A Guidebook for Using Automatic Passenger Counter Data for National Transit Database (NTD) Reporting

DISCLAIMER

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation or the U.S. Department of Transportation.

The report was prepared in cooperation with the State of Florida Department of Transportation and the U.S. Department of Transportation.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
A Guidebook for Using Automatic Passenger Counter Data for National Transit Database (NTD) Reporting	December 2010
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Xuehao Chu	NCTR778-03, FDOT BDK85 977-04

9. Performing Organization Name and Address	10. Work Unit No.
National Center for Transit Research (NCTR) University of South Florida 4202 E Fowler Ave., CUT100, Tampa, FL 33620-5375	
	11. Contract or Grant No. USDOT DTRS98-G-0032

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Research and Innovative Technology Administration U.S. Department of Transportation, Washington, D.C. 20590	Final Report August 2008 - June 2010
Florida Department of Transportation 605 Suwannee Street, MS 30, Tallahassee, FL 32399	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the Florida Department of Transportation and the U.S. Department of Transportation

16. Abstract

This document provides guidance for transit agencies to use data from their automatic passenger counters (APCs) for reporting to the National Transit Database (NTD). It first reviews both the traditional data requirements on the data items to be reported and the statistical criteria that the reported data must satisfy and the new APC requirements if APCs are used to obtain the data items. More importantly, the document provides guidance on the following three areas:

- Options for Meeting NTD Requirements - Whether agencies with APCs are allowed to use their APC data for NTD reporting; whether agencies want to use APC data for NTD reporting if allowed; and how agencies should use their APC data for NTD reporting.
- Meeting the Data Requirements - How agencies should obtain 100% counts from APC data; how agencies should obtain estimates through using all usable APC data; and how agencies should obtain estimates through random sampling.
- Meeting the APC Requirements - How agencies should develop the benchmarking plan and the maintenance plan; and how agencies should conduct the benchmarking study and the annual maintenance study.

The use of the guidance is expected to help agencies better meet NTD requirements and to reduce the possibility of reported APC data not being included in the apportionment for Section 5307.

17. Key Words	18. Distribution Statement
Automatic Passenger Counters, National Transit Database, Unlinked Passenger Trips, Passenger Miles Traveled	Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161, 703-487-465, and through the NCTR website at http://www.nctr.usf.edu

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	50	

ACKNOWLEDGMENT

EXECUTIVE SUMMARY

Background

Using electronic infrared beams or mechanical treadle mats, automatic passenger counters (APCs) have the ability to count transit passengers as they board and alight transit vehicles at individual stops. When coupled with stop location information, archived APC data can be post-processed to generate disaggregate data in both time and space. While APCs are mainly used to gather data for service and operations planning, APC data can also potentially be used for reporting to the National Transit Database (NTD).

To be eligible for the Urbanized Area Formula Grant Program (i.e., Section 5307), transit agencies must meet a set of data requirements. Specifically, they must report annual data on unlinked passenger trips (UPT) and passenger miles traveled (PMT) to NTD for each mode and service type (purchased versus directly operated). They must report a 100% count of each quantity if it is available and reliable. If a reliable 100% count is not available for a quantity, they must estimate it through random sampling, and the obtained estimate must meet the minimum 10% precision level at the 95% confidence level.

To use their APC data for NTD reporting, however, agencies must meet additional requirements. They must submit a benchmarking plan and a maintenance plan before they start using APC data for NTD reporting. Using parallel sample data collected using both APCs and manual ride checks from the same sample of one-way vehicle trips, they must conduct a benchmarking study during the first year of using APC data for NTD reporting. Using parallel sample data collected from at least 100 one-way vehicle trips using both APCs and manual ride checks, they must conduct an annual maintenance study to calibrate the APCs after the benchmarking year. They must adjust 100% counts or estimates of UPT and PMT from APC data for likely missed data and data errors in the APC data.

Transit agencies face many statistical and other technical issues in meeting the new APC requirements and in meeting the traditional data requirements when using APC data. There is no detailed guidance to help transit agencies better deal with these issues. It is important for agencies to properly deal with these issues to avoid not meeting NTD requirements.

Objectives

The objective of this project was to develop guidance for transit agencies to properly deal with these statistical and other technical issues when they try to use their APC data for NTD reporting.

Findings and Conclusions

In addition to an introduction and a review of both the traditional data requirements and the new APC requirements, the guidance is organized into three sections:

- Section 3 presents guidance on whether agencies may use APC data for NTD reporting and on how APC data may be used for NTD reporting. Conditions are presented under which agencies may not use APC data for NTD reporting. In addition, factors are discussed that agencies should consider in deciding if they want to use APC data for NTD reporting when allowed. Finally, options are presented that agencies may select once they have decided to use APC data for NTD reporting.

- Section 4 presents guidance to help transit agencies meet the traditional data requirements. The guidance covers procedures for obtaining 100% counts from APC data. It also covers procedures for obtaining estimates through random sampling using APCs. More important, it covers procedures for obtaining estimates through using all usable APC data. In addition, the guidance covers procedures for obtaining adjustment factors for missed data. It also covers procedures to account for both missed data and data errors in developing plans for data collection.

- Section 5 presents guidance to help transit agencies meet the new APC requirements. The guidance covers procedures for testing the statistical equivalence in the average passenger trip length between sample data from APCs and sample data from manual ride checks for the benchmarking plan and the benchmarking study. In addition, the guidance covers procedures for determining the minimum sample size for the annual maintenance study. The guidance also covers procedures for obtaining adjustment factors for errors in APC data for both the benchmarking year and annual maintenance year.

Benefits

The use of the guidance in this document is expected to help transit agencies better meet the requirements for using APC data for NTD reporting, to reduce the reporting burdens on many transit agencies through a more informed process to determine whether they should use APC data for NTD reporting, and to avoid the possibility of estimates obtained from APC data not being included in the Urbanized Area Formula Program apportionment.

This project was conducted by Dr. Xuehao Chu of the University of South Florida. For more information, contact Mr. Daniel Harris, FDOT Transit Planning Project Manager, at (850) 414-4532, daniel.harris@dot.state.fl.us.

TABLE OF CONTENTS

List of Figures

List of Tables

1. INTRODUCTION

This document contains guidance for using data from automatic passenger counters (APC) for reporting the use of transit services to the National Transit Database (NTD). The guidance does not represent new standards on how transit agencies should use their APC data for NTD reporting. Rather, it aims at helping transit agencies meet existing requirements in the 2009 *NTD Reporting Manual.*

To receive Federal Urbanized Area Formula Program grants, transit agencies must meet NTD's data requirements. Specifically, they must report annual data on unlinked passenger trips (UPT) and passenger miles traveled (PMT) to NTD. Unlinked passenger trips are passenger boardings that are counted each time a passenger boards a transit vehicle in revenue service, no matter how many vehicles the passenger uses to travel from the origin to the destination. Passenger miles traveled, on the other hand, measure the total distance traveled by all passengers. If a man and his wife get on a bus at Stop A and get off at Stop B one mile down the road, they have traveled 2 passenger miles.

In addition to these data items, the data requirements of NTD also specify that agencies must meet specific accuracy levels with the reported annual data. If available and reliable, they must report a 100% count. Otherwise, they must obtain the annual data through random sampling that meet 95% confidence and 10% precision levels.

Obtaining annual data on these data items, particularly for fixed-route services, can be costly to agencies. Passenger miles traveled for each one-way vehicle trip are typically calculated as the distance-weighted sum of passenger loads between consecutive stops of passenger activities. This calculation typically requires detailed data on passenger boarding and alighting activities at individual stops and distances between consecutive stops for individual one-way vehicle trips.

Traditionally, transit agencies use manual ride checks to collect the necessary data on boarding and alighting activities. With manual ride checks, ride checkers are sent to ride in transit vehicles while in revenue service to observe and record passenger boardings and alightings at each stop for a random sample of one-way vehicle trips throughout a year. Increasingly, however, transit agencies are seeking to take advantage of APCs that they have installed on their vehicle fleet for internal data purposes. With APCs, passenger activities are determined automatically with counters that count people, analyzers that translate signals from the counters into numerical counts, and computers that record the numerical counts.

In addition to the data requirements, there are additional requirements for using APC data for NTD reporting. APC data are unique in many ways. If no adequate care is taken in using APC

data, these unique characteristics can represent significant challenges for APC data to meet NTD's data requirements. To ensure minimum care being taken with APC data used for NTD reporting by all agencies, the additional APC requirements focus on initial benchmarking during the first year of using APC data for NTD reporting and annual calibrating of APC data thereafter.

This guidebook is designed to help transit agencies meet NTD's data requirements and APC requirements. The rest of the guidebook consists of four sections, covering NTD's requirements, options for meeting NTD requirements, guidance on meeting the data requirements, and guidance on meeting the APC requirements.

2. NTD REQUIREMENTS

The Federal Transit Administration (FTA) has a set of requirements for transit agencies to report data on the use of transit to NTD. These requirements relate to the specific items of transit use, the nature of a reported data item (100% counts versus estimates), the minimum confidence and precision levels for any estimate of a data item, and the additional requirements for using APCs to collect data.

The NTD requirements described below are specified in the Service Module of the *Current Annual Reporting Manual* for 2009. While these requirements have been stable over time, it is possible that they may be modified. Transit agencies should always follow the most current requirements, which are available at http://www.ntdprogram.gov/ntdprogram/.

2.1   Data Requirements

The NTD data requirements can be summarized into three categories: data items, methods of determining these data items, and precision and confidence levels.

2.1.1 Data Items

FTA requires transit agencies to report two items of service-consumed data – unlinked passenger trips (UPT) and passenger miles traveled (PMT). UPT gives the number of passengers who board transit vehicles in revenue service that are counted each time they board a revenue vehicle, no matter how may vehicles they use to travel from their origin to their destination. PMT refers to the total distance traveled by all passengers of transit vehicles in revenue service. For a few cases, such as simple ferry services, PMT can be cumulated across all passengers if the distance traveled by every passenger is known and recorded. Otherwise, PMT for fixed-route services typically is determined with passenger loads between consecutive locations of passenger boarding or alighting weighted by the distance between these locations for individual one-way vehicle trips.

The requirements of reporting these two data items vary by mode and between monthly and annual reporting. For monthly reporting, FTA requires monthly total UPT for all modes. For annual reporting, FTA requires the following:

- For all modes, annual total for both UPT and PMT. A report year for a given agency is defined as its fiscal year.

- For all modes, average daily UPT and PMT by type of service days. There are three types of service days – weekdays, Saturdays, or Sundays. For scheduled services, service days in a report year are classified according to the schedule

operated on that day. If a weekday that is a holiday is served with a Sunday
schedule, that weekday is considered to be a Sunday. For non-scheduled
services, service days are the actual days of a week regardless of whether they
are a holiday or not.

- For commuter rail (CR), heavy rail (HR), and light rail (LR), annual total by
weekday time period for UPT. There are four weekday time periods – Weekday
AM Peak, Weekday Midday, Weekday PM Peak, and Weekday Other. Agencies
define these time periods.

This guidebook focuses on annual reporting. The data in the monthly and annual reports serve
different purposes. It is the annual data that are used in allocating federal transit funds. It also
is the annual data that are required to meet specific requirements, including methods of
determination and statistical criteria on confidence and precision levels.

2.1.2 Methods of Determination

The NTD requirements specify three methods of determining annual data.

100% *Counts*

For UPT, a 100% count involves counting passengers each time they board a transit vehicle in
revenue service. For PMT, it involves recording the distance traveled by all passengers. A
100% count of PMT is typically only possible for systems that have only two stops, for rail
systems that record entry and exit from the system, or for rail systems that rely upon
destination-based tickets. FTA requires the use of 100% counts whenever they are available
and reliable.

Estimation through Random Sampling

Estimation through random sampling involves multiplying a sample ratio by a corresponding
expansion factor for the current report year. A sample ratio is the ratio of the sample total for
one measure of service consumed over the sample total for another measure of service
consumed. For example, the ratio of the sample total for PMT over the sample total for UPT
gives the sample average passenger trip length (APTL). It may be calculated for the entire
annual sample, or by type of service days, or for individual service groups if a sampling plan is
based on service grouping (i.e., stratification).

An expansion factor is a measure of services actually operated or consumed during a given
period. It is used to convert a sample average to the total of service-consumed data during that
period. It varies with sampling plans. The total number of one-way bus trips operated during an
entire report year is an example of an expansion factor; when multiplied by the sample average
PMT per one-way bus trip derived from an annual NTD sample, it yields a measure of annual
total PMT.

Using APTL from Last Sampling Year to Estimate PMT for Current Year

The NTD recognizes two types of a report year. Every year is a mandatory report year for medium and large transit agencies, while every third year is a mandatory report year for small transit agencies. Non-mandatory report years for small agencies are called intermediate report years. The annual *NTD Reporting Manual* gives definitions of mandatory and intermediate years, which are not repeated here.

The third method, referred to below as Recent APTL, is applicable to intermediate years only. It requires that agencies collect and report 100% counts of UPT for the current report year, and determine PMT for the current report year as the product of their 100% UPT from the current report year and the implied APTL from their most recent sampling year. The implied APTL for a given report year can be obtained as the ratio of the annual total PMT reported to NTD over the annual total UPT reported to NTD for that report year. Example 2.1 below illustrates how the implied APTL can be derived from reported data.

Example 2.1 Deriving the Implied APTL for a Report Year

Problem: Spring Hill Transit collects and reports a 100% count of UPT but estimates PMT through random sampling for its directly-operated motorbus service. It is in the process of estimating annual total PMT for FY2010, which is an intermediate year. It plans to estimate the annual total PMT for FY2010 using its implied APTL from FY2008, which was its most recent sampling year. Spring Hill Transit reported 42,133,908 passenger miles traveled and 8,233,005 unlinked passenger trips for FY2008.

Solution: Dividing the reported PMT by the reported UPT for FY2008 gives 5.12 miles. Spring Hill Transit's implied APTL for FY2008 is 5.12 miles.

Criteria for Method Selection

FTA has specific conditions under which transit agencies may use any of these methods. These conditions include the type of a report year and whether 100% counts are available. Table 2.1 summarizes these methods and conditions.

Table 2.1. Selection of Methods

Method	Condition	Type of Report Year
100% Counts	Must use if available and reliable	Any year
Estimation through Random Sampling	Must use if not reporting 100% counts	Mandatory
	May use if not reporting 100% counts	Intermediate
Recent APTL	May use if reporting 100% UPT counts	Intermediate

For mandatory years, transit agencies must report 100% counts if they are available and reliable; otherwise, they must report estimates through random sampling. For intermediate years, transit agencies may also report 100% counts or estimates through random sampling if they choose; otherwise, they must estimate PMT by using the implied APTL from the annual service-consumed data they reported to NTD for their most recent sampling year and the 100% count of UPT from their current report year.

2.1.3 Confidence and Precision

When determining the annual totals of service-consumed data for annual reporting through random sampling, the reported service-consumed data must meet minimum statistical requirements. Specifically, the annual totals of both UPT and PMT must meet the 10% precision level at the 95% confidence level. If an agency were to sample a large number of times according to a sampling plan to estimate PMT for a given report year, these statistical requirements mean that 95% of those estimates of PMT should fall within 10% of the true value.

2.2 APC Requirements

Some transit agencies want to use automatic passenger counters (APC) for collecting UPT and PMT data for motorbus. FTA encourages the use of APC data for NTD reporting, but to ensure the reliability of the UPT and PMT data, FTA requires the following:

- an approved (by FTA) APC benchmarking plan and an APC maintenance plan before an agency starts using APC data for NTD reporting, and
- conduct of benchmarking and maintenance studies and submission of results.

Failure to obtain prior FTA approval may result in APC-derived PMT data from a transit agency not being included in the Urbanized Area Formula Program apportionment. In addition, failure to conduct the benchmarking and maintenance studies and submit their results to FTA may also result in APC-derived data from an agency not being included in the apportionment.

These requirements for the use of APCs in obtaining UPT and PMT for motorbus also are valid for the use of APCs on other modes such as light rail service.

2.2.1 Benchmarking Plan

The benchmarking plan must contain procedures for conducting the benchmarking study during the first year for which an agency wants to use APC data for NTD reporting. These procedures cover three aspects:

1. Selecting a random sample of one-way vehicle trips covering a full year according to a sampling plan. This sampling plan must be designed for the

resulting annual PMT and UPT data from both APCs and ride checkers to meet FTA's minimum 10% precision and 95% confidence levels.
2. Testing the statistical equivalence in the sample APTL between the parallel APC data and manual ride-check data from the same sample.
3. Adjusting the APC data for UPT and PMT to replicate the data produced by the manual ride-check data.

2.2.2 Maintenance Plan

The maintenance plan must include procedures for calibrating APCs every year after the first year of using APC data for NTD reporting. These procedures cover three aspects:

1. Selecting a sample of at least 100 vehicle trips using ride checkers to collect the UPT and PM data. The trips in the sample do not need to be randomly distributed by route, by day, and time of day.
2. Estimating and comparing the UPT and PMT data collected by ride checkers to APC-derived UPT and PMT data and the statistical variance between the two data sets.
3. Regularly checking and reconciling UPT counts from APCs against trip counts and fares from registering fare boxes.

2.2.3 Studies

The benchmarking study is to be conducted during the first year of an agency using APC data for NTD reporting. Agencies must submit to FTA documentation of the results after the benchmarking plan has been implemented. A maintenance study is to be conducted annually after the first year. They must annually submit to the FTA documentation of the results of the maintenance study.

This page intentionally left blank.

3. OPTIONS FOR MEETING NTD REQUIREMENTS

This section describes the conditions under which APC data may not be used for NTD reporting. It also provides guidance for transit agencies to consider several factors in deciding whether it actually wants to use APC data for NTD reporting when it is an option available to them. It finally presents three general options for meeting NTD requirements when APC data are used for NTD reporting.

3.1   APC Data Not to be Used

Using APC data for NTD reporting may not be an option for some agencies with APCs. Three factors determine if using APC data for NTD reporting is an option.

3.1.1   Statistical Equivalency

Agencies must not use their APC data for NTD reporting if the APTL from the APC data is not statistically equivalent to the APTL from the manual ride-check data. The details of testing statistical equivalency are presented in Section 5 on meeting NTD's APC requirements. The equivalency test is conducted as part of the benchmarking study. Transit agencies are not allowed to use APC data for NTD reporting if the APTL from the APC data is statistically different at the 95% confidence level from the APTL from the manual ride-check data.

3.1.2   Adjustment Factors for Data Errors

Errors often exist in both UPT and PMT when obtained from APC data, and errors in UPT may differ from errors in PMT. This document refers to these errors as the adjustment factors for data errors because UPT and PMT data from APCs must be adjusted to account for these errors. Procedures for measuring these adjustment factors are presented in Section 5. Agencies should not choose to use their APC data for reporting estimates of UPT or PMT if the measured adjustment factor for data errors with respect to the corresponding measure is greater than 9%. Agencies should work on improving their APC system to reduce the errors in APC data before considering the use of APC data for NTD reporting again.

3.1.3   Adjustment Factors for Missed Data

Some vehicle trips with an APC may not provide any APC data. In addition, the APC data from some vehicle trips may not be usable for NTD reporting. The adjustment factors for missed data capture both sources of missed data and are stated as a percentage. Specifically, it is the vehicle trips with an APC that did not provide usable APC data as a percent of all vehicle trips with an APC. Agencies with a 100% penetration of APCs on their fleet should not choose to use their APC data for reporting a 100% count of UPT or PMT if the measured adjustment factor for missed data with respect to the corresponding measure is more than 10%. Agencies should work on improving their APC system to reduce the degree of missed data before using their

APC data for reporting 100% counts. Alternatively, agencies can consider using their APC data to obtain estimates if the relevant adjustment factor for data errors is smaller than 9%.

3.2   Factors to Consider

When agencies are allowed to use APC data for NTD reporting, they should carefully consider two factors before making the decision. One relates to their ability to automatically post-process raw APC data without manually checking them for individual one-way vehicle trips. The other relates to the sample size requirement for annual maintenance studies when APC data are used versus the sample size requirement for obtaining sample data when manual ride checks are used.

3.2.1 Post-Processing of Raw APC Data

Agencies should think carefully about whether they want to use APC data for NTD reporting if they do not yet have the ability to post-process the large amount of raw APC data automatically without manually checking the APC data for individual one-way vehicle trips. This is the case regardless of the rate of APC penetration, whether planning to report 100% counts or estimates, or whether planning to use all usable APC data or to use a subset of all APC data through random sampling. The amount of raw APC data to be processed would be significantly smaller for the case of random sampling than for the case of using all usable APC data. Because of the typical use of vehicle blocks as sampling units and the need to account for both missed data and data errors in a sampling plan for APCs, the amount of raw APC data to be post-processed can still be substantially higher if an agency uses random sampling than what needs to be processed if it uses manual ride checks.

3.2.2 Effective Minimum Sample Size for Annual Maintenance Study

For some agencies that choose to obtain estimates through random sampling, they may want to consider using ride-check data as a better alternative to using APC data, at least for NTD reporting. The following are the essential steps they should consider:

1. Understanding the alternative.
2. Determining the sample size for this alternative.
3. Determining the effective sample size for the annual maintenance study.
4. Making the decision.
5. Having the sampling plan certified for this alternative.

   Understanding the Alternative

This alternative relates to FTA's requirements of collecting manual ride-check data in addition to collecting APC data from the same sample of at least 100 one-way vehicle trips during each maintenance year for which APC data are used for NTD reporting.

Agencies may want to consider this alternative if the minimum sample size for using manual ride-check data is smaller than the effective sample size for using APC data. If using APC data, the effective sample size would typically be more than 100 one-way vehicle trips, because APC data may not be recovered for some trips and recovered APC data may not be usable for additional trips. If usable APC data are available from only one half of the one-way vehicle trips actually operated with an APC, for example, the effective sample size would be 200 rather than 100. If obtaining estimates through random sampling with manual data collection alone requires fewer than the effective sample size for using APC data, using APC data would not make sense. As a result, agencies may choose to use sample data from manual ride checks to estimate PMT or both UPT and PMT.

This alternative is likely to be available to agencies that report a 100% count of UPT but an estimate of PMT by multiplying this 100% count of UPT with an estimated APTL from a random sample. Experience indicates that the minimum sample size required to meet FTA's confidence and precision levels is significantly lower for estimating PMT on the basis of APTL than otherwise. Experience also indicates that the minimum sample size required to meet FTA's confidence and precision levels is often smaller than 200, 150, or even 100 one-way vehicle trips for many agencies when PMT is estimated on the basis of APTL.

It is important to point out that this alternative would not be available to agencies for a particular measure of service consumed if they have a reliable 100% count for this measure. These agencies must report this 100% count for that measure rather than its estimates.

Determining the Sample Size for the Alternative

To determine if using manual ride-check data is potentially a better alternative, agencies need to determine the sample size for manual ride checks that would be required to meet FTA's confidence and precision levels if they choose this alternative.

Agencies should determine the sample size for the alternative through considering a range of sampling techniques appropriate for their circumstances. Experience indicates that the minimum sample size for manual ride checks required for meeting FTA's confidence and precision levels varies significantly across different sampling techniques for a given situation. For detailed guidance for considering alternative sampling techniques, agencies are referred to Section 50, "Template Sampling Plans," in the *National Transit Database Sampling Manual* at http://www.ntdprogram.gov/ntdprogram/sampling.htm.

Agencies may determine the initial sample size for their service in any way they choose. While agencies must have the sampling plan certified by a qualified statistician if they actually use that sampling plan for NTD reporting, they do not need to have the initial sample size certified at this point; they can hire a qualified statistician to accomplish that later at a monetary price. A more reasonable approach would be for agencies to use a ready-to-use Excel template. As part of

the *National Transit Database Sampling Manual*, such a template is available at
http://www.ntdprogram.gov/ntdprogram/misc/The_NTD_Sampling_Template_FINAL.xlsm.
Agencies are referred to Subsection 54 of the above document for guidance for using this
template.

Determining the Percent of Vehicle Trips without Usable APC Data

The percent of vehicle trips with APCs but without usable APC data should be based on all
vehicle trips that were actually operated during a period. When this percentage should be
measured depends on when an agency wants to make a decision about this alternative:

- If it wants a decision before they need to develop and submit the benchmarking and
  maintenance plans, it should measure the percentage before the benchmarking and
  maintenance plans need to be submitted for FTA's approval in case it wants to start
  using APC data for NTD reporting during the benchmarking year. In this case, the
  period for measuring the percentage may be shorter than a full year.

- If it wants a decision after the benchmarking year, it should measure the percentage
  using data collected during the benchmarking year. In this case, the percentage will be
  referred to as the adjustment factor for missed data. It should be measured for the full
  year. But the data for measuring this percentage will need to be collected anyway for
  the benchmarking year.

In either case, several steps are involved to determine the percentage of vehicle trips with APCs
but without usable APC data for PMT during a given period. The unit of tracking and calculation
is in terms of one-way vehicle trips rather than individual APCs.

Step 1. Keep track of the status of every one-way vehicle trip that had an APC and was
actually operated during that period:

- Whether any APC data were recovered from it
- Whether the recovered APC data were usable for obtaining PMT for it

Step 2. Count the number of one-way vehicle trips for each status:

- n_1 = one-way vehicle trips for which no APC data were recovered
- n_3 = one-way vehicle trips for which APC data were recovered but were not usable
  for PMT

Step 3. Count the total number of one-way vehicle trips that were equipped with APCs and
were actually operated and denote it as n.

Step 4. Calculate the percentage as follows:

$$f_M = 100 \cdot \left[\frac{n_1 + n_3}{n} \right]$$ (3.1)

Example 3.1 below illustrates the calculations involved.

Example 3.1 Computing % Vehicle Trips with APCs but without Usable APC Data

Problem: Terrace Transit directly operates fixed-route motorbus service. It has partial penetration of APCs on its vehicle fleet.   The APC vehicles made a total of 150,000 one-way vehicle trips during FY2008, its benchmarking year. It tracked the status of APCs for all of these vehicle trips and obtained the following counts:

# trips for which no APC data were recovered	= 52,500
# trips for which APC data were recovered but not usable for PMT	= 10,500

Solution: The above information means that n = 150,000, n_1 = 52,500, and n_3 = 10,500. Plugging these values into equation (3.1) results in the following percentage for FY2008:

$$f_M = 100 \cdot \left[\frac{52,500 + 10,500}{150.000} \right] = 42\%$$

Determining the Effective Sample Size for Annual Maintenance

Agencies can use the percent of vehicle trips with APCs but without usable APC data to determine the effective sample size. Specifically, the effective sample size for meeting NTD's annual maintenance requirements would be given by the integer rounded up from equation (3.2):

$$S_e = \frac{100}{1 - f_M}$$ (3.2)

Example 3.2 on the following page illustrates the use of this formula.

Example 3.2 Determining the Effective Sample Size for Annual Maintenance

Problem: Terrace Transit directly operates fixed-route motorbus service. It has partial penetration of APCs on its vehicle fleet. The APC vehicles operated a total of 150,000 one-way vehicle trips during FY2008, its benchmarking year. It has determined during the benchmarking year that 42% of these one-way vehicle trips did not provide usable APC data for PMT. To make a decision, it wants to know the effective sample size for meeting NTD's annual maintenance requirements for the first maintenance year.

Solution: The above information means $f_M = 42\%$. Plugging this value as a fraction into the formula in equation (3.2) gives

$$S_e = \frac{100}{1 - 0.42} = 172.4$$

The effective sample size in this case would be 173 one-way vehicle trips, significantly greater than 100.

Making the Decision

If the sample size for manual ride checks is smaller than the effective minimum sample size required by FTA's APC requirements, agencies may want to consider the savings and costs in determining whether they want to use the alternative.

One obvious saving is the smaller number of one-way vehicle trips from which passenger boarding and alighting activities would need to be collected every year. The size of this saving is determined by the difference in the sample size between manual ride checks and APC data collection. Another saving is that agencies no longer need to conduct the analyses required by the annual APC maintenance plan.

Choosing this alternative, however, also involves some cost that would not occur otherwise. One such cost is the need to select the sample for manual ride checks randomly according to the corresponding sampling plan. In contrast, the current NTD rules do not require the at least 100 one-way vehicle trips to be selected at random. This cost results not only from the action of sample selection but also from the added logistical and staffing difficulties of collecting data from one-way vehicles that are selected at random. This cost, however, will likely to be much smaller than the savings from using manual ride-check data if the effective sample size for annual maintenance is greater than the minimum sample size for manual ride checks.

Using the alternative for one year does not mean that an agency has to use it forever. When conditions change, both internally within an agency or externally, agencies may want to reevaluate this alternative. As an example of an internal change, the APC system may grow from a partial fleet to a full fleet, and a reliable 100% count may become available. In addition,

the APC system may become more reliable, resulting in a much smaller percentage of vehicle trips with APCs but without usable data for PMT, which, in turn results in a much smaller effective sample size for the annual maintenance study. Externally, NTD rules may change.

Having the Sampling Plan for the Alternative Certified

If the decision is to use the alternative, the sampling plan for manual ride checks must be certified to meet NTD's 10% precision at the 95% confidence level by a qualified statistician. If an agency has been using this sampling plan before it started considering using APC data for NTD reporting, it would have already been certified. Otherwise, it would need to be certified as a new sampling plan.

3.3　Using APC Data

Three general options are available to transit agencies for annual reporting of UPT and PMT. These options depend on whether reliable 100% counts are available for either measure of service-consumed data.

3.3.1　Option 1 – Reporting Estimates for Both UPT and PMT

This option represents the circumstances where reliable 100% counts are unavailable for either UPT or PMT from APCs or any other source.

Agencies must report estimates for both UPT and PMT with this option. They obtain these estimates by multiplying average UPT and PMT per service unit by the total number of service units actually operated in a year. A service unit is an amount of revenue travel by a single transit vehicle, a set of transit vehicles (e.g., a train), or a component of a transit vehicle (e.g., the upper deck of a bi-level passenger car). For non-scheduled services, it is typically one vehicle day. For scheduled bus services, it is typically either a one-way bus run or a round-trip bus run. For rail services, it is a one-way car run, a one-way train run, a round-trip car run, or a round-trip train run.

Agencies have a number of options to get the averages of UPT and PMT from their APC data. These options depend on the percent of their fleet with an APC (100% or under 100%), whether they choose to use all of their APC data or use some of the APC data through random sampling, and whether they choose to sample before (pre-sampling) or after data collection (post-sampling) from APCs. The relationship tree among these factors in Figure 3.1 shows these options.

For agencies that consider obtaining these averages through pre-sampling, they should determine whether they are willing and able to assign their transit vehicles to any service units by random sampling. The option of random sampling requires APC data on boarding and alighting activities at individual stops or stations from service units to be selected at random

according to a sampling plan that is designed to meet FTA's confidence and precision levels. They should not use pre-sampling if they are not willing or not able to assign their transit vehicles to any service units at random throughout a report year.

As long as APC vehicles are circulated in a manner that covers the entire schedule regularly throughout a year, obtaining estimates through using all APC data is clearly superior to obtaining estimates through random sampling. Using all APC data is feasible, however, only if procedures for processing APC data have been developed to a point where the APC data from each one-way vehicle trip no longer need to be manually checked. Otherwise, agencies should choose to obtain estimates through random sampling. One can argue that an agency that does not yet have the capability to automatically check and process all of its APC data probably should not be using APC data for NTD reporting purposes.

Figure 3.1. Options for Using APC Data to Obtain Estimates
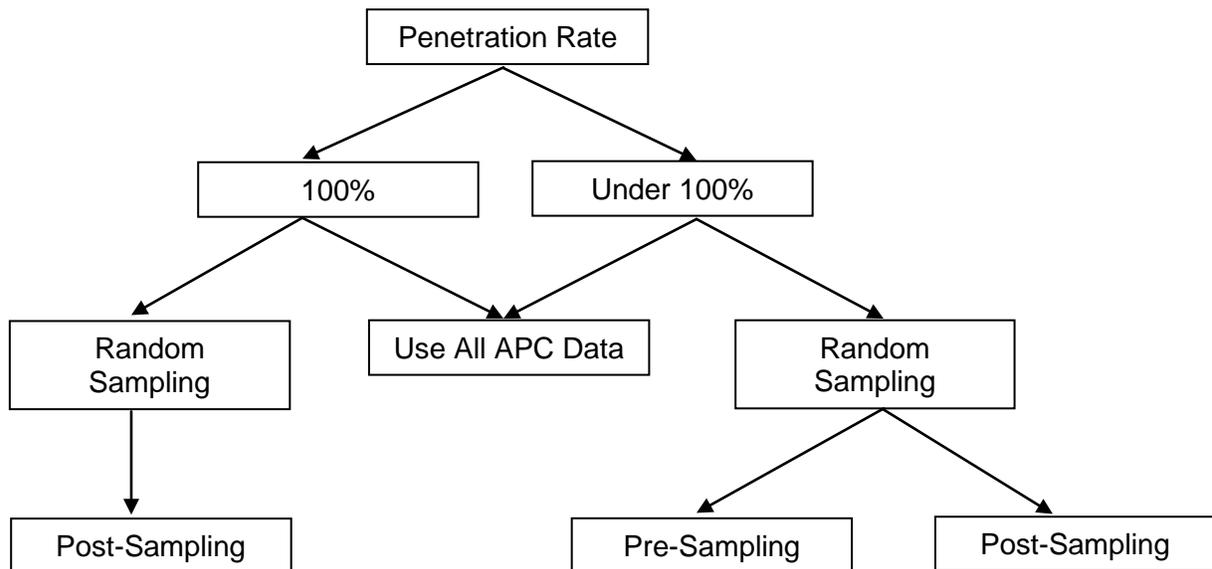
This option is typically used by agencies for their new light rail service because the fare system is open and does not track the number of boardings. It also is used by many agencies for their bus system because the fare boxes did not yet give a reliable 100% count of boardings when they started using APC data for NTD reporting.

3.3.2 Option 2 – Reporting 100% Counts for UPT but Estimates for PMT

This option represents the circumstances where a reliable 100% count of UPT is available but a reliable 100% count of PMT is unavailable. Agencies must report the 100% count of UPT but an estimate of PMT with this option. Agencies may obtain the 100% count of UPT from their APCs or a non-APC source such as fare boxes. Agencies obtain the estimate PMT by multiplying the 100% count of UPT with an estimate of the APTL from their APC data.

Similar to Option 1, agencies may obtain and use their APC data to estimate the APTL in different ways, using all APC data, pre-sampling, or post-sampling.

This option is widely used among agencies that use APC data for NTD reporting. When it is used, however, the 100% count of UPT typically comes from a non-APC source.

3.3.3 Option 3 – Reporting 100% Counts for both UPT and PMT

This option represents the circumstances where reliable 100% counts are available for both UPT and PMT. Agencies must report a 100% count for both UPT and PMT. They obtain these 100% counts from their APCs. This option is rarely used.

This page intentionally left blank.

4. MEETING THE DATA REQUIREMENTS

The guidance in this section assumes that FTA has approved an agency to use its APC data for NTD reporting and that the agency has decided to do so. The guidance is organized into the following sub-sections:

- Obtaining Adjustment Factors
- Obtaining 100% Counts from APCs
- Obtaining Estimates through Using All APC Data
- Obtaining Estimates through Random Sampling
- Certifying Collection and Estimation Plans
- Intermediate Years

Most of the guidance presented here focuses on using APC data for NTD reporting for mandatory years. The last sub-section briefly describes procedures for intermediate years.

Agencies may obtain 100% counts of UPT from a non-APC source such as fare boxes. The guidance for obtaining 100% counts of UPT in this document focuses on obtaining such counts from APCs. However, changes in procedures will be noted if a 100% count of UPT from a non-APC source is used.

4.1 Obtaining Adjustment Factors

4.1.1 Types of Adjustment

There are two general types of adjustments for obtaining UPT or PMT from APC data:

- Adjusting for missed data. Some vehicle trips with an APC may not provide any APC data. In addition, the APC data from some vehicle trips may not be usable for NTD reporting. The adjustment factors for missed data are stated as a percentage. Specifically, it is the vehicle trips with an APC that did not provide usable APC data as a percent of all vehicle trips with an APC. The adjustment factor for missed data is denoted as $f_{M,UPT}$ for UPT and as $f_{M,PMT}$ for PMT. The need to adjust for missed data in obtaining 100% counts is one of NTD's data requirements.

- Adjusting for data errors. Errors often exist in both UPT and PMT when obtained from APC data, and errors in UPT may differ from errors in PMT. The adjustment factor for errors is denoted as $f_{E,UPT}$ for UPT and as $f_{E,PMT}$ for PMT. The need to adjust for data errors in obtaining estimates from APC data is one of NTD's APC requirements.

An adjustment factor for missed data is conceptually similar to the percentage of vehicle trips with APCs but without usable APC data described in Section 3 for determining the effective sample size for annual maintenance studies. They both measure the degree of missed data. However, they differ in several ways:

- Purpose – The percentage of vehicle trips with APCs but without usable APC data is used to determine, before the benchmarking year, the effective sample size for annual maintenance studies. In contrast, the adjustment factors are used determine, after the benchmarking year, the effective sample size for annual maintenance studies. In addition, the adjustment factors are used to adjust direct 100% counts if a 100% count is to be reported.

- Penetration rate – The percentage of vehicle trips with APCs but without usable APC data is used for all levels of APC penetration. For obtaining 100% counts, the adjustment factors are measured only when the APC penetration is 100%. For determining the effective sample size for annual maintenance studies after the benchmarking year, however, the adjustment factors also are used for all levels of APC penetration.

- Frequency – The percentage of vehicle trips with APCs but without usable APC data is measured only when an agency wants to consider whether it should use APCs or manual ride checks for data collection before the benchmarking year. The adjustment factors, on the other hand, must be obtained every year during which such a factor is used to obtain a 100% count of UPT or PMT from APC data for NTD reporting. Even if a 100% count from APC data is not to be reported to the NTD, they still need to be obtained every year for determining the effective sample size for the maintenance study next year.

- Duration – The percentage of vehicle trips with APCs but without usable APC data may be measured for a period shorter than a year. In contrast, the adjustment factors for missed data should be based on all vehicle trips that were actually operated during an entire report year.

4.1.2 Potential Values

The adjustment factors for missed data, $f_{M,UPT}$ and $f_{M,PMT}$, should never be negative. In most cases, $f_{M,UPT}$ would not exceed $f_{M,PMT}$ because some vehicle trips may provide usable data on UPT but not on PMT.

The adjustment factor for data errors in UPT, $f_{E,UPT}$, will typically be negative. Available evidence has shown that UPT data from APCs tend to understate the full UPT. The adjustment

factor for data errors in PMT, $f_{E,PMT}$, can be either positive or negative – positive if PMT from the APC data are greater than the true value, but negative if PMT from the APC data are smaller than the true value. Both adjustment factors for data errors would be zero if the APC data are error-free.

If the value for any of these adjustment factors is zero, no adjustment would be made to the unadjusted value. If the value for an adjustment factor for data errors is positive, the adjustment would factor down the unadjusted value. If the value for an adjustment factor for data errors is negative, the adjustment would factor up the unadjusted value.

4.1.3   Obtaining Adjustment Factors for Missed Data

Several steps are involved to determine the adjustment factors for missed data during a report year. The basic steps are the same as those for obtaining the percentage of vehicle trips with APCs but without usable APC data on PMT. The unit of tracking and calculation is in terms of one-way vehicle trips rather than individual APCs.

Step 1. Track the status of every one-way vehicle trip that was equipped with an APC and was actually operated during that year:

- Whether any APC data were recovered from it
- Whether the recovered APC data were usable for obtaining UPT for it
- Whether the recovered APC data were usable for obtaining PMT for it

Step 2. Count the number of one-way vehicle trips for each of these statuses:

- N_1 = one-way vehicle trips for which no APC data were recovered
- N_2 = one-way vehicle trips for which APC data were recovered but were not usable for UPT
- N_3 = one-way vehicle trips for which APC data were recovered but were not usable for PMT

Step 3. Count the total number of one-way vehicle trips that were equipped with APCs and were actually operated and denote it as N.

Step 4. Calculate the adjustment factors as follows:

$$f_{M,UPT} = 100 * \left[\frac{N_1 + N_2}{N} \right] \tag{4.1}$$

$$f_{M,PMT} = 100 * \left[\frac{N_1 + N_3}{N} \right] \tag{4.2}$$

Example 4.1 illustrates the calculations involved.

Example 4.1 Computing Adjustment Factors for Missed Data

Problem: Beach Transit directly operates fixed-route motorbus service. It has a 100% penetration rate of APCs on its vehicle fleet and it wants to report a 100% count for both UPT and PMT. It operated a total of 350,000 one-way vehicle trips during FY2008. It tracked the status of APCs for all of these vehicle trips and obtained the following counts:

# trips for which no APC data were recovered	= 20,000
# trips for which APC data were recovered but not usable for UPT	= 14,000
# trips for which APC data were recovered but not usable for PMT	= 17,500

Solution: The above information means that $N = 350,000$, $N_1 = 20,000$, $N_2 = 14,000$, and $N_3 = 17,500$. Plugging these values into equations (4.1) and (4.2) results in the following adjustment factors for missed data for FY2008:

$$f_{M,UPT} = 100 * \left[\frac{20,000 + 14,000}{350,000}\right] = 9.7\%$$

$$f_{M,PMT} = 100 * \left[\frac{20,000 + 17,500}{350,000}\right] = 10.7\%$$

4.1.4 Obtaining Adjustment Factors for Data Errors

The adjustment factors for data errors need to be obtained every year during which APC data are used to obtain estimates of UPT or PMT for NTD reporting:

- For the benchmarking year, they need to be derived as part of the benchmarking study and they are used for obtaining estimates for the benchmarking year as an NTD report year.
- For each maintenance year thereafter, they need to be derived as part of the maintenance study, and they are used for obtaining estimates for that maintenance year as an NTD report year.

The guidance on how these adjustment factors should be derived as part of the benchmarking and maintenance studies is discussed in Section 5 on meeting NTD's APC requirements.

4.1.5 Usage

The primary usage of these adjustment factors is to adjust UPT and PMT for missed data and data errors:

- Obtaining estimates of UPT, PMT, or both would need to consider both types of adjustment factors.
- Obtaining 100% counts of UPT, PMT, or both would need to consider only the adjustment factor or missed data.

The adjustment factors for missed data also are used to determine the effective minimum sample size for the parallel sample of one-way vehicle trips for which both APC data and manual ride-check data are collected for the annual maintenance study:

- For the first maintenance year after the benchmarking year, the adjustment factor for missed data should be obtained during the benchmarking year.
- For any other maintenance year, the adjustment for missed data should be obtained during the previous maintenance year.

The adjustment factors for data errors also are used in developing the plan to collect APC data for obtaining estimates of UPT or PMT. Details on how these adjustment factors affect data-collection plans are presented in later sub-sections on obtaining estimates through using all APC data or through random sampling.

4.2   Obtaining 100% Counts from APCs

To report a 100% count of UPT or PMT, agencies must determine that a 100% count is available and reliable. The annual *NTD Reporting Manual* (2009 version) does not provide criteria to determine the availability or reliability of a 100% count. The guidance reflects our best judgment on best practices that would meet FTA requirements on reporting 100% counts.

4.2.1   Evaluating Availability

Evaluating the availability of a 100% count involves three elements that are incremental in nature:

1. The first element is intuitive and includes two dimensions:
    - The count must have 100% coverage of all units of revenue service actually operated. This requires working APCs on every revenue vehicle.
    - The count must have 100% coverage of all passengers. This requires working APCs at every door used for revenue service.

2. Having APCs on every vehicle of revenue service and on every door is not enough, however. The second element deals with counting procedures. There must be an established and tested procedure that is designed to correctly count and record every boarding and alighting.

3. Even with 100% coverage of service units and vehicle doors and an established procedure, counts frequently are not available from a large portion of the service units actually operated. The third element deals with these missed counts. The adjustment factor for missed data must not exceed 10% of the total actually operated. That is, $f_{M,UPT}$ ≤ 10% if for UPT and $f_{M,PMT}$ ≤ 10% if for PMT.

UPT and PMT must be evaluated separated for the availability of a 100% count with all three elements. Under these evaluation elements, there are three possible outcomes:

1. A 100% count is unavailable for either UPT or PMT
2. A 100% count is available for UPT but not for PMT
3. A 100% count is available for both UPT and PMT

If a 100% count is unavailable for UPT, or PMT, or both UPT and PMT, there is no need to evaluate reliability. In these cases, agencies will need to obtain estimates for the corresponding data either through using all APC data or using a sample of their APC data.

4.2.2   Evaluating Reliability

Evaluating the reliability of an available 100% count involves three components – the direct count from the service units with usable APC data, the adjustment factors for missed data and for data errors, and the adjustment of the direct count with the adjustment factors.

Direct Count

The direct count is considered reliable if the following three conditions are satisfied:

1. The agency must have obtained approval from FTA for the benchmarking plan and the maintenance plan.
2. The sample APTL has been proven to be statistically equivalent between APC data and manual ride-check data during the benchmarking study.
3. The random error in the APC data must be within 10% at the 95% confidence level.

Adjustment Factors

The adjustment factors for missed data must be obtained through tracking the APC and data status of all one-way vehicle trips actually operated during an entire year. They must not exceed 10%. The adjustment factors for data errors must be derived during the benchmarking study or the annual maintenance study, as described in the previous section. They should not exceed 9%.

Adjustments

The direct count must be adjusted using the adjustment factors. The following formula shows how this adjustment should be done in the case of obtaining a 100% count of UPT:

$$\text{Adjusted 100\% UPT} = \frac{1}{1 - f_{M,UPT}} \left[\frac{\text{Unadjusted 100\% UPT}}{1 + f_{E,UPT}} \right] \qquad (4.3)$$

where:

$f_{M,UPT}$ = Adjustment factor for missed data on UPT.
$f_{E,UPT}$ = Adjustment factor for data errors in UPT.

Please note that the adjustment factors in equation (4.3) must be entered as fractions rather than as percentages. The equation should have shown $(f_{M,UPT})/100$ and $(f_{E,UPT})/100$. To avoid clutter in the equation, however, $f_{M,UPT}$ and $f_{E,UPT}$ are shown instead. The same is true with other equations later where these adjustment factors are used.

Example 4.2 illustrates the use of this equation for adjusting direct counts.

Example 4.2 Adjustments for 100% Counts

Problem: Paradise Transit directly operates fixed-route motorbus service in the city of Paradise. It has a 100% penetration rate of APCs on its vehicle fleet. Based on the criteria described above, it has determined that a reliable 100% count is available from its APCs for UPT but not for PMT. However, it needs to adjust the direct count of UPT both for missed data and for data errors in the direct count. The following is known:

Direct count of UPT   =   5,000,000
Adjustment factor for missed data   =   +9.5%
Adjustment factor for data errors   =   -7.0%

Solution: The above information means that Unadjusted 100% UPT = 5,000,000; $f_{M,UPT}$ = 0.095; and $f_{E,UPT}$ = -0.07. Using these values in equation (4.3) results in the following:

$$\frac{1}{1 - 0.095}\left[\frac{5,000,000}{1 + (-0.07)}\right] = \frac{1}{0.905}\frac{5,000,000}{0.93} = \frac{5,377,344}{0.905} = 5,940,712$$

Paradise Transit should report this final 100% count of 5,940,712 as its annual UPT.

4.3   Obtaining Estimates through Using All APC Data

If reliable 100% counts are not available from APCs, agencies may choose to obtain estimates either through using all APC data or through using a random sample of all APC data. This sub-section focuses on using all APC data. Guidance is provided on the proper procedures for obtaining estimates through using all APC data in three related elements:

- Managing APCs
- Stratifying One-Way Vehicle Trips with Usable APC data
- Obtaining Estimates

4.3.1 Managing APCs

The objective of properly managing APCs is to avoid biases in final estimates. When all APC data are used, biases in the final estimates may result from improperly managing APCs in two ways:

1. APC data are not available for major segments of the service (i.e., the coverage problem).
2. APC data are available for too many one-way vehicle trips for certain segments of the service but for too few one-way vehicle trips for other segments (i.e., the sampling rate problem).

When all APC data are used to obtain estimates, getting proper coverage of all service units is more critical than getting proper sampling rates across all segments of the service. When a service segment is covered but with an improper sampling rate during data collection, any potential bias in the final estimates can largely be avoided during the estimation process through post stratification. When a service segment is not covered during data collection, on the other hand, there is no second chance to avoid potential biases during estimation.

The procedures for properly managing APCs differ between a fleet with 100% penetration and a fleet with partial penetration. The discussion for these two cases assumes that agencies do not have particular policies or practices that prevent APC-equipped vehicles from being used on certain segments of the service. The third case discusses procedures for agencies that do have such policies or practices.

100% APC Penetration

The problems of inadequate coverage and improper sampling rates can still exist for agencies with 100% APC penetration. These problems result from missed data, i.e., APC data not being recovered from some one-way vehicle trips and from recovered APC data not being usable for NTD reporting.

The key is to constantly monitor the status of APCs on individual one-way vehicle trips actually operated throughout a report year. This monitoring activity could be carried out simultaneously with the tracking activity for obtaining the adjustment factors for missed data. As described in the early section on obtaining adjustment factors, agencies should track the following three APC and data status to obtain the adjustment factor for missed data:

- Whether any APC data were recovered from it
- Whether the recovered APC data were usable for obtaining UPT for it
- Whether the recovered APC data were usable for obtaining PMT for it

Once these statuses are identified for each one-way vehicle trip operated, the agency must identify the reasons for these statuses and take immediate corrective action.

One way to identify systematic problems is to monitor patterns of one-way vehicles trips in terms of their APC status for each combination of individual routes and time periods, including Saturdays, Sundays, and morning peak, midday, afternoon peak, and other for weekdays. Table 4.1 shows how these patterns may be summarized and compared with all vehicle trips actually operated with an APC.

Table 4.1. Monitoring Cumulative Vehicle Trips with 100% APC Penetration

Status	Time Period	Routes			
		1	2	......	n
Actually operated vehicle trips	Morning Peak	$N^A_{1, MP}$	$N^A_{2, MP}$		$N^A_{n, MP}$
	Midday	$N^A_{1, MD}$	$N^A_{2, MD}$		$N^A_{n, MD}$
	Afternoon Peak	$N^A_{1, AP}$	$N^A_{2, AP}$		$N^A_{n, AP}$
	Other	$N^A_{1, OT}$	$N^A_{2, OT}$		$N^A_{n, OT}$
	Saturday	$N^A_{1, SA}$	$N^A_{2, SA}$		$N^A_{n, SA}$
	Sunday	$N^A_{1, SU}$	$N^A_{2, SU}$		$N^A_{n, SU}$
Vehicle trips without any APC data recovered	Morning Peak	$N^R_{1, MP}$	$N^R_{2, MP}$		$N^R_{n, MP}$
	Midday	$N^R_{1, MD}$	$N^R_{2, MD}$		$N^R_{n, MD}$
	Afternoon Peak	$N^R_{1, AP}$	$N^R_{2, AP}$		$N^R_{n, AP}$
	Other	$N^R_{1, OT}$	$N^R_{2, OT}$		$N^R_{n, OT}$
	Saturday	$N^R_{1, SA}$	$N^R_{2, SA}$		$N^R_{n, SA}$
	Sunday	$N^R_{1, SU}$	$N^R_{2, SU}$		$N^R_{n, SU}$
Vehicle trips with APC data recovered but with no usable UPT data	Morning Peak	$N^U_{1, MP}$	$N^U_{2, MP}$		$N^U_{n, MP}$
	Midday	$N^U_{1, MD}$	$N^U_{2, MD}$		$N^U_{n, MD}$
	Afternoon Peak	$N^U_{1, AP}$	$N^U_{2, AP}$		$N^U_{n, AP}$
	Other	$N^U_{1, OT}$	$N^U_{2, OT}$		$N^U_{n, OT}$
	Saturday	$N^U_{1, SA}$	$N^U_{2, SA}$		$N^U_{n, SA}$
	Sunday	$N^U_{1, SU}$	$N^U_{2, SU}$		$N^U_{n, SU}$
Vehicle trips with APC data recovered but with no usable PMT data	Morning Peak	$N^P_{1, MP}$	$N^P_{2, MP}$		$N^P_{n, MP}$
	Midday	$N^P_{1, MD}$	$N^P_{2, MD}$		$N^P_{n, MD}$
	Afternoon Peak	$N^P_{1, AP}$	$N^P_{2, AP}$		$N^P_{n, AP}$
	Other	$N^P_{1, OT}$	$N^P_{2, OT}$		$N^P_{n, OT}$
	Saturday	$N^P_{1, SA}$	$N^P_{2, SA}$		$N^P_{n, SA}$
	Sunday	$N^P_{1, SU}$	$N^P_{2, SU}$		$N^P_{n, SU}$

Partial APC Penetration

When an agency has APCs only on part of its fleet, the problems of inadequate coverage and improper sampling rates can result not only from missed data, but also from some vehicles not being equipped with APCs. The issues and strategies for dealing with data recovery and data usability are the same as those discussed previously for the case of 100% APC penetration. This section focuses on the issue of some vehicles not being equipped with APCs.

Priorities. Unlike the case of 100% APC penetration, partial penetration potentially could create conflicts between the internal need for APC data within an agency and the external need for NTD reporting. These conflicts exist for at least two reasons:

- A far greater sample size is often required for meeting internal data needs. One example of internal needs is scheduling adherence and service quality monitoring in terms of extreme values of passenger load. These extreme values reflect the impacts of load variability and service regularity as well as frequency and better reflect the quality of service as felt by passengers. On the other hand, meeting NTD data needs requires averages, and the FTA confidence and precision levels are applied only to estimated averages for a whole year and an entire system. Estimating extreme values requires a far greater sample size than estimating averages.

- Monitoring extreme values often means the concentration of APC-equipped vehicles on heavy-demand routes, heavy-demand periods, etc. at the expense of low-demand routes, low-demand periods, etc. Such concentration of APC-equipped vehicles may mean missed coverage of certain routes, time periods, etc., or inadequate sampling rates when covered. Meeting NTD data needs, on the other hand, prefers proper coverage and adequate sampling rates for all routes, time periods, etc.

If the rate of partial APC penetration is large enough and APC-equipped vehicles are properly rotated, both types of data needs can be met adequately and conflicts are resolved. Otherwise, agencies must carefully consider their priorities:

- If they decide that meeting their internal data needs is more important after careful consideration, they should not use APC data for NTD reporting and instead focus on the internal data needs.
- If they decide that meeting the NTD data needs is more important, on the other hand, they must resolve these conflicts of data needs by putting their first priority on meeting NTD data needs.

Meeting the NTD data needs requires proper coverage and adequate sample size, and adequate sample size, in turn, depends on the three general options presented in Section 3 for meeting NTD's data requirements and the adjustment factor for data errors in PMT.

Proper coverage means that APC-equipped vehicles are rotated in a manner that covers the entire schedule regularly throughout a year. There are two basic approaches to accomplishing this proper coverage:

- Covering every timetable trip at least once every year, and
- Covering most timetable trips at least once every year.

In both cases, all routes and all time periods must be covered adequately. In the case of covering every timetable trip at least once, adequate sample size is in terms of the total number of one-way vehicle trips to be covered. In the case of covering most timetable trips, adequate sample size is in terms of the fraction of all timetable trips to be covered. The determination of adequate sample size is discussed separately for these two cases.

Covering Every Timetable Trip. Determining adequate sample size for this case involves five basic steps. Based on Figure 4.1, the first three steps are used to determine adequate sample size for weekdays. The last two steps are used for Saturdays and Sundays, respectively, if service is provided on each day type.

1. Determine the total number of timetable trips in the weekday schedule.
2. Use the adjustment factor for data errors in PMT from an agency's benchmarking study to determine another number from Figure 4.1. Use part *a* if the factor ranges from 0% to 8% and use part *b* if the factor is between 8% and 9%.
3. Determine the sample size for weekdays by comparing the two numbers from Steps 1 and 2 and picking the larger of these two numbers.
4. Determine the number of timetable trips for Saturdays; this is the sample size for Saturdays if service is provided.
5. Determine the number of timetable trips for Sundays; this is the sample size for Sundays if service is provided.

Example 4.3 illustrates these steps with an example.

Figure 4.1. Sample Size Determination Based on Adjustment Factor of Data Errors

a. Range of Adjustment Factor of Data Errors in PMT [0%, 8.0%]
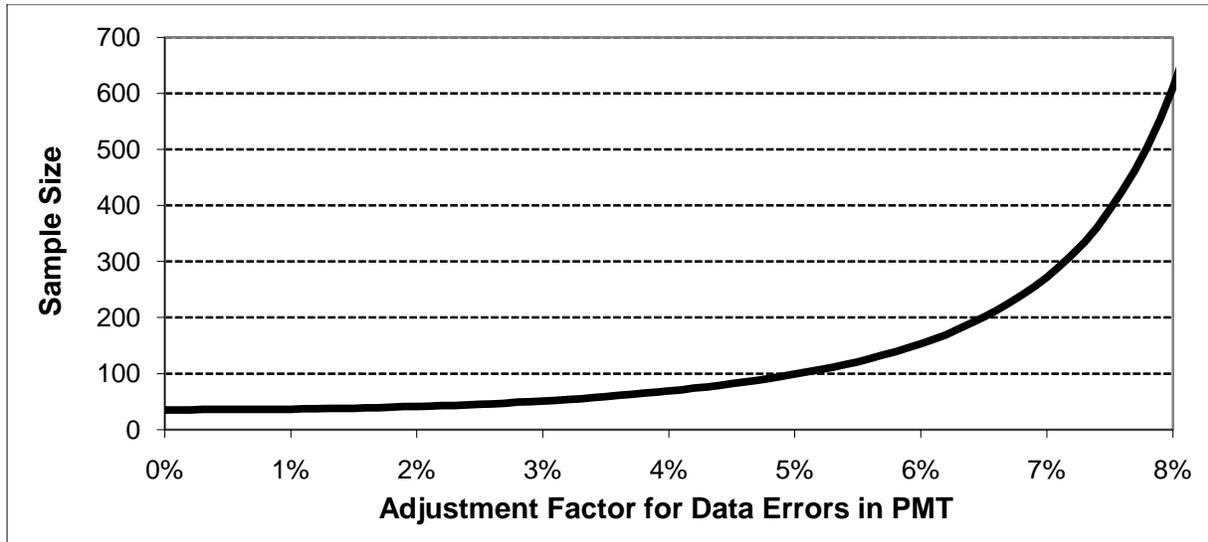
b. Range of Adjustment Factor for Data Errors in PMT (8%, 9%]
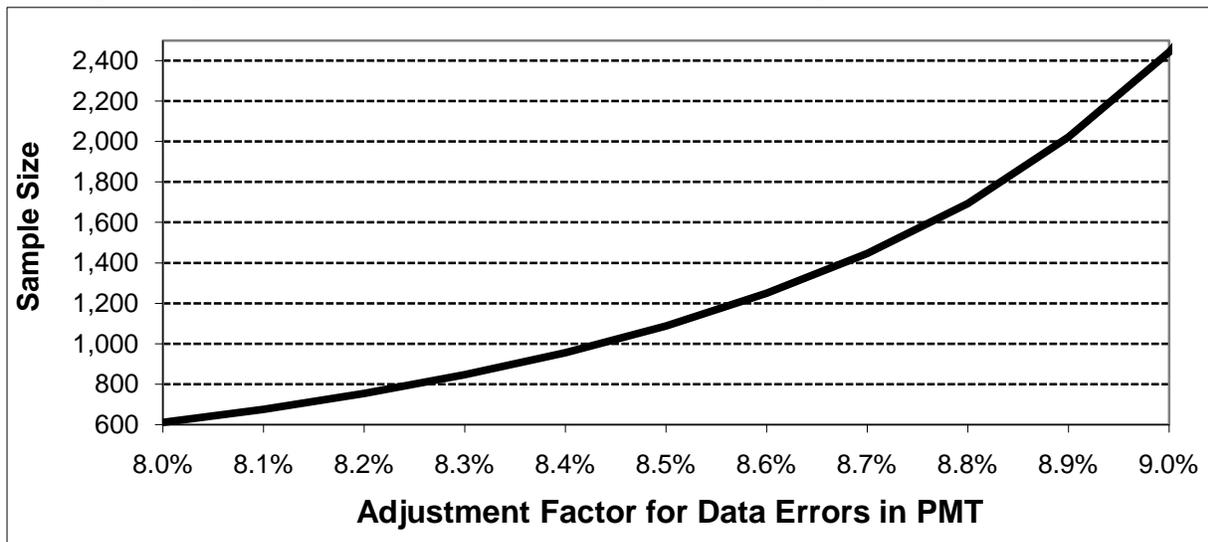
Covering Most Timetable Trips. Logistics and data recovery problems can frustrate plans to observe every timetable trip. An alternative plan would be to cover a fraction of all timetable trips for weekdays, Saturdays, and Sundays. The following provides guidance on how this fraction should be determined for weekdays. The same fraction is used for Saturdays and Sundays. Agencies may use Figure 4.2 to determine the appropriate fraction for their conditions in terms of three factors:

- the effective penetration rate of APCs on their fleet
- the number of timetable trips for weekdays
- the adjustment factor for data errors in their PMT

The effective penetration rate of APCs is the penetration rate of APCs in terms of installation multiplied by $f_{M,PMT}$ or the adjustment factor for missed data in PMT.

Example 4.3 Determining Sample Size when Covering All Timetable Trips

Problem: Wonderland Transit directly operates a 30-route fixed-route motorbus system in the city of Wonderland, with a total of 10,000 timetable trips on the weekday schedule, 1,500 timetable trips on the Saturday schedule, and 1,000 timetable trips on the Sunday schedule. It has a 40% penetration rate of APCs on its vehicle fleet. It has submitted its benchmarking and annual maintenance plans to NTD, and FTA has approved these plans. In addition, it has conducted its benchmarking study during the first year of NTD reporting. This benchmarking study shows an 8.5% error in the annual PMT obtained from its APC data. It wants to use all of its APC data for NTD reporting and wants to cover all timetable trips at least once during a year. It needs to determine adequate sample size for weekdays, Saturdays, and Sundays.

Solution: We solve this problem by following the five steps individually.

Step 1: 10,000.
Step 2: Locate 8.5% in part b of Figure 4.1 on the horizontal axis. Determine the corresponding value on the vertical axis to be around 1,100.
Step 3: Compare the weekday timetable trips from Step 1 to the number identified from Step 2. The larger of these two is 10,000. This is the weekday sample size.
Step 4: The Saturday schedule has 1,500 timetable trips; this is the Saturday sample size.
Step 5: The Sunday schedule has 1,000 timetable trips; this is the Sunday sample size.

These numbers mean that it would be adequate for Wonderland Transit to cover every timetable trip for weekdays, Saturdays, and Sundays at least once a year.

Figure 4.2 contains six graphs that correspond to six different levels of the adjustment factor for data errors in PMT – 4%, 5%, 6%, 7%, 8%, and 9%, respectively. The particular level of errors is labeled at the bottom of each graph. Within each graph, the horizontal axis measures the total number of timetable trips on weekdays, and the vertical axis shows the fraction of the total number of timetable trips to be selected. Each graph has four lines that correspond to four different levels of the effective penetration rate of APCs – 1%, 2%, 4%, and 100%. To determine the appropriate fraction, agencies should use Figure 4.2 in the following steps:

Figure 4.2. Selecting Fraction of Timetable Trips for Weekdays
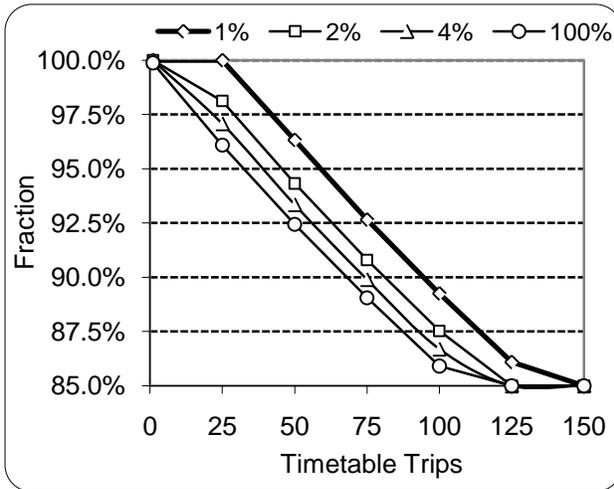
A. Adjustment Factor = 4%
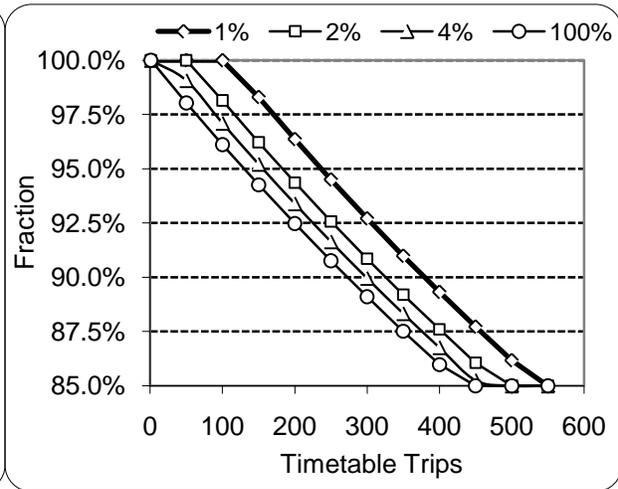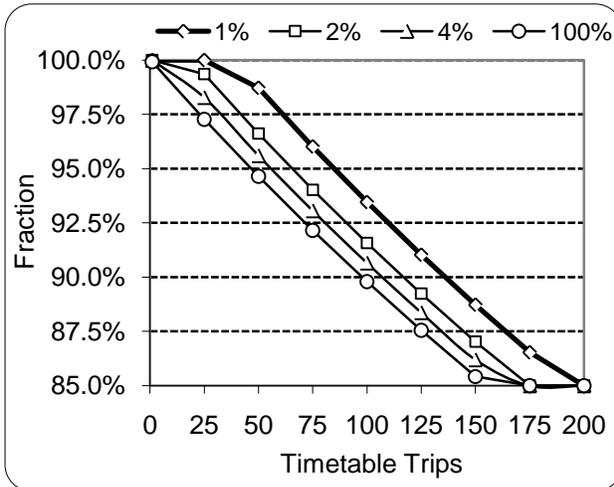
B. Adjustment Factor = 5%

C. Adjustment Factor = 6%

D. Adjustment Factor = 7%
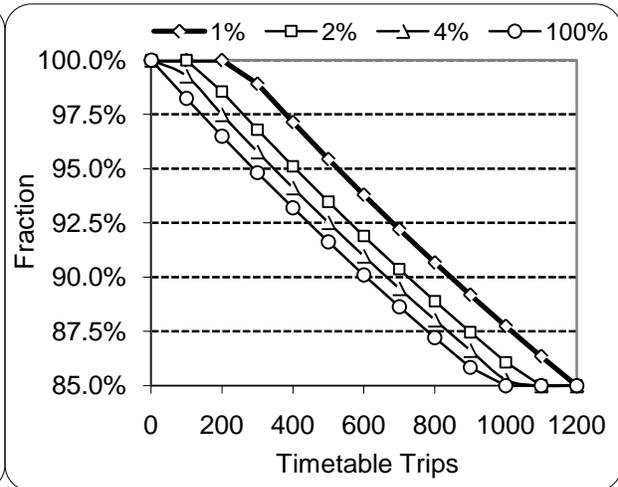
E. Adjustment Factor = 8%
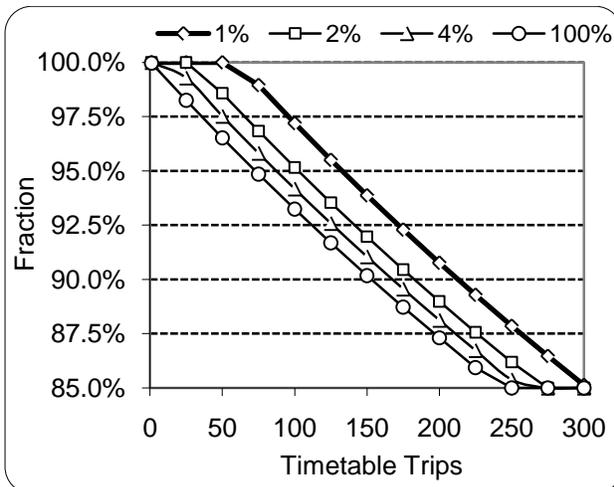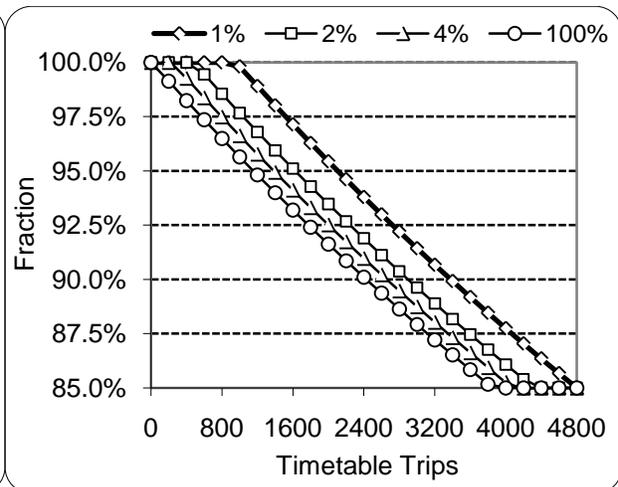
F. Adjustment Factor = 9%

1. Select 85% as the fraction if the adjustment factor for data errors in PMT is less than 4%.

2. Pick the appropriate graph for their level of the adjustment factor for data errors in PMT. Round up to the next graph if their adjustment factor for data errors in PMT is between any of the explicit levels in Figure 4.2. For example if the adjustment factor is 4.4%, use the graph for 5% rather than 4%.

3. Pick the appropriate line within the picked graph for their effective penetration rate. If the effective penetration rate happens to be 1%, 2%, 4%, or 100%, use the corresponding line. If the effective penetration rate falls between two of these levels, use the line corresponding to the lower level. For example, if the effective penetration rate is 3%, use the 2% line rather than the 4% line. Similarly, use the 4% line if the effective penetration rate is above 4% but below 100%.

4. Determine the fraction along the selected line using the total number of timetable trips on weekdays. Use 85% if the total number of timetable trips is greater than the maximum used in the graph.

Unique Service Segments

APC data may not be collected at all for certain unique service segments (e.g., certain routes) even when all vehicles or most of the vehicles for other service segments are equipped with APCs. This occurs because these unique segments use vehicles that are unique to them. One example is the case where one or more bus rapid transit (BRT) lines use only their uniquely-branded vehicles. For historical reasons, APCs are not installed on these BRT vehicles.

When such unique service segments exist, passenger boarding and alighting data at individual stops must be collected with a method of data collection other than APCs. The most reasonable approach would be to collect passenger activity data with manual ride checks from a sample of one-way vehicle trips that are selected at random according to a data collection plan that covers these unique service segments and all other service segments.

4.3.2   Stratifying One-Way Vehicle Trips with Usable APC Data

Before obtaining estimates with all usable APC data for any of the above three cases, agencies should analyze all one-way vehicle trips with usable APC data. The objective of this analysis is to determine the coverage and sampling rates of all one-way vehicle trips with usable APC data. The information from this analysis would be the basis to determine the best post-stratification strategy for estimation.

Agencies need to determine the best stratification strategy after the APC data have been processed because coverage and sampling rates in the final usable data may not always match

those intended during the data collection process. This mismatch can occur even with an agency's best efforts in constant monitoring and making corrective actions throughout a year.

Mismatching can occur in both directions. With partial APC penetration and an APC management plan of covering every timetable trip at least once in a year, some timetable trips may not have any usable APC data at the end. In that case, the stratification strategy would not be by individual timetable trips as originally intended with the APC management plan. Rather, it likely would be by individual routes and time periods (including Saturdays, Sundays, and morning peak, midday, afternoon peak, and other for weekdays). On the other hand, every timetable trip may have usable APC data at the end, even though the APC management plan calls for covering every route but only a fraction of all timetable trips. In this case, the best stratification strategy would be at the level of individual timetable trips rather than at the route level as originally intended by the APC management plan.

If unique service segments exist and APC data are not collected from them by design, the best stratification strategy must include these segments as one or more strata.

4.3.3   Obtaining Estimates

Once the best stratification strategy is determined with all usable APC data, obtaining estimates is straightforward. While the details vary by whether a 100% count of UPT is to be reported to NTD, the key to both cases is the following:

1. Estimation is done individually for each stratum first, and the sum of the estimates across all strata gives the annual total estimates.
2. Estimation must be done separately for each stratum as part of any unique service segments, and the results must be aggregated with those from the other strata to get the annual total estimates.
3. Adjustments are done to account for missed data.
4. Adjustments are done to account for data errors.

The following procedures apply to cases with 100% penetration of APCs as well as to cases with partial penetration of APCs. For ease of presentation, some of the procedures are presented for the case where the stratification is by route, day type, and time period. In addition, the procedures are presented for the case where there are no special service segments where APCs are excluded by design.

Not Reporting 100% UPT

When an agency does not report a 100% count of UPT to NTD, it estimates both UPT and PMT. In this case, the estimates can be obtained in six steps:

1. Compute average UPT and PMT per one-way vehicle trip from all usable APC data for each stratum.
2. Determine the annual number of one-way vehicle trips actually operated for each stratum. This step includes all vehicle trips with or without APCs.
3. Compute estimates of annual UPT and PMT for each stratum by multiplying the average UPT and PMT per one-way vehicle trip from Step 1 with the annual number of one-way vehicle trips actually operated from Step 2.
4. Sum the results across all relevant strata for each day type to get the unadjusted estimates of annual UPT and PMT by day type.
5. Divide the unadjusted estimates of annual total UPT and PMT by day type by the number of days for each day type to get unadjusted average daily.
6. Adjust the unadjusted estimates of annual total and average daily of UPT and PMT by the adjustment factors of data errors - $f_{E,UPT}$ and $f_{E,PMT}$.

Table 4.2 illustrates these steps for the case of estimating both average daily and annual total UPT with stratification by route and time period.

Reporting 100% UPT

When an agency does report a 100% count of UPT to NTD, it estimates PMT only. In this case, the PMT is estimated in six steps:

1. Compute the APTL from all usable APC data for each stratum.
2. Determine the 100% count of UPT for each stratum.
3. Compute estimates of annual PMT for each stratum by multiplying the APTL from Step 1 with the 100% count of UPT from Step 2.
4. Sum the results across all strata to get the unadjusted estimate of annual total PMT.
5. Divide the unadjusted estimates of annual total PMT by day type by the number of days for each day type to get unadjusted average daily PMT.
6. Adjust the unadjusted estimates by the adjustment factors for missed data in UPT (i.e., $f_{M,UPT}$) and by the adjustment factor for data errors in PMT (i.e., $f_{E,PMT}$) to get adjusted estimates of annual total and average daily PMT.

Table 4.3 illustrates these steps for the case of stratification by route and time period and with the 100% count of UPT from APC data. If the 100% count of UPT comes from a non-APC source, the adjustment for missed data may be determined differently but would still need to be made.

If the 100% count of UPT is from APCs, it will need to be adjusted as well both for missed data ($f_{M,UPT}$) and for data errors ($f_{E,UPT}$). Specifically, the adjusted 100% count of UPT can be determined from the unadjusted 100% count of UPT in equation (4.3).

Table 4.2. Estimating Annual UPT with Stratification by Routes and Time Periods

Steps	Time Periods	Routes			
		1	2	...	n
1. Average UPT per one-way vehicle trip from usable APC data	Morning Peak	$UPT_{1, MP}$	$UPT_{2, MP}$		$UPT_{n, MP}$
	Midday	$UPT_{1, MD}$	$UPT_{2, MD}$		$UPT_{n, MD}$
	Afternoon Peak	$UPT_{1, AP}$	$UPT_{2, AP}$		$UPT_{n, AP}$
	Other	$UPT_{1, OT}$	$UPT_{2, OT}$		$UPT_{n, OT}$
	Saturday	$UPT_{1, SA}$	$UPT_{2, SA}$		$UPT_{n, SA}$
	Sunday	$UPT_{1, SU}$	$UPT_{2, SU}$		$UPT_{n, SU}$
2. Annual number of one-way vehicle trips actually operated	Morning Peak	$N_{1, MP}$	$N_{2, MP}$		$N_{n, MP}$
	Midday	$N_{1, MD}$	$N_{2, MD}$		$N_{n, MD}$
	Afternoon Peak	$N_{1, AP}$	$N_{2, AP}$		$N_{n, AP}$
	Other	$N_{1, OT}$	$N_{2, OT}$		$N_{n, OT}$
	Saturday	$N_{1, SA}$	$N_{2, SA}$		$N_{n, SA}$
	Sunday	$N_{1, SU}$	$N_{2, SU}$		$N_{n, SU}$
3. Estimates of annual UPT by stratum	Morning Peak	$UPT_{1, AM} * N_{1, AM}$	$UPT_{2, AM} * N_{2, AM}$		$UPT_{n, AM} * N_{n, AM}$
	Midday	$UPT_{1, MD} * N_{1, MD}$	$UPT_{2, MD} * N_{2, MD}$		$UPT_{n, MD} * N_{n, MD}$
	Afternoon Peak	$UPT_{1, PM} * N_{1, PM}$	$UPT_{2, PM} * N_{2, PM}$		$UPT_{n, PM} * N_{n, PM}$
	Other	$UPT_{1, OT} * N_{1, OT}$	$UPT_{2, OT} * N_{2, OT}$		$UPT_{n, OT} * N_{n, OT}$
	Saturday	$UPT_{1, SA} * N_{1, SA}$	$UPT_{2, SA} * N_{2, SA}$		$UPT_{n, SA} * N_{n, SA}$
	Sunday	$UPT_{1, SU} * N_{1, SU}$	$UPT_{2, SU} * N_{2, SU}$		$UPT_{n, SU} * N_{n, SU}$
4. Unadjusted estimates of total UPT	Weekdays	Sum across all weekday cells in Step 3 to get Unadjusted UPT_{WK}			
	Saturdays	Sum across all Saturday cells in Step 3 to get Unadjusted UPT_{SA}			
	Sundays	Sum across all Sunday cells in Step 3 to get Unadjusted UPT_{SU}			
	Annual Total	Sum across all cells in Step 3 to get Unadjusted UPT_{AN}			
5. Unadjusted estimates of average daily UPT	Weekdays	Unadjusted Average UPT_{WK} = Unadjusted UPT_{WK} / Weekdays			
	Saturdays	Unadjusted Average UPT_{SA} = Unadjusted UPT_{SA} / Saturdays			
	Sundays	Unadjusted Average UPT_{SU} = Unadjusted UPT_{SU} / Sundays			
6. Adjusted estimates of UPT	Weekday Average	$\text{Adjusted Average UPT}_{WK} = \dfrac{\text{Unadjusted Average UPT}_{WK}}{\left(1 + f_{E,UPT}\right)}$			
	Saturday Average	$\text{Adjusted Average UPT}_{SA} = \dfrac{\text{Unadjusted Average UPT}_{SA}}{\left(1 + f_{E,UPT}\right)}$			
	Sunday Average	$\text{Adjusted Average UPT}_{SU} = \dfrac{\text{Unadjusted Average UPT}_{SU}}{\left(1 + f_{E,UPT}\right)}$			
	Annual Total	$\text{Adjusted UPT}_{AN} = \dfrac{\text{Unadjusted UPT}_{AN}}{\left(1 + f_{E,UPT}\right)}$			

4.4   Obtaining Estimates through Random Sampling

Obtaining estimates through random sampling differs in several ways between using APC data and using manual ride-check data. This section considers several unique factors in developing a sampling plan for data collection from APCs. In addition, it also shows the essential steps for obtaining estimates from APC data collected through random sampling.

Table 4.3. Estimating Annual PMT with Stratification by Routes and Time Periods

Steps	Time Period	Routes			
		1	2	...	n
1. Average passenger trip length from usable APC data by stratum	Morning Peak	$APTL_{1,MP}$	$APTL_{2,MP}$		$APTL_{n,MP}$
	Midday	$APTL_{1,MD}$	$APTL_{2,MD}$		$APTL_{n,MD}$
	Afternoon Peak	$APTL_{1,AP}$	$APTL_{2,AP}$		$APTL_{n,AP}$
	Other	$APTL_{1,OT}$	$APTL_{2,OT}$		$APTL_{n,OT}$
	Saturday	$APTL_{1,SA}$	$APTL_{2,SA}$		$APTL_{n,SA}$
	Sunday	$APTL_{1,SU}$	$APTL_{2,SU}$		$APTL_{n,SU}$
2. Annual 100% UPT by stratum	Morning Peak	$UPT_{1,MP}$	$UPT_{2,MP}$		$UPT_{n,MP}$
	Midday	$UPT_{1,MD}$	$UPT_{2,MD}$		$UPT_{n,MD}$
	Afternoon Peak	$UPT_{1,AP}$	$UPT_{2,AP}$		$UPT_{n,AP}$
	Other	$UPT_{1,OT}$	$UPT_{2,OT}$		$UPT_{n,OT}$
	Saturday	$UPT_{1,SA}$	$UPT_{2,SA}$		$UPT_{n,SA}$
	Sunday	$UPT_{1,SU}$	$UPT_{2,SU}$		$UPT_{n,SU}$
3. Estimates of annual PMT by stratum	Morning Peak	$APTL_{1,AM} * UPT_{1,AM}$	$APTL_{2,AM} * UPT_{2,AM}$		$APTL_{n,AM} * UPT_{n,AM}$
	Midday	$APTL_{1,MD} * UPT_{1,MD}$	$APTL_{2,MD} * UPT_{2,MD}$		$APTL_{n,MD} * UPT_{n,MD}$
	Afternoon Peak	$APTL_{1,PM} * UPT_{1,PM}$	$APTL_{2,PM} * UPT_{2,PM}$		$APTL_{n,PM} * UPT_{n,PM}$
	Other	$APTL_{1,OT} * UPT_{1,OT}$	$APTL_{2,OT} * UPT_{2,OT}$		$APTL_{n,OT} * UPT_{n,OT}$
	Saturday	$APTL_{1,SA} * UPT_{1,SA}$	$APTL_{2,SA} * UPT_{2,SA}$		$APTL_{n,SA} * UPT_{n,SA}$
	Sunday	$APTL_{1,SU} * UPT_{1,SU}$	$APTL_{2,SU} * UPT_{2,SU}$		$APTL_{n,SU} * UPT_{n,SU}$
4. Unadjusted estimates of total PMT	Weekdays	Sum across all weekday cells in Step 3 to get Unadjusted PMT_{WK}			
	Saturdays	Sum across all Saturday cells in Step 3 to get Unadjusted PMT_{SA}			
	Sundays	Sum across all Sunday cells in Step 3 to get Unadjusted PMT_{SU}			
	Annual Total	Sum across all cells in Step 3 to get Unadjusted PMT_{AN}			
5. Unadjusted estimates of average daily	Weekdays	Unadjusted Average UPT_{WK} = Unadjusted PMT_{WK} / Weekdays			
	Saturdays	Unadjusted Average UPT_{SA} = Unadjusted PMT_{SA} / Saturdays			
	Sundays	Unadjusted Average UPT_{SU} = Unadjusted PMT_{SU} / Sundays			
6. Adjusted estimates of PMT	Weekday Average	$\text{Adjusted Average PMT}_{WK} = \dfrac{\text{Unadjusted Average PMT}_{WK}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$			
	Saturday Average	$\text{Adjusted Average PMT}_{SA} = \dfrac{\text{Unadjusted Average PMT}_{SA}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$			
	Sunday Average	$\text{Adjusted Average PMT}_{SU} = \dfrac{\text{Unadjusted Average PMT}_{SU}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$			
	Annual Total	$\text{Adjusted PMT}_{AN} = \dfrac{\text{Unadjusted PMT}_{AN}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$			

4.4.1   Unique Considerations in Developing a Sampling Plan

The basics of developing a sampling plan for APCs are the same as those of developing a sampling plan for manual ride checks. One needs to analyze sample data collected with the respective method to determine the statistical variability of the parameters relevant for a specific estimation method. The relevant parameter would be APTL, for example, if a 100% count of UPT is collected and reported. One then needs to determine the minimum sample size based

on the statistical variability. There are several factors, however, that are unique to developing a sampling plan for APCs.

Timing of Sampling

Whether random sampling is practical also depends on the timing of sampling done. When sample data are collected through manual ride checks, agencies must draw a random sample before sending people out for ride checks. With APCs, on the other hand, agencies may draw random samples either before APC data are collected (pre-sampling) or after APC data have already been collected and archived (post-sampling).

- Pre-sampling for collecting data from APCs is similar to sampling for collecting data from manual ride checks. Both involve developing a sampling plan, selecting a sample of vehicle trips, sending ride checkers to sampled vehicle trips or assigning APC-equipped vehicles to sampled vehicle trips, processing the collected sample data, and obtaining estimates from the sample data.

- Post-sampling for getting data from archived APC data is similar to post-sampling for getting data from archived video records of passenger boarding and alighting activities for individual vehicle trips. Both involve developing a sampling plan, selecting a sample of vehicle trips from the archive, processing the archived video for the selected vehicle trips, and obtaining estimates from the sample data.

Unit of Sampling

For either approach, random sampling still may be carried out differently in terms of the unit of sampling used. The unit of sampling using manual ride checks typically is in one-way vehicle trips for bus service. The unit of pre-sampling using APCs typically is in vehicle blocks. One has some flexibility in the unit of sampling for post-sampling with APCs. The most practical unit for post-sampling still would be one-way vehicle trips, especially APC data require manual checking.

Data Errors in PMT

Agencies must use a precision level that is higher than 10% if there are errors in estimated PMT from APC data. Figure 4.3 shows how the higher precision level on the vertical axis relates to any particular degree of errors in estimated PMT on the horizontal axis measured by the adjustment factor for data errors in PMT. The minimum precision is 10% when the adjustment factor is 0%, increases with increases in the adjustment factor, and is 1.2% when the adjustment factor reaches 9%. Agencies should not use APC data for NTD reporting if the adjustment factor for data errors in PMT is greater than 9%.

Figure 4.3. Errors in Unadjusted PMT and Minimum Precision
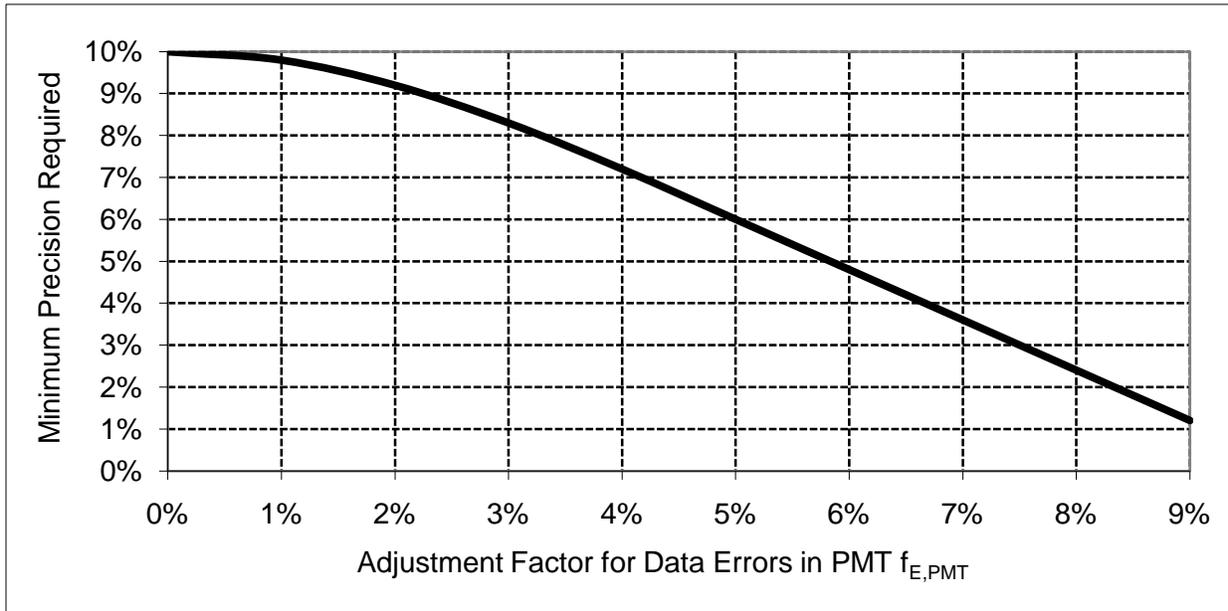
If the adjustment factor for data errors in PMT for an agency is large and the agency does not yet have the capability to automatically process its APC data, it should carefully consider whether it wants to obtain estimates through random sampling. The sample size would be significantly larger for large errors in estimated PMT. Specifically, Table 4.4 shows the ratio of initial sample sizes between the case with a particular positive value for the adjustment factor and the case with a zero value for the adjustment factor. For example, the initial sample size with a 5% adjustment factor would be 2.78 times as large as the initial sample size with a zero adjustment factor. The ratio jumps to 17.36 if the adjustment factor for data errors in PMT is 8%.

Table 4.4. Potential Impact of Errors in PMT on Initial Sample Size

Errors in PMT	Minimum Sample Size Relative to 0% Error in PMT
0%	1
1%	1.04
2%	1.18
3%	1.45
4%	1.93
5%	2.78
6%	4.34
7%	7.72
8%	17.36
9%	69.44

Adjustment for Missed Data

After an initial sample size is determined taking into account both the statistical variability and any errors in estimated PMT, agencies must also adjust it to account for vehicle trips without any usable APC data:

$$\text{Adjusted Final Sample Size} = \frac{\text{Unadjusted Initial Sample Size}}{\left(1 - f_{M,PMT}\right)} \tag{4.4}$$

Consider a case where the unadjusted initial sample size is 300 and $f_{M,PMT}$ = 20%. The adjusted final sample size in this case would be 375, which is 25% higher than the unadjusted initial size.

4.4.2 Obtaining Estimates

The general considerations and procedures for obtaining estimates through random sampling are the same as those for obtaining estimates using all APC data. The general considerations are whether a 100% count of UPT is to be reported, whether stratification is used in sampling and in estimation, and the need to adjust for missed data and data errors. The procedures of estimating annual totals are illustrated in Tables 4.5 and 4.6.

Table 4.5 illustrates the steps for estimating annual PMT when a 100% count of UPT is available and to be reported to NTD for the case with two strata:

1. Compute the APTL from all usable APC data for each stratum.
2. Determine the 100% count of UPT for each stratum.
3. Compute unadjusted estimates of annual PMT for each stratum by multiplying the APTL from Step 1 with the 100% count of UPT from Step 2.
4. Sum the results across all strata to get the unadjusted estimate of annual total PMT.
5. Adjust the unadjusted estimates by the adjustment factor for missed data in UPT (i.e., $f_{M,UPT}$) and by the adjustment factor for data errors in PMT (i.e., $f_{E,PMT}$) to get adjusted estimates of annual total and average daily PMT.

Table 4.5. Estimating Annual PMT Using a 100% Count of UPT

Steps	*Not Stratified*	*Stratified*		
		Stratum 1	*Stratum 2*	*Total*
1. Sample APTL	APTL	APTL$_1$	APTL$_2$	
2. Unadjusted 100%UPT	UPT	UPT$_1$	UPT$_2$	
3. Unadjusted	APTL * UPT	APTL$_1$ * UPT$_1$	APTL$_2$ * UPT$_2$	**4**. Unadjusted PMT
5. Adjusted	$\dfrac{\text{Unadjusted PMT}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$			$\dfrac{\text{Unadjusted PMT}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,PMT}\right)}$
Note: Adjustments need to be made to the unadjusted 100% UPT: $\dfrac{\text{Unadjusted 100\% UPT}}{\left(1 - f_{M,UPT}\right)\left(1 + f_{E,UPT}\right)}$				

Table 4.6 illustrates the steps of estimating annual PMT without a 100% count of UPT for a case with two strata:

1. Compute average UPT and PMT per one-way vehicle trip from all usable APC data for each stratum.
2. Determine the annual number of one-way vehicle trips actually operated for each stratum with or without an APC.
3. Compute estimates of annual UPT and PMT for each stratum by multiplying the average UPT and PMT per one-way vehicle trip from Step 1 with the annual number of one-way vehicle trips actually operated from Step 2.
4. Sum the results across all strata to get the unadjusted estimate of annual total PMT.
5. Adjust the unadjusted estimates of annual total PMT by the adjustment factor for data errors in PMT - $f_{E, PMT}$.

Table 4.6. Estimating Annual PMT without a 100% Count of UPT

Steps	Not Stratified	Stratified		
		Stratum 1	Stratum 2	Total
1. Sample Averages	pmt	pmt_1	pmt_2	
2. Vehicle Trips	V	V_1	V_2	
3. Unadjusted	V * pmt	$V_1 * pmt_1$	$V_2 * pmt_2$	**4**. Unadjusted PMT
5. Adjusted	$\dfrac{V * pmt}{(1 + f_{E,PMT})}$			$\dfrac{\text{Unadjusted PMT}}{(1 + f_{E,PMT})}$
Note: The same steps may be used to estimate annual UPT.				

4.5   Certification of Data-Collection and Estimation Plans

Four elements for data collection and estimation require the certification by a qualified statistician if any is used in the procedures of an agency. These elements are as follows:

1. Adjustments for missed data – If any adjustment factor for missed data, $f_{M,UPT}$ or $f_{M,PMT}$, is greater than 2%, the agency must have a qualified statistician approve the methodology for determining the adjustment factor and factoring the data to account for the missed data.

2. Random errors in 100% counts – A 100% count of UPT, PMT, or both, is obtained from APC data. A qualified statistician must certify that the random errors in such a 100% count meet FTA's 10% precision at the 95% confidence level. In addition to systematic errors, APC data also contain random errors. Again, there is an important difference in random errors between data from APCs and data from manual ride checks. With manual ride checks, less than 100% sampling is assumed to be the only source of random errors. With APCs, however, random errors can also result from 100% counts.

3. Plans for using all APC data – Using all APC data to obtain estimates is not based on random sampling. But agencies still must have its plan to use all APC data certified by a qualified statistician to meet FTA's 10% precision at the 95% confidence level.

4. Plans for using a sample of APC data from random sampling – Just like using a sample of manual ride-check data from random sampling, an agency must have its plan for sampling and using APC data for obtaining estimates certified by a qualified statistician to meet FTA's 10% precision at the 95% confidence level.

Depending on how agencies plan to obtain UPT and PMT, agencies must describe how they plan to obtain certification for each of these four elements. Obtaining certification for these elements is essential to ensure that APC data for UPT and PMT meet FTA standards for at least 95% confidence and 10% precision.

4.6 Intermediate Years

The data and APC requirements apply to both mandatory years for all agencies and intermediate years for small agencies. For an intermediate year, an agency is allowed to multiply its 100% count of UPT with the implied APTL from its most recent mandatory year if that 100% count of UPT is obtained from a source other than APCs. The same agency still can use the implied APTL from its most recent mandatory year, but the 100% count of UPT must be adjusted both for missed data and for data errors if it is obtained from APCs:

- If the adjustment factor for missed data in UPT = $f_{M,UPT}$ and the adjustment factor for data errors in UPT = $f_{E,UPT}$, the adjusted 100% count of UPT can be determined from the unadjusted 100% count of UPT as in equation (4.3).
- The adjustment factors must be obtained from a benchmarking study if the intermediate year happens to be the benchmarking year or from a maintenance study if it happens to be a maintenance year.
- It is the adjusted 100% UPT that is used along with the implied APTL from the most recent sampling year to obtain the annual total PMT.

5. MEETING THE APC REQUIREMENTS

Meeting NTD's APC requirements involves developing and submitting the benchmarking and maintenance plans before the first year of using APC data for NTD reporting and conducting the benchmarking study in the first year and the maintenance study annually thereafter.

5.1   The Plans

The benchmarking plan and the maintenance plan share several components. To avoid redundancy in the presentation of the guidance, the various components from both plans are summarized in Table 5.1 and are presented individually below. Seven components are covered, with three being common across the plans and two being unique to each plan.

Table 5.1. Components of the Plans

Components	*Benchmarking Plan*	*Maintenance Plan*
Option for Estimation and Reporting	√	√
Sampling Plan	√	
Trip Selection Plan		√
Equivalency Testing	√	
Adjustment Factors	√	√
Adjusting APC Data	√	√
Cross Checks		√

5.1.1   Option for Estimation and Reporting

Agencies should describe their methods for obtaining UPT and PMT and how APC data are used in these methods. They should at least describe which of the three general options described in Section 3 for obtaining annual totals of UPT and PMT will be used.

5.1.2   Sampling Plan

Agencies must detail their plan for collecting the data on passenger activities at individual stops with both APCs and manual ride checks from a common annual random sample of one-way vehicle trips for the benchmarking year. This plan must be based on pre-sampling and must take into account the fact that APC data would not be recovered at all for some one-way vehicle trips and that recovered APC data would not be usable for determining UPT or PMT for some one-way vehicle trips. In addition, agencies must confirm that the plan has been certified by a qualified statistician and that annual total UPT and PMT obtained from the sample data will meet FTA's 10% precision at the 95% confidence level for both APCs and manual ride checks.

5.1.3   Trip Selection Plan

The trip selection plan for annual maintenance studies consists of the determination of the effective sample size and the selection of individual vehicle trips.

Agencies must use the percentage of vehicle trips with APCs but without usable APC data on PMT to determine the effective sample size for the annual maintenance study. For the first maintenance year after the benchmarking year, the percentage should be obtained during the benchmarking year. For other maintenance years, the percentage should be obtained during the previous maintenance year. The procedures for obtaining this percentage and for determining the effective sample size during a particular year are described in Section 3 and are not repeated here.

The trips in the annual maintenance sample do not need to be randomly distributed by route, day, and time of day. However, they still should be selected to cover the full range of conditions that reflect at least the following elements of a service:

- Vehicle types
- Divisions, if available
- Route lengths
- Passenger loads
- Day type

5.1.4   Testing Equivalency

Agencies must demonstrate to FTA that the sample APTLs from using APCs and from using manual ride checks are statistically equivalent at the 95% confidence level. Agencies must describe how they will conduct this demonstration in the benchmarking plan. The following steps describe a process for such a demonstration.

Step 1. Compute the sample UPT and PMT for the APC data and for the manual data, respectively, for each one-way vehicle trip in the parallel sample:

- APC Data: $UPT_{i,apc}$ and $PMT_{i,apc}$
- Manual Data: $UPT_{i,man}$ and $PMT_{i,man}$

Figure 5.1 shows the sample UPT and PMT for individual one-way vehicle trips for an illustrative example. Column A counts the number of one-way vehicle trips. The sample UPT and PMT from APC data are in columns B and C. The sample UPT and PMT from manual data are in columns D and E. Only the first five and the last five one-way vehicle trips are shown.

Figure 5.1. Showing Sample UPT and PMT in a Spreadsheet

	A	B	C	D	E
1	No.	APC UPT	APC PMT	Manual UPT	Manual PMT
2	1	31	146.93	15	23.44
3	2	8	53.44	32	167.29
4	3	6	53.31	39	95.66
5	4	108	458.59	9	36.15
6	5	26	62.29	15	141.39
497	496	44	150.49	62	331.70
498	497	14	76.86	33	118.55
499	498	1	12.83	11	17.00
500	499	19	36.55	14	37.28
501	500	43	225.90	35	76.57

Step 2. Determine the number of one-way vehicle trips in the parallel sample and denote it as m. With the example in Figure 5.1, m = 500.

Step 3. Compute the correlation coefficient between the sample UPT and PMT across all one-way vehicle trips for the sample data collected from APCs and from manual ride checks, respectively. One way to compute this correlation coefficient would be to use the CORREL() function in Excel. For the example in Figure 5.1, the correlation coefficient is given by CORREL(B2:B501,C2:C501) for the APC data and by CORREL(D2:D501,E2:E501) for the manual data. Specifically:

- APC Data: R_{apc} = CORREL(B2:B501,C2:C501) = 0.8245
- Manual Data: R_{man} = CORREL(D2:D501,E2:E501) = 0.8060

Step 4. Add up the sample UPT and PMT across all one-way vehicle trips in the parallel sample and divide the sum by the number of one-way vehicle trips in the sample to get the average UPT and PMT for the APC data and the manual data, respectively:

- APC Data: UPT_{apc} = 21.41 and PMT_{apc} = 111.67
- Manual Data: UPT_{man} = 21.12 and PMT_{man} = 108.50

Step 5. Computing the APTL for the APC data and the manual data, respectively:

- APC Data: $APTL_{apc}$ = PMT_{apc} / UPT_{apc} = 111.67 / 21.41 = 5.22
- Manual Data: $APTL_{man}$ = PMT_{man} / UPT_{man} = 108.50 / 21.12 = 5.14

Step 6. Compute the standard error of the sample UPT and PMT for the APC and manual data, respectively, across the one-way vehicle trips in the parallel sample. One way to compute the standard error in Excel is to use the variance function (i.e., VAR()) and the square root function (i.e., SQRT()). For the example in Figure 5.1, the standard error for the sample UPT from APC data can be computed as SQRT(VAR(B2:B501)):

- APC Data: $SE(UPT_{apc})$ = 18.77 and $SE(PMT_{apc})$ = 116.82
- Manual Data: $SE(UPT_{man})$ = 17.36 and $SE(PMT_{man})$ = 111.66

Step 7. Compute the sample coefficient of variation for both UPT and PMT and for APC data and manual data, respectively. For a given quantity (e.g., UPT) and a given source of data (e.g., APCs), the sample coefficient of variation is defined by the ratio of the standard error of the quantity over the corresponding average:

- $CV(UPT_{apc})$ $= SE(UPT_{apc}) / UPT_{apc}$ $= 18.77 / 21.41$ $=$ 0.8767
- $CV(PMT_{apc})$ $= SE(PMT_{apc}) / PMT_{apc}$ $= 116.82 / 111.67$ $=$ 1.0461
- $CV(UPT_{man})$ $= SE(UPT_{man}) / UPT_{man}$ $= 17.36 / 21.12$ $= 0.8220$
- $CV(PMT_{man})$ $= SE(PMT_{man}) / PMT_{man}$ $= 111.66 / 108.50$ $= 1.0291$

Step 8. Compute the standard error for APTL for APC data and manual data, respectively:

$$SE(APTL_{apc})$$
$$= \frac{APTL_{apc}}{\sqrt{m}} \sqrt{[CV(UPT_{apc})]^2 + [CV(PMT_{apc})]^2 - 2 \cdot R_{apc} \cdot CV(UPT_{apc}) \cdot CV(PMT_{apc})}$$
$$= \frac{5.22}{\sqrt{500}} \sqrt{0.8767^2 + 1.0461^2 - 2 \cdot 0.8245 \cdot 0.8767 \cdot 1.0461} = 0.1381$$

$$SE(APTL_{man})$$
$$= \frac{APTL_{man}}{\sqrt{m}} \sqrt{[CV(UPT_{man})]^2 + [CV(PMT_{man})]^2 - 2 \cdot R_{man} \cdot CV(UPT_{man}) \cdot CV(PMT_{man})}$$
$$= \frac{5.14}{\sqrt{500}} \sqrt{0.8220^2 + 1.0291^2 - 2 \cdot 0.8060 \cdot 0.8220 \cdot 1.0291} = 0.1400$$

Step 9. Compute the statistic for testing the statistical equivalency of $APTL_{apc}$ and $APTL_{man}$:

$$T_{APTL} = \frac{|APTL_{apc} - APTL_{manual}|}{\sqrt{[SE(APTL_{apc})]^2 + [SE(APTL_{manual})]^2}} = \frac{ABS(5.22 - 5.14)}{\sqrt{0.1381^2 + 0.1400^2}} = 0.41$$

Step 10. Test the statistical equivalency of APTL$_{apc}$ and APTL$_{man}$:

- APTL$_{apc}$ and APTL$_{man}$ are statistically equivalent if T$_{APTL}$ < 1.96.
- APTL$_{apc}$ and APTL$_{man}$ are not statistically equivalent if T$_{APTL}$ ≥ 1.96.

For the example in Figure 5.1, T$_{APTL}$ = 0.41 < 1.96, implying that the APTL is statistically equivalent between the APC data and the manual data at the 95% confidence level.

5.1.5 Adjustment Factors

Agencies must develop the adjustment factors for data errors from using the sample PMT and UPT from APCs and manual ride checks, respectively. Denote the sample averages as UPT$_{apc}$, UPT$_{man}$, PMT$_{apc}$, and PMT$_{man}$. One can use the AVERAGE() function in Excel to compute these averages. For the example in Figure 5.1, UPT$_{apc}$ = AVERAGE(B2:B501). The following gives the adjustment factor for data errors in PMT:

$$f_{E,PMT} = 100 * \frac{PMT_{apc} - PMT_{man}}{PMT_{man}}$$

Similarly, the adjustment factor for data errors in UPT is given by:

$$f_{E,UPT} = 100 * \frac{UPT_{apc} - UPT_{man}}{UPT_{man}}$$

For the example in Figure 5.1, f$_{E,PMT}$ = 2.9% and f$_{E,UPT}$ = 1.4%.

Agencies also should describe how they plan to obtain the adjustment factors for missed data in UPT (f$_{M,UPT}$) and in PMT (f$_{M,PMT}$). Although these adjustment factors would not be obtained from the sample data collected during the benchmarking study or an annual maintenance study, they still need to be obtained during the benchmarking year or the corresponding maintenance year. Section 4 on obtaining adjustment factors describes guidance for such procedures. They are not repeated here.

5.1.6 Adjustments for APC data

Agencies must describe the planned procedures to adjust both for missing data and for data errors, where appropriate. The exact adjustments depend on the option that an agency chooses for meeting NTD's data requirements. Specifically, unadjusted estimates from APC data must be adjusted as follows in general:

- Option 1: Estimates of UPT and PMT must be adjusted for data errors.

- Options 2 and 3: Estimates of UPT must be adjusted both for missed data ($f_{M,UPT}$) and for data errors ($f_{E,UPT}$). Estimates of PMT must be adjusted both for missed data ($f_{M,PMT}$) and for data errors ($f_{E,PMT}$).

The detailed procedures for these adjustments interact with how UPT and PMT are obtained and have been described in the guidance for meeting NTD's data requirements in Section 4. Agencies should describe these detailed procedures relevant to their specific circumstances. They are not repeated here.

5.1.7   Cross Checks

The APC requirements suggest that UPT counts from APCs for fixed-route bus services should be regularly checked and reconciled against trip counts and fares from registering fare boxes.

Trip Counts

Cross checks against trip counts would be meaningful only when such UPT counts from APCs are used for reporting a 100% count of UPT.

Checking UPT counts from APCs against trip counts should be done as part of tracking the status of each one-way vehicle trip that is equipped with an APC:

- Whether any APC data were recovered from it
- Whether the recovered APC data were usable for obtaining UPT for it

Reconciling UPT counts from APCs against trip counts should be done through adjusting direct UPT counts from APCs by the adjustment factors for missed data. The results from the tracking process during an entire year can be used to develop adjustment factors for missed data. These adjustment factors for missed data can then be used to adjust the direct UPT counts from APCs.

Section 4 provides detailed guidance for tracking the status of APCs, developing adjustment factors for missed data, and adjusting APC data by these adjustment factors for getting a 100% count. It is not repeated here.

Fare Revenues

Temporal changes in fare revenues from registering fare boxes could be used as an indicator to check the reasonableness of UPT counts from APCs in terms of temporal changes.

- UPT counts from APCs may be considered reasonable if the percent change in these counts is similar to the percent change in fare revenues from fare boxes during a given period. UPT counts from APCs may not be considered reasonable, however, if these two percent changes differ significantly or if they are in opposite directions.

- Fare revenues from registering fare boxes may be used as a reasonableness indicator only for a 100% count of UPT if annual data on fare revenues from fare boxes are available only for a whole system. If fare revenues from fare boxes can be separated for those one-way vehicle trips with usable APC data on UPT, they may also be used to conduct cross checks for an estimate of UPT from APCs.
- Fare revenues from registering fare boxes would not be a sound indicator, however, for a period during which fare policies change and this change significantly impacts the collection of fare revenues from registering fare boxes.

Fare revenues from registering fare boxes, however, would not provide a sound basis for checking the reasonableness of UPT counts from APCs for a single year or for reconciling UPT counts from APCs. Fare revenues from registering fare boxes account for an increasingly smaller share of the total fare revenues of a fixed-route bus service. In addition, fare revenues from registering fare boxes have become less correlated with boardings across individual one-way vehicle trips over time as fare passes become more popular.

5.2   The Studies

Agencies should follow the approved plans to conduct the studies for a given year (i.e., the benchmarking year or a maintenance year). Once a study is complete, they must document the results from the study for submission to NTD as part of the submission process for the study year as an NTD report year. If the study was conducted differently in some aspects from the plan, the differences should be documented along with the results.

For both studies, agencies must present and document the results on the adjustment factor for data errors in UPT ($f_{E,UPT}$) and the adjustment factor for data errors in PMT ($f_{E,PMT}$). In addition, agencies should also present and document the results on the adjustment factor for missed data on UPT ($f_{M,UPT}$) and the adjustment factor for missed data on PMT ($f_{M,PMT}$). The document should discuss changes in these adjustment factors from those for earlier years.

For the benchmarking study, agencies must present and document the results on whether their sample APTL is statistically equivalent between data from the APCs and data from the manual ride checks at the 95% confidence level.

For an annual maintenance study, agencies also should present and document results from comparing annual total UPT obtained from APC data with annual total one-way vehicle trips actually operated and with fares from registering fare boxes.

This page intentionally left blank.