number of important activity centers, it may be desirable to conduct counts at a number of points. Such routes might have several points in different areas with loads at or near the peak load. In these cases, any one of several points might dictate the schedule for the entire route and the frequency on short-turn segments. Occasionally point checks are taken at the ends of the routes to also provide a check on running time.

Figure 3.2

Sample Form for Point Checks

ROUTE(S) ______________________ DAY ______________ DATE ______________

AT ______________ DIRECTION ______________ WEATHER ______________ TEMP ______

MAX LOAD ☐ ARRIVING ☐ LEAVING ☐ CHECKER ______________

Day	Route Number	Direction	Garage/Run Number	Bus Number	Bus Stop Number	Arriving Time		Passengers	
						Scheduled	Actual		
1	2 3 4	5 6	7 8 9 10		17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32 33 34 35	
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		
						M	M		

Point checks are typically taken from the street, but one variation is to have the checker briefly board the bus. This practice may become more common as more buses with tinted windows are purchased, since such windows greatly reduce the ability of checkers to see into the bus during daylight hours.

If point checkers board each bus briefly, they can also take farebox readings (treated as a separate technique: Section 3.8). In this way, if checkers are stationed at both ends of the route, they could measure revenue per trip. If they are stationed at one point, they could measure revenue per round trip.

3.3 Boarding Counts

Boarding counts involve the counting of boarding passengers by fare category. Boarding counts are distinguished from riding checks in that the data are often recorded by trip and not by stop. (See Figure 3.3 for a sample boarding count form.)

Boarding counts are most often conducted by vehicle operators using mechanical counters. Operators are often in a better position than checkers to determine fare category, because they can more easily see the fare deposited. In some properties which use operators to perform the counts, the counters are attached to the fareboxes.

Boarding counts generally do not involve the collection of arrival time data. However, if a checker is performing the count, arrival time data at selected stops can be recorded. Checkers might be asked to perform a boarding count (recording boarding passengers by fare category by stop) rather than a ride check (recording passengers on/off by stop). This may be desirable if fare category information is more important than information on passenger alightings and vehicle loads, and if both cannot be obtained simultaneously (e.g., because routes are heavily utilized).

Figure 3.3

Sample Form for Boarding Counts

ROUTE ______________________ BUS # ______________ DATE ______________

DAY	ROUTE NO.	DIR.	GAR/RUN NO.	TIME LV. TERM
1	2 3 4	5 6	7 8 9 10	11 12 13 14 15
		0		M

WEATHER ______________ TEMP ______

CHECKER ______________________

Bus Stop	Bus Stop Number	Arriving Time		Passengers Boarding						
		Scheduled	Actual	Full	Full + Transfer	Reduced	Reduced + Transfer	Transfer Received	All Passes	
	17 18 19 20	21 22 23 24 25	26 27 28 29 30	31 32	33 34	35 36	37 38	39 40	41 42	
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							
		M	M							

3.4 Farebox Readings

Registering fareboxes keep a running total of the amount of money that is collected on-board a bus. (See Figure 3.4 for a sample farebox reading form.) These registers are often used to compute route revenue on a daily or even per trip basis. Register readings are almost always taken at the beginning and end of each day. If a bus remains on the same route all day (i.e., no interlining), these readings can be used to obtain total route revenue. Some properties require drivers to read the boxes at the beginning and end of their shifts. If there is no interlining, this data can also be used to compute route revenue.

Ideally, farebox readings can be taken on a trip-by-trip basis by vehicle operators, so that interlining poses no problem and revenue by time-of-day can also be calculated. As mentioned in previous sections, farebox readings on each trip generally can be recorded by checkers when they perform ride checks or boarding counts.

In the past few years, a number of properties have installed fareboxes which electronically register boardings by fare category (and hence revenue). These fareboxes require operators to register each fare as it is deposited. To use these fareboxes, the operators, in effect, must perform boarding counts.

3.5 Revenue Counts

All properties record total revenue, generally on a daily basis. In some properties, revenue is counted by route every day, or on a sample basis. These revenue counts differ from farebox readings in that the farebox vault must be removed from the bus and the individual coins counted using special equipment at a bus garage or other facility. If buses are interlined on two or more routes, it is difficult to compute accurate route revenue using revenue counts.

Figure 3.4

Sample Form for Farebox Readings

ROUTE ________________ BUS # ________________ DATE ________________

DAY	ROUTE NO.	DIR.	GAR/RUN NO.	TIME LV. TERM
1	2 3 4	5 6	7 8 9 10	11 12 13 14 15
		0		M

WEATHER ________________ TEMP ________

CHECKER ________________

Route Number	Direction	Trip Begin Time	Begin Farebox	End Farebox	Trip End Time
17 18 19	20 21	22 23 24 25 26	27 28 29 30 31 32	33 34 35 36 37 38	39 40 41 42 43
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m
		m			m

3.6 Transfer Counts

Many properties use transfer tickets which may or may not indicate the route of origin. In these properties, it is possible to collect and process the tickets on a sample basis. If transfer tickets do not include the route of origin, a special transfer ticket (perhaps color-coded for a number of intersecting routes) can be distributed, collected and counted for several days to obtain transfer rates by route or origin on a sample basis.

3.7 Passenger Surveys

In "passenger surveys," the passengers are asked to provide information. Transit surveys are generally conducted while passengers are on-board the bus. In longer surveys, passengers may be given the option of mailing back the surveys, which are printed on postage-free mailback forms. On-board surveys may be handed out by operators, by checkers or by special survey administrators. The person distributing the forms helps answer questions and may ask some or all of the questions (in particular if surveying the elderly).

Surveys are the only way to obtain information on passenger travel patterns, characteristics and attitudes. Complete on-board surveys generally include questions of use for general transportation planning as well as those specifically geared to transit management. Typically, questions cover the following topics:

- route on which survey occurs
- fare paid
- other routes used on trip
- origin and destination
- access mode and distance
- trip purpose
- time-of-day of travel
- frequency of use
- age and sex

- occupation or income level
- auto availability

On-board surveys can also be used to count ridership if sequentially numbered survey forms are handed out to all passengers and forms refused by passengers are discarded.

Some properties periodically conduct special purpose surveys to collect limited data. These should not be substituted for the baseline phase survey described above, but can be used to supplement these surveys to acquire accurate data subsequently. Examples of special purpose surveys include:

1. Passholders Survey: On systems with significant (and changing) pass usage, it may be desirable to obtain directly ridership patterns of passholders through a survey. This survey can be conducted when passes are issued or through the mail. These data can then be combined with revenue figures at the route level to update ridership estimates. For systems with growing pass usage these data will allow projection of total revenue for budget planning purposes.

2. Origin-Destination Survey: A survey of travel patterns can be conducted by direct interview on board the bus. One technique used to ensure a complete picture of origin-destination pairs at the stop level is to have the checker record the origin stop when the passenger boards, hand the questionnaire to the passenger to record some other information, and then collect the form and record the drop-off stop when the passenger alights. (This approach requires that the rear door not be used.)

3. Transfer Survey: If two routes are being considered for throughrouting, or monitoring indicates a substantial change in the number of transfers, it may be desirable (in systems which do not issue transfer tickets) to conduct a special transfer survey of certain routes. One way this might be accomplished is to station an interviewer at the stop where two routes intersect, where he/she would ask passengers whether they are transferring. An alternative is to issue coded transfer cards to all boarding passengers on the route in question; the cards are then collected on the second bus.

3.8 Comparison of the Principal Data Collection Techniques

The seven principal data collection techniques listed in Table 3.1 provide a range of different data items depending on individual property and route characteristics. Table 3.2 specifies the data items which can be obtained by using each technique.

Ride checks provide the most complete set of data, especially if boarding passengers can be recorded by fare category. Ride checks, boarding counts and farebox readings all provide reliable and complete data when they are performed by traffic checkers. If drivers are used to collect the same data, experience shows that the results may be less reliable since data collection is secondary to their primary responsibility of operating the vehicle.

Point checks provide reasonably accurate, but more limited, data. Multiple point checks (on the same route) increase the usefulness of this technique by providing information at more than just the peak load point, especially on longer routes which serve more than one activity center. The utility of point checks may decrease somewhat, however, when buses with tinted windows become more common, since tinted windows prevent easy estimation of passenger loads.

Passenger surveys provide a wide range of data items; however, some problems exist in ensuring accurate and unbiased results using survey data (see Section 4.4). Surveys generally should not be used to obtain data items which can be directly observed using alternative technique(s) because of the potential problems with determining the accuracy of survey results.

Revenue and transfer counts provide information on a limited number of data items for those properties with operating characteristics allowing the use of these techniques.

Table 3.2

Data Items Obtained by Seven Principal Techniques

	Technique(1)						
Data Item	Point Check	Ride Check	Boarding Count	Farebox Reading	Revenue Count	Transfer Count	Survey(2)
Load (peak or other)	√	√					
Bus arrival time	√	√	(3) √	(3) √			
Passenger-trips	(4) √	√	√	(5) √			(6) √
Revenue		(7) √	(7) √	(8) √	√		√
Passenger-trips (or revenue) by fare category		(7) √	(7) √	(5) √			√
Passengers on-off by stop		√					√
Transfer rates						(9) √	√
Passenger characteristics, travel patterns, and attitudes							√
Unlinked trips		√	√	(5) √			(6) √
Passenger-miles		√					√
Unlinked trip travel time		√					√
Linked trips		(9) √	(9) √			√	√

Key: √ = applicable
blank = not applicable

(1) Techniques as defined in Table 3.1.

(2) For all survey-collected data other than total passengers, the quality of the data depends on the representativeness of the response.

(3) If time can be recorded.

(4) For "pure" feeder and express routes only.

(5) If electronic multiple fare registering boxes are available.

(6) If surveys are numbered consecutively and distributed to all passengers.

(7) If boarding passengers are recorded by fare category. This typically can only be done with riding checks if boardings are relatively low.

(8) If revenue can be counted by route, this can be substituted for farebox readings although time-of-day data are sacrificed.

(9) If transfer tickets are distributed, collected on terminating route, and identifiable by initial (and intermediate) route(s).

3.9 How to Select Appropriate Combinations of Techniques

Different combinations of techniques can be used to collect the data items listed in Table 2.1. Selecting the best combination of techniques depends on many factors, including the characteristics of individual routes and the system as a whole.

The route structure of a property can influence the relative desirability of point and ride checks. A radial route structure is likely to have points at which a number of routes converge, enabling several routes to be observed by a single checker. Grid systems are less likely to have a single maximum load point, and thus a single point check for each route is less appropriate.

The relative efficiency of the different techniques will also depend in part on the number of buses and level of patronage on a route. Ride checks become more expensive as the number of buses increases. Conversely, point checks become relatively less attractive as the number of buses decreases. If a route is heavily patronized, boarding and riding checks become more difficult to reliably perform. Ride checks can be used to measure ridership by fare category only if boarding passengers can be counted and recorded by fare category. While this may be possible on a lightly patronized route, it is much more difficult, and subject to greater error, on a high ridership route. Nonetheless, it is often important to perform ride checks to obtain detailed boarding and alighting counts for heavily used routes since scheduling and dispatching strategies such as turnbacks and branching can often improve the efficiency of such routes.

The operating policies of a property directly influence the feasibility of certain data collection techniques. For example, properties that do not issue transfer tickets (i.e., have no free or reduced fare transfers) have no mechanism to directly count route-to-route transfers. These properties may have to rely on a passenger survey to determine transfer rates.

There are two operating policies, however, which constrain the selection of appropriate combinations of techniques to a small number. The first constraint is the ability of vehicle operators to record reliable data. Reliable driver-collected data can reduce the cost of a data collection program dramatically. It allows a property to obtain a much larger amount of data than could be afforded if traffic checkers had to be used. The reduced cost and higher sample sizes must be weighed, however, against the possible reduced accuracy of the data obtained by drivers. The possible second constraint is the availability of registering fareboxes. Registering fareboxes allow a driver, on-board checker or even a street checker to monitor route revenue and, indirectly, total ridership. Regular farebox readings may provide accurate route revenue figures and could provide a check on total ridership figures generated from driver trip sheets.

3.10 Recommended Techniques for Baseline Phase of Data Collection

Several options for combining data collection techniques are preferred for common property characteristics. These are presented below along with a brief discussion of other alternatives. While these recommendations generally yield the complete set of data at the lowest cost for a typical property, specific local characteristics might make other combinations more desirable. For this reason, a property should select its own combination of techniques. The following recommendations and discussion are intended to provide guidance for this choice.

For the initial baseline data collection phase, the following set of techniques is recommended:

- ride checks (plus possible supplementary point checks);
- farebox readings or boarding checks;
- on-board surveys.

The ride check is included in the baseline phase in order to obtain boardings and alightings by stop and, thus, average loads on each route segment. Supplementary point checks are needed only when the sample required for load data exceeds that required for total boardings (since it is less costly to gather additional peak load data by using a single point checker than by using on-board checkers). Farebox readings or boarding checks provide complete route revenue information, although only the latter breaks down ridership and revenue by fare category (and probably should be included by any property which can reliably use operators to perform such counts). Finally, the on-board survey provides a variety of passenger information which cannot be collected in any other way.

3.11 Recommended Techniques for On-going Monitoring

The recommended techniques for the on-going monitoring phase depend more heavily on property and route characteristics. If a property can use drivers to collect total boardings, the following combination of techniques is recommended:

- point checks;
- boarding counts (by operator);
- farebox readings (if registering farebox available).

Properties which cannot depend on drivers to obtain reliable data have several options. The best combination often includes direct monitoring of peak load, total boardings and farebox revenue through:

- ride checks (plus possible supplementary point checks);
- farebox readings (if registering fareboxes available).

However, for routes which exhibit a strong relationship between peak load or revenue and total boardings (as measured during the baseline phase), route performance can be monitored simply by using point checks. (It is assumed that a street checker at a busy stop could also board the bus and obtain a farebox reading, if available.) Although using a load or revenue conversion factor to estimate total boardings requires larger sample sizes than does measuring load or revenue alone (see Section 4.5), often the overall expense of this option is less since on-board checkers are not required. The key to using this option is the test of the relationship between the data items, as described in Section 4.5. In these cases, the least costly data collection program is determined by comparing the relative sample sizes, as described in the following chapter.

CHAPTER 4

SAMPLING

Once the techniques to be used have been selected, it is necessary to determine the amount of data required. A combination of quantity of data (i.e., sample size) and timing of data collection is called a "sampling plan." A sampling plan is a reflection of two factors: the desired accuracy and the inherent variability of the data. The greater the accuracy desired and the higher the variability of the data item, the greater the amount of data which must be collected.

The concept of sampling is introduced and the statistical and practical issues related to determining sample size are discussed in this chapter. Detailed procedural steps for determining sample size are specified in Chapter 6 and Appendix A. The various options available for the timing of data collection efforts are also described, so that a property can easily develop several alternative sampling plans for which total cost can be estimated.

In this chapter, Section 4.1 discusses the concept of sample accuracy and the implications of selecting specific route and system accuracy levels. Section 4.2 then discusses data variability and provides a basis for two measures of variation in the determination of final sample size and sampling plans. In Section 4.3, the method for determining sample size for the direct collection of data is described. This is followed by discussions of modifications to this method when a property chooses to perform on-board passenger surveys (Section 4.4) and to use conversion factors (Section 4.5). Sampling plan (i.e., timing) and sample selection considerations are discussed in Section 4.6. Special sampling considerations for UMTA-required Section 15 data are discussed in Section 4.7. Finally, statistical interpretation and use of the sample data are discussed in Section 4.8.

4.1 Accuracy

Most data are collected using some type of sampling strategy, since 100 percent coverage of all routes every day is generally infeasible. In any sampling strategy, there is some uncertainty about how well the sample data represent the true values of the underlying data. For this reason, it is important that an appropriate level of accuracy be chosen.

Accuracy has two components: an error range ("tolerance") and a probability ("confidence") level. The tolerance indicates the range around the observed value within which the true value of the data item is likely to lie. For example, for Section 15, the sample is based on the true value being within ±10% of the observed value. The confidence level indicates the probability that the true value is within the tolerance range around the observed value. For Section 15, a confidence level of 95% is specified. Thus, for Section 15 data, there is a 95 percent chance that the true value of the data item is within ±10% of the observed value.

Changes in tolerance and confidence levels can impact transit operating decisions. For example, suppose a property has a service standard that states that a bus is added if the peak load exceeds an average of 75 persons during any 15 minute period. If this property chooses to measure load to within ±10%, and the average peak load is measured at 60 passengers, we know (with a certain probability) that the true value is 60 ± 6, or between 54 and 66. In this case, decreasing the error range to ±5% would provide no further useful information (since it is clear that the standard of 75 is not being violated). In fact, a tolerance of ±20% would still be acceptable, since we would know the true value is between 48 and 72. In this case, the only time a more accurate estimate is required is if the standard is very close to being violated. For example, if the measured load was 70 ± 10%, or between 63 and 77, a smaller error range would be appropriate the next time the route was checked to ensure that, in fact, the loading standard is not being violated. In general, the selected tolerance becomes more important as the measured value of the data item increases.

4.1.1 Recommended Accuracy Levels

The accuracy levels selected by a property should be influenced by the data item being measured, the type of route (i.e., capacity constrained or not) and the time of day being analyzed. It is generally important to monitor route performance more accurately during peak periods than during other periods since a higher than proportional percentage of a system's resources are allocated to provide peak period service. Similarly, a capacity-constrained route (i.e., with standing loads in the peak period) generally requires greater accuracy since vehicle and manpower allocation decisions typically are made based on a loading standard. Recommended tolerances for the basic data items discussed in Chapter 2 are presented in Table 4.1 for various time periods and route types. These recommendations are based on uses of the different data items by the industry and an analysis of the sample size requirements assuming alternative error ranges.

It is also recommended that a confidence level of 90% be used at all times for collecting route-level monitoring data. The 90% confidence level provides a balance between obtaining highly accurate route-level measurements and the overall cost of the collection program.

4.1.2 Route Versus System Level Accuracy for Section 15

All transit properties are required to report certain systemwide data under Section 15 of the UMTA Act. (These data items were identified in Section 1.4 of this manual.) Generally, systemwide data which are obtained by aggregating route level data will be more accurate than the individual route data. Equation 4.1 defines systemwide tolerance in terms of the required Section 15 systemwide confidence level and selected route level tolerance and confidence levels:

Table 4.1

Recommended Tolerances for Basic Data Needs

Data Item	Time Periods	Route Type	Recommended Tolerance
Route Level			
Load, Bus Arrival Time, Total Boardings, Revenue	Peak	Capacity-Constrained	±10%
	Peak	Not Capacity-Constrained	±15%
	Midday	All	±15% to ±20%
	Evenings, Owl & Weekends	All	±30% to ±50%
Boardings (revenue) by fare category	Peak, Midday	All	±20%
	Evenings, Owl & Weekends	All	±20%
Boardings and alightings by stop	All	All	±50%
Transfer rates between routes	All	All	±30%
Passenger characteristics, attitudes, & travel patterns	All	All	±30%
Systemwide			
*Unlinked passenger trips, *Passenger-miles, *Average unlinked passenger travel time, Linked passenger trips	All	-	±10%*

*Required by Section 15 (at 95% confidence level); if route level data are obtained at the tolerances recommended here, systemwide tolerance will generally be within ±10% (see Section 4.1.2).

$$T_s = \frac{t_s T_r \sqrt{\sum_r (B^2)}}{t_r \sum_r B} \qquad (4.1)$$

where T_s = systemwide tolerance level (e.g., $\pm$ .03)

T_r = route tolerance level (e.g., $\pm$ .15)

t_s = t-value for systemwide confidence level (e.g., 1.96 for 95% confidence)

t_r = t-value for route confidence level (e.g., 1.645 for 90% confidence)

B = average daily boardings on each route r

r = 1, 2, 3, ..., R, where there are R routes in the system.

Thus, one need only know the total daily boardings on each route and the desired accuracy for route-level sampling to estimate the accuracy of systemwide data obtained from summing the route sample means. If a particular property has roughly the same total daily boardings on each route in the system, this equation is simplified even further to:

$$T_s = \frac{t_s T_R}{t_r \sqrt{R}} \qquad (4.2)$$

Using equation 4.2, systemwide tolerances have been computed for varying system sizes (i.e., number of routes) and route tolerances. As shown in Table 4.2, for large systems, quite high accuracy levels are achieved by aggregating route level data if boardings are approximately the same on all routes.

Equation 4.2 gives a lower bound for the systemwide tolerance; if total boardings vary greatly among routes, then the real systemwide accuracy will be less than suggested by this second equation.

Table 4.2

Systemwide Tolerances
Achieved Using Route Level Data*

System Size (# of routes)	Route Level Tolerance ± 10%	± 15%	± 20%	± 30%
5	± 5.3%	± 8.0%	± 10.7%	± 15.0%
10	± 3.8%	± 5.7%	± 7.5%	± 11.3%
20	± 2.7%	± 4.0%	± 5.3%	± 8.0%
50	± 1.7%	± 2.5%	± 3.4%	± 5.1%
75	± 1.4%	± 2.1%	± 2.8%	± 4.1%
100	± 1.2%	± 1.8%	± 2.4%	± 3.6%
150	± 1.0%	± 1.5%	± 1.9%	± 2.9%

* Route confidence level assumed to be 90% and system confidence level assumed to be 95%; total boardings assumed the same on all routes in a system.

Table 4.3
Systemwide Tolerance Achieved
Using Actual Route Level Data
From Two Properties*

System Size (# of routes)	System Tolerance (assuming ± 15% route tolerance)
5	± 9.7% to 12.3%
10	± 6.9% to 8.6%
20	± 4.8% to 6.1%
50	± 3.1% to 3.8%
134**	± 1.9%
165***	± 2.1%

* Route confidence level assumed to be 90% and system confidence level assumed to be 95%; total boardings distributed from route to route as in actual data from Chicago CTA and Los Angeles SCRTD.

** Chicago CTA case

*** Los Angeles SCRTD case

The actual distribution of boardings among routes in the Chicago Transit Authority and Los Angeles SCRTD were used to determine systemwide tolerance using the exact equation (4.1). As shown in Table 4.3, different boardings among routes do not increase the systemwide tolerance substantially over the constant boardings case in Table 4.2 at the ±15 percent route tolerance level. These tables strongly support using route level data to estimate system totals for purposes such as Section 15 reports. The accuracy achieved using route level data exceeds that required for Section 15 reports except for operators with fewer than about 10 routes.

4.2 Inherent Data Variability

All the data items which are of interest to properties are variable, and through sampling, the aim is to estimate the true mean for that data item. The more variable a data item is, the larger the sample that is required for accurate estimation of the mean. For example, if every passenger on a bus route were counted for one day, the sum would be the actual route ridership for that day. However, this value may not be equal to the average daily route ridership since ridership will vary from day to day. For example, total boardings on a particular Wednesday might be 10% lower than on Monday and 5% higher than on Friday of the same week.

This type of variation, known as "between-day" variation, must be estimated to determine sample size (as discussed in the next section). The greater the between-day variation, the lower the probability that a single day value is close to the true mean. Thus, when the between-day variation is high, more days must be sampled to obtain an "accurate" estimate of average daily ridership.

If all passengers are not counted on a single day, then within-day variation must also be considered. If the ridership on every trip were exactly the same, an accurate estimate of the total ridership for the day or time period could be obtained by counting ridership on only one trip. However, as the variation in ridership from trip to trip becomes larger, more trips must be sampled to accurately estimate total daily (or time period) ridership. This type of variation is known as "within-day" or "within-period" (if estimates are needed for ridership in different periods) variation and is also an input into the sample size determination in the next section.

An extensive analysis of transit data variability from several different cities was performed for this study. This analysis did not identify any easily applied rules-of-thumb concerning data variation. Therefore, it is recommended that each property collect (or assemble from existing data) at least three days of route-specific data to determine variation measures for individual routes. This "pretest" sample can be assembled for any data item, but generally, the easiest to collect and most representative data item is load or total boardings. (This procedure is discussed further in Section 4.3.2.)

Since it may be difficult for a large property to collect new data on all routes, it could use a shortcut to minimize initial data collection. This involves applying the variability measures calculated for one route in a system to other routes of a similar type. To do this, a property must develop a route classification scheme, wherein routes are classified according to factors which are likely to influence variability. These factors include such route characteristics as headway, length, and functional type (e.g., feeder, express, crosstown). Appendix B provides a description of a general route classification scheme.

The degree to which a property can usefully classify routes with similar data variability depends on local operating conditions and knowledge of route ridership patterns. A simple, yet potentially effective classification scheme is based on headway characteristics. For example, three route headway categories which might be used are:

1. less than or equal to 10 minutes (i.e., routes with heavy demand for which passengers do not necessarily schedule their trips to coincide with a particular bus);

2. between 10 and 30 minutes (i.e., routes with moderate demand for which passengers generally schedule their trips to catch a particular bus);

3. 30 minutes or greater (i.e., routes with policy headways for which service frequency is not determined by demand).

The boundaries for each headway classification could be adjusted based on local conditions; for example, 15 minutes could be used instead of 10 minutes between the first and second categories. This type of classification scheme is recommended here because it is simple and because evidence obtained during this study suggests that data variability is related to route headway. However, a property can apply the sample size procedures discussed in the next section to any locally developed classification scheme.

The route classification approach may not always provide a good solution to the problem of determining route level variation measures. It is recommended that a property calculate the variability measures for several routes in a number of possible classification schemes to determine if they are similar enough to support the classification approach. If the calculated measures in the same class are substantially different (and thus suggest significantly different sample sizes), the proposed classification categories should be discarded.

4.3 Sample Size Determination for Direct Measurement Techniques

This section discusses the procedures for determining the sample size for data items using direct measurement techniques. The following sections modify these basic procedures to determine sample sizes for on-board surveys and to account for the use of "conversion factors" to minimize the cost of monitoring.

The procedure for determining sample size involves three basic steps:

1. Determine route characteristics (number of round trips and number of buses assigned in each time period) for all routes in the system;

2. Determine statistical inputs for sample size calculations for at least one data item - preferably load or total boardings;

3. Use sample size tables in Volume 2 of this manual to select a sampling plan.

4.3.1 Determine Route Characteristics

Individual route characteristics must be compiled for each time period during the day. The individual route characteristics which are needed include the number of vehicle round trips, and the number of buses assigned to the route during each time period. Only the number of round trips are needed for sample size calculations, but the number of buses assigned to each route as well as information such as the number of load check points on each route are needed to estimate checker requirements and total costs as discussed in Chapter 5.

It has been assumed that most properties will choose to obtain separate sample data for the four basic weekday time periods (a.m. peak, base or midday, p.m. peak, and night) as well as all day Saturday and all day Sunday. Sample sizes for any other time period (including all day) can be determined using these procedures. If this is done, the route characteristics outlined above should simply be compiled for the time period of interest.

4.3.2 Determine Statistical Inputs

Statistical inputs for desired accuracy and inherent data variability are necessary to estimate sample sizes. The choices available for desired accuracy (i.e., confidence level and tolerance range) were described in Section 4.1. For ease of use, the sample size tables contained in Volume 2 have been limited in order to keep the total number of sample size tables manageable. For all the sample size tables, the confidence level of 90 percent has been adopted. Each table shows the sample required for different tolerance ranges: ±10 percent, ±15 percent, ±20 percent and ±30 percent. If a property determines that these accuracy limits are unacceptable for its particular purposes, detailed formulae are also given in Appendix A to calculate sample sizes for different confidence and tolerance levels.

The measures of the inherent variability of the data to be monitored are key inputs to the sample size determination. For each route (or route classification if similar routes demonstrate similar variability characteristics), measures must be calculated for "within-day" variation and "between-day" variation for the time period of interest. These measures are expressed in terms of variances or "coefficients of variation," which are precisely defined in Appendix A. Separate coefficients should generally be calculated for different time periods (a.m., base, p.m., night) since total sample sizes are likely to be minimized by grouping ("stratifying") trips which exhibit lower overall variability. (Data for trips within a single time period vary less than data collected throughout a day.)

If data cannot easily be stratified by time period, however, coefficients of variation and corresponding sample sizes can be determined using all-day data. Coefficients of variation should also be calculated separately for each direction of travel on a route, since data variation is likely to be different for the peak (i.e., higher ridership) and off-peak directions, and operations planning decisions must be made primarily based on route performance in the peak direction.

To calculate coefficients of variation for each route, at least three days of data (for at least 75 percent of the trips) for each time period of interest must be analyzed. This three-day sample should be collected in a special "pretest," compiled from existing data (which had been collected within a single three month period), or be a combination of recent data and newly collected samples. If many days of data are available (e.g., from driver counts taken every day), it is recommended that at least five and up to ten days of data selected randomly from a three month period be used to calculate these coefficients of variation.

Obtaining this amount of data for each data item listed in Table 2.1 may be difficult for many transit operators. An analysis of the variability of these data during this study showed that load data typically exhibit the same or higher variances than total boardings and most of the other data items. Therefore, it is recommended that at least three days of existing or newly collected load data (which are generally obtained using a point check for 100% of the trips during any given time period) be used to calculate the coefficients of variation for each route. It should be noted, however, that use of load variability measures may result in larger than necessary sample sizes for measuring total boardings. If a property feels that the load sample sizes suggest an unreasonable burden for collecting total boardings data (i.e., a three or more day sample), one of two courses of action should be pursued: 1) three-day samples could be collected on several routes in each route classification to determine separate sample sizes for total boardings data; and 2) the baseline phase could be completed using three-day boardings data samples for all routes for which load variation measures indicate that three or more day data are required, after which new between-day variation measures can be calculated for the boardings data.

For properties which have total boardings data available from driver counts by trip, it is recommended that these be used in lieu of load data to determine these coefficients of variation. Coefficients of variation calculated from total boardings data should be inflated by about 30 percent to ensure that accurate load data samples are obtained for the same routes, since load coefficients of variation can be 30 percent higher than the corresponding boardings coefficients.

A few individual data items may exhibit greater variability than both load and total boardings. This means that the sample size determined using the coefficients of variation for either of these two data items will provide less accurate estimates for those other items with higher variability than the selected tolerance. In most cases, this will not be significant since the items with higher coefficients of variation (generally passengers by fare category and passengers on/off by stop) need not be obtained at the same (high) levels of accuracy as load and boardings. In any case, the procedures outlined here and in Appendix A can be easily applied to any data item by simply substituting the appropriate coefficients of variation for the desired data item in place of the load or boardings coefficients.

Detailed worksheets and instructions for calculating the two coefficients of variation for any data item, along with the formulae which define them, are included in Appendix A. The calculations require several simple steps. The total time required is reasonable and can be substantially shortened if a programmable calculator or computer is used to execute the basic calculations.

4.3.3 Determine Sample Size From Tables in Volume 2 of this Manual

The procedure for using the sample size tables in Volume 2, Sample Size Tables, is quite simple. For each data item for which data are available, the within-day and the between-day coefficients of variation are calculated for each time period of interest (during the day), as discussed above. In most cases, the coefficients which were calculated for the peak (higher ridership) direction during a given time period should be used to determine the sample size for both directions. (This is because it is generally more important to have accurate data in the peak direction for which the level-of-service is primarily determined.) If both directions on a route have similar numbers of passengers during a given period (e.g., midday), it is recommended that the coefficients with the higher values be used to ensure the desired accuracy in either direction.

The set of tables corresponding to the value of the within-day coefficient (or the next higher value given) is identified first. A property then uses the value of the between-day coefficient and the number of scheduled round trips in the period to identify the table which lists the appropriate sample sizes for a range of tolerance levels. The desired tolerance is then selected and a property is provided with several different combinations of trips and days which will all provide data at the desired tolerance and 90% confidence level. For example, Table 4.4 shows the Volume 2 table for a within-day coefficient of .40, a between-day coefficient of .08 and 15 scheduled trips. An operator selecting ±15 percent tolerance using this table would be provided eight different sampling plan options.

A property may have to adjust the sample sizes and strategies (i.e., trips and days) determined here to account for detailed operating issues unique to the property. For example, a property should eliminate any combinations of trips and days which require a larger checking staff on any one day than is readily available (e.g., a property with 10 checkers could not ride-check 100 percent of the trips on a route which requires 15 buses during the peak periods). Generally, a property should avoid any sampling plan which requires a ride check of virtually 100 percent of the trips in a period (because of the possibility of missed trips by either the vehicle or the checker). Also, checking 100 percent of the trips should be avoided if a route has a lot of interlining, since a vehicle may make only a one-way trip on a route and checker hours would be wasted.

Constraints such as maximum swing, night, and weekend work policies for checkers are not explicitly dealt with here. Each property must determine how these affect the cost of data collection and must adjust the sampling plan accordingly.

Table 4.4

Typical Sample Size Table

STATISTICAL INPUTS

WITHIN-DAY COEFFICIENT	.400
BETWEEN-DAY COEFFICIENT	.080
NUMBER OF SCHEDULED TRIPS	15

SAMPLING PLAN OPTIONS

+/- 10 PERCENT TOLERANCE		+/- 15 PERCENT TOLERANCE		+/- 20 PERCENT TOLERANCE		+/- 30 PERCENT TOLERANCE	
NUMBER OF DAYS	NUMBER OF TRIPS PER DAY	NUMBER OF DAYS	NUMBER OF TRIPS PER DAY	NUMBER OF DAYS	NUMBER OF TRIPS PER DAY	NUMBER OF DAYS	NUMBER OF TRIPS PER DAY
44	1	20	1	12	1	5	1
22	2	10	2	6	2	3	2
14	3	7	3	4	3	2	3
10	4	5	4	3	4	1	5
8	5	4	5	2	5		
7	6	3	6	1	9		
6	7	2	8				
5	8	1	13				
4	9						
3	11						
2	14						

4.4 Sample Size Determination for On-Board Passenger Surveys

On-board passenger surveys must be conducted during the baseline phase since they are the only source of data on passenger travel patterns and passenger characteristics and attitudes. In this section, the special sampling considerations associated with surveys are discussed.

The principal purpose of an on-board survey is to estimate the proportion of the total passengers using a particular route who have a specific characteristic, e.g., are elderly or transfer passengers. As the number of passengers who are surveyed increases, the margin for error in estimating this proportion decreases. The margin of error cannot be reduced without limit for two very important reasons:

1) the response rate is inevitably less than 100%; and

2) the whole population cannot, in practice, be identified.

The response rate limits the total sample which is obtained even if an attempt is made to survey all passengers on a given day. In addition, since the total population of users of a specific route do not all ride the bus on any single day, it is extremely difficult to determine the total population (whether it is defined in terms of passengers or passenger trips). Complicating these limitations, a survey must be conducted on a single day since passengers generally are not willing to fill out the same survey more than once. As a result, infrequent users are underrepresented in the response to a one-day survey.

In conducting an on-board survey, several other issues must be considered:

1. Should the questionnaire be hand-in or mail-back?

   In order to maximize the response rate and to avoid bias, it is suggested that both options be provided to the passenger. Response rates are usually higher on hand-in surveys; however, on crowded buses, very few people are willing and able to complete a long questionnaire en route. Similarly, response is biased towards those boarding early enough to get a seat.

2. <u>Should the survey be conducted inbound (or outbound) only?</u>

This is a common method for avoiding asking the same person to fill in the survey twice; however, it fails to provide information on timing and even routing of the return trip. If this approach is adopted, it is advisable to request limited information on other transit trips made that day and to ask that a person complete only one questionnaire.

3. <u>How is a sampling plan developed?</u>

As with other data collection techniques, one need not survey every passenger in order to obtain adequate data. The size of the sample needed depends upon desired confidence and tolerance levels, the size of the population, and the expected distributions for the data item of interest (see equation below). Once the sample size is determined, the sampling plan is typically developed by determining the number of bus trips to be sampled, given a conservative estimate of expected return rates (see below). Surveys would be handed out to all passengers on selected trips.

4. <u>What is the expected response rate?</u>

Not everyone fills out a survey form. The response rate depends on such factors as crowding, route length, and survey length. Transit properties around the country have experienced response rates from 15% to 90%. It is always best to be conservative in projecting response rate (i.e., project a low level of response), since the cost of handing out more surveys than necessary is not likely to be great, and it is not necessary to process all surveys returned if the response rate exceeds expectations.

One problem with response rates is that not all segments of the population are likely to respond in the same proportion. This may bias the results, as discussed below.

5. <u>How can bias be dealt with?</u>

The problem of bias is always present in surveying. It exists when the survey responses are not representative. Sampling design can be very effective in reducing bias. Any device to reduce the probability of differential response rates should be used, including:

1. Offering questionnaires to all passengers on a bus, to avoid bias introduced through the selection of passengers by the checkers.

2. Providing a mail-back option to avoid higher response rates from those obtaining seats (not a random selection of all passengers).

3. Keeping the questionnaire simple so that everyone can understand it.

4. Making foreign language versions available in heavily ethnic neighborhoods.

5. Selecting buses on which to survey either randomly or uniformly from the time period of interest.

6. Obtaining control totals at fine enough levels of disaggregation to allow use of expansion factors as described below.

Once the survey has been completed, the processing phase should be established to account for differential response rates by different segments of the population. This can be done by defining expansion factors at the finest possible level of detail consistent with the number of responses obtained. For example, suppose that the survey results for a route show two distinct response rates for different segments of the route. Expansion factors should then be estimated for each separately, rather than for the route as a whole. This will not eliminate bias completely, but should reduce it substantially. An alternative method which should also be considered is developing expansion factors by fare category if different response rates are observed on this basis.

Given these considerations, the number of surveys required in any particular case is given by the following equation:

$$n = \frac{t^2 p(1 - p)}{d^2 r} \qquad (4.3)$$

where n = number of passengers to be surveyed;

t = t-statistic for desired confidence level (t=1.645 for recommended 90% confidence level);

p = expected proportion of the passenger characteristic or data item of interest (for the worst case or largest sample, assume p = 0.5);

d = tolerable margin of error as a percentage of the mean value;

r = expected response rate.

Since many different data items are typically included in one survey and the proportion for each item is not generally known before conducting the survey, it is recommended that the value used for "p" in the above equation be 0.5, which gives a worst-case sample size. In practice, surveys are usually conducted by handing out questionnaires to all passengers boarding selected trips. Once "n" is determined using Equation 4.3, the number of bus trips to be surveyed should be estimated by dividing n by the expected number of boardings per trip.

Once the survey has been conducted, the actual number of responses may differ from the calculated value of "n". The margin of error associated with a specific proportion once the surveys have been analyzed can be determined by rearranging the sample size equation to:

$$d = \sqrt{\frac{t^2 p(1 - p)}{nr}} \qquad (4.4)$$

where all definitions are as above, except that the actual values of n and p can be inserted for any data item.

Because of the limited, and to some extent unpredictable and uncontrollable nature of survey accuracy, it is not recommended that surveys be used to obtain data which can be reliably estimated using an alternative technique. Similarly, properties should use care in acting upon results of a survey which are not supported by other evidence.

4.5 The Use of Conversion Factors

Conversion factors can be used to reduce the total resources required for data collection in the on-going monitoring phase provided that specific conditions, which will be defined in this section, are met. Conversion factors are most useful for estimating data items which are important, but expensive to measure directly. The primary example is the estimation of total boardings per trip from peak load counts or farebox readings.

Conversion factors are the constants in an equation which relate the value of a data item which is measured directly to another data item which has not been measured. For example, in the equation:

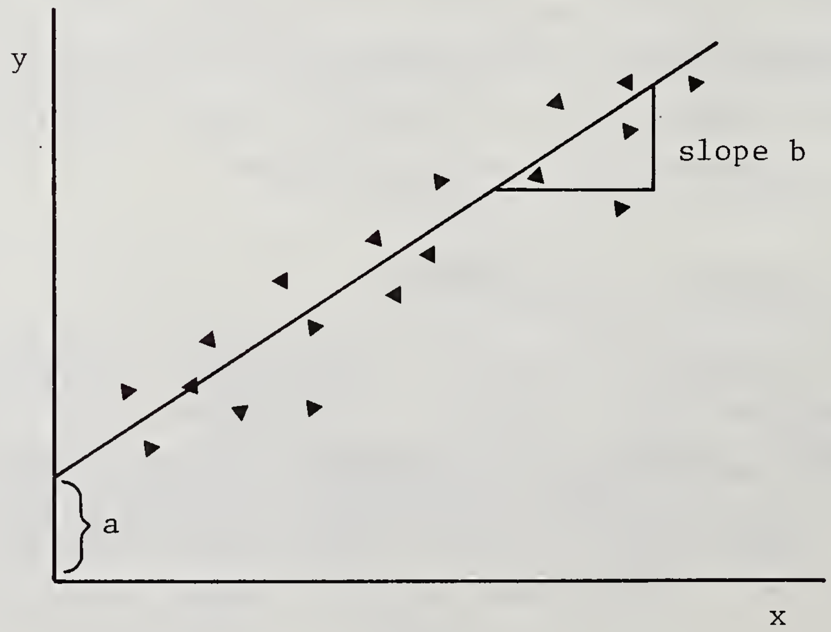
$$y = a + bx, \tag{4.5}$$

a and b are conversion factors which allow y to be estimated based on a measured value of x. In this case the factors a and b are estimated from a sample of paired data for x and y, as shown in Figure 4.1. A line is fitted to the data points which minimizes the sum of squares of the distance of each point from the line.

The technique for determining the best line is known as ordinary least squares regression (regression for short) and standard packages exist for applying it on all programmable calculators, and many pocket calculators. One standard output from the regression is the variance associated with the

Figure 4.1

Conversion Factors



equation, referred to by s^2. Higher values of the variance mean that the best line does not closely fit the sample of data.

The variance is an important measure of the "goodness of fit" of the line to the data and, hence, of the strength of the relationship between the two variables x and y. Specifically s^2 can be used to define a confidence interval around the mean value of y, as follows:

$$c = \frac{t\sqrt{s^2}}{\bar{y}\sqrt{n}} \qquad (4.6)$$

where c = interval at 90% confidence level (as percent of the mean);

t = t-statistic for desired confidence level (t=1.645 for the recommended 90% confidence level);

s^2= variance from the regression;

n = number of data points input to the regression;

$\bar{y}$ = mean value of the sample y input to the regression.

This confidence interval specifies the range of uncertainty which would be associated with using the equation to estimate the value of y at a given value of x. If this confidence interval is larger than the accuracy desired for y, then the equation cannot be used, and it is necessary to collect data y directly, rather than estimate it. On the other hand, if the confidence interval is small compared with the accuracy desired for y, then the equation is a satisfactory basis for estimating y.

There are three distinct aspects to the use of conversion factors:

1) developing conversion factors

2) sampling using conversion factors

3) monitoring using conversion factors.

In the ensuing discussion, it will be assumed that the relationship of interest is between peak load counts (observed) and total boardings (to be estimated if possible).

4.5.1 Developing Conversion Factors

In order to determine whether there is a strong relationship between the two data items, it is necessary to gather a sample of both data items. Specifically, in the baseline phase for each route and time period for a number of bus trips, the total boardings and corresponding peak load counts must be obtained. Regression is then used to estimate the best linear equation between the data items in the form of Equation 4.5, where y is boardings per trip, x is peak load count, and a and b are parameters estimated by the regression.

Equation 4.6 is then used to determine the confidence interval associated with the regression equation, where s is obtained from the regression package, n is the number of data points used to estimate the relationship, and y is the mean boardings per trip in that data set. If the resulting confidence interval is greater than the accuracy level desired for boardings per trip, then conversion factors cannot be used for this route and time period because the estimates of boardings would be too unreliable. It may be possible to improve the quality of the equation by gathering additional data on both boardings and peak loads, but otherwise, the monitoring phase should be designed to collect boardings per trip directly.

If the confidence interval is smaller than the desired accuracy level, then the next step is to determine the sample size required to use the conversion factors for estimating boardings from directly measured peak loads.

4.5.2 Sampling Using Conversion Factors

If an acceptably small confidence interval exists for the regression equation, the operator now has the option of using the conversion factors and measuring peak loads to estimate boardings per trip. Here the question is one of cost effectiveness: it may be less expensive to conduct boarding counts even though a good regression equation has been estimated, because a smaller sample is always required for direct measurement than for estimates using conversion factors.

To determine the sample size required when using conversion factors, it is necessary to add s^2, the variance of the regression estimation, to the variance of the population of the data item being estimated. Since we are estimating the boardings per trip, the total variance for sample size calculation s_t^2, is the sum of the variance of the distribution of total boardings[1] and the variance from the regression. The total number of trips to be sampled can then be obtained from the following equation:

$$n = \frac{t^2 s_t^2}{d^2 \bar{y}^2} \qquad (4.7)$$

where: n = total number of trips to be sampled;

t = t-statistic for desired confidence level ($t = 1.645$ for 90% confidence level);

s_t^2 = total variance associated with boardings;

d = tolerance range (as a fraction of the mean boardings);

$\bar{y}$ = mean boardings per trip.

To determine the number of days of peak load counts needed, it is assumed that on each day sampled, all trips will be counted. (This is certainly most efficient for peak load counts.) Using this assumption, the number of days to be

---

[1] The variance of the distribution of total boardings is estimated by adding the between-day variance to the within-day variance (or the square of the between-day coefficient of variation multiplied by the overall mean boardings to the square of the within-day coefficient of variation multiplied by the overall mean boardings, i.e., the quantity $\left[(CV_{BD} \cdot \text{mean})^2 + (CV_{WD} \cdot \text{mean})^2\right]$.

sampled equals the total trips to be sampled, n, divided by the number of trips operated daily within the time period of interest (and rounding up the result to the next whole day). The sampling plan using conversion factors then consists of making counts (either load or revenue) on all trips for the indicated number of days.

The resulting sampling plan may or may not be less expensive than that developed for directly monitoring boardings per trip. However, results of the Chicago field tests indicate that, for many routes, monitoring by using conversion factors is likely to be less costly than directly counting boardings.

4.5.3 Monitoring Using Conversion Factors

A property which chooses to use conversion factors can easily estimate total boardings based on measured peak loads. This would be done by calculating the mean peak load which has been measured for each time period and inserting this value in the equation which was derived from the baseline phase data. For example, let us assume that the equation for estimating boardings on a particular route which was developed during the baseline phase is:

$$y = 10 + 1.5x \qquad (4.5a),$$

where x is peak load and y is total boardings. Then, if the mean peak load measured for an a.m. peak period is 50 passengers per trip during a subsequent monitoring phase, mean boardings for the same period would be estimated to be 85 passengers per trip. As discussed previously, the regression equation can be derived and used in this way to predict boardings from farebox readings or, for that matter, to predict any data item from another with which a strong enough relationship exists.

4.6 Sample Selection, Seasonal Considerations and Timing

4.6.1 Random Versus Systematic Sampling

For each sampling plan determined by a property, the desired accuracy is only achieved if the final sample is selected randomly. Random sampling refers to a method of selection whereby each possible sample has an equal chance of being chosen.[1]

For example, if the procedures described in this chapter call for 15 out of 20 vehicle trips to be sampled for two days to obtain an estimate of the average total passenger-trips for the a.m. peak period for a season within ±10%, the tolerance level only applies if the 15 vehicle trips are selected randomly for each of two randomly selected days during the season. If the first 15 trips during the a.m. period were selected for two consecutive days, the true average passenger trips may well fall outside the indicated range.

There is no easy way to determine what level of accuracy is actually achieved if a nonrandom sample is selected. A property should strive to select as random a sample as possible. Sampling the same route on consecutive days should be avoided whenever possible. The selection of different, widely scattered days (if more than one is needed) during a season helps to ensure the representativeness of the sample

---

[1] Various techniques exist for selecting a random sample and are described in most standard statistics texts. One standard technique is to use random number tables. To do this, number each day in the season or other sampling period consecutively starting with the number "1". Then systematically go down the lists of numbers in any section of the random numbers table, writing down the numbers (and rejecting any numbers higher than the highest number of days in the sampling season) until the number of days to be sampled has been identified. The property would sample the days corresponding to the numbers listed from the scan of the table. Once the days are chosen, the same procedure can be repeated with trips if less than 100% of the trips need be sampled.

data. Truly random selection of trips within a given day or time period may wreak havoc on checker schedule assignments. For this reason, it is recommended that trip selection within a specific time period be more systematic, perhaps by selecting random driver runs instead of vehicle trips. The major criterion in selecting trips within a time period should be to spread the sample over the entire period. (Note that if headway distribution data are important to a property, the sample should include groups of at least three consecutive driver runs so that differences in the headway between consecutive buses can be directly computed.)

4.6.2 Seasonal Considerations

The timing and frequency of conducting the baseline and monitoring data collection phases with regard to the season of the year is highly dependent on the characteristics of the individual property and its routes. A property should initially conduct the baseline phase during any season of its preference. For at least one year after the baseline phase has been completed, however, it is recommended that monitoring phases be conducted corresponding to those periods of the year for which route level-of-service (i.e., scheduling) changes can be made. If schedule changes are not normally made during the year (as in many small properties), it is recommended that all routes be monitored during two seasons (one when schools are in session and one when they are not in session) during the first year.

This procedure will allow the property to determine the extent of route-level seasonal variation as well as to flag routes which exhibit significant ridership growth or shrinkage trends. Some simple rules-of-thumb are recommended to determine if measured ridership changes over the first year of monitoring indicate significant seasonal variation (which would require separate conversion factors, if used in monitoring, to be derived) a significant overall change in ridership (which would indicate the desirability of redoing the full baseline phase):

1) if total boardings on a route changes by more than 25 percent over the first full year of monitoring (i.e., when comparing the baseline phase figure to a monitoring phase measurement during the same season one year later), an overall trend should be assumed and the complete baseline phase should be redone on that route;

2) if total boardings on a route do not change by more than 25 percent over the first full year of monitoring, but do change (from the baseline phase) by more than 25 percent during any intervening season during the first year, a significant seasonal variation should be assumed.

Detected seasonal variation of the magnitude indicated above is important from two perspectives. First, it indicates those seasons during which a monitoring phase should be regularly conducted on an ongoing basis in addition to the season during which the baseline phase was conducted. Second, for those routes for which a property wishes to use conversion factors to decrease the cost of ongoing monitoring, it indicates those seasons for which separate conversion factors should be derived in order to reliably estimate during each season the data item for which the conversion factor was originally developed.

After the first year of monitoring, the frequency of monitoring phase cycles should depend on the property's need for up-to-date data on each route. At a minimum, however, it is recommended that monitoring phases be conducted during the season of the most recent baseline phase and any season showing a significant variation as outlined above.

4.6.3 Redoing the Baseline Phase

A property should seriously consider redoing the entire baseline phase if significant changes occur in a route's alignment, fare structure or ridership. Any change in routing will impact ridership patterns, on-off profiles and travel time which should be updated by redoing the baseline collection techniques. Similarly, a fare change will usually change many of the data items measured during the baseline phase. In addition, when regular monitoring during the baseline phase season (or any other season which does not exhibit significant seasonal variation) indicates a change in total boardings of 25 percent or more from the baseline phase, the baseline phase should be redone. This procedure is recommended because a change in ridership of more than 25 percent may mean that ridership profiles (e.g., on-off by stop, passenger characteristics, and other baseline phase data items) may have changes in directions not necessarily proportional to the initial baseline phase distributions.

4.7 Section 15 Data Requirements

The data collection approach proposed in this manual insure that a property will satisfy the UMTA Section 15 "Transit Service Consumed Schedule." The Section 15 requirement covers three items: unlinked passenger trips, passenger miles, and average time per unlinked passenger trip. These items must be reported annually on a systemwide basis for specified periods during an average week.

The procedures recommended by UMTA for gathering the required data are based on conducting ride checks on a sample of all bus trips made during the year. The total sample size is selected so that the true value is within 10% of the sample estimate with 95% probability. In order to ensure that the sample selected is representative of all bus trips, the sampling plan requires that ride checks be conducted at regular intervals of between one and six days throughout the year. Depending on the number of days selected, a number of bus trips, ranging from two to fifteen, must be chosen randomly from all trips made on the selected day. The randomly selected bus trips are then ride-checked to yield the sample data for unlinked passenger trips, passenger miles, and average time per unlinked passenger trip. Because of the random sample, expansion of the sample data to produce annual figures is quite straightforward.

For the data collection approach proposed in this manual to satisfy Section 15, the sampling plan must provide a level of accuracy as great as, or greater than, that required by Section 15. This requires that the data collection program be defined in greater detail, particularly in terms of weekend and seasonal sampling and the use of conversion factors. Each of these topics is addressed below.

4.7.1 Section 15 Sampling without Using Conversion Factors

If the monitoring program adopted by the property is based on ride checks, all data items required for Section 15 are measured directly and the question of systemwide accuracy is simply one of the adequacy of the sample size and the acceptability of the sampling plan. Section 4.1.2 showed that for systems with 10 or more routes, the suggested route level tolerance of ±15% is consistent with the desired systemwide tolerance of ±10%. For smaller systems, it may be necessary to reduce the route level tolerance to ±10% to achieve the desired systemwide accuracy.

However, even for properties with a very small number of routes, it is recommended that sampling be conducted to achieve ±15% accuracy at the route level. After the data have been collected, the actual tolerance can be determined by applying the technique described in Section 4.8, and, if appropriate, additional data can then be collected to attain the desired systemwide tolerance of ±10%. In the great majority of cases it will not prove necessary to collect these additional data. Given the adequacy of the sample size, the question of the acceptability of the sampling plan remains.

The effect of seasonal variation on Section 15 data derived using route-level data will be minimal and can be ignored as long as two conditions are met:

1) that the property follow the procedure (outlined previously in Section 4.6) of monitoring every route during each "schedule" period (or at least twice) for one year following the baseline phase

to determine if a significant (i.e., greater than 25% change) seasonal variation exists and, if so indicated, continues to monitor during the baseline season as well as all seasons which exhibited a 25% change in total boardings; and

2) that route-level monitoring activity is spread throughout the year so that routes which are monitored only once a year (i.e., show no significant seasonal variation) are monitored during different schedule periods throughout the year.

Systemwide annual passenger and passenger mile totals are obtained by expanding the seasonal statistics from the bus trip level. This expansion must recognize that different sampling rates may be applied for different periods of the day. Annual systemwide estimates of unlinked passenger and passenger miles for a average weekday, by period of the day, are then computed using the following equation:

$$p_s = \frac{1}{N} \sum_i \sum_j \left(\frac{p_{ij}}{s_{ij}} \right) \qquad (4.8)$$

where: p_s = annual systemwide estimate of passengers (passenger-miles) for a time period of an average weekday;

p_{ij} = total passengers (passenger-miles) observed on sampled trips on route i during season j in time period;

s_{ij} = sampling rate for trips on route i during season j, which is defined as the ratio of sampled trips on route i during season j in a time period to all trips operated on route i during season j in that time period;

N = the total number of weekdays in the year.

The systemwide annual average passenger travel time by time period is derived from the estimates for each route in each season using the following equation:

$$t_s = \frac{1}{p_s N} \sum_i \sum_j \left(\frac{t_{ij}}{s_{ij}} \right) \qquad (4.9)$$

where: t_s = annual average systemwide unlinked passenger travel time by time period;

t_{ij} = total unlinked passenger travel time for all passengers on route i during season j by time period;

p_s = annual systemwide estimate of passengers for a time period (from Eq. 4.8);

s_{ij} = as defined above;

N = as defined above.

As discussed earlier, care should be taken to ensure that the set of days to be sampled is selected randomly from all weekdays in the season. Similarly, the trips to be checked on a selected day should be selected randomly from all trips operated during the period of interest. This two-stage sample will then yield acceptable, unbiased estimates of the Section 15 data items.

Turning now to the problem of estimating weekend statistics for the annual systemwide reports, it must first be recognized that passengers, passenger-miles and passenger trip times will be quite different from the weekday figures and also contribute much less to annual systemwide figures. There is no evidence to suggest that significant seasonal variation occurs for weekend performance compared with normal between-day variation. Hence it is suggested that weekends be analyzed treating each route over a single year-long "season," with Saturdays and Sundays, of course, treated separately. Either of the following two methods is acceptable for estimating Section 15 data for weekends:

1. Sampling 75% of all trips on at least one randomly selected Saturday and one randomly selected Sunday for each route in the system; or

2. Random selection of 260 total trips (or 3 trips per day) from all Saturday and Sunday trips operated systemwide during the year (the existing Section 15 sampling requirements for weekends).

While the second method will be less costly, the first method will provide substantially more information to transit planners and managers. Clearly, ride checks are required to produce the desired Section 15 data for weekends. Equations 4.8 and 4.9 can be applied to Saturdays and Sundays (with N = 52) to yield separate estimates of annual averages for each day.

Before leaving the issue of the sampling plan's relationship to systemwide totals, the treatment of holidays needs some discussion. At present, the recommended Section 15 sampling plan results in holidays being included in weekday, Saturday and Sunday statistics depending on where sampled holidays fall. In the approach recommended here, holidays are classified on the basis of the type of schedule which is operated by the property as weekday, Saturday, or Sunday. The holiday will be included in the population of the appropriate type of day and is then subject to the manual random sampling procedures. This is an important distinction because the resulting Section 15 reports will have a different treatment of holidays than the existing Section 15 reports.

4.7.2 Section 15 Sampling Using Conversion Factors

An analysis of numerous bus routes in Chicago and other cities suggests that average passenger trip length and average time per passenger trip on a specific route are quite stable over long periods of time. This is true as long as neither the service provided on the route, nor the route ridership changes substantially (i.e., by more than 25%). This indicates that stable conversion factors can be developed which would relate total boardings, peak load or trip revenue to passenger miles and average passenger travel time. Such conversion factors would be developed using baseline phase data as outlined in Section 4.5 and the regression confidence interval would be calculated from Equation 4.6 to determine the route level tolerance. If the confidence interval is ±15 percent or smaller, the route level Section 15 data would be consistent with the desired systemwide accuracy. Conversion factor data would be acceptable for use for Section 15 reports, then, if two conditions are met:

1) if separate conversion factors are developed and used with the appropriate seasonal data for routes which exhibit significant seasonal variation as outlined in Section 4.6.2,

2) if the baseline phase is redone and all conversion factors recalculated if significant route changes occur as recommended in Section 4.6.3 (i.e., when route alignment or fare structure modifications are made or when ridership changes by 25 percent or more).

If a property makes use of conversion factors to estimate systemwide Section 15 data items, the regression equations should be used to estimate the total passenger trips, passenger miles and passenger trip times at the route level. These are then aggregated to produce systemwide estimates using equations 4.8 and 4.9. Some properties may be able to use conversion factors for weekdays, but have to perform ride checks for weekends because of inadequate data to demonstrate that satisfactory interrelationships exist.

4.8 Interpretation of the Sampled Data

Several statistical procedures can be used to help interpret the results from the data collection program. These procedures are of two types:

1. the calculation of confidence intervals for each data item and time period; and

2. the test of whether one sample mean is significantly different from an earlier sample mean on the same route.

4.8.1 Calculation of Confidence Intervals

Once a sample of data has been collected, a property may want to calculate the actual confidence interval (about the mean) for each data item. This is similar to, but not the same as, the tolerance range used to determine sample size, because actual sample variances are used to calculate the interval rather than variances developed from the pretest or previously collected data. The confidence interval defines a range within which the manager can be, say, 90 percent confident that the true mean value lies. By slightly modifying the calculation, a manager can raise the confidence level (and thus widen the interval) or lower the confidence level (and thus narrow the interval). If a manager decides that the interval is too wide at a particular confidence level, he can enlarge the sample and narrow the range within which the true mean lies.

The confidence interval is determined from the following equation:

$$d = \frac{tD}{\bar{x}} \qquad (4.10)$$

where t = the normal t-statistic for the desired confidence;

D = the standard error of the sample;

$\bar{x}$ = the mean value of the data item;

d = the accuracy (or interval) obtained expressed as a fraction of the mean.

The exact definition of D is given in Appendix A, along with further explanation of the confidence interval or sample accuracy concept.

By varying the confidence level (by adjusting the t-statistic in the above equation corresponding to different levels of confidence), a manager can also estimate the probability (confidence) that the mean value lies above or below a certain service standard or policy. This is done by changing the width of the interval to coincide on either end with the standard or policy value, calculating the value of d corresponding to the new end value $\left[d = (\text{end value} - \bar{x})/\bar{x}\right]$, and solving Equation 4.10 for t. Using the calculated t-value, a manager would then consult a standard t-distribution table ($n = \infty$) to determine the confidence or probability level which most closely corresponds to his t-value. This would then be the probability that the mean value exceeds (or falls under) a given service standard value.

4.8.2 Difference of Means Test

The "difference of means" test is another procedure which may be useful in interpreting the data collected over time in a comprehensive monitoring program. This procedure will test whether two independent samples (i.e., taken at two distinct times) do, in fact, have significantly different average values. This test should be especially useful to managers who need to know whether a change has occurred or if, in fact, the apparent change in means is simply a result of the inherent variability of the data and the sample sizes. (For example, the test might show that the difference between an observed mean load of 55 and 60 is due only to normal data variation and not to a real change in loading.) Detailed definitions and worksheets to perform the difference of means test are included in Appendix A for use when comparing two samples collected at different times.

Chapter 5

CHECKER REQUIREMENTS AND COST ESTIMATION

The previous chapter described procedures for determining sample sizes on a route by route basis for each time period of the day. The next step is to translate these sample size requirements into checker requirements and total data collection costs. While the cost of data collection will vary among different properties, the procedures discussed here involve identification of only the basic component costs and, therefore, can be adapted to most operating environments.

5.1 Estimating Checker Requirements

The largest single item in any data collection budget is the manpower needed to collect data on-board buses or on the street. Checking practices vary substantially across the industry,[1] both in terms of unit cost (i.e., wage rates) and work policies (e.g., in some cases, non-union part-time workers can be used for data collection, while in other properties, full-time traffic checkers form a bargaining unit within the transit workers union). Checker costs depend greatly on the ability of management to assign personnel to odd shifts and have the same personnel perform varied duties (related to different data collection techniques).

The translation of route-by-route sampling plans into total checker requirements generally begins with the sample size required for each data collection technique selected. Equation 5.1 is a general equation for determining checker requirements based on the sample sizes required for load and total boardings for each route and on the selected techniques for each data collection phase. The specific form which Equation 5.1 takes

---

[1] See Interim Report #1, Data Requirements and Collection Techniques, Bus Transit Monitoring Study, prepared for UMTA by Multisystems, Inc. and ATE Management and Service Co., Inc., April 1979, NTIS No. PB80-161409.

is guided by a set of four "decision" rules which define the exact terms and values within the equation. The equation is appropriate for: (a) any sampling plan which requires load data only at a number of points on a route, (b) any sampling plan based on boardings data obtained using a ride check, and in many cases, (c) a combination of (a) and (b) when both point and ride checks are required. The general form of the equation is shown here:

$$\begin{array}{l} \text{Checkers} \\ \text{required} \\ \text{for each} \\ \text{time} \\ \text{period} \end{array} = \left[\begin{array}{l} \text{Days} \\ \text{sampled} \\ \text{(load)} \end{array} \times \begin{array}{l} \text{Number} \\ \text{of} \\ \text{points} \end{array} \right] + \left[\begin{array}{l} \text{Days} \\ \text{sampled} \\ \text{(board-} \\ \text{ings)} \end{array} \times \begin{array}{c} \text{Sampled} \\ \text{trips} \\ \div \\ \text{Total} \\ \text{trips} \end{array} \times \begin{array}{l} \text{Number} \\ \text{of} \\ \text{buses} \end{array} \right] \qquad (5.1)$$

The various terms of the equation will vary depending on the data collection techniques used and the sample sizes required. The detailed rules guiding the use of this equation are explained fully in Step 7 in Chapter 6.

Using an individual property's policies and work rules, the individual time period checker requirements determined by this equation can be transformed into checker assignments. If a point check is included for a number of routes, the total checker assignments can be adjusted to account for the possibility that several routes might be counted by one checker at the same maximum load point. Once total checker shifts are pieced together to most effectively utilize checker time and to meet the required sampling plan for each route, a property can determine how long a complete sampling cycle will take using the existing checker staff. If this cycle is too long for either the baseline or monitoring phases (for example, more than six to nine months), a property should either consider increasing its checker force or decreasing the accuracy on which the sample sizes are based.

5.2 Costs of Other Data Collection Techniques

The discussion to this point has focused on the direct cost of collecting transit performance data using traffic checkers.

This focus is appropriate since the major cost component of the comprehensive bus transit data collection program is the cost of checkers whose sole job is data collection. However, several data collection techniques discussed in Chapter 3 do not directly involve traffic checkers, at least in the traditional sense of counting passengers or noting bus arrival times. These techniques include:

1) operator-collected boarding counts and/or farebox readings,

2) revenue counts by bus,

3) transfer ticket counts,

4) on-board surveys.

The cost of operator-collected data is straightforward: multiply any premium (or extra) hourly cost for operator performance of such tasks by the number of pay hours associated with the activity. Total pay hours can be calculated by multiplying the maximum number of sample days required for total boardings for any weekday time period by the number of weekday pay hours on the particular route. Similarly, weekend pay hours can be obtained by multiplying the number of sample days for both Saturday and Sunday by the pay hours for each day.

The cost of revenue counts varies widely among properties depending on their procedures for counting farebox revenues. In some cases, a property may already be set up to count and record revenue by bus run; in others, such a procedure may require the assistance of one or more additional personnel per route, garage, or other operating entity. A property should examine its current operating procedures to determine what level of additional cost may be involved.

A transfer ticket count is a straightforward technique involving the manual counting and recording of transfer tickets collected by originating route and, possibly, by time of day. The cost of this technique is directly dependent on the number of transfers collected on each route and the ease of determining the originating route. A typical cost should be determined by each property through the actual performance of a transfer ticket count using a 3-day accumulation of transfers (to ensure statistical reliability) on an average route.

The cost of an on-board passenger survey varies with the method of survey distribution and return, the complexity of the survey, the processing methods, the sample size and the return rate. A survey can be distributed to passengers in a variety of ways: by the vehicle operator, by an on-board checker, at a major terminal boarding point, and through a combination of these methods. (See Section 3.7.) A property should determine the cost of the distribution method deemed most feasible for its particular operating environment. The on-board checker method is probably the most costly since a checker has to ride each trip being surveyed. However, that checker may also be able to conduct a ride (or board) check on many routes. (This depends on the level of patronage and whether the survey requires the checker to explain how to complete any item.)

Other significant survey costs include the coding, keypunching, and data processing of the completed returns. These costs vary from $0.15-$2.00 (with typical figures of $0.75-$1.00) per return, depending on the amount and type of survey coding needed and the length of the survey. If survey returns are to be geocoded (i.e., origin and destination zones identified), the costs vary with the density of the service area, the size of the zones used, and whether any automated (computer) address files are available.

5.3 Other Program Costs

Two other cost categories, program planning and overall data processing, impact the overall data collection program costs. Again, it is difficult to provide hard guidelines to estimate these costs since they depend heavily on the current operating environment and resources available to an individual property. The factors which influence these costs most significantly are discussed in general terms below.

5.3.1 Program Design and Planning Costs

Program design and planning includes the determination of data needs, the level of effort to be assigned to each of the data collection stages, selection of the appropriate combination of data collection techniques, and sample size determination. The trade-off between data collection costs and the quality (reliability and accuracy) of the data needed must be resolved primarily at this stage of the project.

Costs of this type fall into three main categories:

1) the overall design and planning of the data collection process;

2) the calculation of sample size requirements; and

3) the detailed scheduling of checker work assignments.

Costs in the first category are determined by the amount of management time required, which, in turn, is a function of the size and complexity of the system and the sophistication of current data collection procedures.

The cost of the second category, sample size determination, depends on whether available data or pre-test data are used to estimate the variability of different data items and the attendant sample size requirements. If existing data can be used, costs are reduced to the time required to compile the data, to calculate the between- and within-day variability of each route or route classification, and to use the appropriate table to determine the required sample size. If a pre-test is necessary, an additional data collection cost is incurred, which would include all the cost components identified in the previous two sections (although it would generally be conducted using only point checks as discussed in Section 4.3).

The cost of the third category, scheduling of traffic checkers, depends primarily on the number of traffic checkers involved and the flexibility of management in assigning checker work at different times during the day. Large properties may have checker work policies which limit assignments to a specific duration (similar to driver work rules). Scheduling of checkers becomes more complex for on-board techniques when many buses operate on the same route and/or interlining is common. However, once monitoring sample sizes are determined, checker schedules can be developed for each route in the system initially and used each time a data collection cycle is performed.

5.3.2 Overall Data Processing Costs

Data processing costs depend on the amount of data collected, and the availability of computer support and staffing for the technical analysis. When a computer is used, processing costs fall into the following categories:

- software per type of data collected
- editing incomplete data
- coding
- keypunching
- computer time
- analysis time

If appropriate software is available, the software costs are the costs of acquisition. If it is necessary to create the software, costs will be considerably higher. Many properties (e.g., SCRTD in Los Angeles, MTC in Minneapolis-St. Paul, MBTA in Boston) have already developed software to analyze point and ride check counts, as well as some other types of transit performance data (such as revenue). These programs are usually made available to other properties at no cost upon request.

Several interactions exist between cost components in both the collection and processing stages. Good software, for example, can reduce analysis time. The use of self-coding forms, punch cards, mark sense, or character-recognition forms can all reduce costs by reducing or eliminating the need for separate manual coding and keypunching of checker and survey data.

A "ballpark" estimate of total data processing costs when using computer processing would range from $0.10-$0.30 per computer card (80 column) processed. This would translate into $2.00-6.00 per traffic checker field form (Figures 3.1 - 3.4) assuming an average of 20 computer cards of data from each completed checker sheet. Keypunching (including verification) costs alone for typical point and ride check records (i.e., the point check arrival time and load for one bus or the on/off activity at one stop for a ride check) average about $60-$70 per thousand computer cards at a commercial service. (These costs may be higher or lower when done "in-house," depending on wage rates and the skill of the key operator.) Data cleaning and editing costs in preparation for keypunching can often approach total keypunching costs, depending on the condition of the checker recorded data. By comparison, an optical character recognition-commercial service (used by MARTA in Atlanta) which automatically reads checker forms costs about $50 per thousand records, including normal editing and keypunch correction of misread forms.

When computers are not used, processing costs include:

- editing incomplete data
- analysis time

Smaller properties often develop standard ready-to-use tables on which data are manually compiled from sheets completed in the field. Checkers can be assigned to office duty to compile the data collected for one or more hours per day or, alternatively, one day a week. Using the summarized tables, the manager (or an analyst) responsible for operations would then examine the data to determine if any changes have occurred and if management action is in order.

Chapter 6

STEP-BY-STEP PROCEDURES FOR DESIGNING A DATA COLLECTION PROGRAM

This chapter describes the detailed steps and procedures which can be used to design a data collection program tailored to the characteristics of any transit property. Figure 6.1 shows how each of the steps outlined in this chapter fits into the overall process of data collection program design and implementation. The procedures allow a property to select the most cost-effective collection techniques based on individual property and route characteristics. For each selected set of techniques, the procedures determine the required sample sizes and the estimated costs. These steps cover the establishment of comprehensive base conditions in a baseline data collection phase and the periodic monitoring of route performance. However, if a property has already established base conditions or chooses for some other reason to forego the baseline phase, the procedures outlined here can also be followed to design a continuing monitoring program.

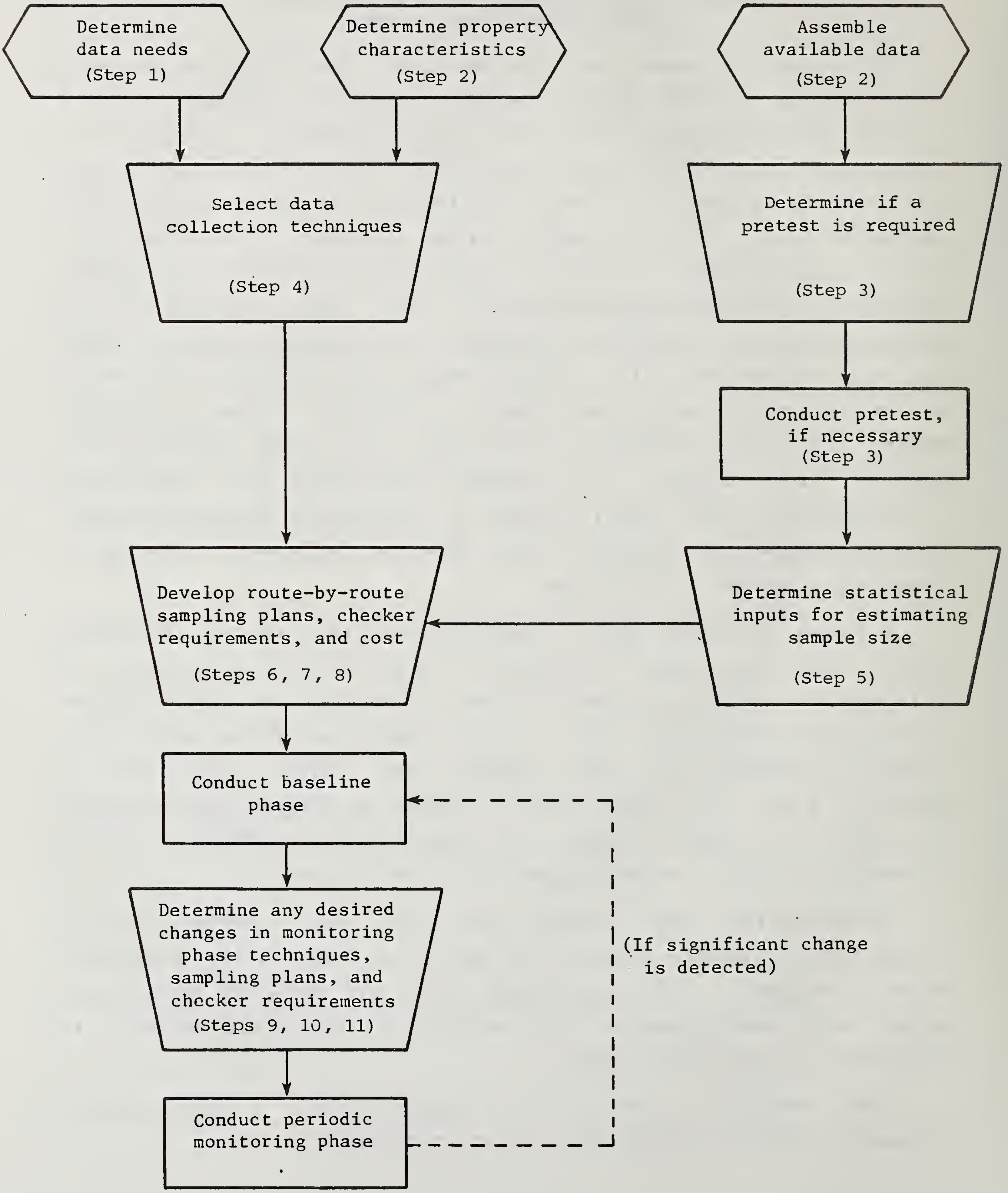
A total of eleven major steps are outlined on the following pages. For each step, reference is made to prior sections in this manual which provide further discussion of the relevant issues. In addition, some of the steps refer to additional sampling procedures, work sheets and tables contained in Appendix A and the accompanying Volume 2, Sample Size Tables. In order to fully design a comprehensive program, it is necessary to refer to both Appendix A and Volume 2.

Accompanying each major step in this chapter is an illustrative example of how it might be applied in a typical transit property. For easy reference, the detailed procedural steps have been placed on the left pages, while the example is explained on the right pages.

The example is based on "Property A," a hypothetical property with the following characteristics:

Figure 6.1

Summary of Data Collection Program Design and Implementation



- 500 buses;
- 75 routes;
- 8 traffic checkers;
- regular point check program;
- no other route level data collection currently.

Procedure

STEP 1: <u>DETERMINE DATA NEEDS</u>
(Chapter 2)

Based on the needs of the various management functions and departments, the staff person responsible for designing a monitoring program should develop a list of required service performance data, including Section 15 requirements. Those responsible for the planning, scheduling, financial, transportation, and general management functions should be consulted before developing this list. Table 2.1 (p. 15) provides a recommended list of data needs which were reported by most properties contacted during the course of this study. Along with the required data, managers should be asked to estimate how frequently each required data item need be directly monitored.

Example

STEP 1: An operations planner has been assigned to develop a comprehensive monitoring program for Property A. After consulting with the appropriate managers, (s)he has determined that a program should be designed to obtain the list of data needs shown in Table 2.1. Furthermore, peak load, bus arrival times and total boardings should be monitored directly at least four times a year corresponding to the schedule changes which are implemented on a seasonal basis.

Procedure

STEP 2: ASSEMBLE AVAILABLE DATA AND ROUTE CHARACTERISTICS (Sections 4.3.1 and 4.3.2)

Two types of data should be gathered as inputs to the data collection design process:

1. Recently collected load and/or total boardings data for each route in the system, broken down by the time periods of the day of interest to the property:

   - Data should not initially be aggregated; actual load and boardings data per vehicle trip must be used to calculate measures of data variability.

   - Only data collected during the last six months to a year (depending on overall system ridership and service changes) should be used, excluding any data gathered during low ridership periods (e.g., summer). Weekend and/or holiday data should be compiled and analyzed separately. Data from as many different days (up to 10) as available should be compiled; as this will ensure accurate measurement of data variability. If appropriate data are not currently available, go on to Step 3.

2. Route Characteristics:

   - For each route, the number of scheduled round trips and buses assigned during each time period to be sampled must be compiled. Also, a listing of all desired load check locations including the maximum load points in the system (along with the routes passing each point) should be compiled.

Example

STEP 2: Since property A currently has a regular point check program, 1-2 complete days of load checks made during the past six months (excluding summer) are available for each route in the system.

This property has already compiled the detailed route characteristics (scheduled round trips and buses assigned) for each of six time periods (am peak, base, pm peak, evening/owl, Saturday, Sunday) as part of an operating statistics information sheet. A list of point check locations (and the routes observed at each point) is also available from the existing checking staff. The characteristics of two routes are listed below:

<u>Route 16</u> (1 load check location)

	6-9am	9-3pm	3-6pm	6pm-Midnight	Sat	Sun
Round Trips	20	18	20	12	24	24
Buses Assigned	10	5	10	3	3	3

<u>Route 48</u> (1 load check location shared with one other route)

	6-9am	9-3pm	3-6pm	6pm-9pm	Sat
Round Trips	9	18	9	4	24
Buses Assigned	5	5	5	3	3

Procedure

STEP 3: <u>DETERMINE IF A "PRETEST" IS NECESSARY TO GATHER ADDITIONAL INPUT DATA AND, IF SO, CONDUCT IT</u> (Section 4.3.2)

If three days of data on each route are not available, a property has two options:

1. Conduct a "pretest" (i.e., a preliminary data collection effort aimed at determining data variability) consisting of peak load or boarding counts for three full days or the number of days which, together with other recently collected data, add up to three days of data on each route; or

2. Develop a route classification scheme (see Appendix B) and conduct the pretest by collecting three days of load or boardings data on 2-3 routes in each route category. If any route category includes 3 or fewer routes, collect pretest data on each route in that category. The route classification scheme should group routes according to similar data variability characteristics and may be based on several factors:

   - Functional type of route (e.g., feeder, express, crosstown, shuttle, suburban, etc.)
   - Route length
   - Headway
   - Total boardings
   - Ridership productivity (e.g., passengers per vehicle mile or vehicle hour)
   - Peak load factor (e.g., percentage of available seat capacity)

   Note: If data variability calculated (see Step 5) from several routes in each route classification prove dissimilar, reclassify routes or discard classification and conduct pretest on all routes.

Example

STEP 3: Since 1-2 days of load data already exist (for each route), Property A has decided to perform the necessary load checks to obtain a minimum of three days data for each route. In this way, route specific variation measures can be calculated.

Procedure

STEP 4: <u>SELECT APPROPRIATE DATA COLLECTION TECHNIQUES</u> (Section 3.8)

Based on the characteristics of its system, a property should select appropriate data collection techniques for the initial baseline data collection phase. This choice should also include preliminary selection of monitoring phase techniques.

Table 3.2, reproduced on page 94, summarizes the available data collection techniques and the data provided by each technique. The following combination is recommended for the baseline phase:

- ride checks (plus possible supplementary point checks)
- farebox readings or board checks
- on-board passenger surveys.

For the monitoring phase, if a property can rely on operator-collected data, the following combination of techniques is recommended:

- point checks
- boarding counts (by operator)
- farebox readings (if registering fareboxes are available).

If operator data are not available, the following combination of techniques is recommended:

- ride checks (plus possible supplementary point checks)
- farebox readings (if registering fareboxes available).

In the latter case, if the use of conversion factors proves feasible, the ride checks can be largely replaced by point checks.

(Step 4 procedure continued on page 94.)

Table 3.2

Data Items Obtained by Seven Principal Techniques

	Technique (1)						
Data Item	Point Check	Ride Check	Boarding Count	Farebox Reading	Revenue Count (1)	Transfer Count	Survey (2)
Load (peak or other)	√	√					
Bus arrival time	√	√	(3) √	(3) √			
Passenger-trips	(4) √	√	√	(5) √			(6) √
Revenue		(7) √	(7) √	(8) √	√		√
Passenger-trips (or revenue) by fare category		(7) √	(7) √	(5) √			√
Passengers on-off by stop		√					√
Transfer rates						(9) √	√
Passenger characteristics, travel patterns, and attitudes							√
Unlinked trips		√	√	(5) √			(6) √
Passenger-miles		√					√
Unlinked trip travel time		√					√
Linked trips		(9) √	(9) √			√	√

Key: √ = applicable
blank = not applicable

(1) Techniques as defined in Table 3.1.

(2) For all survey-collected data other than total passengers, the quality of the data depends on the representativeness of the response.

(3) If time can be recorded.

(4) For "pure" feeder and express routes only.

(5) If electronic multiple fare registering boxes are available.

(6) If surveys are numbered consecutively and distributed to all passengers.

(7) If boarding passengers are recorded by fare category. This typically can only be done with riding checks if boardings are relatively low.

(8) If revenue can be counted by route, this can be substituted for farebox readings although time-of-day data are sacrificed.

(9) If transfer tickets are distributed, collected on terminating route, and identifiable by initial (and intermediate) route(s).

Example

STEP 4: Property A cannot use drivers to collect any form of passenger data but registering fareboxes are installed on all buses. Based on these characteristics and a desire to obtain accurate transfer data (for use in systemwide route restructuring), Property A selects the following techniques for use during the baseline phase:

- point checks (if needed to obtain more load samples than those provided by ride checks);
- ride checks;
- farebox readings (by ride checker at the beginning and end of each trip);
- on-board surveys;
- transfer ticket counts.

For the monitoring phase, Property A tentatively chooses the combination of ride checks (with supplemental point checks) and farebox readings, but hopes to make use of conversion factors to replace the ride checks with point checks.

Procedure

STEP 5: DETERMINE STATISTICAL INPUTS FOR SAMPLE SIZE ESTIMATION (Sections 4.1, 4.2, 4.3)

For each route and time period, select an appropriate tolerance, based on expected use of load and total boardings data on a route level. Based on analysis of actual data and planning uses in several properties, the following tolerance ranges are recommended:

Type of Route	Time Period	Data Item	Recommended Tolerance
Capacity Constrained	Peak Periods	Load, Boardings	± 10%
Non-capacity Constrained	Peak Periods	Load, Boardings	± 15%
All types	Midday	Load, Boardings	± 15% to ± 20%
All types	Evening Owl & Weekends	Load, Boardings	± 30% to ± 50%

Use detailed instructions and work sheets in Appendix A (Section A.3.1) along with route data previously assembled or collected during the pretest to calculate the within-day (i.e., within-time-period) and between-day coefficients of variation. These coefficients should be calculated for each route (or for several routes within each route classification if these are defined by a larger property), and for each time period during the day. For peak periods, only data from the peak direction should be used to calculate the coefficient of variation; for off-peak periods or for routes with no peak direction all day, coefficients of variation should be calculated for both directions and the higher coefficients used for determining sample size. If three days of weekend data are unavailable, use calculated evening coefficients for weekend sampling inputs.

If load data are used to calculate the coefficients of variation, use same results for calculating sample size for both load and total boardings.

If boardings data are used to calculate the coefficients of variation, use results directly for sample size inputs for total boardings sampling plan, but inflate the results by 30% for input into load sample size determination.

Example

STEP 5: Property A has decided to slightly modify the recommendations regarding tolerance ranges (since total boardings data are not needed at the same tolerance level as load data on capacity-constrained routes). Twenty-seven of its seventy-five routes have peak headways of less than 10 minutes or are long-haul express routes and are at capacity (since the headways are set according to observed demand). For these twenty-seven routes, load will be sampled at ±10 percent and boardings at ±15 percent during the peak periods. For all other routes, both load and total boardings will be sampled at ±15 percent during peak periods. For all routes in the system, midday loads and total boardings will be sampled at ±20 percent and evening, owl and weekend period data will be sampled at ±30 percent.

Using the available load data, the worksheets in the accompanying Sampling Volume have been used by Property A to calculate the coefficients of variation for each route and time period in the system. The results for two typical routes are presented here:

Route 16 (capacity-constrained)

	6-9am	9-3pm	3-6pm	6pm-Midnight	Sat/ Sun
Load Tolerance	.10	.20	.10	.30	.30
Boardings Tolerance	.15	.20	.15	.30	.30
Within-Day Coef.	.37	.48	.34	.50	.50
Between-Day Coef.	.06	.15	.05	.19	.19

Route 48 (non-capacity-constrained)

	6-9am	9-3pm	3-6pm	6pm-9pm	Sat
Load Tolerance	.15	.20	.15	.30	.30
Boardings Tolerance	.15	.20	.15	.30	.30
Within-Day Coef.	.48	.50	.43	.59	.59
Between-Day Coef.	.10	.15	.06	.20	.20

Procedure

STEP 6: <u>USE SAMPLE SIZE TABLES AND STATISTICAL INPUTS TO DETERMINE TYPICAL SAMPLE SIZES</u> (Section 4.3)

For each route and time period, use load and/or total boardings variation factors, selected tolerance ranges, and the number of scheduled trips in the time period to determine sample size using the tables in Volume 2, <u>Sample Size Tables</u>. (If the coefficients of variation were calculated for different route classifications instead of individual routes, use these with the individual route number of trips to determine route specific sample sizes.)

Finding the correct table involves three steps for each <u>route</u> and <u>time period</u> of interest:

1. First, using the dark tabs which separate the volume on its right edge, locate the set of tables corresponding to the within-day coefficient of variation (or the next higher value listed) for each route.

2. Within this set of tables, locate the smaller subset of tables corresponding to the between-day coefficient of variation (or the next higher value listed).

3. Within this subset of tables, turn to the page and individual table corresponding to the number of scheduled trips (or the next higher value listed).

After locating the correct table, scan the columns for the desired tolerance (i.e., ±10%, ±15%, ±20%, ±30%). All of the sampling plan combinations of <u>days</u> and <u>trips</u> included in each column will provide data estimates accurate to the indicated tolerance range. A property should select the most appropriate sampling plan from the appropriate column based on the collection technique being used, the size of the available checker staff, the ability to sample several routes at one time, etc.

<u>Note</u>: If the number of sample trips per day (given at the bottom of each column) exceeds the number of scheduled trips in the period (because the actual number of scheduled trips is not included in the tables), adjust the number of sample trips per day down to the number of scheduled trips, leaving the number of days to sample unchanged. For each data item for which

(Step 6 procedure continued on page 99.)

coefficients of variation are available, a property should list the selected sampling plan for each route and time period.

Note: If load data were used to calculate coefficients of variation for both load and total boardings data items, and the sample size tables call for more than three days of data for the selected total boardings tolerance range, the load coefficients of variation may not be reliable for use in determining sample sizes for monitoring total boardings. In this case, drop sampling requirement for boardings to three days in the baseline phase (after which the coefficients can be recalculated directly for boardings data).

Example

STEP 6: Using the route characteristics defined in Step 2 and the coefficients of variation calculated in Step 5, Property A refers to the sample size tables in the accompanying sampling volume to determine the following route sampling plans for both the load and total boardings data items. For convenience in scheduling its checkers, this property has generally chosen the sample requiring the minimum number of days on each route (except for cases when the ride check requirement will exceed 8 buses during one period):

Route 16 Sampling Plan

	6-9am		9-3pm		3-6pm		6pm-Midnight		Sat/ Sun	
	D*	T**	D	T	D	T	D	T	D	T
Load	1	20	2	13	1	20	2	6	2	7
Total Boardings	1	13	2	13	1	13	2	6	2	7

Route 48 Sampling Plan

	6-9am		9-3pm		3-6pm		6pm-9pm		Sat	
	D*	T**	D	T	D	T	D	T	D	T
Load/ Total Boardings	2	8	2	13	1	8	2	3	2	9

* Number of days to be sampled
** Number of trips to be sampled per day.

Since route-to-route transfer rates can be expected to be somewhat more variable than either load or total boardings data, Property A decides to count transfer tickets for 3 days systemwide during the baseline phase.

STEP 7: <u>DETERMINE DETAILED CHECKER REQUIREMENTS FOR EACH ROUTE</u> (Section 5.1)

Using the selected data collection techniques for the baseline phase (Step 4), the sample sizes (i.e., days and trips) determined in Step 6 for the load and total boardings data items, and the individual route characteristics (i.e., scheduled trip and buses assigned), calculate the number of checkers required during each time period using the following equation:

$$\text{Checkers required for each time period} = \overbrace{\left[\text{Days sampled (load)} \times \text{Number of points}\right]}^{\text{Term 1}} + \overbrace{\left[\text{Days sampled (boardings)} \times \text{Sampled trips} \div \text{Total trips} \times \text{Number of buses}\right]}^{\text{Term 2}}$$

Apply the following rules to use the equation for the selected technique combination. Begin with rule (1) and continue down <u>only until the selected combination first meets the condition(s) of a rule.</u> Proceed to use the equation according to the instruction of the first rule for which conditions are met.

1. If a combination does not include a point check, omit the first term in the equation and set DAYS SAMPLED (BOARDINGS) equal to the greater of DAYS SAMPLED (LOAD) and DAYS SAMPLED (BOARDINGS).

2. If a combination does not include a ride check or a checker-performed board check, omit the second term of the equation.

3. If a combination includes both a point check and ride check <u>and</u>:

   if DAYS SAMPLED (LOAD) is less than DAYS SAMPLED (BOARDINGS), omit the first term of the equation;

   if DAYS SAMPLED (LOAD) is equal to DAYS SAMPLED (BOARDINGS) <u>and</u> LOAD TRIPS is less than or equal to BOARDING TRIPS, omit the first term of the equation;

   if DAYS SAMPLED (LOAD) is equal to DAYS SAMPLED (BOARDINGS) <u>and</u> LOAD TRIPS is greater than BOARDINGS TRIPS, set DAYS SAMPLED (LOAD) equal to "1" and use equation as is;

   if DAYS SAMPLED (LOAD) is greater than DAYS SAMPLED (BOARDINGS) set DAYS SAMPLED (LOAD) equal to DAYS SAMPLED (LOAD) minus DAYS SAMPLED (BOARDINGS) and use equation as is.

4. If a combination does not fit any of the above rules, use equation as is.

Example

STEP 7: Property A has calculated the checker requirements for Routes 16 and 48, assuming point and ride check combinations for both the baseline and monitoring phases. Therefore, rule (3) on the opposite page applies when using the checker requirement equation. (Note that, for the Route 16 peak periods, load data samples were required for 20 trips while boardings samples were required for only 13 trips. In this case, a point check was introduced to measure an additional 7 trips since it would be less expensive than using ride checks.)

Time Period	Days Sampled (load)		# of Points		Days Sampled (boardings)		Sampled Trips		Total Trips		# of Buses		Checkers Required
Route 16													
6-9 am	(1	x	1)	+	(1	x	13	÷	20	x	10)	=	8
9-3 pm					(2	x	13	÷	18	x	5)	=	8 (4 per day)
3-6 pm	(1	x	1)	+	(1	x	13	÷	20	x	10)	=	8
6 pm - midnight					(2	x	6	÷	12	x	3)	=	3 (1.5 per day)
Sat/Sun					(2	x	7	÷	24	x	3)	=	2 (1 per day)
Route 48													
6-9 am					(2	x	8	÷	9	x	5)	=	10 (5 per day)
9-3 pm					(2	x	13	÷	18	x	5)	=	8
3-6 pm					(1	x	8	÷	9	x	5)	=	5
6-9 pm					(2	x	3	÷	4	x	3)	=	5 (2.5 per day)
Sat					(2	x	9	÷	24	x	3)	=	3 (1.5 per day)

Procedure

STEP 8: <u>ESTIMATE OVERALL COSTS OF BASELINE PHASE</u> (Chapter 5)

Several cost components should be summed to estimate total cost of the baseline phase effort:

1. For all routes in the system use checker requirements by time-of-day to develop total checker hours or days either by multiplying the requirements by the number of hours applicable or by using checker work rules and policies to develop checker assignments for each route. Multiply total hours or days by prevailing wage and overhead rate. (Using the systemwide total checker days, i.e., shifts, it would also be informative to determine how long it would take to perform the baseline phase using existing checker resources. If more than six months, it is recommended that additional resources be sought or target accuracy levels be reduced.)

2. For technique combinations relying on operator-collected data for which a property must pay a premium, determine the total number of sample days on which total boardings need to be counted and multiply the applicable premium hourly rate by the total sample days and total number of pay-hours allocated to each route. Sum this cost for each route to obtain system totals.

3. For combinations which include revenue counts, determine the incremental cost of counting vault revenue by route.

4. For baseline phase combinations which include a survey, determine the cost of survey distribution, collection and processing as described in Section 5.3 and based on the survey sampling procedures discussed in Section 4.4.

5. For baseline phase combinations which include a transfer count, determine the cost of counting all transfers collected by origin route for 3 days. This will depend greatly on the number of transfers collected each day.

6. Determine data processing costs for checker or operator-collected data depending on the amount of data being processed and the type of processing (computer or manual). A discussion of these costs is contained in Section 5.3, but it is expected that they will vary widely from property to property.

Example

STEP 8: 1. For Route 16, Property A has applied its checker work rules and determined that the time-of-day requirements calculated in Step 7 will require 15 checker days for weekday counts and 6 checker days for weekend counts (obtained roughly by multiplying the checkers required by the number of hours needed). At $100 per checker day, this amounts to $2,100 for both the baseline and monitoring phases for this route. Repeating these calculations for every route, Property A estimates that 125 work days or six months (or about 1000 checker days) would be required to complete a data collection cycle assuming all eight existing checkers were dedicated to the task. Property A is willing to accept this time frame for the baseline phase, but would like to monitor all routes more frequently.

2. For the baseline phase, Property A will conduct an on-board survey on one weekday using operators to simply hand-out surveys and collect completed returns. A $0.30 per hour premium rate has been negotiated with the operators' union to be paid for each hour during which surveys are distributed. Since Route 16 has 129 weekday operator pay-hours, the cost of distributing the survey for this route will be about $36.

3. Property A will not conduct revenue counts.

4. About 400 returned surveys are expected as they will be handed out to all inbound riders. Half of the responses are expected to be returned by mail for an additional cost of $30. Coding, keypunching and computer processing are expected to cost about $0.75 per completed survey form, so processing costs for this route will total about $300. Thus, the total estimated cost of an on-board survey for Route 16 will be approximately $366. Systemwide totals estimated in a similar fashion result in a total survey cost of about $25,000.

5. Since Route 230 is a feeder route to Property A's rapid transit and has some transfers to other bus routes on its outer end, it has been estimated that a transfer count for three days will cost about one checker or $100. The systemwide transfer ticket count cost was estimated to total approximately $6,500.

6. Finally, Property A estimates that data processing costs (exclusive of the on-board survey) will total approximately $20,000 (primarily for editing, keypunching and producing route profile reports from the point and ride check data for each intensive or monitoring cycle performed through the system).

The total estimated cost of the baseline phase (1,4,5,6) comes to $151,500.

Procedure

STEP 9: <u>TEST POTENTIAL CONVERSION FACTOR USE IN THE MONITORING PHASE</u> (Section 4.5) (Optional)

In order to determine whether less costly monitoring techniques can be used (especially if the only feasible method to monitor total boardings is by a ride check), a property may want to test the relationship between total boardings and peak load or trip revenue. Two simple steps are required to perform this test and, if appropriate, adjust the techniques selected and sample sizes for the monitoring phase:

1. Use a standard linear regression program (on many pocket calculators and computer statistical packages) to estimate an equation that predicts, for example, total boardings per trip (the dependent variable) based on peak load <u>or</u> revenue per trip (the independent variable). Use baseline data collected during the phase to estimate the equation for each route. One equation based on all-day data is usually accurate for all time periods. Data from as many days as available should be used to estimate the equation.

2. Using the regression variance (or "regression error mean square," an output of most standard packages) calculate the confidence interval for total boardings per trip as a percent of the mean using the equation:

$$c = \frac{t\sqrt{s^2}}{\bar{y}\sqrt{n}}$$

where c = confidence interval, expressed as a fraction;

t = t-statistic for desired confidence level ($t = 1.645$ for recommended 90% confidence level);

s^2 = regression variance;

$\bar{y}$ = mean boardings per trip;

n = total observations input to estimate regression equation.

If c is less than or equal to the selected tolerance range for the total boardings data item, the estimated relationship can be used as a conversion factor in the monitoring phase.

If c is greater than the selected tolerance range, the conversion factor cannot be used and the monitoring program should be as developed in Step 6.

Example

STEP 9: Property A elects to test the relationships between total boardings and both peak load and trip revenue (as recorded on the farebox). Regression equations are estimated separately for these two independent variables for each route. The following results were obtained for Route 16, am peak period:

Peak Load C.F.	Revenue C.F.
$B = 1.38(L) + 2.56$	$B = 2.68(R) + 3.21$
$s_r^2 = 432$	$s_r^2 = 298$
$s_w^2 = 858$*	$s_w^2 = 858$
$s_b^2 = 31$**	$s_b^2 = 31$
$t = 1.645$	$t = 1.645$
$n = 78$	$n = 78$
$\bar{y}_{am} = 71$	$\bar{y}_{am} = 71$
$c = \frac{t\sqrt{s_r^2}}{\bar{y}\sqrt{n}}$	$c = \frac{t\sqrt{s_r^2}}{\bar{y}\sqrt{n}}$
$c = .055$	$c = .045$

Both regressions produced boardings estimated well within $\pm15\%$ of the true mean; therefore, either could be used if they prove less costly for on-going route monitoring.

---

* s_w^2 is the within-day variance of total boardings calculated from baseline phase data and used in Step 10 to calculate conversion factor sample size.

** s_b^2 is the between-day variance of total boardings calculated from baseline phase data and used in Step 10 to calculate conversion factor sample size.

Procedure

STEP 10: DETERMINE SAMPLE SIZES FOR USE OF CONVERSION FACTORS IN MONITORING PHASE (Section 4.5) (Optional)

In order to determine the sample size required for use of the conversion factor (i.e., the sample size required for measurement of the independent variable - peak load or revenue), the following equation should be used:

$$n = \frac{t^2 s_t^2}{d^2 \, \overline{y}^2}$$

where n = total number of trips to be sampled using, e.g., a point check or farebox reading;

s_t^2 = total variance associated with the dependent variable, total boardings, (equal to the sum of the regression variance (s_r^2), the within-day variance (s_w^2), and the between-day variance (s_b^2) for each time period);

$\overline{y}$ = mean of the dependent variable (total boardings) in the baseline phase for the given time period;

t = 1.645 for 90% confidence level;

d = desired tolerance range (as fraction of the mean).

In order to determine the sampling plan for each time period, n is simply divided by the number of scheduled trips in each period and rounded up to obtain the sample days for each time period.

Example

STEP 10: For Route 16, am peak period, Property A calculates the respective sample sizes for load and revenue conversion factors as follows:

$$n = \frac{t^2 s_t^2}{d^2 \bar{y}^2} \qquad n = \frac{t^2 s_t^2}{d^2 \bar{y}^2}$$

$$= \frac{(1.645)^2 (1321)}{(.15)^2 (71)^2} \qquad = \frac{(1.645)^2 (1187)}{(.15)^2 (71)^2}$$

$$= 32 \qquad = 29$$

Since there are 20 round trips on Route 16 during the am peak period, either conversion factor would require a two day sample. In the case of Property A, both the load and revenue data will be obtained from the same technique (i.e., a point check where the checker boards each bus to read the farebox). Thus, either conversion factor (or an average of the two) could be used.

Procedure

STEP 11: DETERMINE IF USE OF CONVERSION FACTORS IN MONITORING PHASE IS LESS COSTLY THAN DIRECT MEASUREMENT (Section 4.5) (Optional)

To determine whether use of conversion factors in the monitoring phase is less costly, compare the cost (in terms of checker days) of directly measuring total boardings (simply the baseline phase cost of performing ride checks) with the cost of performing load checks or farebox readings for the number of days determined in STEP 10 above.

Based on the this comparison, select monitoring phase techniques for each route and adjust the sampling plan accordingly. Note: Different techniques could be used on different routes, as conversion factors may be appropriate only on some routes (e.g., the higher frequency routes) in the system.

Example

STEP 11: Property A compares the cost of directly measuring total boardings on Route 16, am peak period, to the cost of using load or revenue conversion factors during the baseline phase. While the baseline phase (and direct measurement of total boardings during the monitoring phase) would require 24 checker hours for the am peak period (8 checkers for 3 hours for 1 day), use of either conversion factor will require only 6 checker hours in the monitoring phase (1 checker for 3 hours for 2 days).

After similar analysis for all routes and time periods, Property A found that it could use conversion factors during the monitoring phase on 57 of its 75 routes, and could save almost 50 percent of the total checker requirements estimated for the baseline phase. Thus, using its existing checker force, Property A could accurately monitor all of its routes four times each year, at a cost of about $70,000 per monitoring phase cycle (including about $50,000 for checker resources and $20,000 for data processing).

Appendix A

SAMPLING THEORY AND WORK SHEETS

This appendix briefly describes sampling theory and the various technical inputs for sample size determination, and, in addition, provides work sheets for the calculation of these inputs.

In Sections A.1 to A.3, the theory of sampling on which the procedures in the manual are based is briefly discussed, including the nature of sampling, how to sample, and how much to sample. Section A.3 includes instructions and work sheets for calculating the between- and within-day coefficients of variation, and instructions for specifying the level of accuracy and confidence desired. The formulae for calculating the sample sizes based on these inputs, which were used to calculate the sample size tables contained in Volume 2, are then presented and explained.

In Section A.4, procedures are presented for calculating the accuracy of previously collected data and for specifying a confidence interval about the mean. Sections A.5 and A.6 include a discussion of and provide work sheets for performing a difference-of-means test to determine whether or not a statistically significant change has occurred in the data being collected.

A.1 The Nature of Sampling

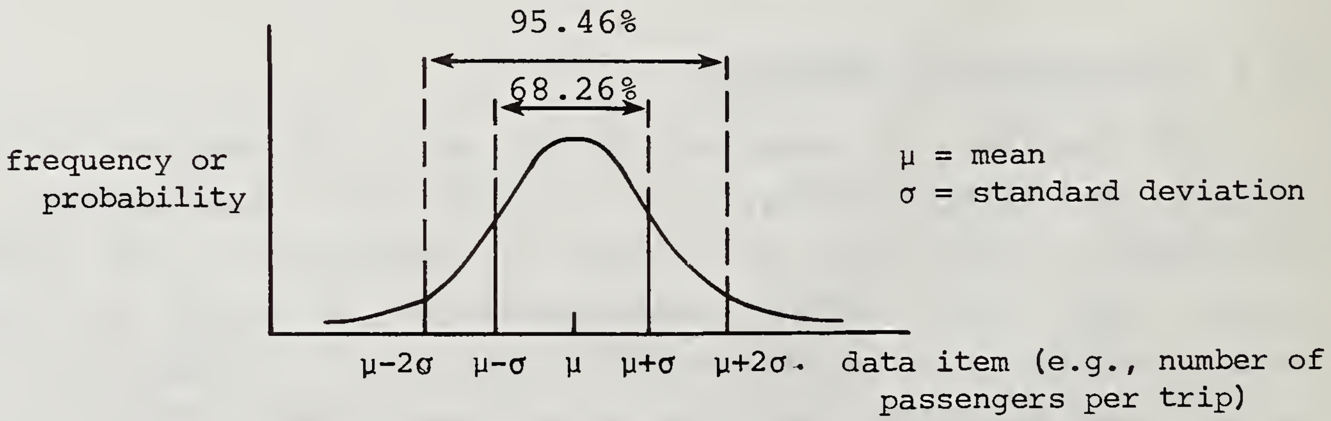
The purpose of sampling is to gain information about the nature or "distribution," of a particular population. This population describes the total of passengers, bus trips or other data item under investigation, in terms of certain characteristics or attributes of interest. A sample is simply a subgroup selected from this total population. Since the form of a population distribution is often unknown, the sample data must be used to estimate the characteristics of the total population. The basic logic of estimation is comparison between the observed sample data and the results one would predict given various possible forms of the underlying distribution.

Note that the sample data also have a distribution. Thus, it is possible to calculate statistics about the sample data, such as the sample mean and sample variance. The sample mean is an estimate of the most likely value - often termed the "expected value" - of the population mean. The sample variance indicates how widely spread out the sample data are. A third statistic, the sample standard deviation, is another measure of the dispersion about the mean, and is simply defined as the square root of the sample variance.

As previously mentioned, in order to make inferences regarding how representative these sample statistics are of the population from which the sample was taken, it is necessary to make assumptions about the form of the underlying distribution. One commonly assumed distribution is the normal distribution, shown below in Figure A.1.

Figure A.1

The Normal Distribution



Source: H. M. Blalock, Jr., Social Statistics, p. 99.

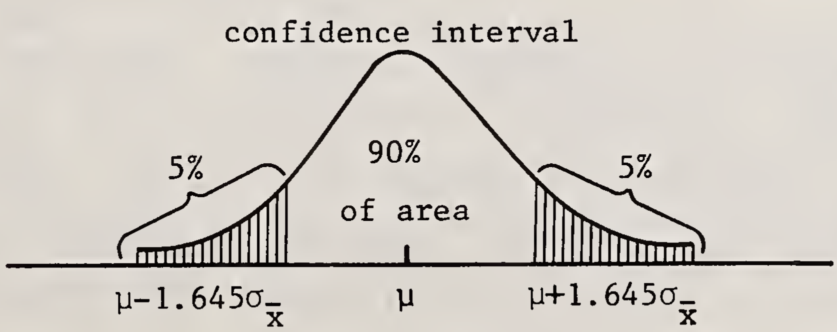
The normal distribution is important because a large number of populations are found to be approximately normally distributed, and because it serves as the basis for most of the statistical tests described in subsequent sections of this appendix.

The normal distribution has the important property that, regardless of the particular mean or standard deviation a normal curve may have, the same proportion of cases always lies between the mean and a point along the horizontal axis that is a given distance from the mean. Figure A.1 shows, for example, that 68.26% of the cases always fall within one standard deviation on either side of the mean. This can also be expressed as a 68.26% probability that a particular case (e.g., a particular bus trip) falls within one standard deviation of the population mean.

One problem with the normal distribution for statistical tests based on sample data, however, is that the normal distribution assumes that the true population mean and standard deviation are known. Another distribution, the "t" distribution, allows the use of the mean and standard deviation computed from the sample data. The t distribution, like the normal distribution, is a bell-shaped symmetrical distribution that for small samples (e.g., with fewer than 50 or 60 observations) is flatter than a normal curve, but as the sample size increases, approaches the normal curve (Figure A.2). Thus for a large sample, a t value of 1.645 indicates that 90%

Figure A.2

t-Distribution Showing a 90% Confidence Interval



of the total area under the curve is contained within 1.645 standard deviations of the mean, as shown by the unshaded area in Figure A.2. This can also be thought of as a probability distribution in which there is a 90% likelihood that the sample mean falls in the unshaded area, and a 10% chance it falls in one of the shaded areas.

A.2 How to Sample

The important issue in data collection design is what is the most feasible (cost-effective) way to define the population and to take a sample from it. In determining this sample design, two major principles are: 1) avoid bias in the selection procedure, and 2) achieve maximum precision for a given outlay of resources. An estimate is unbiased if the expected value of the estimate (e.g., of the mean or variance) is the same as the true population parameter.

If a sample is biased, e.g., if too many observations are taken from one segment of a population and too few from another, the sample mean and variance do not accurately characterize the underlying distribution. Thus, if too many heavily patronized bus trips are sampled, the estimate based on the sample of average passengers per trip will not accurately represent bus utilization for the population in question. Note that any one sample may yield an inaccurate estimate of the population value, even though the estimator is unbiased. The difference is that an unbiased estimator on average produces an accurate estimate. On the other hand, if there is an inherent bias in the sample selection process, repeated samples will not produce an accurate estimate of the true population mean.

Following these principles, the type of sampling contained in this manual is called "cluster sampling." In cluster sampling, the population is divided into a large number of groups, and samples are taken from them. The objective in cluster sampling is to select clusters that are as heterogeneous as possible to reflect the whole range of the characteristics under consideration.

The primary advantage of cluster sampling is that it reduces the cost of data collection by allowing a concentration of effort. For example, instead of selecting trips at random throughout all days of the year, taking samples from bus trips within a small group of days allows checkers to be more efficiently scheduled. Random sampling within clusters will produce unbiased estimates of the population characteristics, meaning that if enough samples are taken, the values collected will, on average, rest on the "true" value. An important disadvantage, however, is that each sampling stage contributes to the total sampling error. Thus, cluster sampling tends to be more variable than pure random sampling, although it is less costly to conduct.

A.3 How Much to Sample

The decision of how much data to collect depends directly on how variable the data are, how accurate the estimates need to be, and how confident one wants to be that they fall within certain accuracy limits. A trade-off exists between the amount of data that are collected and the accuracy and confidence that can be obtained. First, in Section A.3.1, the relevant measures of data variation are described. Work sheets for calculating the variances and coefficients of variation are included at the end of Section A.3.2. A discussion of accuracy and confidence levels follows in Section A.3.3. Section A.3.4 concludes this section on how much to sample by presenting the actual sample size formulae used to calculate how many days and trips per day should be sampled to obtain the desired accuracy.

A.3.1 Coefficients of Variation

For the two-stage sampling framework used in this manual, two components of variation are important: the between-day variance and the within-day (or within-period) variance. The between-day variance is the weighted average of the square of the difference between the average ridership per trip for each day and the overall average. It is the variance of the mean values from day to day. The formula for this is:

$$s_b^2 = \frac{n}{K(n-1)} \sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2 k_i$$

where s_b^2 = the between-day variance;

K = the total number of trips that are counted over all n days during the time period of interest;

k_i = the number of trips counted on day i during the time period of interest;

$\bar{x}_i$ = the average number of passengers boarding per trip on day i during the time period of interest;

$\bar{x}$ = the average number of passengers boarding over all days during the time period of interest;

n = the total number of days for which data were collected;

$\sum_{i=1}^{n}$ = the summation across all n days sampled.

The between-day coefficient of variation can then be calculated from the between-day variance simply by taking the square root of the between-day variance and dividing by the overall mean, as shown below:

$$v_b = \frac{\sqrt{s_b^2}}{\bar{x}}$$

where v_b = the between-day coefficient of variation; and all other symbols are as previously defined.

The advantage of the coefficient of variation is that it does not reflect the overall level of the data item (e.g., boardings) but just the variability. In other words, by dividing by the overall mean, it is possible to standardize the scale of each variance to enable comparisons among time periods, routes, or data items.

The second component of variation that must be calculated in order to estimate the necessary sample size is the within-day (or within-period) variance. The within-day variance is the average variance for the time period under investigation. As with the between-day variance, the within-day variance also is a weighted average, but in this case of the variances in ridership from trip to trip for the specified periods within each day. As before, these are weighted by the number of trips each day, summed, and divided by the total number of trips sampled to arrive at an average daily variance.

This is expressed mathematically as:

$$s_w^2 = \frac{1}{K} \sum_{i=1}^{n} s_i^2 k_i$$

where s_w^2 = the within-day variance;

s_i^2 = the variation from trip to trip for each individual day i;

and all other symbols are as defined previously.

As before, the within-day (or within-period) coefficient of variation is calculated by taking the square root of the within-day variance and dividing by the overall mean, as follows:

$$v_w = \frac{\sqrt{s_w^2}}{\bar{x}}$$

where v_w = the within-day (or within-period) coefficient of variation;

and all other symbols are as defined previously.

A.3.2 Instructions and Work Sheets for Calculating Coefficients of Variation

Work sheets are included to calculate the mean and variance for each time period, the between and within-day (or within-period) variances, and the between and within-day coefficients of variation. Instructions are presented on the left-hand page, with the corresponding work sheets on the right-hand page. These work sheets can be reproduced and used continually. In addition, a sample set of work sheets completed for the load data item follow at the end of this section.

Instruction

Step One. Calculate the MEAN and each day and time period

The mean is simply the average value for a time period on a given day, and is calculated using the following equation:

$$\text{mean} = \bar{x} = \frac{\Sigma x}{n}$$

The "Σ" is a summation sign and simply instructs you to add the x's together. (Each x corresponds to a value.) If, for example, you have five trips during your morning peak, then the mean ridership on a given day would be:

$$\frac{\text{Value 1} + \text{Value 2} + \text{Value 3} + \text{Value 4} + \text{Value 5}}{5}$$

Using this procedure (or a calculator with a function key for calculating the mean):

a) CALCULATE THE MEAN FOR EACH DAY AND TIME PERIOD. ENTER THE MEANS IN TABLE 1.

Work Sheet

Step One

Table 1

Mean Values for Each Day and Time Period

DAY / TIME PERIOD	1	2	3	4	5	...
1						
2						
3						
4						
5						
⋮						

Instruction

Step Two. Calculate the VARIANCE for each day and time period.

The variance is a measure of how all the values are distributed about the mean. A low variance occurs if all the values are close to the mean. A high variance occurs if the values are spread widely apart. The variance is calculated using the following equation:

$$\text{variance} = s^2 = \overbrace{\frac{\Sigma x^2}{n - 1}}^{\text{Term 1}} - \overbrace{\left(\frac{\Sigma x}{n - 1}\right)^2}^{\text{Term 2}}$$

Again, each x corresponds to a value, and n is the number of values.

In the above equation, the first term instructs you to square each x, add the squared values together, and then divide by n-1. The second term instructs you to add the x's together, divide by n-1, and then square the result. Finally, subtract the second term from the first term.

Using this procedure (or a calculator with a function key for calculating the mean):

a) CALCULATE THE VARIANCE FOR EACH DAY AND TIME PERIOD. ENTER THE VARIANCES IN TABLE 2.

Work Sheet

Step Two

Table 2

Variances for Each Day and Time Period

DAY / TIME PERIOD	1	2	3	4	5	...
1						
2						
3						
4						
5						
⋮						

Instruction

Step Three. Calculate the BETWEEN-DAY VARIANCE AND COEFFICIENT for each time period.

Use Table 3 and Lines 1 through 9 to help calculate the between-day coefficient of variation for each of the time periods identified in Table 1. This means that Table 3 will be completed once for each time period.

The first column of Table 3 lists the days for which data are available. It is assumed to be no more than five in the worksheets; if data are available for more than five days, simply add extra rows. The rest of Table 3 and Lines 1 through 9 are filled out as follows:

a) For each time period of the day, COPY THE ENTRIES FROM THE APPROPRIATE ROW OF TABLE 1 INTO THE SECOND COLUMN OF TABLE 3.

b) IN COLUMN 3, RECORD THE NUMBER OF TRIPS OBSERVED ON EACH DAY.

c) Add the entries in Column 3. ENTER THE TOTAL ON LINE 1.

d) IN COLUMN 4, MULTIPLY THE ENTRIES IN COLUMNS 2 and 3. Add the entries in Column 4. ENTER THE TOTAL ON LINE 2.

e) COMPUTE THE OVERALL MEAN USING LINE 3. This value combines information from all days for which data are available. RECORD ANSWER IN COLUMN 5. Note that the same number will be entered for each day.

f) IN COLUMN 6, SUBTRACT THE ENTRIES IN COLUMN 5 FROM THE ENTRIES IN COLUMN 2.

g) IN COLUMN 7, SQUARE THE ENTRIES IN COLUMN 6.

h) IN COLUMN 8, MULTIPLY THE ENTRIES IN COLUMNS 3 and 7.

i) Add the entries in column 8. ENTER THE TOTAL ON LINE 4.

j) ENTER THE TOTAL NUMBER OF DAYS SAMPLED IN LINE 5. This is the total number of days for which rows are filled out in Column 1 of Table 3.

(Step Three Instruction continued on page 128.)

(Step Three Instruction, continued)

k) COMPUTE FACTOR USING LINE 6.

l) COMPUTE THE BETWEEN-DAY VARIANCE USING LINE 7.

m) COMPUTE THE BETWEEN-DAY STANDARD DEVIATION USING LINE 8.

n) COMPUTE THE BETWEEN-DAY COEFFICIENT OF VARIANCE USING LINE 9.

Work Sheet

Step Three.

Table 3

Calculation of Between-Day Variation

Time Period: ________

1	2	3	4	5	6	7	8
DAY	MEAN FOR DAY	# TRIPS FOR DAY	COL. 2 X COL. 3	OVERALL MEAN	COL. 2 - COL. 5	COL. 6 X COL. 6	COL. 3 X COL. 7
1							
2							
3							
4							
5							
		Total Column 3	Total Column 4				Total Column 8

Line #1) TOTAL OF COLUMN 3 = ________ (#1)

Line #2) TOTAL OF COLUMN 4 = ________ (#2)

Line #3) OVERALL MEAN = (________ (#2) ÷ ________ (#1)) = ________ (#3)

Line #4) TOTAL OF COLUMN 8 = ________ (#4)

Line #5 TOTAL NUMBER OF DAYS SAMPLED = ________ (#5)

Line #6 ________ (#5) ÷ [________ (#1) x (________ (#5) - 1)] = ________ (#6)

Line #7) BETWEEN-DAY VARIANCE = (________ (#4) x ________ (#6)) = ________ (#7)

Line #8) BETWEEN-DAY STANDARD DEVIATION = $\sqrt{\text{\_\_\_\_\_\_\_\_ (\#7)}}$ = ________ (#8)

Line #9) BETWEEN-DAY COEFFICIENT OF VARIATION = (________ (#8) ÷ ________ (#3)) = ________ (#9)

Step Four. Calculate the WITHIN-DAY (WITHIN-PERIOD) VARIANCE and COEFFICIENT OF VARIATION for each time period.

Use Table 4 and Lines 10 through 13 to calculate the within-day (or within-period) coefficient of variation for each of the time periods identified in Table 1. As in Table 3, the first column lists the days for which data are available. The rest of Table 4 and Lines 10 through 13 are filled out as follows:

a) For each time period of the day, COPY THE ENTRIES FROM THE APPROPRIATE ROW OF TABLE 2 INTO THE SECOND COLUMN OF TABLE 4.

b) IN COLUMN 3, RECORD THE NUMBER OF TRIPS OBSERVED ON EACH DAY. Note that these entries are the same as those recorded in Column 3 of Table 3.

c) IN COLUMN 4, MULTIPLY THE ENTRIES IN COLUMN 2 and 3.

d) Add the entries in Column 4. ENTER THE TOTAL ON LINE 10.

e) COMPUTE THE WITHIN-DAY VARIANCE USING LINE 11.

f) COMPUTE THE WITHIN-DAY STANDARD DEVIATION USING LINE 12.

g) COMPUTE THE WITHIN-DAY COEFFICIENT OF VARIATION USING LINE 13.

Work Sheet

Step Four.

Table 4

Calculation of Within-Day Variation

Time Period: ______________

1	2	3	4
DAY	VARIANCE FOR DAY	# TRIPS FOR DAY	COL. 2 X COL. 3
1			
2			
3			
4			
5			
			Total Column 4

Line #10) TOTAL OF COLUMN 4 = ________ (#10)

Line #11) WITHIN-DAY VARIANCE = (________ (#10) ÷ ________ (#1)) = ________ (#11)

Line #12) WITHIN-DAY STANDARD DEVIATION = $\sqrt{\_\_\_\_\_\_\_\_}$ (#11) = ________ (#12)

Line #13) WITHIN-DAY COEFFICIENT OF VARIATION (________ (#10) ÷ ________ (#3)) = ________ (#11)

Work Sheet EXAMPLE:

Calculation of Coefficients of Variation for Chicago RTA Route 210 (Load data, 3 time periods, 4 days of data available)

Step One

Table 1

Mean Values for Each Day and Time Period

TIME PERIOD \ DAY	1	2	3	4	5	...
1	34.5	33.7	33.3	28.8		
2	12.4	15.0	13.7	14.3		
3	21.0	19.3	20.3	13.3		
4						
5						
⋮						

Work Sheet EXAMPLE (continued)

Step Two

Table 2

Variances for Each Day and Time Period

TIME PERIOD \ DAY	1	2	3	4	5	...
1	57.3	25.6	16.2	54.5		
2	61.8	36.0	29.9	16.6		
3	36.0	63.2	95.7	10.9		
4						
5						
⋮						

Work Sheet EXAMPLE (continued)

Step Three.

Table 3

Calculation of Between-Day Variation

Time Period: 1

1	2	3	4	5	6	7	8
DAY	MEAN FOR DAY	# TRIPS FOR DAY	COL. 2 X COL. 3	OVERALL MEAN	COL. 2 - COL. 5	COL. 6 X COL. 6	COL. 3 X COL. 7
1	34.5	6	207	32.6	1.9	3.6	21.6
2	33.7	6	202.2	32.6	1.1	1.2	7.2
3	33.3	6	199.8	32.6	0.7	0.5	3.0
4	28.8	6	172.8	32.6	-3.8	14.4	86.4
5							
		24 (Total Column 3)	781.8 (Total Column 4)				118.2 (Total Column 8)

Line #1) TOTAL OF COLUMN 3 = 24 (#1)

Line #2) TOTAL OF COLUMN 4 = 781.8 (#2)

Line #3) OVERALL MEAN = (781.8 (#2) ÷ 24 (#1)) = 32.6 (#3)

Line #4) TOTAL OF COLUMN 8 = 118.2 (#4)

Line #5 TOTAL NUMBER OF DAYS SAMPLED = 4 (#5)

Line #6 4 (#5) ÷ [24 (#1) x (4 (#5) - 1)] = 0.056 (#6)

Line #7) BETWEEN-DAY VARIANCE = (118.2 (#4) x 0.056 (#6)) = 6.62 (#7)

Line #8) BETWEEN-DAY STANDARD DEVIATION = $\sqrt{6.62}$ (#7) = 2.57 (#8)

Line #9) BETWEEN-DAY COEFFICIENT OF VARIATION = (2.57 (#8) ÷ 32.6 (#3)) = 0.079 (#9)

Work Sheet EXAMPLE (continued)

Step Four.

Table 4

Calculation of Within-Day Variation

Time Period: 1

1	2	3	4
DAY	VARIANCE FOR DAY	# TRIPS FOR DAY	COL. 2 X COL. 3
1	57.3	6	343.8
2	25.6	6	153.6
3	16.2	6	97.2
4	54.5	6	327.0
5			
			921.6
			Total Column 4

Line #10) TOTAL OF COLUMN 4 = 921.6 (#10)

Line #11) WITHIN-DAY VARIANCE = (921.6 (#10) ÷ 24 (#1)) = 38.4 (#11)

Line #12) WITHIN-DAY STANDARD DEVIATION = $\sqrt{38.4}$ (#11) = 6.20 (#12)

Line #13) WITHIN-DAY COEFFICIENT OF VARIATION (6.20 (#10) ÷ 32.6 (#3)) = 0.19 (#11)

A.3.3 <u>Accuracy and Confidence: The Standard Error</u>

Before calculating the number of days and trips per day to be sampled, it is also necessary to specify an acceptable standard error, in terms of a percent of the mean and the desired confidence expressed as the "t value" corresponding to the appropriate confidence level. Each of these components is discussed below, followed by the formula for calculating the standard error.

In the standard error, the accuracy refers to the error range which is the range in which the true value of the statistic (in this case the mean) may be around the observed value (the value calculated from the sample). For example, the error range specified in Section 15 is ±10% of the observed value. In calculating the sample size using the formula presented in Section A.3.4, it is possible to set the level of accuracy desired for a given data item, keeping in mind the trade-off between the level of accuracy and the sample size required.

The confidence level, on the other hand, indicates the probability that the true value will be contained within the specified error range. In the manual, a confidence level of 90% is specified for data at the route level. Thus, there is a 90% chance that the true value of the data item is within the specified error range (e.g., ±10%) of the observed value. As previously mentioned, the desired confidence is expressed as the "t value" corresponding to the appropriate confidence interval.

The equation used to determine the desired standard error is:

$$D = \frac{d\bar{x}}{t}$$

where d = the desired accuracy expressed as a fraction of the mean (e.g., ±.10, ±.15 or ±.20 or ±.30)

$\bar{x}$ = the population mean defined, for example, as the total number of passengers ÷ the total number of trips.

t = the t value associated with large samples for the desired confidence interval, e.g. 1.645 for a 90% confidence interval or 1.960 for a 95% confidence interval. (Note that 1.645 is used throughout the manual.)

To calculate the standard error, simply multiply the desired accuracy by the overall mean and divide by the t statistic for the desired level of confidence.

A.3.4 The Sample Size Formula

Using the calculated values of s_b^2, s_w^2 and D, the next step is to calculate the combinations of days and trips per day that will fulfill the accuracy requirements previously specified. (Note that convenient sample size tables are included in Volume 2 of the manual, which can be used instead of these formulas for determining sample sizes. For a discussion of these tables, see Section A.3 of this appendix.) The formula for this is:

$$\text{trips per day} = k = \frac{TNs_w^2}{n\,[D^2TN + Ts_b^2] + Ns_w^2 - TNs_b^2}$$

where T = number of scheduled trips in the time period;

N = number of days in the season being analyzed;

k = number of trips sampled per day;

n = number of days sampled;

D = the standard error = dx/t (as defined previously);

s_w^2 = within-day variance;

s_b^2 = between-day variance.

This formula can be transformed to one using the within and between-day coefficients of variation (rather than the variances) as shown below:

$$k = \frac{TNv_w^2}{n\left[\frac{d^2}{t^2} TN + T v_b^2\right] + Nv_w^2 - TNv_b^2}$$

where $v_w = s_w / \bar{x}$ = the within-day coefficient of variation;

$v_b = s_b / \bar{x}$ = the between-day coefficient of variation;

and all other terms are as defined above.

A further transformation can be made to allow one to solve for the number of sample days as shown below:

$$n = \frac{v_w^2(T - k)\ N + v_b^2\ TNk}{Tk\left[\frac{d^2}{t^2} N + v_b^2\right]}$$

where all terms are as defined above.

By solving these equations for different numbers of days, or conversely different numbers of trips per day, one can identify a set of combinations of days and trips per day that will provide the required sample.

A.4 Calculating a Confidence Interval From Sample Data

While in Section A.3.3 the procedure for specifying the standard error prior to the data collection was explained, it is also possible to determine the accuracy of a previously collected sample. This is done by calculating the standard error of the sample data, and solving for "d", the accuracy obtained. Further, once the accuracy of the sample is known, a confidence interval about the mean can be calculated. The confidence interval shows the range within which the true mean should fall with a specified level of confidence (e.g., 90% of the time).

The standard error, D, is calculated given the number of days and trips per day of the actual sample. The mathematical expression for the standard error is as follows:

$$D = \sqrt{\frac{s_w^2 \, (T - k)}{Tkn} + \frac{s_b^2 \, (N - n)}{Nn}}$$

where: D = standard error;

k = number of trips sampled per day;

n = number of sampled days;

s_w^2 = within-day variance of the sample data;

s_b^2 = between-day variance of the sample data;

T = number of scheduled trips per day;

N = number of days in a season being analyzed.

Note: When fewer than three days of sample data are available, the original (or pretest) between-day variance must be used, as the sample data will be unstable.

Once the value D is known, it is possible to solve for "d", the accuracy obtained, using the following formula:

$$d = \frac{tD}{\bar{x}}$$

where: d = the accuracy obtained expressed as a fraction of the mean.

t = the normal statistic z for the desired confidence, e.g., 1.645 for 90% confidence.

D = the standard error from above

$\bar{x}$ = the population mean

Finally, the confidence interval about the mean can be defined as follows:

$$\bar{x} \pm (d\bar{x})$$

A.5 The Difference of Means Test

The difference of means test is used to determine if a change has occurred in the data being collected. For example, this test can be used to determine whether or not average ridership has increased significantly from one sample to another. The question addressed here is whether the difference is "statistically significant," or whether or not it could have occurred by chance.

In the difference of means test, a "null hypothesis" is tested that the mean of the one sample is equal to the mean of the second sample. This is usually written as H_o: $\mu_1 = \mu_2$, or H_o: $\mu_1 - \mu_2 = 0$. To test this hypothesis, the "t-statistic" for the sample data is calculated using the following formula:

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}}$$

where $\bar{x}_i$ refers to the sample mean of each sample (i = 1,2)

$\hat{\sigma}_i^2$ is calculated for each sample (i = 1,2) using the formula below:

$$\hat{\sigma}_i^2 = \frac{s_{wi}^2 (T - k_i)}{Tk_i n_i (k_i n_i - 1)} + \frac{S_{bi}^2 (N - n_i)}{Nn_i (n_i - 1)}$$

where all of the terms are defined as in Section A.4.

The absolute value of "t_{calc}" is calculated (i.e., ignore any negative sign), and then compared with the t value associated with the desired level of confidence from Table A.1.

TABLE A.1

Confidence	80%	85%	90%	95%
Value Associated with confidence	1.28	1.45	1.65	1.96

If t_{calc} is greater than t_{table}, then the null hypothesis that the means are the same is rejected, and it is concluded that the means are statistically different at the level of confidence selected. Thus, for example, if $t_{calc} = 1.8$ and $t_{table} = 1.645$, then it is appropriate to conclude that one can be 90% confident that the means from the two samples are in fact different. If however, t_{calc} is less than t_{table}, then the means are not statistically different.

A.6 Instructions and Work Sheets for Performing the Difference of Means Test

Work sheets are presented for performing the difference of means test on the following pages. As before, instructions are given on the left hand page, with the corresponding work sheets for performing the calculations provided on the right hand page.

Instruction

Step One. Calculate the WITHIN-DAY (WITHIN-PERIOD) and BETWEEN-DAY VARIANCE (if not currently available).

The difference of means test uses the within-day (or within-period) and between-day variances for each sample. If only the coefficients of variation are available, the variances can be calculated by multiplying each coefficient of variation by the population mean and squaring the product. (Note that Step One will be completed twice, once for each sample.) If the variances are currently available, skip to instruction "b" below.

a) CALCULATE THE WITHIN-DAY AND BETWEEN-DAY VARIANCES USING LINES #1 THROUGH #5 FOR EACH SAMPLE.

b) If not calculated in "a" above, ENTER THE WITHIN-DAY VARIANCE IN LINE #4 FOR EACH SAMPLE (from Line #11, page 131 of the previous set of work sheets).

c) If not calculated in "a" above, ENTER THE BETWEEN-DAY VARIANCE IN LINE #5 FOR EACH SAMPLE (from Line #7, page 129, of the previous set of work sheets).

Step Two. Calculate the SQUARE OF " $\hat{\sigma}_i$ ".

For each sample, calculate the square of " $\hat{\sigma}_i$ " (an adjusted form of the standard error) based on the within-day and between-day variances, the number of scheduled trips, the number of sampled trips, the number of days in a season, and the number of days sampled. Note that Step Two will also be completed twice, once for each sample.

a) ENTER THE NUMBER OF SCHEDULED TRIPS PER DAY IN LINE #6.

b) ENTER THE AVERAGE NUMBER OF SAMPLED TRIPS PER DAY IN LINE #7.

c) ENTER THE NUMBER OF DAYS IN A SEASON (FOR WHICH SERVICE IS PROVIDED) IN LINE #8.

d) ENTER THE AVERAGE NUMBER OF DAYS SAMPLED IN LINE #9.

e) Using these inputs, CALCULATE THE SQUARE OF THE ADJUSTED STANDARD ERROR WITH LINES #10 THROUGH #12.

Work Sheet

Sample: ______

Time Period: __________

Step One.

Line #1) WITHIN DAY COEFFICIENT = ____ (#1)

Line #2) BETWEEN DAY COEFFICIENT = ____ (#2)

Line #3) POPULATION MEAN = ____ (#3)

Line #4) WITHIN-DAY VARIANCE = (____ X ____) X (____ X ____) = ____
(#1, #3, #1, #3, #4)

Line #5) BETWEEN-DAY VARIANCE = (____ X ____) X (____ X ____) = ____
(#2, #3, #2, #3, #5)

Step Two.

Line #6) NUMBER OF SCHEDULED TRIPS PER DAY = ________ (#6)

Line #7) AVERAGE NUMBER OF SAMPLED TRIPS PER DAY = ________ (#7)

Line #8) NUMBER OF DAYS IN A SEASON = ________ (#8)

Line #9) AVERAGE NUMBER OF DAYS SAMPLED = ________ (#9)

Line #10) [(____ - ____) X ____] ÷ [((____ X ____) - 1) X
(#6, #7, #4; #7, #9)

(____ - 1) X ____ X ____] = ____
(#6, #7, #9, #10)

Line #11) [(____ - ____) X ____] ÷
(#8, #9, #5)

[(____ - 1) X (____ - 1) X ____] = ____
(#9, #8, #9, #11)

Line #12) SQUARE OF THE ADJUSTED STANDARD ERROR = ____ + ____ = ____
(#10, #11, #12)

Instruction

Step Three. Calculate a "t-statistic" for the DIFFERENCE OF MEANS TEST.

The t-statistic for the difference of means test is calculated by subtracting the mean of the second sample from that of the first sample, and dividing by the square root of the adjusted standard error. The results of Steps One and Two are summarized, and the calculations performed according to the following instructions:

a) ENTER THE MEAN OF EACH SAMPLE IN THE APPROPRIATE COLUMN OF LINE #13,#14 (from Step One, Line #3, or from Line #3, page 129 , of the previous set of work sheets).

b) ENTER THE SQUARE OF THE STANDARD ERROR OF EACH SAMPLE IN THE APPROPRIATE COLUMN OF LINE #15,#16 (from Step Two, Line #12).

c) CALCULATE THE t-statistic WITH LINES #17 THROUGH #19.

Step Four. Select the "t" value that corresponds to the desired level of confidence.

Select the "t" value that corresponds to the desired level of confidence from the following table:

Confidence	80%	85%	90%	95%
"t" Value Associated with confidence	1.28	1.45	1.65	1.96

a) ENTER THE SELECTED t VALUE IN LINE #20.

Step Five. Perform the DIFFERENCE OF MEANS TEST.

Determine whether the difference in the mean values of sample one and sample two is statistically significant by comparing the calculated t values, "t_{calc}", and the t value selected from the table. To do this:

a) ANSWER QUESTION IN LINE #21.

Step Three.

		Sample One	Sample Two
Line #13,#14.	MEAN (FROM LINE #3) =	______ #13	______ #14
Line #15,#16.	SQUARE OF STANDARD ERROR (FROM LINE #12) =	______ #15	______ #16

Line #17. (______ #13 - ______ #14) = ______ #17

Line #18. $\sqrt{\underset{\#15}{\_\_\_\_\_\_} + \underset{\#16}{\_\_\_\_\_\_}}$ = ______ #18

Line #19. t_{calc} = ______ #17 ÷ ______ #18 = ______ #19

Step Four.

Line #20. SELECTED t value = ______ #20

Step Five. Perform the DIFFERENCE OF MEANS TEST

If one "ignores" whether t_{calc} is positive or negative:

Line #21. IS ______ #19 GREATER THAN ______ #20 ?

IF YES, THE DIFFERENCE OF MEANS IS STATISTICALLY SIGNIFICANT AT THE LEVEL OF CONFIDENCE SELECTED IN STEP FOUR.

IF NO, THE DIFFERENCE OF MEANS IS NOT STATISTICALLY SIGNIFICANT AT THE LEVEL OF CONFIDENCE SELECTED IN STEP FOUR. ANY OBSERVED DIFFERENCE COULD HAVE OCCURRED BY CHANCE.

Appendix B

ROUTE CLASSIFICATION

The purpose of a route classification scheme is to group routes with similar characteristics into categories that can be used to streamline the data collection process. The categorization of routes facilitates the data collection process in two related ways: 1) it allows pretest variances for a particular route type and data item to be applied to all routes in a category, and 2) it simplifies the development of a sampling strategy for a transit property in cases where certain data collection techniques are more appropriate for some route types than for others.

In the case of feeder routes, for example, load at the transfer point may exhibit relatively low variances, and at the same time a fairly stable relationship between load at that point and total route ridership. Thus, load counts at the transfer point could be used to estimate total ridership using conversion factors. Sample sizes would be calculated using characteristic coefficients of within- and between-day variation for feeder routes of the appropriate route length and/or headway.

Several different stratifications may be useful in route classifications for sampling purposes: route type, length, headway, total boardings, average load factor, and productivity (e.g., passengers per mile). Possible route categories include radial, crosstown, feeder, express, shuttle, intra-suburban, satellite suburban, and inter-suburban. These are discussed further below.

Route length categories may be useful for properties with a wide range of route lengths. Headway categories help to differentiate routes and time periods with frequent headways such that riders do not schedule departures on particular trips, routes and time periods for which riders do schedule their trips (e.g., with headways ten minutes or longer), and routes and time periods operating on policy headways (e.g., where buses depart at regular intervals to achieve levels of service determined independently of ridership during these periods). Stratification by total boardings, average load factor and productivity measures all are aimed at taking advantage of the relationship, if any, between patronage levels and overall data variability.

All types of route stratification are characterized by different types of variances and optimal sampling strategies. The categories may not be mutually exclusive, however, as, for example, routes may serve both feeder and crosstown functions. In such cases, it is recommended that a route simply be assigned to the predominant category, or if it seems appropriate, divided into segments that conform to one type or another. Another alternative is to create a separate classification if a significant number of routes serve the same combination of functions.

A number of route categories are described in general terms below, along with the implications for data collection strategies to be employed. More precise definitions, of course, depend on the characteristics of a particular property.

1. Radial - These are routes that run from outward sectors of the city to the central business district or major activity centers. These routes may be primarily short-haul or local in nature, or may carry longer trips from outlying areas. As this configuration suggests, these routes are likely to have a single maximum load point in the CBD or activity center, increasing the attractiveness of point checks as a data collection technique.

2. Crosstown - Crosstown routes avoid the CBD areas, and typically distribute trips along the entire length of the route. As such, these routes have many origin and destination points, with no obvious maximum load point to facilitate the data collection process.

3. Feeder - These routes distribute passengers between residences and commuter rail or rapid transit stations, primarily during the a.m. and p.m. peak periods. For

data collection purposes, the central characteristic of a true feeder is a fairly stable relationship between route ridership and load at the transfer point. The load at this point is likely to be the maximum load point, increasing the attractiveness of point checks as a possible data collection technique.

4. Express - Express routes differ from other route categories as they typically are long routes that travel over major portions of each trip without picking up or discharging passengers. Due to the nature of the route, a count of passengers conducted just prior to the express portion of a route frequently can yield a good estimate of express ridership.

5. Shuttle - Shuttle routes distribute passengers throughout downtown or other employment and recreational districts. Shuttles often exhibit different peaking characteristics than commuter-oriented routes and can be highly variable, thus complicating the data collection process.

6. Intra- or Satellite Suburban - This category includes many types of small systems, centered around different types of traffic generators within a particular town or suburban area. Although these routes tend to be short, relatively infrequent, and homogeneous within any one system, on the whole, different systems exhibit a wide range of route types and configurations in this category. If a focal point exists, as in the CBD or commuter rail station, then point checks at these centers can be of use. However, in general, the optimal sampling strategy depends on the characteristics of the particular system, as opposed to any generic route in this category.

7. Inter-Suburban - These routes link towns in the urban fringe, and as such tend to be relatively long but infrequent routes, linking CBD's, industrial parks, schools, and other activity centers. Again, no particular sampling strategy is suggested separate from the characteristics of particular routes, although the length of those routes often results in widely distributed loads and origin-destination patterns.

The route types described here illustrate a number of the most common categories, but are not all-inclusive. The objective of categorization by route type as well as other classification schemes is to group routes in categories about which certain generalizations can be made. The most appropriate scheme of classification depends on the nature and complexity of the system. It is strongly recommended that each property develop its own route classification scheme based on the principles noted above. In this way, a property can take full advantage of local knowledge regarding service characteristics and ridership patterns.

☆U.S. GOVERNMENT PRINTING OFFICE: 1981— 341-428:1339

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

This document is being distributed through the U.S. Department of Transportation's Technology Sharing Program.

DOT-I-81-30

U.S. Department of Transportation

Office of the Secretary of Transportation

400 Seventh St., S.W.
Washington, D.C. 20590

DOT-I-81-30

TECHNOLOGY SHARING

SPECIAL STUDIES IN TRANSPORTATION PLANNING (SSTP)

PROGRAMS OF THE U.S. DEPARTMENT OF TRANSPORTATION