NCTR
NATIONAL CENTER for TRANSIT RESEARCH

Final Report

Guidebooks for Estimating Total Transit Usage through Extrapolating Incomplete Counts

September 2016

PREPARED FOR

U.S. Department of Transportation
Federal Transit Administration

Contract DTRT-12-G-UTC22 Mod 2

National Center for Transit Research
Center for Urban Transportation Research
University of South Florida
www.nctr.usf.edu

Guidebooks for Estimating Total Transit Usage through Extrapolating Incomplete Counts

Contract DTRT-12-G-UTC22 Mod 2

Prepared for:

U.S. Department of Transportation
Federal Transit Administration

Maggie Schilling, NTD Program Manager

Prepared by:

USF Center for Urban Transportation Research

Xuehao Chu, Ph.D.

Final Report

September 2016

Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Transit Administration.

Metric Conversion

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC

Technical Report Documentation

1. Report No.

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

Guidebooks for Estimating Total Transit Usage through Extrapolating Incomplete Counts

5. Report Date

September 2016

6. Performing Organization Code

7. Author(s)

Xuehao Chu

8. Performing Organization Report No.

NCTR 79060-17

9. Performing Organization Name and Address

National Center for Transit Research
Center for Urban Transportation Research
University of South Florida
4202 East Fowler Avenue, CUT100, Tampa, FL 33620

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

U.S. DOT DTRT-12-G-UTC22 Mod 2

12. Sponsoring Agency Name and Address

Research and Innovative Technology Administration
U.S. Department of Transportation, Mail Code RDT-30
1200 New Jersey Ave., SE, Room E33, Washington, DC 20590

Florida Department of Transportation, Research Center
605 Suwannee Street, MS 30, Tallahassee, FL 32399

13. Type of Report and Period Covered

Final Report
August 2013–September 2016

14. Sponsoring Agency Code

15. Supplementary Notes

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

16. Abstract

This report provides guidance for transit agencies to estimate transit usage for reporting to the National Transit Database (NTD) when their counting procedure that is designed to perform full counts misses some trips. Transit usage refers to unlinked passenger trips (UPT), passenger miles traveled (PMT), and average passenger trip length (APTL) in terms of annual totals and annual average daily by schedule type for annual reporting and monthly total UPT for monthly reporting. The guidance is provided in two self-contained guidebooks for bus and rail, respectively. Bus service includes all four fixed-route bus modes defined in the NTD: motor bus (MB), commuter bus (CB), bus rapid transit (RB), and trolleybus (TB). Rail includes light rail (LR), streetcar rail (SR), and hybrid rail (YR). For both mode types, the guidance focuses on data from automatic passenger counters (APC).

The guidance details a methodology for determining transit usage for each mode type through stratified extrapolation of incomplete counts rather than intentional sampling with APCs. The guidance views the total transit usage determined from the methodologies as estimates rather than 100% counts. It also views each methodology as an alternative sampling technique. The guidance identifies the conditions under which transit agencies may estimate annual total transit usage with this methodology as a pre-certified alternative sampling technique. For example, agencies must pass an equivalence test by demonstrating that their APC data are statistically equivalent to paired manual data within ±7.5% at the 95% confidence level. The guidance provides detailed steps for agencies to conduct the equivalence test in an Excel environment. When agencies meet the identified conditions and follow the guidance, they may use Appendix A in each guidebook as the document of certification by a qualified statistician for the alternative sampling technique.

For the NTD program, the guidance fills a gap in current NTD guidance and will result in more accurate UPT and PMT data. For agencies, it prevents under-reporting and saves the need to hire a qualified statistician for certifying the methodology as an alternative sampling technique.

17. Key Words

National Transit Database (NTD), automatic passenger counter, unlinked passenger trips, passenger miles traveled, sampling manual, sampling plans, bus, light rail, streetcar rail

18. Distribution Statement

Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161, (703) 487-4650, http://www.ntis.gov, and through the NCTR website at http://www.nctr.usf.edu/

19. Security Classification (of this report)
Unclassified

20. Security Classification (of this page)
Unclassified

21. No. of Pages

103

22. Price

Acknowledgments

The project and this report greatly benefited from discussions with and insights and different views provided by Keith Gates, formerly with the Federal Transit Administration (FTA); Brian McCollom with McCollom Management Consulting; Matt Bonzek with BCG Transportation, Inc.; Andeas Rakebrandt with INIT; and Tom Kowalski with Urban Transportation Associates, Inc. This report also benefited from comments and edits by Dr. Victoria Perk, Dr. Steve Polzin, Patricia Ball, and Joel Volinski with the Center for Urban Transportation Research. Several transit agencies provided paired data collected from the same sample of trips with both automatic passenger counters and human ridecheckers.

Table of Contents

Part I – Introduction

Table of Contents – Part I

1.00 Report Organization

The report is presented in three parts for agencies to estimate and report transit usage to the National Transit Database (NTD) through extrapolating all valid but incomplete data from automatic passenger counters (APC) installed on the entire fleet of transit vehicles. Part II is a bus guidebook, and Part III is a rail guidebook. Bus service includes all four fixed-route bus modes in NTD: commuter bus (CB), motor bus (MB), bus rapid transit (RB), or trolley bus (TB). Rail service includes light rail (LR), streetcar rail (SR), and hybrid rail (YR). The Federal Transit Administration (FTA) created YR as a new NTD mode in 2011 to include rail service that is operated primarily on the national system of railroads but whose operating characteristics resemble those of light rail. Before 2011, YR was reported either as LR or commuter rail service. Four agencies have reported YR service since 2011: Metrorail in Austin, Sprinter in San Diego, Hudson-Bergen Light Rail in New Jersey, and Westside Express Service in Portland.

As of the publication date of this report, FTA has not officially adopted these guidebooks. However, the estimation methodology in these guidebooks has been certified by the author of this report to meet FTA's statistical requirements for estimates of annual total transit usage. This estimation methodology represents an alternative sampling technique. FTA requires that any alternative sampling technique must be certified by a qualified statistician to meet FTA's statistical requirements. If agencies follow the guidance in each guidebook, they can use this estimation methodology as an alternative sampling technique and use Appendix A in each guidebook as the corresponding certification document. The author of this report developed the *National Transit Database Sampling Manual* that FTA has officially adopted since 2011 and has certified alternative sampling techniques for many agencies and a range of modes.

This section, Part I—Introduction, is for those who are interested in issues related to the reporting of transit usage to the NTD. For those who may not actually report such data to the NTD, its intent is to increase awareness of a statistically-sound methodology for determining transit usage for bus and rail services. For those who do report transit usage to the NTD, this will help determine if they want to consider using the methodology presented in these guidebooks. For both groups, it provides justifications for key elements of the guidance, particularly pre-certification of the methodology in these guidebooks as an alternative sampling procedure for estimating annual total transit usage.

This Introduction provides background, describes the problem, summarizes the methodology included in these guidebooks, compares the statistical features between this methodology and two approaches to sampling, and justifies the certification of this methodology.

2.00 Background

This section summarizes the basic requirements for reporting transit usage data to NTD by individual transit agencies, describes the traditional sampling approach to using manual data for estimating transit usage, and discusses the intentional sampling approach to using APC data for estimating transit usage.

2.01 NTD Requirements

To be eligible for the Urbanized Area Formula Grant Program (Section 5307), transit agencies must report to FTA the following measures of transit usage:

- Annual total unlinked passenger trips (UPT)
- Annual total passenger miles traveled (PMT)
- Annual average daily UPT by schedule type (weekday, Saturday, Sunday)
- Annual average daily PMT by schedule type
- Monthly total UPT

For rail services, agencies must also report annual average weekday UPT by time period, including the AM Peak period, Midday period, PM Peak period, and Other period.

To obtain annual total UPT or PMT, agencies must choose one of two approaches. If available and reliable, a 100% count must be reported. If a reliable 100% count is not available or is available but considered not reliable, annual totals must be estimated through sampling. The obtained estimate of annual total UPT or PMT from sampling must meet the minimum 10% precision level at the 95% confidence level. When sampling, agencies use either sampling techniques pre-approved by FTA or other sampling techniques (referred to as alternative sampling techniques) certified by a qualified statistician. Data on annual average daily UPT and PMT, annual average weekday UPT, and monthly total UPT are not subject to these statistical requirements.

UPT measures 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 origin to destination. Increasingly, agencies report 100% count UPT from onboard electronic registering fareboxes for traditional bus services. However, this is not the case for new RB lines or modern LR lines, which typically rely on off-board proof-of-payment fare collection, requiring passengers to purchase fare media off-board the transit vehicle and show proof-of-payment upon random inspection. Such fare collection systems do not produce 100% UPT counts.

It is costly for agencies to obtain 100% PMT. PMT measures the total distance traveled by all passengers. For both bus and rail, PMT for each one-way trip typically is calculated as the distance-weighted sum of passenger loads between consecutive stops/stations. This calculation typically requires detailed data on passenger boarding and alighting activities at individual stops or stations and distances between consecutive stops or stations for one-way trips. For this reason, agencies rarely collect and report 100% PMT for bus and rail services; instead, they use the only alternative approach available, i.e., estimation through sampling.

2.02 Traditional Sampling with Manual Ridechecks

Traditional sampling with manual ridechecks has been accepted as the “gold standard” for estimating PMT through sampling. For each sampled trip, one or more human ridecheckers ride with passengers on the sampled trip to observe and record the on-off passenger activities through every door at every stop. Although errors can occur in the raw data on the recorded passenger activities, they can largely be corrected through balancing the boardings and alightings for each sampled trip. This is especially true when additional data are collected from the field and used in trip balancing:

- Inherited passengers from previous trip at first stop
- Passengers continuing to next trip at last stop
- Passenger loads between each consecutive pair of stops

2.03 Intentional Sampling with APCs

The most prominent problem with traditional sampling with manual ridechecks is the significant labor cost of collecting and processing the sample data. Although transit agencies have installed APC units on their fleets, primarily for internal planning purposes, increasingly they are taking advantage of these systems for NTD reporting. Using electronic infrared beams or mechanical treadle mats, APCs can convert electronic signals from the beams or mats to counts of passengers as they board and alight transit vehicles. An APC also includes a dedicated onboard computer, an APC analyzer, which converts sensor information into passenger counts. Each time a vehicle leaves a stop, the APC analyzer closes out a record and transmits the on-off counts at that stop to a general onboard computer, the in-vehicle logit unit. When coupled with stop location information, archived APC data can be post-processed to generate both UPT and PMT data potentially for each trip operated.

The installation of APC units on buses has increased significantly since 2004 when only an estimated 3.6% had APC units. By January 2014, the percentage had reached to almost 40% (Figure I-1).

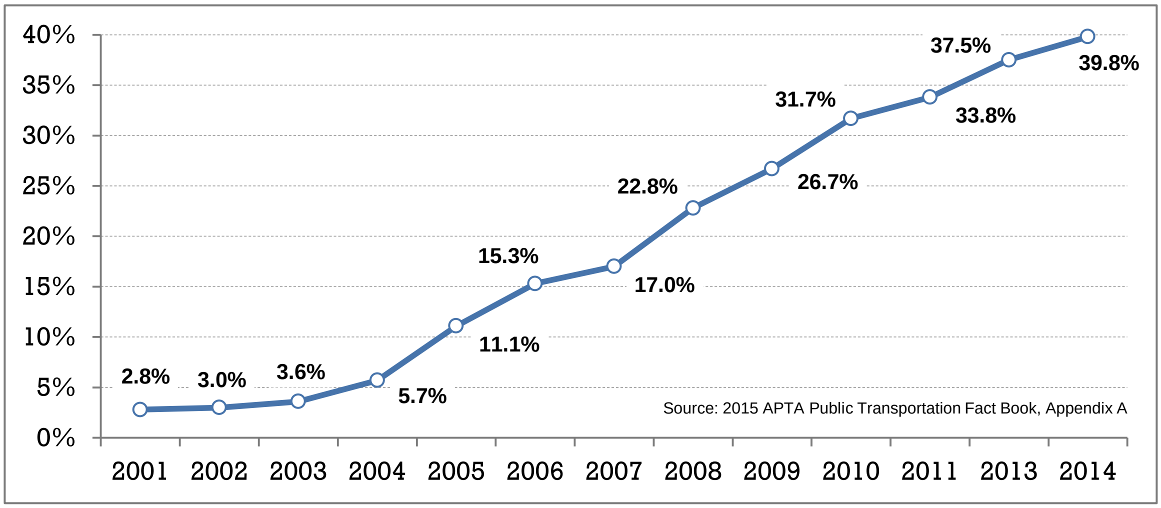

Figure I-1. Share of Bus Vehicles with APC Units

In addition, many rail passenger cars had APC units by January 2014—34.1% for light rail, 31.0% for streetcar rail, and 75% for hybrid rail (Figure I-2).

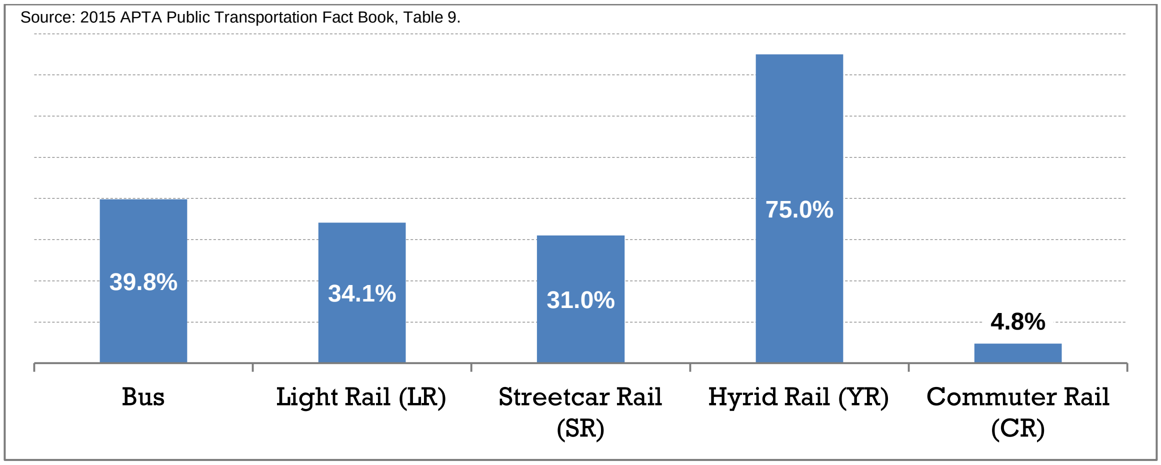

Figure I-2. Penetration of APC Units by Mode by January 2014

When the entire fleet is fully equipped with APC units, it is possible, theoretically, to obtain a 100% count of both annual total UPT and PMT from APC data. In reality, however, getting reliable 100% counts of UPT and PMT from an APC system is difficult to achieve. Because of this, to use APC data for NTD reporting, agencies must estimate annual total UPT and PMT through intentional sampling with APCs. "Intentional sampling with APCs" is a term used in this report to contrast it with traditional sampling with manual ridechecks and with using all valid APC data, which is a form of "full-population" sampling and is the focus of these guidebooks. Intentional sampling with APCs can take one of two forms:

- **Pre-sampling** takes place before a trip is actually operated. Pre-sampling is common for an agency with a relatively small share of its fleet equipped with APC units. Pre-sampling is necessary so that APC-equipped vehicles can be assigned to the sampled trips. Pre-sampling also is applicable to other cases, and it must account for the expected data recovery rate in the sample size used. As with traditional sampling with manual ridechecks, pre-sampling selects trips from the full list of all trips to be operated.
- **Post-sampling** happens after a trip is operated and the APC data have been processed. This is common for an agency with a full coverage of APC units on all doors of all vehicles. Different from either traditional sampling or pre-sampling, post-sampling selects trips from all trips with valid APC data.

2.04 APC Certification by FTA

Regardless of how agencies use their APC data for estimating transit usage, approval by FTA to use APC data for NTD reporting must be obtained beforehand. FTA must certify the APC system initially and periodically. To be certified each time, agencies must pass a benchmarking test. For this test, agencies use paired APC and manual data collected from the same sample of vehicle trips and must demonstrate that differences between the paired APC and manual data are within ±5% for both UPT and PMT. This benchmarking test is required every fiscal year that is evenly divisible by three (2019, 2022, 2025, etc.).

3.00 The Problem of Intentional Sampling with APCs

3.01 Introduction

For NTD reporting, estimating transit usage through intentional sampling with APCs avoids the significant labor cost necessary with traditional sampling with manual ridechecks. However, intentional sampling has at least three major shortcomings:

- It does not mitigate any potential bias in the estimates from missing data.
- It does not necessarily take into account systematic errors in valid APC data.
- It does not take advantage of the significant amount of valid APC data available.

3.02 Missing Data

3.02.1 Reasons for Missing Data

Valid APC data can be missing for some trips because they were not recovered or they were discarded due to significant errors in the raw data or in the processed data. Many aspects of the problem of missing data are shared by APC systems on both bus and rail:

- Raw data may fail to be recovered from individual APC units. One reason is general hardware malfunction, including sensors in the doorways failing to send signals to the onboard computer, the onboard computer failing to convert sensor information into passenger on-off counts, and the onboard computer failing to transmit and store on-off counts. The effect of this problem can be reduced to some degree through a continuous monitoring and maintenance program.

- Raw data may be successfully recovered from individual APC units, but serious errors can exist in the recovered raw data that require the data to be discarded through processing procedures:
  - In many cases, raw APC data are recovered, but they fail to match in space and time to the actual service provided; in these cases, the recovered raw APC data are not usable.
  - In many other cases, raw APC data are recovered and matched to actual service, but errors in the recovered raw data fall out of range based on pre-set screening criteria for differences in on-off counts for individual blocks; in these cases, the matched raw APC data are not usable for purposes focusing on boarding counts. When matched raw APC data fall out of range, the raw APC data are rejected for purposes focusing on load or passenger miles.

Most modern light rail systems in the US have unique features that are not shared by most bus systems, including low-floor boarding and double-wide doors. These two features combined allow more than one passenger to board and alight at the same time. Such complex passenger activities can be a challenge for APC systems. In addition, the number of APC units on fully-equipped trains can be significantly greater than the typical 2 units on most buses. A train with 4 cars may have a total of 16 APC units with 2 wide doors on each car and 2 APC units on each door. Even with 1 or a few of these APC units failing to recover any raw data, it would be difficult to balance the on-off data for a whole train trip to a degree that can pass pre-set screening criteria for further consideration.

At the same time, however, some unique features of light rail are positive in reducing the degree of missing data. Many modern light rail systems have a simple network with just one or a few lines. This network simplicity helps avoid some problems due to complexities of a transit network; specifically, it is less likely to have APC data failing to match in space and time to the actual service provided.

3.02.2 Potential Bias from Missing Data

The data recovery rate for an APC system refers to the percent of all one-way trips operated during a specific period that provide valid data on trip-level UPT and PMT. This rate has been reported to be as low as 60–75% for bus APC systems. For rail modes, little data are available. It could be higher because APC systems for rail typically are newer and, thus, probably provide higher data recovery rates; it could be lower because of low-floor boarding, double-wide doors and the larger number of APC units on a single train.

As with all valid APC data, valid sample data from intentional sampling with APCs also suffer from missing data. When the data recovery rate is relatively low and the degree of missing data is high, it is highly likely that certain segments of the whole service are either no longer represented or are disproportionately represented in the valid sample from intentional sampling. If not properly mitigated, bias results in the estimates. Mitigating such bias, however, is almost impossible because of the limited amount of data from intentional sampling. Even in the case of disproportionate representation, estimation for such service segments is highly unreliable due to the small remaining sample size for each segment.

3.03 Systematic Error

3.03.1 Sources of Error

In almost all cases in which raw APC data are recovered, matched, and accepted for further consideration, on-off counts will have errors. Unlike with manual data, these errors cannot be corrected in determining trip-level UPT and PMT because APCs collect on-off counts but do not collect additional information on the number of inherited passengers from previous trips, the number of continuing passengers to the next trip, and stop-to-stop loads.

Whereas screening based on on-off differences avoids substantial errors in on-off totals, substantial errors still can develop in calculated loads and passenger miles, even with small errors in raw on-off counts due to a phenomenon called “drift.” An effective way to control drift is to parse a vehicle block’s data stream of automated counts at points of known load. Most routes have natural known-load points at terminals and layover points. To control drift through parsing, blocks are divided at known-load points into sections, usually single trips. On-off counts within each section then need to be balanced so the calculated loads match the known loads at each end of the section.

Some unique features of light rail systems may lead to smaller magnitudes in potential systematic errors in APC data. Whereas a bus driver may get on or off a vehicle multiple times through the same doors for passenger boarding and alighting, a light-rail train operator is much less likely do to so. In addition, it is much less likely that passengers stay onboard from one train trip to the next without getting off. This avoids a source of potentially serious errors in calculating passenger loads and passenger miles that many bus APC systems experience.

3.03.2 Need to Consider Systematic Error in Sampling

Sampling plans for traditional sampling with manual ridechecks are designed to meet the 10% precision and 95% confidence levels without considering any potential systematic error in the manual data for estimation. For historical reasons, manual data through human ridechecks are assumed to represent the true value. As discussed earlier, raw manual data can have errors, sometimes serious errors. With additional field data and proper balancing, however, such errors in the raw data can be corrected. Also, any remaining error is likely to be mostly random rather than systematic.

For this reason, sampling plans designed for intentional sampling with APCs frequently can ignore potential systematic error in the APC data. Even if the APC data for a given APC system suffer 8% systematic error at 95% confidence, for example, such sampling plans still would be designed to meet the 10% precision and 95% confidence levels. In this case, the precision reached by estimates from such sampling plans could be as low as 18% rather than 10%.

If the systematic error is greater than 10%, the APC data cannot be used for NTD reporting. If the systematic error is under 10%, it is possible to account for it in designing a sampling plan for intentional sampling. With 8% systematic error at 95% confidence, for example, the plan for intentional sampling should be designed to meet 2% precision and 95% confidence. A consequence is that the required minimum sample size would be 25 times as big as a sampling plan designed for meeting 10% precision. When requiring such a large sample size, the general advantage of estimation through sampling is significantly reduced.

Paired APC and manual data from several individual agencies indicate that the degree of systematic measurement error in valid APC data can be as large as 12.5% and as small as <1%.

3.04 Benefits of Using All Valid Data

The potential bias from missing data can be largely mitigated by using all valid APC data. The problem of missing data still can lead to certain service segments being disproportionately represented, but it is highly unlikely to have certain service segments not represented. When a service segment is disproportionately represented, the significantly large amount of valid data allows relatively reliable estimation for this segment. In addition to helping mitigate the potential bias from missing data, the large amount of all valid APC data will result in estimates with much better precision than those from intentional sampling.

3.05 Need for an Alternative Approach to Using APC Data

As a result of these issues related to intentional sampling with APCs, estimates from this approach can seriously violate the 10% precision and 95% confidence requirements. An alternative approach is needed to using APC data for estimating annual total usage that is designed to solve the problem of intentional sampling with APCs.

There is an alternative approach for agencies whose counting procedure covers all trips but misses some of them. This approach uses all valid APC data and proportionally expands the incomplete APC data. If an agency knows that the trips with missing data are 1.99% of all trips for an entire year, for example, it can simply multiply the incomplete count with a factor equal to 1/(1-0.0199) = 1.020. NTD rules allow this approach only for agencies whose degree of missing data does not exceed 2%. This approach solves the problem with intentional sampling and is pre-approved by FTA. Unfortunately, this approach is rarely applicable with APC data. The degree of missing data accounts for not only trips without any raw APC data recovered but also trips with recovered raw data being thrown out. It is highly unlikely that the degree of missing data for an APC system stays within 2%; it is possible for electronic registering fareboxes but not for APCs.

What is needed is an alternative approach to using APC data for estimating transit usage that does all of the following:

- Solves the problems of intentional sampling.
- Is applicable to APC systems with degrees of missing data exceeding 2%.
- Is pre-certified to meet the 10% precision and 95% confidence levels.

4.00 Guidebook Methodology

4.01 Stratified Extrapolation

The methodology takes an approach of stratified extrapolation of all valid but incomplete APC data for estimating total transit usage. Specifically, it:

- Divides a whole service into a set of mutually-exclusive service segments (e.g., a trip on the weekday schedule, a route, weekday AM peak, Saturdays, etc.).
- Estimates total transit usage by extrapolating all valid but incomplete APC data within each service segment.
- Sums the estimates of segment-level total transit usage across all these service segments to get total transit usage for the whole service.

4.02 General Conditions

To use the methodology in this guide for estimating and reporting annual UPT and PMT, an agency must first meet all of the following general conditions:

- **FTA certification** – The agency has obtained approval by FTA for using its APC data for NTD reporting.
- **Ability to keep track of trips with missing data** – The APC system must allow the agency to know the number of vehicle trips actually operated during an entire year with missing data and its percentage of all vehicle trips operated.
- **Ability to use all usable data** – The agency must be able to use a significant amount of APC data at a disaggregated level; this is especially relevant for a large bus system.
- **2% Rule** – The trips with missing data should represent greater than 2% of all trips actually operated during an entire year. This methodology may be applied to cases in which trips with missing data are less than 2% of all trips. In these cases, in fact, it is likely to produce estimates with higher accuracy levels than proportional expansion if missing data are not random. Since NTD rules allow proportional expansion when the degree of missing data does not exceed 2%, proportional expansion would be the better option for the agency.

4.03 Suggested Segmentation

A critical element of the stratified approach to extrapolating incomplete APC data is the careful segmentation of a whole service into service segments. To help an agency that is new to the methodology, a specific set of service segments is suggested, as follows:

- Individual trips on a weekday schedule
- Individual routes for Saturdays
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight

- Individual routes for Sundays
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight

This suggested segmentation meets the requirement of being mutually-exclusive and adding up to the whole service, uses schedule type as one attribute, and balances all three objectives for defining service segments: minimizing variation, adequate sample size, and relative convenience. The agency may choose the level of segmentation and should define the service segments according to these criteria.

4.04 Data Requirements

Ideally, an agency should have an annual database of trip-level information for all one-way trips actually operated during a full year. This database would accumulate such information as individual trips are being operated and as the raw APC data are being processed from the first day to the end of the year. At a minimum, this database should have the following data items for each one-way trip actually operated:

- Date on which the trip was operated
- Whether it was an atypical day
- Route
- Schedule type
- Start time
- Service segment to which the trip belongs
- Whether raw APC data were returned
- Whether trip-level valid APC data are available
- Trip-level UPT if valid APC data are available
- Trip-level PMT if valid APC data are available

These minimum data items are based on the need to identify trips with the suggested set of service segments and the need for estimating monthly total UPT and annual average daily UPT and PMT by schedule type. Additional data items may be included for alternative ways of defining service segments. The number of cars on a train may be needed for light rail so that estimation could be done on a per-car basis rather than on a per-train basis.

4.05 Nature of Methodology

This methodology is based on a form of full-population sampling with non-response (i.e., missing data) and without intentional sampling. It treats the results as estimates from sampling rather than as 100% counts. NTD rules allow agencies to proportionally expand incomplete counts without certification by a qualified statistician if trips with missing data do not exceed 2% of all trips operated. However, the same rules require certification by a qualified statistician for any method of extrapolating incomplete counts when trips with missing data exceed 2% of all trips operated. NTD rules are unclear about whether such extrapolated totals can always be treated as a 100% count. Treating the results from the methodology as estimates also helps avoid the gray area between estimates and 100% counts. NTD rules allow either 100% counts or estimates through sampling and, therefore, are inclusive of this methodology as an alternative sampling technique.

4.06 Certification

Appendix A of each guidebook serves as the document of certification for this alternative sampling technique for estimating annual total UPT and PMT. An agency should report on its CEO Certification Form the method used as an “alternative sampling procedure that meets 95% confidence and ±10% precision levels as determined by a qualified statistician (estimated data).” Section 5.00, Comparison, provides the statistical justification for this certification.

Certification of the alternative sampling technique is conditional on an agency meeting the following conditions. Some of these conditions may appear restrictive but are necessary for pre-certifying this procedure for general application.

4.06.1 Coverage, Operations, and Data Recovery

The entire fleet of transit vehicles must be fully-equipped with APC units. For bus service, APC units must be installed at every door of every bus in the fleet. For rail, APC units must be installed at every door of every passenger car in the fleet.

In general, 100% coverage of a counting procedure is not necessary for estimating total transit usage through extrapolating incomplete APC data. Without 100% coverage, an agency can rotate its APC-equipped vehicles freely, which produces additional sources of potential bias in the final estimates that is more difficult to control through any methodology, including the methodology in these guidebooks. As a result, stratified extrapolation without 100% APC coverage cannot be pre-certified for wide applications.

To ensure at least 500 trips with valid APC data, using the methodology also requires at least 1,000 trips operated annually, with at least 50% of them with valid APC data. An agency with 4 trips daily for 250 days a year would operate 1,000 trips annually. This extreme case is rare but possible for a weekday CB service with 2 morning trips and 2 afternoon trips. This extreme case, however, is unlikely to happen for rail services.

4.06.2 Use of All Valid APC Data

The agency must be able to use a significant amount of APC data at a disaggregated level. This is especially relevant for a large bus system. One effective approach to having this ability by an agency is that it should have an annual database of trip-level information for all one-way vehicle trips actually operated during a full year. This database would accumulate such information as individual trips are being operated and as the raw APC data are being processed from the first day to the end of the year.

4.06.3 Continuous Monitoring

An agency must have a process in place to monitor and maintain individual APC units and the APC system continuously. An effective process, for example, would immediately notify the agency when an APC unit on a trip failed to acquire any raw APC data on the on-off activities for the trip and would indicate the source of the equipment failure problem. Once equipment failure occurs, the same effective process would have an ongoing partnership with vendors for quick maintenance and repair. Reducing the number of trips with missing data due to equipment failure benefits not only NTD reporting but also internal service and other planning efforts within the agency.

4.06.4 Statistical Equivalence

Initial Testing

A transit agency must conduct a test of statistical equivalence for its APC data and must pass the test. Detailed guidance is provided in each guidebook on conducting this test. The agency also must collect paired APC and manual data from the same sample of trips selected at random during a period that is at least one week long.

- To pass the test of statistical equivalence, the agency must demonstrate that its APC data are statistically equivalent to the paired manual data within ±7.5% at the 95% confidence level.
- This condition of statistical equivalence is imposed to ensure that the estimate of annual total transit usage meet the 10% precision and 95% confidence levels. The equivalence bounds are set at ±7.5% to allow up to ±2.5% for errors in the estimates due to any remaining random measurement error and bias from missing data. Refer to Section 5.00, Comparison, for more discussions in this.
- The requirement of the paired APC and manual data coming from a random sample is necessary to assess the level of confidence achieved for the test of statistical equivalence.
- A minimum duration of one week for the testing period is specified to minimize the burden on the agency while non-seasonal variation is mostly being captured.

Figure I-1 shows what it takes to pass the test of statistical equivalence. The horizontal axis measures the systematic measurement error in the APC data, which is given by the mean difference between APC and manual data in percent terms. The vertical axis measures the random measurement error of the APC data, which is given by the standard error of the mean difference in percent terms. The large triangular area under the two thick lines gives the combination of random and systematic measurement errors that are required to pass the test. This triangular area may be referred to as the “zone of statistical equivalence.” It is clear that both random and systematic errors need to be small enough to pass the test.

The nine dots represent actual paired APC and manual data on PMT from nine agencies. Seven of these would pass the test. Also, these limited data points suggest that APC data tend to overstate PMT. Among these agencies, six of them have positive mean differences, indicating that APC data are bigger. In addition, the difference tends to be small when APC data understate PMT. For the three agencies with negative mean differences (located to the left of the vertical line at 0% of the horizontal axis), the differences are all extremely close to 0%.

Relationship to NTD’s Benchmarking Test

The equivalence being tested here is referred to as statistical equivalence because it accounts for both systematic errors and random errors. In contrast, the equivalence being tested by NTD’s benchmarking test may be referred to as numerical equivalence because it accounts for systematic measurement errors but not random measurement errors. This is the conceptual difference between them.

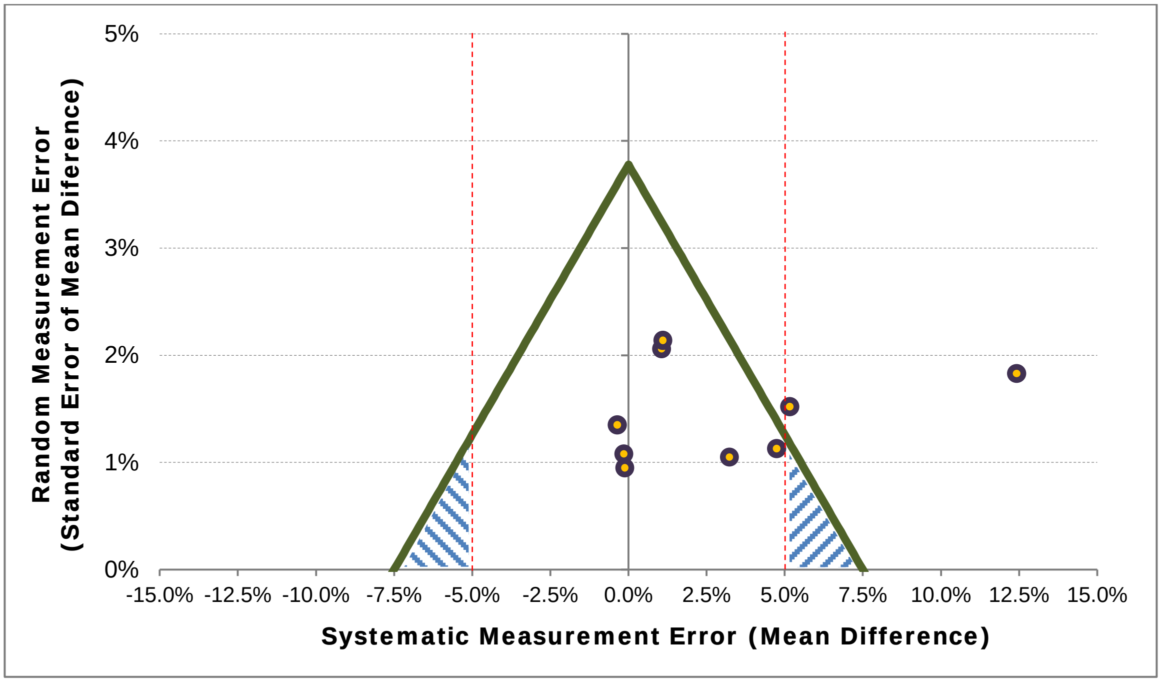

Figure I-1. Zone of Statistical Equivalence

Figure I-1 may be used to see how this conceptual difference leads to different testing outcomes. Also drawn in the figure are two vertical red dash lines at -5% and 5%, respectively. The entire vertical area between these two red lines defines the "zone of numerical equivalence" for the benchmarking test. Regardless of how large the random measurement error might be, agencies can pass the test when their systematic error is within ±5%. One way to see how the testing outcomes of these two tests differ is to examine the degree to which these two zones of equivalence overlap. The following calculations assume that agencies are uniformly distributed in this two-dimensional space:

- A total of 88.9% of all agencies that pass the test of statistical equivalence would also pass the benchmarking test, with the remaining 11.1% representing agencies in the two small shaded triangular areas on the two sizes of the zone of statistical equivalence. These remaining agencies tend to have relatively larger systematic measurement errors (-7.5% to -5% or 5% to 7.5%) but relatively small random measurement errors (0% to 1.26%). It is unfair to reject these remaining agencies from using their APC data for NTD reporting when it is certain that their estimates will meet the 10% precision and 95% confidence levels if they use the methodology in these guidebooks and follow the guidance.
- Assuming that random measurement errors never exceed 3.78% (the peak of the zone of statistical equivalence), 75% of all agencies that pass the benchmarking test would also pass the test of statistical equivalence. The remaining 25% agencies tend to have systematic errors within ±5% or random errors exceeding 1.26%.

Future Testing

Once the agency has passed the test of statistical equivalence once, it stays valid unless the agency fails FTA's benchmarking test. As mentioned earlier, the benchmarking test is required every year that is evenly divisible by three. When that happens, the agency would need to do the test of statistical equivalence again and pass it. For a new test, the agency should use the paired APC and manual data from the most recent test of statistical equivalence to determine the proper sample size. Each guidebook provides detailed guidance for the agency to determine the proper sample size.

4.07 Benefits

The most significant benefits of the methodology in these two guidebooks relate to more accurate UPT and PMT data from agencies with 100% coverage of APC units but without 100% counts. The methodology:

- Uses all valid APC data. Implicitly, NTD requires agencies to use the best data available. Estimates from all valid APC data are better in all cases than estimates from intentional sampling; in fact, estimates from this methodology will achieve significantly higher precision than those from intentional sampling.
- Is able to mitigate potential bias from non-random missing data by using the large amount of all valid APC data and the stratified extrapolation approach.
- Accounts for the systematic and random measurement errors in APC data by requiring agencies to demonstrate that their APC and paired manual data from a random sample are statistically equivalent within ±7.5% at the 95% confidence level.
- Estimates of annual total usage meet the 10% precision and 95% confidence levels.

The guidance also benefits the NTD program by filling a significant gap in its current rules and guidance on estimating total transit usage through extrapolating incomplete counts when the counting procedure designed to perform full counts misses some trips. This additional guidance also benefits individual transit agencies:

- An agency is less likely to violate NTD rules by choosing to report a 100% count of UPT or PMT if it actually failed to obtain 100% counts even though their data collection method has 100% coverage of an entire service.
- An agency is less likely to under-report annual total UPT or PMT if it uses estimates from this methodology rather than an incomplete 100% count. When PMT is under-reported, the agency would get a smaller share of Section 5307 funds than what it would deserve for its actual PMT.
- When an agency’s counting procedure is designed for full counts but misses more than 2% of all trips operated, the guidance saves it from having to have a qualified statistician develop and certify an alternative extrapolation method for filling the missing data.
- If an agency finds this methodology desirable as an alternative sampling technique, it can include testing statistical equivalence and passing the test as system acceptance requirements for procuring APCs.

5.00 Comparison

5.01 Introduction

This section compares how three approaches to estimating total transit usage compare for each of four types of errors in the resulting estimates. It also compares these approaches for the chance that these estimates do not meet the minimum 10% precision and 95% confidence levels. This comparison assumes that the same measure of transit usage is being estimated with each approach. To be specific for the following presentation, PMT is assumed. In addition, this comparison assumes that the same method is used in estimating PMT. The following are two common methods for estimating PMT:

1. Expanding average PMT per trip from sampling by all trips operated.
2. Expanding APTL from sampling by 100% UPT.

Table I-1 summarizes this comparison. The middle rows show the four types of errors potentially present in estimates of total transit usage from using APC data. The top row lists the three approaches already discussed earlier and being compared in this section: the methodology in the guidebooks, intentional sampling with APCs, and traditional sampling with manual ridechecks. The bottom row summarizes the comparison on the chance that estimated annual total usage does not meet 10% precision and 95% confidence levels.

Table I-1. Comparison of Statistical Quality of Estimates

Error Type	Guidance Methodology	Intentional Sampling with APCs	Traditional Sampling with Manual Ridechecks
1) Random sampling error	None	Present but controlled by target precision (p)	Present but controlled by target precision (p)
2) Random measurement error	Negligible with the large sample from using all valid data	Can be significant	None
3) Systematic measurement error (bias)	Present and controlled with ±7.5% equivalence bounds	Present and controlled by ±5% certification bounds	None
4) Systematic error from missing data (bias)	Effectively mitigated with stratified extrapolation	Can be present and significant but cannot be mitigated	None
Chance of not meeting 10% precision & 95% confidence	Negligible due to ±2.5% error allowable for 2) and 4)	Can be significant due to all four types present; lower if p = 5%	Negligible

5.02 Definition of Each Error Type

Errors in estimates are either random or systematic. Random errors can become negligible when the sample size is large enough, but systematic errors would become smaller from larger samples. Systematic errors are often referred to as bias.

Measurement errors refer to the difference in the APC or manual data from the true value and have both a random component and a systematic component. Systematic measurement errors represent the average of the differences across all trips operated. Random measurement errors, on the other hand, represent how the difference varies across trips. For an APC system, both types of measurement errors can be assessed through paired APC and manual data from the same sample of trips.

Random errors can result not only from measurement by APCs or manual ridechecks but also from sampling. They also can result from other reasons. Errors from manually entered data, for example, can exist and typically are random with traditional sampling with manual ridechecks. Such sources of random errors are not included in this comparison.

Systematic errors also can result from the sampled trips not being representative of all trips operated due to trips with missing data.

5.03 Traditional Sampling with Manual Ridechecks

Random sampling errors are always present but are effectively controlled by the target precision level used in designing the sampling plan used. The typical target precision is 10% in response to the minimum requirements of 10% precision and 95% confidence levels. Using this lowest target precision allowed within the required precision level makes sense because all other three types of errors are assumed to be absent. When traditional sampling is done with care, the resulting estimated annual total PMT is assumed to meet the 10% precision and 95% confidence levels.

5.04 Intentional Sampling with APCs

In contrast with traditional sampling with manual ridechecks, all four types of errors can be present in estimated annual total PMT from intentional sampling with APCs:

1) **Random sampling errors** are always present but are controlled by the target precision level used in designing the sampling plan used. Random sampling errors decline with using a higher target precision level (e.g., 5% rather than 10%).
2) **Random measurement errors** are always present and vary significantly across agencies.
3) **Systematic measurement errors** are always present and vary significantly across agencies. They will be controlled by the ±5% bounds by FTA's benchmarking test.
4) **Systematic errors from missing data** depend on the overall data recovery rate of an APC system, how transit usage varies across service segments, how the recovery rate varies across these segments, and how trips with missing data are distributed across these segments. Once a sample is drawn, this error can hardly be mitigated.

The overall precision reached by the estimated annual total PMT from intentional sampling is the sum of all four types of errors:

- The estimated PMT will not meet the 10% precision and 95% confidence levels if the target precision used for designing the sampling plan is 10%. This is because using 10% as the target precision level leaves no room for the other three types of errors.
- A higher target precision for the sampling plan will help improve the chance but does not ensure meeting 10% precision and 95% confidence levels. It depends on the sum of the other three types of errors.

5.05 Guidance Methodology

5.05.1 Presence of Various Errors

Sampling errors are absent in estimated annual total PMT from the methodology in these guidebooks. All other three types of errors can be present with this methodology:

1) **Random measurement errors** become negligible in estimated annual total PMT due to the extremely large sample from using all valid APC data in most cases. The zone of statistical equivalence in Figure I-1 peaks at 3.78%, effectively preventing agencies with larger random measurement errors from using their APC data for NTD reporting.
2) **Systematic measurement errors** are present but controlled by the required ±7.5% bounds for equivalence testing.
3) **Systematic errors from missing data** are effectively mitigated with the approach of stratified extrapolation at a disaggregated level. The problem of missing data with an APC system is the same as the problem of nonresponse with a traditional survey. The most effective approach to mitigating potential bias from nonresponse is estimation with stratification and weighting.

With systematic measurement errors controlled at ±7.5% with 95% confidence, up to ±2.5% errors are allowable for remaining random measurement errors and systematic errors from missing data. The next two sub-sections examine the potential magnitude for each of the two remaining errors.

5.05.2 Potential Magnitude of Random Measurement Error

Table I-2 shows how random measurement errors would vary with both the degree of variation in the difference between APC (measured by the standard deviation in percent terms) and manual data and the sample size, i.e., the annual total number of trips with valid APC data.

Table I-2. Variation of Random Measurement Error by Sample Size

Daily Trips	Annual Trips with Valid APC Data	Random Measurement Error by Range of Standard Deviation of APC and Manual Differences			
		5%	***10%***	***15%***	***25%***
		0.71%	***1.41%***	***2.12%***	***3.54%***
4	500	0.22%	0.45%	0.67%	1.12%
6	750	0.18%	0.37%	0.55%	0.91%
8	1,000	0.16%	0.32%	0.47%	0.79%
10	1,250	0.14%	0.28%	0.42%	0.71%
20	2,500	0.10%	0.20%	0.30%	0.50%
50	6,250	0.06%	0.13%	0.19%	0.32%
100	12,500	0.04%	0.09%	0.13%	0.22%
250	31,250	0.03%	0.06%	0.08%	0.14%
500	62,500	0.02%	0.04%	0.06%	0.10%
1,000	125,000	0.01%	0.03%	0.04%	0.07%
2,000	250,000	0.01%	0.02%	0.03%	0.05%

Note: A 50% data recovery rate and 250 days of service annually are assumed in determining annual trips with valid APC data.

The range for the standard deviation is capped at 25%, adequately covering all reasonable cases that can pass the test of statistical equivalence. Just below the row for the range of standard deviation, Table I-2 also shows the corresponding random measurement error if the sample size were just 50, which is a plausible one for a sample used in a test of statistical equivalence. For the other rows, the sample size is stated in terms of both daily trips and annual trips with valid APC data, assuming a low 50% recovery rate and 250 days of service annually. The range for daily trips is as low as 4 trips a day and as high as 2,000 trips a day. For a given standard deviation, the random measurement error for daily trips greater than 2,000 will be smaller than those for daily trips being 2,000. The extremely low numbers of 4 daily trips are possible for a commuter bus service with a single route. Across these ranges for the standard deviation and daily trips, it is clear that the likely random measurement errors are well below 1% and that it is likely to be well below 0.5% for most cases.

5.05.3 Potential Bias from Missing Data

One problem with potential bias from missing data is that its magnitude is never known. For intentional sampling with APCs, this adds much uncertainty to its chance of meeting the 10% precision and 95% confidence levels. For the methodology in the guidebooks, it will reduce this potential bias but the degree of reduction is never known either. To get some sense on how much this methodology may be able to reduce the potential bias from missing data, a simple simulation was done in terms of PMT for a hypothetical transit service.

Hypothetical Service

Table I-3 summarizes some characteristics of this hypothetical service. This service is operated 360 days per year with the same daily schedule of 50 trips with 10 in the AM peak, 15 in the PM peak, 20 during midday, and 5 during other times. The annual total PMT represents the true value. Similarly, average PMT per trip also represents the true value. It is noted that average PMT varies significantly across the four periods.

Table I-3. Characteristics of a Simple Transit Service

Characteristics	AM Peak	PM Peak	Midday	Night	Overall
Daily Trips on Schedule	10	15	20	5	50
Annual Trips Operated	2,600	3,900	5,200	1,300	13,000
Annual Total PMT	259,930	584,760	181,480	19,280	1,045,450
Average PMT per Trip	100.0	149.9	34.9	14.8	80.4

Simulation

The true annual total PMT and average PMT per trip were derived from trip-level PMT for every trip operated during a year. The true trip-level PMT varies within each period across both days and trips on the schedule. In fact, the trip-level PMT for trips operated within each period was randomly generated on the basis of a uniform distribution within lower and upper bounds pre-selected for the same period. Table I-4 shows these lower and upper bounds for each period.

Table I-4. Range of Trip-Level PMT by Period

Bounds	AM Peak	PM Peak	Midday	Night
Lower bound	80	120	20	5
Upper bound	120	180	50	25

Once the trip-level PMT was generated and fixed for each trip operated, trips with missing data were also randomly determined. A target data recovery rate was first set for each combination of scenarios, levels and periods. Under the scenario of constant recovery rates across periods, the recovery rate was the same for all periods. With the other scenario, it varied across the periods. For each scenario, the targets were set to four levels, representing 50%, 70%, 90%, and 98%, respectively, for the overall recovery rate. Table I-5 summarizes these target recovery rates. For a given combination of a scenario and level, a random number from 0 to 1 was then generated for each trip operated. The APC data for a trip operated were determined to be missing if the random number for this trip exceeded the target recovery rate for the period to which this trip belonged. This was done separately for each of the eight combinations of scenarios and levels.

Table I-5. Target Data Recovery Rates by Period

Scenario	Level	AM Peak	PM Peak	Midday	Night
Constant across Periods	1	50%	50%	50%	50%
	2	70%	70%	70%	70%
	3	90%	90%	90%	90%
	4	98%	98%	98%	98%
Varied across Periods	1	49%	45%	54%	58%
	2	70%	60%	75%	80%
	3	85%	85%	95%	98%
	4	98%	97%	99%	99.5%

Results

Table I-6 shows the results. The five columns in the middle show the realized degree of missing data for each combination of scenarios, levels, and periods. For each combination, the realized degree of missing data differed slightly from the corresponding target value in Table I-5 because of the random process used in generating trips with missing data. The results on bias from missing data are presented in the last two columns on the right side.

Table I-6. Potential Bias from Missing Data

Scenario	Level	Realized Degree of Missing Data					Bias from Missing Data	
		Overall	*AM Peak*	*PM Peak*	*Midday*	*Night*	*Simple Expansion*	*Stratified Extrapolation*
Constant across periods	1	49.4%	49.0%	50.8%	49.0%	47.5%	1.03%	-0.17%
	2	69.7%	69.5%	70.4%	69.6%	68.1%	0.52%	0.04%
	3	89.7%	89.4%	89.8%	89.6%	89.9%	0.12%	0.10%
	4	98.1%	98.3%	98.0%	98.0%	98.4%	-0.02%	0.00%
Varied across Periods	1	49.8%	47.6%	43.2%	54.2%	57.0%	-6.77%	0.03%
	2	70.1%	71.4%	60.1%	75.2%	77.3%	-6.02%	0.05%
	3	90.0%	83.8%	85.4%	94.6%	97.5%	-3.48%	0.03%
	4	98.3%	97.7%	97.4%	99.0%	99.2%	-0.47%	0.04%

Base Bias from Missing Data

The Simple Expansion column in Table I-6 represents the base bias from missing data. With simple expansion, the annual total PMT for the whole service was estimated by multiplying the total number of trips actually operated for all periods (13,000 in Table I-3) by the average PMT per trip from all valid APC data for all periods (not shown). The bias shown was calculated as the percent difference of this estimate from the true annual total PMT in Table I-3. Several observations can be made about the base bias:

- The base bias is higher when the data recovery rate varies across the periods than when the recovery rate is largely the same across the periods.
- The base bias is higher when the data recovery rate is lower.
- While it is not shown, the larger bias from varied recovery rates across periods largely disappeared when the true average PMT per trip was similar across the periods.

More generally, base bias from missing data can become significant when all of the following unfavorable conditions are true:

- Trip-level PMT varies significantly across service segments.
- Data recovery rates are extremely low.
- Data recovery rates vary significantly across the service segments.

Remaining Bias

The Stratified Expansion column in Table I-6 shows the remaining bias from missing data after the methodology in these guidebooks was applied to estimating annual total PMT for the whole service. Specifically, simple expansion was first applied to each trip on the schedule separately. That is, simple expansion was done 50 times. The annual total PMT for each trip on the schedule was then summed to get the annual total PMT for the whole service. The remaining bias represents the percent difference of this estimate from the true annual total PMT in Table I-3. It is clear that any remaining bias is negligible for all scenarios and ranges of data recovery rates considered. This means that the methodology in the guidebooks is highly effective in mitigating base bias from missing data.

5.05.4 Summary

The combined total of remaining random measurement errors and bias from missing data does not exceed 1.5% for all of the scenarios considered. Although these scenarios do not represent all possible cases, they cover cases with an overall data recovery rate as long as 50%, the number of daily trips as few as 4, the standard deviation of differences between APC and manual data as high as 25%, and significant differences in trip-level PMT and data recovery rates across time periods of the day. As a result, the allowable ±2.5% of error will be more than enough to cover any remaining random measurement errors and bias from missing data in estimated annual totals from this methodology. It is concluded with confidence that estimated annual totals from using the methodology in these guidebooks will meet the 10% precision and 95% confidence levels.

5.06 The Bottom Line

Traditional sampling ensures 10% precision and 95% confidence levels when it is done with care. The methodology in the guidebooks would mostly likely perform better, however, if the systematic measurement error in APC data does not exceed 5%. The systematic measurement error was within this range for seven of the nine agencies plotted in Figure I-1. In this case, the total error from this methodology would not exceed 5% + 2.5% = 7.5%. The sampling error from traditional sampling, on the other hand, is typically close to 10% because the corresponding sampling plan would have been designed to meet the 10% precision level.

Intentional sampling should not be used with APC data. It has no advantage over either of the other two approaches. All four types of errors are present. Random measurement errors stay significant due to the limited sample size. Systematic measurement errors may be controlled through NTD's benchmarking test but can still be as high as 5%, which is one half of the allowed 10% precision. Potential bias from missing data cannot be mitigated once a sample is drawn, resulting in an uncertain level of error. The potential bias was as high as 6.77% among the wide ranges considered in Table I-6 when estimation was done without stratification. These problems of intentional sampling get even worse when APCs do not cover the entire fleet.

What should agencies do estimate annual totals if they do not want to use all valid data and want to keep sampling?

- Raise the target precision in designing the sampling plan to increase their chance of meeting the 10% precision and 95% confidence levels. The effect of this strategy is limited, however. A target precision of 5% probably is not high enough because the combined total of the other three types of errors is likely to exceed 5% in most cases. An even higher target precision, such as 2%, may work in some cases but will require a sample size 25 times as large as required by the target precision of 10%. Such a large sample size would lose much of the general benefit of sampling over using all APC data.

- Return to traditional sampling with manual ridechecks.

The bottom line is that the only robust approaches to using APC data for estimating annual totals are those that use all valid APC data. Depending on both the coverage of APCs on a fleet and the degree of missing data, these robust approaches are:

- **Partial Coverage of APCs on the Fleet** – Use all valid APC data with a vehicle rotation plan that has been certified by a qualified statistician on a case-by-case basis.

- **Full Coverage of APCs on the Fleet** – The appropriate method also depends on the degree of missing data:
  - *Not Exceeding 2% - Proportional expansion of incomplete* counts without certification by a qualified statistician.
  - *Exceeding 2%* - Two options are available:
    - Stratified extrapolation that is certified by a qualified statistician on a case by case basis.
    - The methodology in these guidebooks that is certified by these guidebooks for general application as long as the agency meets the certification conditions, including passing the test of statistical equivalence.

6.00 Vanpool Reporting

Unique issues with vanpool data make it difficult to develop a generic methodology that is certifiable for the use of potentially all vanpool reporters. Therefore, a vanpool guidebook was not developed.

Unlike reporters of any other transit service, reporters of vanpool service rely on individual vanpools for passenger data. The problem for vanpool is complex:

- For those reporters relying on monthly reports of daily logs, a sizable portion of vanpools may fail to send their monthly reports, leading to missing their data for an entire month. It is also not uncommon for vanpools to not send their monthly reports for several months in a row. Reporters of vanpool have little power of enforcement over individual vanpools.
- Increasingly, reporters of vanpool service develop an online platform through which individual vanpools submit their daily activities. Such platforms typically assume that each member of a vanpool travels both directions on every day of a vanpool's normal commute days. This approach potentially can lead to an upward bias in the reported passenger activities.
- Reporters of other transit services almost always have records of vehicle operations throughout a year. Even when counting procedures miss ridership for some operated trips, such records of vehicle operations still allow them to determine the amount of service provided in terms of vehicle revenue miles and hours. Reporters of vanpool service, however, do not have direct knowledge of vehicle operations. If a vanpool fails to submit a monthly report, the reporter has no basis for knowing if it actually operated during the month. In other words, both usage and service data can be missing. Even with a sound way to extrapolate incomplete daily logs, the resulting total usage would be inconsistent with the incomplete service data.

These issues must be dealt with on a case-by-case basis. Any reporter of vanpool service should have these issues carefully examined relative to its actual conditions and have a strategy developed and certified by a qualified statistician.

Part II – Bus Guidebook

GUIDEBOOK:

Estimating Transit Usage through Extrapolating Incomplete APC Data for Bus Service

Table of Contents – Part II

1.00 Introduction

1.01 Objective

The objective of this bus guidebook is to detail a methodology for transit agencies to estimate transit usage for the National Transit Database (NTD) through extrapolating incomplete data from automatic passenger counters (APC) for bus service. Bus service includes all four fixed-route bus modes defined in the NTD: commuter bus (CB), motor bus (MB), bus rapid transit (RB), or trolley bus (TB). More specifically, the methodology:

- Divides the whole service into a set of mutually-exclusive service segments (e.g., a trip on the weekday schedule, a route, etc.).
- Estimates total transit usage for each service segment by extrapolating all but incomplete valid APC data for that service segment. For each segment, this is to multiply the transit usage per trip among all trips with valid APC data by the annual total number of trips actually operated for that service segment.
- Sums the estimates of segment-level total transit usage across all these service segments to get total transit usage for the whole service.

This methodology can be used to estimate all of the following measures of transit usage:

- Annual total unlinked passenger trips (UPT)
- Annual total passenger miles traveled (PMT)
- Annual average daily UPT by schedule type (weekday, Saturday, Sunday)
- Annual average daily PMT by schedule type
- Monthly total UPT

The methodology represents an alternative sampling technique for estimating annual total UPT and PMT. Agencies can use Appendix A of this guidebook as the document of certification by a qualified statistician. When an agency uses this methodology for estimating annual total UPT, annual total PMT, or both, it should report on the CEO Certification Form the method used as an "alternative sampling procedure that meets 95% confidence and ±10% precision levels as determined by a qualified statistician (estimated data)."

It is critical that this certification is conditional. These conditions are detailed later and are summarized here:

- The agency operates at least 1,000 vehicle trips annually (e.g., 4 trips daily).
- APCs must cover all doors of all vehicles of the fleet.
- The agency should have a continuous monitoring and maintenance program.
- The agency must recover valid APC data from at least 50% of all vehicle trips operated.
- The agency must pass a test of statistical equivalence for its APC data.
- The agency must use all valid APC data for estimation.

Agencies should not use this methodology if trips with missing APC data do not exceed 2% of all trips actually operated during an entire year. NTD rules allow proportional expansion at the system level for this case.

1.02. Directions

This guidebook is organized into the following sections:

- **Section 2.00, NTD Requirements** – summarizes the basic requirements that agencies must meet when reporting transit usage data to NTD and the additional requirements that agencies must meet when they use their APC data for reporting.
- **Section 3.00, Reporting Options** – discusses the various possibilities of estimating annual total UPT and PMT from using the methodology in this guidebook.
- **Section 4.00, Certification Conditions** – presents the requirements that agencies must meet to use the methodology in this guidebook as a pre-certified alternative sampling technique.
- **Section 5.00, Defining Service Segments** – presents guidance on how agencies should divide their whole service into service segments.
- **Section 6.00, Data and Estimation** – presents guidance on the basic data requirements for using this methodology and the detailed steps for using this methodology for estimating transit usage.
- **Section 7.00, Appendix A** – is the document of certification that agencies should keep on file for future reference.
- **Section 8.00, Appendix B** – lists several key definitions and the relevant formulas used in the test of statistical equivalence.

2.00 Reporting Requirements

2.01 Basic Requirements

To be eligible for the Urbanized Area Formula Grant Program (Section 5307), Full Reporters of bus service must report to NTD their data on the following measures of transit usage:

- Annual total unlinked passenger trips (UPT)
- Annual total passenger miles traveled (PMT)
- Annual average daily UPT by schedule type (weekday, Saturday, Sunday)
- Annual average daily PMT by schedule type
- Monthly total UPT

A Full Reporter operates more than 30 vehicles across all modes and types of service, operates fixed guideway and/or high-intensity busway, or is building a new mode of service. A fixed guideway is a roadway that agencies reserve at all times for transit vehicles. A high-intensity busway is a roadway that agencies reserve at some times, but not all, for transit use. As of May 2016, 84% of all CB and MB reporters were Full Reporters, and all RB and TB reporters were Full Reporters.

Data on annual total UPT and PMT and on monthly total UPT must include all services provided on all days during a corresponding year or month. Data on annual average daily UPT and PMT, on the other hand, should include services only on typical days and must exclude services on atypical days. Refer to the *NTD Policy Manual* on determining typical and atypical days. For completion, the NTD definition is provided here:

> Atypical days occur when an agency does not operate its normal, regular schedule; rather, it:
>
> - Provides extra service to meet demands for special events, such as conventions, parades, or public celebrations, or
> - Operates significantly-reduced service because of unusually bad weather (e.g., snowstorms, hurricanes, tornadoes, earthquakes) or major public disruptions (e.g., terrorism).

These data must be reported separately for each service the agency provides, i.e., each combination of modes and service types (directly-operated or purchased). UPT measures 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 origin to destination. PMT, on the other hand, measures the total distance traveled by all passengers.

Obtaining annual data on these usage measures, particularly PMT, can be costly to a transit agency. PMT for each one-way trip typically is calculated as the distance-weighted sum of passenger loads between consecutive stops. This calculation typically requires detailed data on passenger boarding and alighting activities at individual stops and distances between consecutive stops for individual one-way trips.

To obtain the annual total data for either UPT or PMT, a transit agency must choose one of two approaches. If available and reliable, it must report a 100% count. If a reliable 100% count of UPT or PMT is not available or is available but considered not reliable, it must estimate through sampling. The obtained estimate of annual total UPT or PMT from sampling must meet minimum 10% precision at 95% confidence. Data on annual average daily UPT and PMT and monthly total UPT are not subject to these statistical requirements.

When estimating through sampling, agencies must use either sampling plans that have been pre-approved by FTA or alternative sampling techniques that have been certified to meet the 10% precision and 95% confidence levels by a qualified statistician.

2.02 APC Requirements

2.02.1 APC Data

Using electronic infrared beams or mechanical treadle mats, APCs can convert electronic signals from beams or mats to counts of passengers as they board and alight transit vehicles. An APC includes a dedicated onboard computer, the APC analyzer, which converts sensor information into passenger counts. Each time a bus leaves a stop, the APC analyzer closes out a record and transmits the on-off counts at that stop to a general onboard computer, the in-vehicle logit unit. When coupled with stop location information, archived APC data can be post-processed to generate both UPT and PMT data potentially for every vehicle trip operated.

2.02.2 The Problem of Missing Data

When the entire fleet is equipped with APC units, theoretically, it is possible to obtain a 100% count of both annual total UPT and PMT from APC data. In reality, however, getting reliable 100% counts of UPT and PMT from APCs is difficult to achieve.

Part of the problem is that data may fail to be recovered from individual APC units, perhaps caused by general hardware malfunction, including sensors in the doorways failing to send any signals to the onboard computer, the onboard computer failing to convert sensor information into passenger on-off counts, and the onboard computer failing to transmit and store on-off counts. Another reason is incorrect sign-on (route, driver) information. The effect of this problem is reduced, to some degree, through a continuous monitoring and maintenance program.

Also, raw data may be successfully recovered from individual APC units, but serious errors exist in the recovered raw data, causing it to be discarded through the procedures of processing recovered raw data:

- In many cases, raw APC data are recovered but they fail to match in space and time to the actual service provided. Much of this inability comes from the difficulty of correctly identifying the end of a line and the complexity and unpredictability of end-of-line operations; in these cases, the recovered raw APC data are not usable.
- In many other cases, raw APC data are recovered and matched to actual service but errors in the recovered raw data fall out of range based on pre-set screening criteria for differences in on-off counts for individual blocks; in these cases, the matched raw APC data are not usable for purposes focusing on boarding counts. When matched raw APC

data fall out of range based on pre-set load screening criteria, the raw APC data are rejected for purposes focusing on load or passenger miles.

- In almost all cases in which raw APC data are recovered, matched, and accepted for further consideration, on-off counts will have errors. Unlike with data from manual ridechecks, these errors cannot be corrected easily in determining trip-level UPT and trip-level PMT because APCs collect on-off counts but not the additional information on the number of inherited passengers from the previous trip, number of continuing passengers to the next trip, and stop-to-stop loads.

  Although screening based on on-off differences avoids substantial errors in on-off totals, substantial errors still can develop in calculated loads and passenger miles, even with small errors in raw on-off counts due to a phenomenon called “drift.” An effective way to control drift is to parse a vehicle block’s data stream of automated counts at points of known load. Most routes have natural known-load points at terminals and layover points. To control drift through parsing, blocks are divided at known-load points into sections, usually single trips. On-off counts within each section then need to be balanced so the calculated loads match the known loads at each end of the section.

- For many situations, however, passengers stay onboard from one vehicle trip to the next without getting off, and load does not necessarily become zero at the end of a trip due to looping or interlining operations. In these cases, APC processing software must be able to locate artificial points of known load and control drift through parsing blocks into sections at these points. Alternatively, the operator must manually input the actual load at pre-specified points; otherwise, load and trip-level PMT can be seriously biased.
- For all cases, the bus driver may get on or off a vehicle multiple times, which further complicates the problem.

The data recovery rate for an APC system refers to the percent of all one-way trips operated during a specific period that provide valid data on trip-level UPT and PMT. This rate varies across different service segments of a given agency and across agencies. At the system level, it can be as low as 60–75%%.

2.02.3 Additional Requirements

An agency must first obtain the approval of FTA to use its APC data for reporting UPT and PMT data. As part of this approval, the agency must have its APC system certified initially by passing a benchmarking test. At the initial certification, the agency must have its APC system re-certified every fiscal year that is evenly divisible by three (i.e., 2019, 2022, 2025, etc.). To pass the benchmarking test, the agency must demonstrate that its APC data are within ±5% of paired manual data collected from a sample of vehicle trips that may not be selected at random.

3.00 Reporting Options

3.01 Basic Reporting Options

Under the basic reporting requirements, a transit agency has up to three basic methodological options for reporting both annual total UPT and PMT:

1. Report 100% counts for both UPT and PMT.
2. Report 100% UPT, but estimate PMT indirectly through estimating APTL via traditional sampling with manual ridechecks.
3. Estimate both UPT and PMT directly through estimating average UPT and average PMT per trip via traditional sampling with manual ridechecks.

For Option 2 or Option 3, sampling is done according to a pre-specified sampling plan. This sampling plan must be either pre-approved by FTA or be an alternative sampling technique that has been certified to meet the 10% precision and 95% confidence levels by a qualified statistician. With either type of sampling plan, sampling traditionally is done with human ridecheckers riding with the passengers of every sampled one-way trip. This approach to sampling is referred to as "traditional sampling with manual ridechecks."

With Option 2, the agency estimates PMT indirectly as the product of the 100% UPT with the average passenger trip length (APTL) estimated through sampling. Specifically, it first obtains both sample total UPT and sample total PMT from the same sample and then calculates the ratio of sample total PMT over sample total UPT to estimate the APTL.

For Option 3, on the other hand, the agency estimates PMT directly as the product of the total number of one-way trips operated during a full year by the average PMT per one-way trip estimated through sampling. Specifically, it first obtains sample total PMT and the number of one-way trips in the sample and then calculates the ratio of the sample total UPT over the sample size to estimate average UPT per one-way trip. It estimates UPT similarly with the same sample data.

Most agencies reporting CB and MB services get a 100% count of UPT from their electronic registering fareboxes. If they choose to report the available 100% UPT, both Options 1 and 2 may be used.

This is not the case for most agencies reporting RB service. Like modern light rail systems, the new bus rapid transit lines typically rely on off-board proof-of-payment fare collection, which requires passengers to purchase fare media off-board the bus and show proof-of-payment upon random inspection. Such fare collection systems do not produce 100% UPT counts. As a result, Option 1 and Option 2 would not be available to these RB agencies. Option 3 is available to every bus agency.

3.02 Reporting Options with APC Data

If an agency has installed APC units on some or all of the vehicles in its fleet and has obtained FTA certification of its APC system for NTD reporting, it is likely unwilling to continue using traditional sampling with manual ridechecks for either Option 2 or Option 3. Instead, it is likely to choose intentional sampling in one of two forms:

- **Pre-Sampling** takes place before a trip is actually operated. This is common for an agency with a relatively small share of its fleet equipped with APC units. Pre-sampling is necessary so that APC-equipped buses can be assigned to the sampled trips and is applicable to other cases. Pre-sampling must account for the expected data recovery rate in the sample size used. As with traditional sampling, pre-sampling selects trips from the full list of all trips to be operated.
- **Post-Sampling** occurs after a trip has been operated and the raw data have been processed. This is common for an agency with a full coverage of APC units on all doors of all buses. Different from either traditional sampling or pre-sampling, post-sampling selects trips from all trips with valid APC data.

The term "intentional sampling with APCs" is used to contrast it with traditional sampling with manual ridechecks and with using all valid APC data, which is a form of "full-population" sampling and is the focus of this guidebook. With intentional sampling with APCs, an agency has two additional reporting options:

4. Report 100% UPT from a non-APC source, but estimate the APTL through intentional sampling with APCs; this is another version of Option 2.
5. Estimate both UPT and PMT directly through intentional sampling with APCs; this is another version of Option 3.

3.03 Reporting Options with All Valid APC Data

If an agency has obtained FTA certification of its APC system for NTD reporting and also has full coverage of APC units on every vehicle in its fleet, additional reporting options are open to it. These additional options are made available from using all valid APC data through the methodology included in this guidebook. These additional reporting options are:

6. Report 100% UPT from a non-APC source, but estimate the APTL through extrapolating all valid but incomplete APC data on both PMT and UPT; this is another version of Option 2.
7. Report 100% UPT from a non-APC source, but estimate PMT directly through extrapolating all valid APC data on PMT; this is a new option.
8. Estimate both UPT and PMT through extrapolating all valid but incomplete APC data with the methodology in this guidebook; this is another version of Option 3.

3.04 Choices

3.04.1 Having 100% UPT from a Non-APC Source

When available from a non-APC source, it is largely up to an agency to decide whether it reports this 100% UPT or not.

Reporting 100% UPT

If the agency reports the available 100% UPT from a non-APC source, it has four potential options from which to choose for estimating PMT:

- Option 2 – estimate the APTL through traditional sampling with manual ridechecks.
- Option 4 – estimate the APTL through intentional sampling with APCs.
- Option 6 – estimate PMT directly using the methodology in this guidebook through extrapolating all valid APC data on UPT and PMT.
- Option 7 – estimate PMT directly using the methodology in this guidebook through extrapolating all valid APC data on PMT only.

Traditional sampling with manual ridechecks is largely irrelevant for this choice. Once FTA certification of an APC system for NTD reporting has been obtained, an agency will not continue to use traditional sampling due to its high labor cost.

This Methodology vs. Intentional Sampling

This choice is between Options 4 and 6; this methodology (Option 6) will always lead to better estimates than intentional sampling with APCs (Option 4):

- These two options suffer the same problem of missing data to the same degree.
- For both options, the problem of missing data can lead to two undesirable outcomes for certain segments of the service. These service segments are either no longer represented or disproportionally represented in the valid APC data. If not properly mitigated, both outcomes potentially can result in bias in the estimates.
- The potential bias is likely to be realized with intentional sampling. With intentional sampling, the problem of missing data is significantly more likely to lead to certain service segments not being represented or being disproportionately represented. Any bias for a service segment resulting from not being represented in the valid data cannot be mitigated. Even in the case of disproportionate representation, estimation for such service segments is highly unreliable due to the small remaining sample size for each segment.
- The potential bias can be largely mitigated by using all valid APC data through this methodology. The problem of missing data still can lead to certain service segments being disproportionately represented, but is highly unlikely that certain service segments will not be represented. When a service segment is disproportionately represented, the significantly large amount of valid data allows relatively reliable estimation for this segment.
- The large amount of all valid APC data not only can mitigate potential bias from the problem of non-random missing data, but also will result in estimates with much better precision than those from intentional sampling.

APTL vs. Average PMT

This choice is between two different ways of using the methodology in this guidebook. The methodology could be used to estimate both UPT and PMT and the APTL as the ratio of these two estimates for Option 6. In other words, Option 6 expands the APTL from this methodology by the 100% UPT from another source. On the other hand, the methodology could be used to estimate only PMT for Option 7. That is, Option 7 expands the average PMT per trip from this methodology by the total number of one-way trips actually operated.

If the 100% UPT from another source were the true value, an agency should choose Option 6. That is, it should report the annual total PMT that is estimated as the product of the 100% UPT from another source and the APTL from this methodology. The precision reached by this APTL-based estimate will most likely to be better than the average-based estimate. This is simply because APTL varies much less than does PMT.

If the 100% UPT count understates the true value to some degree, the choice is not immediately clear. If the APC data have similar magnitudes of error in APTL versus in average PMT, the agency should choose Option 7. This is because the expansion factor for Option 7 is the number of trips operated and is error-free. The expansion factor for Option 6, on the other hand, is the 100% UPT and understates the true value. As discussed later related to testing the statistical equivalence between APC data and paired manual data, the magnitude of error can be estimated with the paired APC and manual data for both APTL and average PMT. If the magnitude of error is smaller in APTL than in average PMT, the choice is difficult.

Not Reporting 100% UPT

With 100% UPT from another source, an agency still can choose to use the methodology in this guidebook to get an estimate of annual total UPT. The choice in this case is whether it should report the 100% count or the APC-based estimate to the NTD:

- FTA requires that agencies report 100% UPT when it is available and reliable. It should report the APC-based estimate from this methodology if the 100% count is considered not reliable. However, “reliability” is not defined; it is up to the agency to determine whether the 100% count is reliable. If the 100% UPT understates the true value by more than 10%, for example, the agency still can consider it to be available and reliable and, therefore, report it. To meet the 10% precision and 95% confidence levels in this case, however, the agency should report the APC-based estimate.
- From the perspective of the agency, another consideration is that the APC-based estimate is likely to be larger than the 100% count if the latter understates the true value to some degree. In this case, the agency has an incentive to report the APC-based estimate.

3.04.2 No 100% UPT

If an agency does not have a 100% UPT count, it has three potential options for estimating both UPT and PMT:

- Option 3 – Estimate through traditional sampling with manual ridechecks
- Option 5 – Estimate through intentional sampling with APCs
- Option 8 – Estimate through using the methodology in this guidebook

Similar to the choice among Options 2, 4, and 6 discussed earlier, the real choice is between Options 5 and 8. In addition, Choice 8 will always give better estimates than Option 5.

4.00 Certification Conditions

4.01. Introduction

To use the methodology in this guidance as a certified alternative sampling technique, an agency must meet all of the following conditions:

1. The agency operates at least 1,000 vehicle trips annually (e.g., 4 trips daily).
2. APCs must cover all doors of all vehicles of the fleet.
3. The agency should have a continuous monitoring and maintenance program.
4. The agency must recover valid APC data from at least 50% of all vehicle trips operated.
5. The agency must pass a test of statistical equivalence for its APC data.
6. The agency must use all valid APC data for estimation.

4.02 Coverage, Operations, and Data Recovery

The agency must have APC units installed at every door of every bus in the fleet. To ensure a minimum of 500 vehicle trips with valid APC data during a year, the agency is also required to operate at least 1,000 vehicle trips annually and to recover valid APC data from at least 50% of all vehicle trips operated. It is rare but possible for an agency to operate a weekday CB service with a single route that has morning two trips and two afternoon trips.

4.03 Continuous Monitoring

The agency must have a process in place to monitor and maintain individual APC units and the APC system continuously. An effective process, for example, would immediately notify the agency when a trip failed to acquire any raw APC data on the on-off activities for the trip and would indicate the source of the equipment failure problem. Once equipment failure occurs, the same effective process would have an ongoing partnership with vendors for quick maintenance and repair. Reducing trips with missing data due to equipment failure benefits not only NTD reporting but also internal service and other planning efforts within an agency.

4.04 Use of All Valid APC Data

The agency must be able to use a significant amount of APC data at a disaggregated level. This is especially relevant for a large bus system. One effective approach to having this ability by an agency is that it should have an annual database of trip-level information for all one-way vehicle trips actually operated during a full year. This database would accumulate such information as individual trips are being operated and as the raw APC data are being processed from the first day to the end of the year.

4.05 Test of Statistical Equivalence

4.05.1 Introduction

To use the methodology in this guidebook as a certified alternative sampling procedure, a transit agency must conduct a test of statistical equivalence for its APC data and must pass the test. The agency must collect paired APC and manual data from the same sample of trips selected at random during a period that is at least one week long.

- To pass the test of statistical equivalence, the agency must demonstrate that its APC data are statistically equivalent to the paired manual data within ±7.5% at the 95% confidence level.
- This condition of statistical equivalence is to ensure that the estimate of annual total transit usage meets the 10% precision and 95% confidence levels. The equivalence bounds are set at ±7.5% to allow up to ±2.5% of other types of errors in the estimates.
- Requiring the paired APC and manual data coming from a random sample is necessary to access the level of confidence achieved for the test of statistical equivalence.
- The minimum duration of one week for the testing period is specified to minimize the burden on transit agencies while non-seasonable variation is being captured.

If the agency plans to use the pre-certified alternative sampling procedure only for estimating annual total UPT, it should focus the test on the trip-level UPT data from the testing sample. Similarly, if it plans to use the pre-certified alternative sampling procedure only for estimating annual total PMT, it should focus the test on the trip-level PMT data from the testing sample. If the agency wants to use the pre-certified alternative sampling procedure for estimating both annual total UPT and PMT, it should test statistical equivalence for both UPT and PMT.

During the procurement process for an APC system, the agency may want to consider whether the RFP should include the requirement of having the new APC units tested for statistical equivalence and whether its acceptance of the APC system is conditional on it passing the test of statistical equivalence. If the agency plans to use this methodology as a pre-certified alternative sampling procedure, it should include testing statistical equivalence and passing the test as system acceptance requirements.

This section provides detailed guidance for a transit agency to test statistical equivalence on its own in an Excel environment.

4.05.2 Selection of Trips for Testing

The selection of trips for testing statistical equivalence consists of several aspects, including:

- **Duration** – The shortest duration for selecting trips is a full week to capture most non-seasonal variations. It is better to cover a much longer duration so that both seasonal and non-seasonal variations are accounted for.
- **Unit of Trips** – The appropriate unit of trips is one-way vehicle trips.
- **Number of Trips** – Refer to Section 4.05.6, Sample Size for Equivalence Testing, for guidance on the appropriate number of trips to be selected.
- **Method of Selection** – The trips must be selected at random from all scheduled trips during the chosen period. One commonly-used approach to selecting a random sample from a sampling frame of trips is the Excel random number function. Assuming a testing sample of 50 trips, this approach involves the following steps:
    1. Starting at the 2nd row and column B in a blank Excel worksheet, list all one-way trips to be operated during the selected data period. The trips on a weekday

schedule should be repeated for each weekday in the period. Similarly, the trips on the Saturday schedule should be repeated for each Saturday in the period. This list should reflect any case in which the Sunday schedule is operated on a weekday or Saturday holiday.

2. In cell A2, type "=rand()" and press “Enter.” This will produce a random number (decimal) between 0 and 1.
3. Copy and paste the formula in cell A2 down column A to the bottom of the list. One way to do this is to hover the mouse pointer over the lower right corner of cell A2 until it changes to a thin black cross, and then double-click. This will automatically copy the formula to the bottom of the list.
4. Copy all data in column A. Use “Paste Special” to paste values over the random formula.
5. Sort by column A. Highlight the entire table of data, select the data, and sort from the menu. In the “Sort by” window, select column A or the heading for that column, if applicable. Select “Ascending” and click on “OK.”
6. The trips in the first 50 rows of the sorted list will be the random sample.

4.05.3 Processing and Extraction of APC Data

Once the trips have been selected at random during the data period, the agency should make sure the APC units for each selected trip are in working conditions before each trip takes place. This will increase the data recovery rate for the testing sample.

The agency should make sure that the extracted APC data are processed from the raw APC data with procedures identical to those used for processing raw APC data in general.

The processed APC data for each trip must be extracted from the database of all processed APC data. At a minimum, the extracted APC data must have the trip-level UPT and PMT for each trip in the sample. The agency must make sure that the extracted APC data are for the correct trips. The results are trip-level UPT and PMT data for each of the trips in the sample to be paired with trip-level UPT and PMT from the manual data.

4.05.4 Collection and Processing of Manual Data

Collection

Manual data commonly are collected by sending human ridecheckers to ride with passengers on the selected trips. It is also possible to collect manual data through watching recorded videos if enough cameras are properly positioned to capture passenger activities through all doors of the vehicles. For either method of data collection, it is critical that manual data collected for the selected trips are correctly identified so they can be correctly matched with the APC data extracted for these trips.

To ensure that the manual data used in testing represent the true value:

- When using ridecheckers, agencies are highly encouraged to use one ridechecker per door, especially for selected trips that are expected to be heavily used.

- If using recorded videos, agencies should make sure that enough cameras are adequately positioned to capture all passenger activities throughout a bus. In addition, recorded videos may be kept for a limited time; agencies should make sure that data collection from the recorded videos is done before they are deleted.
- When using either method of collecting manual data, include extra columns on the field form for the following passenger activities for each sampled trip:
  - Inherited passengers from previous trip at first stop
  - Passengers continuing to next trip at last stop
  - Passenger loading between each consecutive pair of stops

Although proper processing can correct minor errors in the raw manual data for many cases, these additional data items can help in other cases with bigger errors in the raw manual data.

It is in the best interest of transit agencies to minimize errors in the manual data, as it is likely that errors in the manual data would increase variation in the difference between APC and manual data. This increase in variation not only will mean larger sample sizes for equivalence testing in the future, but also will reduce the chance of the APC data passing the test of statistical equivalence.

Processing

For either method, trip-level UPT and PMT should be calculated for each selected trip only after the boarding and alighting activities have been properly balanced. When on-off activities are properly balanced for a given trip, trip-level boardings and trip-level alightings are equal and passenger load at the last stop is zero. In balancing on-off activities for each trip:

- Passengers from the previous trip at the first stop should be counted as boardings at the first stop.
- Passengers continuing to the next trip at the last stop should be counted as alightings at the last stop.

If using stop-level on-off activities plus passengers from the previous trip and passengers continuing to the next trip fails to balance the trip, the agency should compare the calculated passenger load from these passenger activities with the passenger load recorded in the field.

Having the additional data on passenger loads, passengers from the previous trip, and passengers continuing to the next trip has at least three advantages:

- A high degree of accuracy in the processed manual data through using the additional data in balancing passenger on-off activities and calculating passenger loads.
- Less time and effort to balance passenger on-off activities and calculate passenger loads with the additional data items than without them.
- Fewer sampled trips whose passenger on-off activities cannot be balanced. Having the additional data items significantly reduces the chance of passenger on-off activities not being balanced for individual sampled trips. Sampled trips with unbalanced passenger on-off activities must be excluded for testing statistical equivalence.

The results are trip-level UPT and PMT data for each of the trips in the testing sample to be paired with the corresponding APC data.

4.05.5 Statistical Testing in Excel

Once the trip-level UPT and PMT are ready for both APC and manual data, the agency can easily test their statistical equivalence in Excel. An example is used to illustrate the detailed steps for the test. The example assumes that only PMT data are being tested.

Testing Equivalence in UPT or PMT

Example

The steps for such a test are discussed in the context of actual paired APC and manual data on PMT for a small sample of one-way bus trips from a real transit agency. To put these steps in concrete terms, the screen capture in Figure II-1 is used for illustration.

	A	B	C	D	E	F	G
1	1) APC	2) Manual	3) Diff		Measures	Excel Formulas and Fuctions	Values
2	65.3	145.2	-79.9				
3	55.9	92.3	-36.5	4)	Sample Size	=COUNT(C2:C34)	33
4	136.6	125.1	11.5				
5	107.4	98.0	9.4	5)	Mean of manual data	=AVERAGE(B2:B34)	318.5
6	377.7	367.0	10.7				
7	400.7	330.4	70.3	6)	Mean difference	=AVERAGE(C2:C34)	3.23
8	131.4	109.9	21.5				
9	124.6	120.0	4.6	7)	Standard deviation	=STDEV(C2:C34)	39.25
10	350.2	360.5	-10.3				
11	506.1	429.9	76.2	8)	Standard error	=F9/SQRT(F3)	6.83
12	214.8	241.2	-26.4				
13	339.3	337.7	1.6	9)	Relative mean difference	=F7/F5	1.01%
14	566.3	567.6	-1.3				
15	132.1	142.1	-10.0	10)	Relative standard deviation	=F9/F5	12.32%
16	231.2	262.9	-31.7				
17	417.2	364.1	53.1	11)	Relative standard error	=F11/F5	2.14%
18	297.1	278.2	18.9				
19	342.2	308.9	33.3	12)	t_critical value	=TINV(2*0.05,F3-1)	1.69
20	477.0	409.0	68.1				
21	4.2	4.2	0.0	13)	CI lower bound	=F13-F19*F17	-2.62%
22	368.1	351.2	16.9				
23	363.0	325.4	37.6	14)	CI upper bound	=F13+F19*F17	4.65%
24	404.7	404.0	0.7				
25	493.3	425.5	67.8	15)	CI Lower bound > -7.5%?	=IF(F21>-7.5%,"Yes","No")	Yes
26	352.2	352.2	0.0				
27	409.4	439.6	-30.2	16)	CI Upper bound < 7.5%?	=IF(F23<7.5%,"Yes","No")	Yes
28	586.6	585.9	0.7				
29	490.1	565.7	-75.6	17)	Equivalent?	=IF(AND(F25="Yes",F27="Yes"),"Yes","No")	Yes
30	274.0	339.7	-65.7				
31	470.0	512.6	-42.6				
32	361.3	341.2	20.0				
33	393.3	402.1	-8.8				
34	375.0	372.0	3.0				

Figure II-1. Steps for Equivalence Testing in Excel

The paired APC and manual data are entered into columns A and B; their difference is in column C. For subsequent calculations and testing, Column E lists the relevant statistical measures, Column F shows the Excel formulas and functions, and Column G presents the calculated values for the example data used. For easy identification, the functions are in bold. The reference to cells in the formulas assumes that transit agencies will perform the computations in column F. The steps are numbered from 1) through 17). The 95% confidence level is reflected through Step 12, and the equivalence bounds of ±7.5% are reflected in Steps 15) and16).

Testing Procedures

A transit agency can conduct equivalence testing with the following steps:

1) Enter the processed APC data in column A. For the example, the APC data are in range A2:A34.
2) Enter the processed manual data in column B. For the example, the APC data are in range B2:B34. It is critical that the APC and manual data in the same row come from the same trip.
3) Subtract the manual data from the APC data for each trip in column C.
4) Determine the size of the entered sample by counting the number of rows in range C2:C34 with the **COUNT** function in cell F3.
5) Calculate the mean of the manual data in cell F5 with the **AVERAGE** function.
6) Calculate the mean of the APC and manual differences in cell F7 with the **AVERAGE** function.
7) Calculate the standard deviation of the APC and manual differences in cell F9 with the **STDEV** function.
8) Calculate the standard error of the difference in cell F11 as the ratio of the standard deviation in cell F9 from Step 7 over the square root of the sample size in cell F3 from Step 4 with the **SQRT** function.
9) Calculate the relative mean difference between the APC and manual data in cell F13 as the ratio of the mean of APC and manual differences in cell F7 from Step 6 over the mean of the manual data in cell F5 from Step 5.
10) Calculate the relative standard deviation in cell F15 as the ratio of the standard deviation in cell F9 from Step 7 over the mean of the manual data in F5 from Step 5.
11) Calculate the relative standard error in cell F17 as the ratio of the standard error in cell F11 from Step 8 over the mean of the manual data in cell F5 from Step 5.
12) Determine the critical value at the 95% confidence level for the sample size in cell F3 with the **TINV** function.
13) Determine the lower bound of the confidence interval for equivalence testing in cell F21.

14) Determine the upper bound of the confidence interval for equivalence testing in cell F23.

15) Determine if the lower bound is greater than -7.5% in cell F25

16) Determine if the upper bound is small than 7.5% in cell F27.

17) The agency passes the test of statistical equivalence if “Yes” appears in cell F29. It fails to pass it if “No” is in cell F29.

Testing Equivalence in APTL

To conduct an equivalence test for APTL, the test must be completed for both UPT and PMT. If the agency passes the tests for UPT and PMT, it passes the test for APTL. If the agency fails the test for UPT, PMT, or both, it does not pass the test for APTL.

4.05.6 Sample Size for Equivalence Testing

The proper sample size for testing the statistical equivalence of paired APC and manual data depends on the relative standard deviation in the difference of APC and manual data and the expected recovery rate of valid APC data. Table II-1 shows the minimum sample size for each of five ranges of relative standard deviation and 5% increments of the expected recovery rate of valid APC data.

Table II-1. Sample Size for Equivalence Testing

Expected Recovery Rate of Valid APC Data	Minimum Sample Size in One-Way Trips by Range of Relative Standard Deviation				
	≤5%	>5% & ≤10%	>10% & ≤15%	>15% & ≤20%	>20%
≥95%	7	25	55	98	152
≥90% & <95%	7	26	58	103	160
≥85% & <90%	7	28	61	109	170
≥80% & <85%	8	29	65	116	180
≥75% & <80%	8	31	70	123	192
≥70% & <75%	9	33	75	132	206
≥65% & <70%	9	36	80	142	222
≥60% & <65%	10	39	87	154	240
≥55% & <60%	11	42	95	168	262
≥50% & <55%	12	47	104	185	288

If an agency has a set of paired APC and manual data previously collected, it should compute the relative standard deviation as illustrated in Figure II-1, determine the range in Table II-1 into which its value falls, determine the expected data recovery rate, determine the range of recovery rates in Table II-1, and select the proper sample size accordingly from Table II-1. For the example in Figure II-1, the relative standard deviation for PMT is 12.93%, which falls between 10% and 15%. Therefore, the proper sample size for this case would be 55 trips if the recovery rate is at least 95%, but it would increase to at least 70 if the recovery rate does not exceed 80%.

Potential sources of paired APC and manual data are data collected during the procurement process and paired APC and manual data from a previous effort of equivalence testing.

An agency that has failed a previous equivalence test may want to select a sample size that is larger than what their relative standard deviation suggests from Table II-1. The larger the sample size, the greater the chance of a successful equivalence test.

The relative standard deviation generally differs between UPT and PMT and, in fact, is larger for PMT than for UPT. Consider the following to select the sample size from Table II-1:

- For testing UPT only, use the relative standard deviation for UPT.
- For testing PMT only, use the relative standard deviation for PMT.
- For testing both UPT and PMT, use the relative standard deviation for PMT.

For many cases, however, an agency may not have its own previously-paired APC and manual data to determine the relative standard deviation. Limited evidence indicates that the relative standard deviation typically does not exceed 15% and the expected recovery rate typically is at least 70%. For these cases, a reasonable sample size would be 75.

4.05.7 Future Testing of Statistical Equivalence

Once the agency has passed this test, it stays valid until the agency fails FTA's benchmarking test. The benchmarking test is required in every fiscal year that is evenly divisible by three (i.e., 2019, 2022, 2025, etc.). If that happens, the agency should use the paired APC and manual data from the last test of statistical equivalence to determine the relative standard deviation and use Table II-1 for the proper sample size to use for the new test of equivalence.

5.00 Defining Service Segments

5.01 Introduction

A critical element of this methodology is to divide an agency's whole service during an entire year into individual pieces, referred to as service segments. Each service segment consists of a set of individual one-way vehicle trips actually operated during the year. The individual one-way vehicle trips in all defined service segments add up to all trips actually operated during the year. The following are examples of service segments meeting this requirement:

1. Individual trips on weekday, Saturday, and Sunday schedules
2. Individual trips on weekday schedules, individual Saturday routes, and individual Sunday routes
3. Individual routes
4. Individual routes by schedule type
5. Weekday peaks, weekday midday, weekday night, Saturday, and Sunday

To facilitate the estimation of annual average daily UPT and PMT by schedule type, schedule type must be one attribute of the service segments defined. For example 3, the service segments defined as individual routes are not identified by schedule type and, hence, do not meet this requirement.

5.02 Data Variation

For the purpose of this methodology, one objective in selecting a level of segmentation and defining service segments is to minimize variation within each service segment. To better meet this objective in defining service segments, it is important to understand the nature of variation, which involves several aspects:

- Trip-level UPT and PMT for a given route typically are higher for trips in the peak direction, during peak periods, and on weekdays. Trip-level UPT and PMT also can vary significantly for trips across routes. In addition, trip-level PMT typically is higher for longer routes than for shorter routes. For routes of similar length, trip-level PMT typically is higher for routes whose passengers are more likely to travel from one end to the other than for other routes.
- Trip-level PMT varies more than trip-level UPT in most cases. The additional variation in trip-level PMT results from the variation in the distance traveled by individual passengers.
- APTL varies much less than UPT and PMT across individual trips.
- APTL typically does not vary significantly across individual one-way trips for a given route; however, it can vary significantly across routes of different length.

With this understanding of the nature of variation, selecting individual routes as service segments would not work well for weekdays if an agency wants to estimate annual total UPT or PMT from this methodology. Individual routes, however, would work well as service segments if an agency wants to estimate APTL from this methodology. Individual routes also may work as service segments for weekend days. For services with just a few routes, however, using routes as service segments would be too aggregated.

5.03 Adequate Sample Size

Another objective in selecting the level of segmentation and defining service segments is to make sure that each defined service segment has an adequate number of trips with valid APC data.

An agency also should understand the importance of the data recovery rate of an APC system in defining service segments. For a service segment consisting of 40 trips actually operated during a year, the number of trips with valid APC data would be only 20 trips if the rate is 50%, 28 trips if the rate is 70%, and 36 trips if the rate is 90%. Although an agency can never be sure in advance of the actual recovery rate for the whole APC system or for individual segment segments, it can keep track of missing data for individual trips actually operated throughout the year. At a minimum, it can get a good indication from previous years.

Also important to understand is how the number of one-way trips in individual service segments depends on the level of segmentation:

- A weekday trip on a schedule for an entire year may be operated as many as 260 (=52*5) times during the year.
- A Saturday trip on a schedule may be operated 52 times, whereas a Sunday trip on a schedule can be operated 52 times plus weekdays with Sunday service.
- These numbers drop quickly, however, if the agency changes its schedule either quarterly or bi-monthly.
  - A weekday trip on a schedule may be operated only 65 times during the year with quarterly service changes and only 43 times with bi-monthly service changes.
  - A Saturday trip on a schedule may be operated only 13 times a year with quarterly service changes.

A determining factor is that each service segment defined should have at least 10 one-way trips with valid APC data. To ensure this target sample size and to account for the uncertainty in the actual recovery rate for each service segment, **the minimum number of trips actually operated should be at least 20 for any service segment defined for using this methodology**. For example, service segments defined by trips on a weekday schedule would meet this requirement even if service changes occur bi-monthly. Service segments defined by trips on weekend schedules with quarterly or bi-monthly service changes, however, would not meet this requirement.

5.04 Convenience

The third objective in selecting the level of segmentation and defining service segments is the level of convenience in using the methodology. Service segments defined with generally fewer attributes would be better than those defined with more attributes. Individual routes and individual trips on schedules are two examples that would be considered to meet this convenience requirement. Service segments defined as individual routes by direction, time of day, and day of week would be complex to use for this methodology. However, having an adequate sample size and minimizing variation are more important than relative convenience.

5.05 Frequency of Segmentation

It is possible to modify the service segments defined and used for the previous year using data from the current year. Estimation using this methodology is done after the year has ended and all raw APC data have been collected and processed. The agency can examine the pattern of trips with missing data, evaluate the service segments defined previously, and modify if needed so that each new service segment has an adequate number of trips.

Having a new set of service segments every year is not suggested. Instead, the agency should study the pattern of trips with missing data, define a set of service segments accordingly, and use them for this methodology for a number of years. If its monitoring process indicates that the pattern of trips with missing data has changed significantly from the time when these service segments were defined, it should modify its service segments according to the new pattern.

5.06 Suggested Segmentation

An agency may choose the level of segmentation and define the service segments accordingly. To help agencies that are new to the methodology, a specific set of service segments are suggested:

- Individual trips on the weekday schedule
- Individual routes for Saturday
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight
- Individual routes for Sundays
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight

This suggested segmentation meets the requirement of being mutually-exclusive and adding up to the whole service, uses schedule type as one attribute, and balances all three objectives for defining service segments: minimizing variation, adequate sample size, and convenience.

Column A in Figure II-2 lists all service segments for a hypothetical service as an illustrative example. In this case, there are 26 service segments, including 1 Saturday route, 1 Sunday route, and 20 trips on the weekday schedule. No service changes occur during the year in this example. The service segments in this example follow the suggested segmentation approach.

	A	B	C	D	E	F	G	H	I	J	K
1	**Define Service Segments**	**Schedule Type**	**Time of Day**	**Count Annual Trips Operated**	**Count Annual Trips with Valid APC Data**	**Aggregate Annual UPT from Valid APC Data**	**Aggregate Annual PMT from Valid APC Data**	**Calculate Average UPT**	**Calculate Average PMT**	**Estimate Annual UPT**	**Estimate Annual PMT**
2				1	2	3	3	4	4	5	5
3	1	Saturday	Morning	300	236	10,695	30,235	45.3	128.1	13,596	38,434
4	2	Saturday	Afternoon	350	309	35,000	98,943	113.3	320.2	39,644	112,071
5	3	Saturday	Night	300	206	11,215	31,704	54.4	153.9	16,332	46,171
6	4	Sunday	Morning	300	257	7,299	21,748	28.4	84.6	8,521	25,387
7	5	Sunday	Afternoon	350	309	11,905	35,470	38.5	114.8	13,484	40,176
8	6	Sunday	Night	300	221	5,333	15,891	24.1	71.9	7,240	21,571
9	7	Weekday	Trip 1	255	228	6,623	16,470	29.0	72.2	7,407	18,420
10	8	Weekday	Trip 2	255	168	3,623	13,132	21.6	78.2	5,499	19,932
11	9	Weekday	Trip 3	255	219	22,099	50,918	100.9	232.5	25,732	59,288
12	10	Weekday	Trip 4	255	185	13,756	32,330	74.4	174.8	18,961	44,563
13	11	Weekday	Trip5	255	185	3,220	11,751	17.4	63.5	4,439	16,198
14	12	Weekday	Trip 6	255	183	8,523	40,551	46.6	221.6	11,876	56,505
15	13	Weekday	Trip 7	255	185	7,885	22,295	42.6	120.5	10,868	30,731
16	14	Weekday	Trip 8	255	155	24,444	66,361	157.7	428.1	40,215	109,175
17	15	Weekday	Trip 9	255	156	10,870	35,964	69.7	230.5	17,768	58,787
18	16	Weekday	Trip 10	255	182	20,548	68,849	112.9	378.3	28,790	96,465
19	17	Weekday	Trip 11	255	155	13,668	34,377	88.2	221.8	22,487	56,556
20	18	Weekday	Trip 12	255	166	8,290	27,886	49.9	168.0	12,735	42,837
21	19	Weekday	Trip 13	255	225	8,324	28,682	37.0	127.5	9,434	32,506
22	20	Weekday	Trip 14	255	174	11,173	26,945	64.2	154.9	16,374	39,489
23	21	Weekday	Trip 15	255	238	1,381	5,525	5.8	23.2	1,480	5,919
24	22	Weekday	Trip 16	255	241	6,579	27,855	27.3	115.6	6,961	29,473
25	23	Weekday	Trip 17	255	190	17,857	65,801	94.0	346.3	23,966	88,312
26	24	Weekday	Trip 18	255	157	11,429	21,995	72.8	140.1	18,562	35,724
27	25	Weekday	Trip 19	255	205	3,772	46,707	18.4	227.8	4,693	58,098
28	26	Weekday	Trip 20	255	191	5,374	93,367	28.1	488.8	7,175	124,653
29											
30	**Estimate Annual Total UPT and PMT for the Whole Service** (6)									394,238	1,307,441
31											
32	**Weekday Total UPT and PMT** (7)									295,422	1,023,630
33	**Saturday Total UPT and PMT** (7)									69,572	196,676
34	**Sunday Total UPT and PMT** (7)									29,245	87,134

Figure II-2. Example of Calculations for Estimation Procedure

6.00 Data Requirements and Estimation Procedure

6.01 Data Requirements

Ideally, an agency should have an annual database of trip-level information for all one-way vehicle trips actually operated during a full year. This database would accumulate such information as individual trips are being operated and as the raw APC data are being processed from the first day to the end of the year. At a minimum, this database should have the following data items for each one-way trip actually operated:

- Date on which the trip was operated
- Whether it was an atypical day
- Route
- Schedule type
- Start time
- Service segment to which the trip belongs
- Whether raw APC data were returned
- Whether valid APC data are available
- Trip-level UPT if valid APC data are available
- Trip-level PMT if valid APC data are available

Often, transit agencies operate their Sunday schedule on holidays that fall on Monday through Saturday. An agency should include the data for these holidays under the day for the schedule that it operates (e.g., if operating on a Sunday schedule for a holiday on a Tuesday, the data would be included under Sunday).

These minimum data items are based on the suggested definition of service segments and the need for estimating monthly total UPT and annual average daily UPT and PMT by schedule type. Additional data items may be included for alternative ways of defining service segments. Trip direction would need to be added, for example, if an agency uses it in defining its service segments.

6.03 Estimation Procedure for Annual Totals

Once service segments are defined and the annual database is ready, the agency is ready for estimation using this methodology. As detailed below, the procedure for estimating annual totals using this methodology consists of six steps. The first three steps extract data from the database of trips actually operated; the next two steps involve calculations for each service segment separately; the last step aggregates the results across all service segments. These steps are illustrated with the example in Figure II-2.

1. **Count Annual Trips Operated** – This step counts the annual number of one-way vehicle trips actually operated during an entire year for each service segment. This would simply count the number of rows in the database for each service segment for the whole year. Column D of Figure II-2 shows these numbers for the example.
2. **Count Annual Trips with Valid APC Data** – This step counts the annual number of one-way vehicle trips actually operated for which valid trip-level data on UPT and PMT are

available from the APC system during an entire year. This would count the number of rows in the database with valid APC data for each service segment for the whole year. Column E of Figure II-2 adds these trips with valid APC data for each service segment.

3. **Aggregate Annual UPT and PMT from Valid APC Data** – This step aggregates the trip-level UPT and PMT across all one-way vehicle trips actually operated with valid APC data during an entire year. This requires summing the trip-level UPT and PMT across all rows for each service segment. Columns F and G of Figure II-2 show the aggregated annual UPT and PMT from valid APC data, respectively, for each service segment.
4. **Calculate Average UPT and PMT** – This step calculates the average UPT per one-way vehicle trip and average PMT per one-way vehicle trip by dividing the annual aggregated UPT and PMT from Step 3 by the annual number of one-way vehicle trips operated with valid APC data from Step 2. Columns H and I of Figure II-2 calculate the average UPT and PMT per trip, respectively, for each service segment. In the case of no trips with valid APC data for any specific service segment, the averages for a similar service segment should be used.
5. **Estimate Annual UPT and PMT by Service Segment** – This step estimates annual UPT and PMT for each service segment by multiplying the average UPT and average PMT from Step 4 by the annual number of one-way vehicle trips actually operated from Step 1. Columns H and I of Figure II-2 show the estimated annual UPT and PMT, respectively, for each service segment.
6. **Estimate Annual Total UPT and PMT for the Whole Service** – In this step, agencies estimate the annual total UPT and PMT for the entire service by summing across all service segments the already-estimated annual UPT and PMT for each service segment. Row 30 of Figure II-2 presents the estimated annual totals. Shown in cell J30, the annual total UPT for the whole service is estimated at 394,238. Shown in cell K30, the annual total PMT for the whole service is estimated at 1,307,441.

6.04 Estimating Annual Total UPT and PMT by Schedule Type

6.04.1 No Atypical Days

If an agency had no atypical days during an entire year, it can easily estimate annual total UPT and PMT by schedule type from using the summary data at the level of individual service segments used in estimating annual totals. Specifically, the agency simply sums the estimated annual UPT and PMT across all service segments for each schedule type. In the example, the annual total PMT is 1,023,630 for weekday schedule, 196,676 for Saturday schedule, and 87,134 for Sunday schedule.

Once the annual totals by schedule type are estimated, it is straightforward to estimate annual average daily UPT and PMT by schedule type—simply divide the annual totals for each schedule type just estimated above by the number of days operated for the same schedule type to estimate the annual daily average UPT and PMT by schedule type. The number of days operated by schedule type used here should be consistent with those reported to the NTD.

6.04.2 Having Atypical Days

If an agency did have atypical days during a year, it cannot use the segment-level summary data from estimating annual totals for the whole service. Instead, the six steps need to be followed again without including trips operated on atypical days to estimate the annual totals needed for this purpose by schedule type. With atypical days being identified in the database, excluding them for Steps 1 through 3 is straightforward.

Annual average daily UPT and PMT by schedule type can be estimated similarly relative to the case without atypical days. In this case, however, the number of days operated must exclude atypical days as well.

6.05 Estimating Monthly UPT

Whereas the estimation steps were stated for estimating annual total UPT and PMT, they are just as applicable to estimate monthly total UPT once the period of data is changed from one year to one month. Only total UPT is needed for monthly reporting.

To maintain consistency between the estimated annual total UPT from this methodology for annual reporting and the sum of the monthly UPT values for monthly reporting, it is important that the reported monthly UPT comes from the same methodology.

A potential complication with monthly estimation with this methodology is that the chance is higher for some service segments to have no trips operated with valid APC data during a month. This complication is not an inherent shortcoming of this methodology; in fact, it can occur with intentional sampling as well if the monthly portion of the annual sample is used in monthly reporting. In those cases, the average UPT per trip from a similar service segment from Step 4 should be applied for Step 5.

Agencies should not use the sum of monthly UPT estimated with this methodology as the annual total UPT for annual reporting. Potentially, it is possible to estimate monthly UPT and PMT every month and to add these monthly estimates as the annual total UPT and PMT. The advantage of doing this is that the annual total UPT would be equal to the sum of the reported monthly UPT values. The shortcoming is that the monthly number of trips with valid APC data is likely too small for some service segments. In fact, using monthly sum as annual total implicitly adds another dimension to the already defined service segments for using this methodology. This additional dimension reduces the number of one-way trips operated and the reduced number may not reach the target of 20 annually.

6.06 Estimating APTL

Once an agency has estimated annual total UPT and PMT for the whole service and annual total UPT and PMT for each schedule type, it can easily estimate the corresponding APTL. If it has a reliable 100% UPT count from its electronic registering fareboxes and choose to report it to the NTD, it may want to estimate APTL from the estimated UPT and PMT data with the methodology in this guidebook.

If an agency uses this methodology only for estimating the APTL for the whole service and for each schedule type, it may define the service segments at a more aggregated level than otherwise. Specifically, a reasonable and relatively easy segmentation level is to use the individual routes by schedule type as the service segments. This is reasonable because APTL does not vary significantly across routes. It also is easy—at least for most small and medium agencies— because the number of routes operated is relatively small for MB services. For the other three modes, an agency typically operates only one or just a few routes.

7.00 Appendix A – Certification Document

7.01 Introduction

This appendix serves as the document of certification for the methodology in this guidebook for an agency to estimate annual total unlinked passenger trips (UPT), annual total passenger miles traveled (PMT), or average passenger trip length (APTL) as the ratio of annual total PMT over annual total UPT through using data from its automatic passenger counter (APC) system. It certifies that this methodology meets the 10% precision and 95% confidence levels as long as the agency meets the certification conditions and follows the guidance in this guidebook. This document lists the certification conditions, describes the methodology, justifies the certification and summarizes my qualifications as a qualified statistician.

7.02 Conditions

To use the methodology as a pre-certified alternative sampling technique, the agency must meet all of the following conditions:

- The agency operates at least 1,000 vehicle trips annually (e.g., 4 trips daily).
- APCs must cover all doors of all vehicles of the fleet.
- The agency should have a continuous monitoring and maintenance program.
- The agency must recover valid APC data from at least 50% of all vehicle trips operated.
- The agency must pass a test of statistical equivalence for its APC data.
- The agency must use all valid APC data for estimation.

7.03 Methodology

This methodology consists of a sampling component and an estimation component. The sampling component is based on full-population sampling (i.e., every vehicle trip is sampled) with some degree of nonresponse (i.e., valid APC data are not recovered from some trips operated). To minimize any potential bias from missing APC data, the estimation component requires post-stratification as follows:

- Divides the whole service into a set of mutually-exclusive service segments (e.g., a trip on the weekday schedule, a route, etc.).
- Estimates total transit usage for each service segment by extrapolating all valid APC data for that service segment. For each segment, this is to multiply the average transit usage per trip among all trips with valid APC data by the annual total number of trips actually operated for that service segment.
- Sums the estimates of segment-level total transit usage across all service segments to get total transit usage for the whole service.

7.04 Justification

Since this methodology uses full-population sampling, estimates from this methodology have no sampling error. However, these estimates can still have errors from other sources. As summarized in Table II-2, these errors combined are within ±9.0% at the 95% confidence level.

Table II-2. Precision at 95% Confidence

Error Type	Magnitude
Random sampling error	0%
Random measurement error	≤ ±1.12% with at least 4 daily trips operated on 250 days a year & data recovery rate ≥ 50%
Systematic measurement error	≤ ±7.5%
Systematic error from missing data	< ±0.2%
Total @95% confidence	< ±9.0%

The valid APC data for any vehicle trip can differ from the unknown true value. Systematic measurement errors exist because the average of differences between the valid APC data and the true value cross all trips typically is not zero. Random measurement errors exist because the difference varies randomly across vehicle trips. Systematic errors can result not only from the valid APC data being different from the true value but also from the trips with valid APC data not being representative of all trips operated. The systematic error due to missing data may be referred to as the missing-data bias.

Both types of measurement errors can be assessed through paired APC and manual data from a random sample of trips with an adequate sample size. In fact, this assessment is part of conducting and passing the test of statistical equivalence. Random measurement errors become negligible from using all valid APC data. Random errors are within ±1.12% for the extreme case of operating only 4 trips daily on weekdays only with a 50% recovery rate of valid APC data. Random measurement errors are smaller for all other cases and, in fact, are well within ±0.5% for agencies with at least 20 vehicle trips daily. Systematic errors, on the other hand, do not become smaller from using more data but are controlled to be within ±7.5% by the requirement of passing the test of statistical equivalence. As a result, the two measurement errors combined do not exceed ±8.62% for all cases and are within ±8.0% for most cases.

The missing-data bias is not directly measurable. As part of the effort of developing the guidance for this methodology, a comprehensive simulation was carried out to measure a base level of the missing-data bias when estimation is done without using this methodology and to examine how much of this base bias would disappear once estimation is done with this methodology. This assessment was done for a range of possibilities on how trip-level usage varies, the range of data recovery rates, and how the data recovery rate varies. The results show that the base level of bias can be as high as 7% for some circumstances. One example of such circumstances is when trip-level PMT varies significantly across service segments and data recovery rates are low and vary significantly across these service segments. The same simulation, however, also showed that the large base bias largely disappeared (i.e. within ±0.2%) once estimation is done through this methodology.

7.05 Qualification

I have a Ph.D. in Economics with a concentration in Transportation Economics from the University of California, Irvine. I have developed and certified alternative sampling techniques for many transit agencies. The Federal Transit Administration has adopted the *National Transit*

Database Sampling Manual that I developed to give detailed guidance to NTD reporters for determining annual service-consumed data for all modes.

8.00 Appendix B – Glossary and Formulas

8.01 Glossary

The following definitions are largely based on the *NTD Glossary* and the 2015 *NTD Policy Manual*.

Atypical Days: Atypical days occur when an agency does not operate its normal, regular schedule; rather, the agency:

- Provides extra service to meet demands for special events, such as conventions, parades, or public celebrations, or
- Operates significantly-reduced service because of unusually bad weather (e.g., snowstorms, hurricanes, tornadoes, earthquakes) or major public disruptions (e.g., terrorism)

The concept of atypical days is relevant only for scheduled, fixed-route services such as motorbus (MB), commuter bus (CB), bus rapid transit (RB), trolley bus (TB), and rail modes.

Average Passenger Trip Length (APTL): The average distance ridden for an unlinked passenger trip (UPT) by time period (weekday, Saturday, Sunday), computed as passenger miles traveled (PMT) divided by unlinked passenger trips (UPT). May be determined by sampling, or calculated based on actual data.

Bus Rapid Transit (RB): A fixed-route bus mode in which:

- The majority of each line operates in a separated right-of-way dedicated for public transportation use during peak periods
- Features are included that emulate the services provided by rail fixed guideway public transportation systems, including:
  - Defined stations
  - Traffic signal priority for public transportation vehicles
  - Short headway bidirectional services for a substantial part of weekdays and weekend days
  - Pre-board ticketing, platform level boarding, and separate branding

This mode may include portions of service that are fixed-guideway and non-fixed-guideway.

Commuter Bus (CB): Fixed-route bus systems that connect primarily outlying areas with a central city through bus service that operates with at least five miles of continuous closed-door service. This service may operate motor coaches (over-the-road buses) and usually features peak-scheduling multiple-trip tickets and limited stops in the central city.

Days Operated: The number of days that service was actually operated according to the schedule of service.

High-Intensity Busway: A new category of guideway distinct from fixed guideway, defined by MAP-21 federal legislation. High Intensity Motorbus (HIB) comprises lanes that are exclusive to transit vehicles at some, but not all, times, and lanes that are restricted to transit vehicles, HOV, and HO/T. HIB lanes do not have their own funding tier under UAFP, but do receive State of Good Repair funding once they reach seven years of age.

Motor Bus (MB): A transit mode comprising rubber-tired passenger vehicles operating on fixed routes and schedules over roadways. Vehicles are powered by diesel, gasoline, battery, or alternative fuel engines contained within the vehicle.

Passenger Miles Traveled (PMT): The cumulative sum of the distances ridden by each passenger.

Trolleybus (TB): A transit mode comprising electric rubber-tired passenger vehicles, manually-steered and operating singly on city streets. Vehicles are propelled by a motor drawing current through overhead wires via trolleys, from a central power source not onboard the vehicle.

Vehicles Operated in Annual Maximum Service (VOMS): The number of revenue vehicles operated to meet the annual maximum service requirement. This is the revenue vehicle count during the peak season of the year on the week and day that maximum service is provided. VOMS excludes atypical days or one-time special events.

8.02 Formulas for Equivalence Testing

Those interested in the statistics involved in this test of equivalence may refer to the following symbols, statistical concepts, related formulas, hypotheses, and testing criteria:

- *n* paired observations for a given certification metric (e.g., PMT per trip)
- y_{ai} (APC) and y_{mi} (manual)
- Mean of manual data: $\overline{y_a} = \frac{\sum y_{ai}}{n}$
- Difference for each pair: $d_i = y_{ai} - y_{mi}$
- Mean difference: $\bar{d} = \frac{\sum d_i}{n}$ (or $\frac{\bar{d}}{\overline{y_a}}$ in percent terms)
- Standard deviation of difference: $s_d^2 = \frac{\sum(d_i - \bar{d})^2}{n-1}$
- Standard error of mean difference: $s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$ (or $\frac{s_{\bar{d}}}{\overline{y_a}}$ in percent terms)
- Significance level for statistical testing: α (e.g., 5%)
- Critical value in the t distribution: $t_{n-1,1-\alpha}$
- Population mean difference: μ_d (true value but not observed)
- Equivalence bounds: $\pm\theta$

The null hypothesis is H_0: $\mu_d \leq -\theta$ or $\mu_d \geq \theta$. The alternative hypothesis is H_1: $|\mu_d| < \theta$.

Not Equivalent **Equivalent**

The equivalence of APC and manual data is accepted if the null hypothesis is rejected at the 1-α confidence level. More specifically, equivalence is established if the following two one-sided t-tests both are rejected:

$$t_1 = \frac{\bar{d} - (-\theta)}{s_{\bar{d}}} > t_{n-1,1-\alpha} \text{ AND } t_2 = \frac{\bar{d} - \theta}{s_{\bar{d}}} < -t_{n-1,1-\alpha}$$

Alternatively, equivalence is established if the (1-2α)% confidence interval falls within (-θ, θ).

Part III – Rail Guidebook

GUIDEBOOK:

Estimating Transit Usage through Extrapolating Incomplete APC Data for Rail Service

Table of Contents – Part III

1.00 Introduction

1.01 Objective

This guidebook details a methodology for agencies to estimate transit usage for the National Transit Database (NTD) through extrapolating incomplete data from automatic passenger counters (APC) for three rail modes. These are light rail (LR), streetcar rail (SR), and hybrid rail (YR). YR is operated primarily on the national system of railroads but its operating characteristics resemble those of LR. More specifically, the methodology:

- Divides the whole service into a set of mutually-exclusive service segments (e.g., a trip on the weekday schedule, an LR line, etc.).
- Estimates total transit usage by extrapolating all valid but incomplete APC data within each service segment. For each service segment, this is to multiply the transit usage per car-trip among all train trips with valid APC data by the annual total number of car-trips actually operated for that service segment.
- Sums the estimates of segment-level total transit usage across all these service segments to get total transit usage for the whole service.

This methodology can be used to estimate all of the following measures of transit usage:

- Annual total unlinked passenger trips (UPT)
- Annual total passenger miles traveled (PMT)
- Annual average daily UPT by schedule type (weekday, Saturday, Sunday)
- Annual average weekday UPT by time period (AM peak, midday, PM peak, other)
- Annual average daily PMT by schedule type
- Monthly total UPT

The methodology represents an alternative sampling technique for estimating annual total UPT and PMT. Agencies can use Appendix A of this guidebook as its document of certification by a qualified statistician. When an agency uses this methodology for estimating annual total UPT, annual total PMT, or both, it should report on the CEO Certification Form the method used as an “alternative sampling procedure that meets 95% confidence and ±10% precision levels as determined by a qualified statistician (estimated data).”

It is critical that this certification is conditional. These conditions are detailed later and are summarized here:

- The agency operates at least 1,000 train trips annually (e.g., 4 trips daily).
- APCs must cover all doors of all passenger cars of the fleet.
- The agency should have a continuous monitoring and maintenance program.
- The agency must recover valid APC data from at least 50% of all train trips operated.
- The agency must pass a test of statistical equivalence for its APC data.
- The agency must use all valid APC data for estimation.

Agencies are not suggested to use this methodology if car trips with missing APC data do not exceed 2% of all car trips actually operated during an entire year. NTD rules allow proportional expansion at the system level for this case.

1.02. Directions

This guidebook is organized into the following sections:

- **Section 2.00, NTD Requirements** – summarizes the basic requirements that agencies must meet when reporting transit usage data to NTD and the additional requirements that agencies must meet when they use their APC data for reporting.

- **Section 3.00, Reporting Options** – discusses the various possibilities of estimating annual total UPT and PMT from using the methodology in this guidebook.

- **Section 4.00, Certification Conditions** – presents the requirements that agencies must meet to use the methodology in this guidebook as a pre-certified alternative sampling technique.

- **Section 5.00, Defining Service Segments** – presents guidance on how agencies should divide their whole service into service segments.

- **Section 6.00, Data and Estimation** – presents guidance on the basic data requirements for using this methodology and detailed steps for using this methodology for estimating transit usage.

- **Section 7.00, Appendix A** – is the document of certification for agencies to keep on file for future reference.

- **Section 8.00, Appendix B** – lists several key definitions and the relevant formulas used in the test of statistical equivalence.

2.00 Reporting Requirements

2.01 Basic Requirements

To be eligible for the Urbanized Area Formula Grant Program (Section 5307), agencies reporting LR, SR, or YR service must report to the NTD data for the following measures of transit usage:

- Annual total unlinked passenger trips (UPT)
- Annual total passenger miles traveled (PMT)
- Annual average daily UPT by schedule type (weekday, Saturday, Sunday)
- Annual average daily PMT by schedule type
- Annual average weekday UPT by time period (AM peak, midday, PM peak, other)
- Monthly total UPT

Data on annual total UPT and PMT and on monthly total UPT must include all services provided on all days during the corresponding year or month. Data on annual average daily UPT and PMT and annual average weekday UPT, on the other hand, should include only services on typical days and must exclude services on atypical days. Refer to the *NTD Policy Manual* on determining typical and atypical days. For completion, the NTD definition is provided here:

Atypical days occur when an agency does not operate its normal, regular schedule; rather, it:

- Provides extra service to meet demands for special events, such as conventions, parades, or public celebrations, or

- Operates significantly-reduced service because of unusually bad weather (e.g., snowstorms, hurricanes, tornadoes, earthquakes) or major public disruptions (e.g., terrorism).

These data must be separately reported for each service that agencies provide, i.e., each combination of modes and service types (directly operated or purchased). UPT measures 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. PMT, on the other hand, measures the total distance traveled by all passengers.

Obtaining annual data on these usage measures, particularly for PMT, can be costly to agencies. PMT for each one-way trip typically is calculated as the distance-weighted sum of passenger loads between consecutive stations. This calculation typically requires detailed data on passenger boarding and alighting activities at individual stations and distances between consecutive stations for individual one-way trips.

To obtain the annual total data for either UPT or PMT, an agency must choose one of two approaches. If available and reliable, it must report a 100% count. If a reliable 100% count of UPT or PMT is not available or is available but is considered not reliable, it must estimate it through sampling. The obtained estimate of annual total UPT or PMT from sampling must meet the minimum 10% precision level at the 95% confidence level. Data on annual average daily UPT and PMT and monthly total UPT are not subject to these statistical requirements.

2.02 APC Requirements

2.02.1 APC Data

Using electronic infrared beams or mechanical treadle mats, APCs can convert electronic signals from beams or mats to counts of passengers as they board and alight transit vehicles. An APC includes a dedicated onboard computer, the APC analyzer, which converts sensor information into passenger counts. Each time a bus leaves a stop, the APC analyzer closes out a record and transmits the on-off counts at that stop to a general onboard computer, the in-vehicle logit unit. When coupled with stop location information, archived APC data can be post-processed to generate both UPT and PMT data potentially for every vehicle trip operated.

2.02.1 The Problem of Missing Data

Introduction

When an entire fleet is fully-equipped with APC units, theoretically, it is possible to obtain a 100% count of both annual total UPT and PMT from APC data. In reality, however, getting reliable 100% counts of UPT and PMT from APCs is difficult to achieve. In fact, APC data are missing for some train trips either because they were not recovered or because they were discarded due to significant error. Although the literature does not provide much on the problem of missing data for rail APC systems, a rail APC system shares some common issues with APC systems on other modes and has some unique issues at the same time.

Common Issues

Data may fail to be recovered from individual APC units due to general hardware malfunction, including sensors in the doorways failing to send signals to the onboard computer, the onboard computer failing to convert sensor information into passenger on-off counts, and the onboard computer failing to transmit and store on-off counts. The effect of this problem can be reduced, to some degree, through a continuous monitoring and maintenance program.

Also, raw data may be recovered successfully from individual APC units, but serious errors can exist in the recovered raw data, requiring it to be discarded through the procedures of processing recovered raw data:

- In many cases, raw APC data are recovered but they fail to match in space and time to the actual service provided. In these cases, the recovered raw APC data are not usable.
- In many other cases, raw APC data are recovered and matched to actual service but errors in the recovered raw data fall out of range based on pre-set screening criteria for differences in on-off counts for individual blocks. In these cases, the matched raw APC data are not usable for purposes focusing on boarding counts. When matched raw APC data fall out of range based on pre-set load screening criteria, the raw APC data are rejected for purposes focusing on load or passenger miles.
- In almost all cases in which raw APC data are recovered, matched, and accepted for further consideration, on-off counts will have errors. Unlike with data from manual ridechecks, these errors cannot be easily corrected in determining train-level UPT and train-level PMT. This is because APCs collect on-off counts but not the additional

information on the number of inherited passengers from the previous trip, the number of continuing passengers to the next trip, and stop-to-stop loads.

Although screening based on on-off differences avoids substantial errors in on-off totals, substantial errors still can develop in calculated loads and passenger miles, even with small errors in raw on-off counts due to a phenomenon called "drift." An effective way to control drift is to parse a vehicle block's data stream of automated counts at points of known load. Most routes have natural known-load points at terminals and layover points. To control drift through parsing, blocks are divided at known-load points into sections, which usually are single trips. On-off counts within each section then need to be balanced so that the calculated loads match the known loads at each end of the section.

Unique Issues

Most modern light rail systems in the US have unique features that are not shared by most bus systems, including low-floor boarding and double-wide doors. These two features combined allow more than one passenger to board and alight at the same time. Such complex passenger activities can be a challenge for an APC system. In addition, the number of APC units on fully-equipped trains can be significantly greater than the typical 2 units on most buses. A train with 4 cars may have a total of 16 APC units with 2 wide doors on each car and 2 APC units on each door. Even with one or a few of these APC units failing to produce any raw data, it would be difficult to balance the on-off data for a whole train trip to a degree that can pass pre-set screening criteria for further consideration.

At the same time, however, some unique features of light rail are positive in reducing the degree of missing data. Although a bus driver may get on or off a vehicle multiple times through the same doors used for passenger boarding and alighting, a light rail train operator is much less likely do to so. More significantly, most light rail systems have a simple network, with just one or a few lines. This network simplicity helps avoid some problems due to complexities of a transit network:

- It is less likely to have APC data failing to match in space and time to the actual service provided.
- It is much less likely have passengers stay onboard from one train trip to the next without getting off. This avoids a source of potentially serious error in calculated passenger loads and passenger miles that many bus APC systems experience.

The data recovery rate for an APC system refers to the percent of all one-way train trips operated during a specific period that provide valid data on train-level UPT and PMT. Although this rate has been reported as low as 60–75% for bus APC systems, little data are available on this rate for light rail APC systems. APC systems for light rail typically are new and would provide higher data recovery rates if new APC systems do better than older APC systems.

2.02.3 Additional Requirements

An agency must first obtain the approval of FTA to use its APC data for reporting UPT and PMT data. As part of this approval, the agency must have its APC system certified initially by passing a benchmarking test. At the initial certification, the agency must have its APC system re-certified every fiscal year that is evenly divisible by three (i.e., 2019, 2022, 2025, etc.). To pass the benchmarking test, the agency must demonstrate that its APC data are within ±5% of paired manual data collected from a sample of train trips that may not be selected at random.

3.00 Reporting Options

3.01 Basic Reporting Options

Under the reporting requirements, a transit agency has up to three basic methodological options for reporting both annual total UPT and PMT:

1. Report 100% counts for both UPT and PMT
2. Report 100% UPT, but estimate PMT indirectly through estimating APTL via traditional sampling with manual ridechecks
3. Estimate both UPT and PMT directly through estimating average UPT and average PMT per train trip via traditional sampling with manual ridechecks

For Option 2 or Option 3, sampling is done according to a pre-specified sampling plan. This sampling plan must be either pre-approved by FTA or an alternative sampling technique that has been certified to meet the 10% precision and 95% confidence levels by a qualified statistician. With either type of sampling plan, sampling traditionally is done with human ridecheckers riding with the passengers of every sampled one-way trip. This approach to sampling is referred to as traditional sampling with manual ridechecks.

With Option 2, the agency estimates PMT indirectly as the product of the 100% UPT with the average passenger trip length (APTL) estimated through sampling. Specifically, it first obtains both sample total UPT and sample total PMT through sampling and then calculates the ratio of sample total PMT over sample total UPT to estimate the APTL.

For Option 3, on the other hand, the agency estimates PMT directly as the product of the total number of one-way train trips operated during a full year by the average PMT per one-way train trip estimated through sampling. Specifically, it first obtains sample total PMT and the number of one-way trips in the sample and then calculates the ratio of the sample total PMT over the sample size to estimate average PMT per one-way trip. It estimates UPT similarly with the same sample data.

Modern light rail lines typically rely on off-board proof-of-payment fare collection, which requires passengers to purchase fare media off-board the train and show proof-of-payment upon random inspection. Such fare collection systems do not produce 100% UPT counts. As a result, Option 1 is unlikely to be available to a light rail agency. Option 2 may be available to a few agencies with fare collect systems that count and record every boarding. Option 3 is available to every light rail agency.

3.02 Reporting Options with APC Data

If an agency has installed APC units on all doors of every passenger car and has obtained FTA certification of its APC system for NTD reporting, it is unlikely that it is willing to continue using traditional sampling for either Options 2 or 3. Instead, it will choose intentional sampling with APCs in one of two forms:

- **Pre-sampling** takes place before a train trip is actually operated. It is common for agencies with a relatively small share of its fleet equipped with APC units. Pre-sampling is necessary so that APC-equipped passenger cars can be assigned to the sampled train trips. Pre-sampling is applicable to other cases and must account for the expected data recovery rate in the sample size used. As with traditional sampling, pre-sampling selects trips from the full list of all train trips to be operated.

- **Post-sampling** occurs after a train trip has been operated and the raw APC data have been processed. It is common for agencies with a full coverage of APC units on all doors of all passenger cars. Different from either traditional sampling or pre-sampling, post-sampling selects trips from all train trips with valid APC data.

The term "intentional sampling with APCs" is used to contrast it with traditional sampling with manual ridechecks and with using all valid APC data, which is a form "full-population" sampling and is the focus of this guidebook. With intentional sampling with APCs, an agency has two additional reporting options:

4. Report 100% UPT from a non-APC source, but estimate the APTL through intentional sampling with APCs; this is another version of Option 2.
5. Estimate both UPT and PMT directly through intentional sampling with APCs; this is another version of Option 3.

3.03 Reporting Options with All Valid APC Data

If an agency has not only obtained FTA certification of its APC system for NTD reporting but has full coverage of APC units on all doors of every passenger car, additional reporting options are open to it. These additional options are made available from using all valid APC data in a special way through the methodology included in this guidebook. These additional reporting options are:

6. Report 100% UPT from a non-APC source, but estimate the APTL through extrapolating all valid but incomplete APC data on both PMT and UPT; this is another version of Option 2.
7. Report 100% UPT from a non-APC source, but estimate PMT directly through extrapolating all valid APC data on PMT; this is a new option.
8. Estimate both UPT and PMT through extrapolating all valid but incomplete APC data with the methodology in this guidebook; this is another version of Option 3.

3.04 Choices

3.04.1 Having 100% UPT from a Non-APC Source

When available from a non-APC source, it is largely up to the agency to decide if it reports 100% UPT.

Reporting 100% UPT

If an agency reports the available 100% UPT from a non-APC source, it has four potential options from which to choose for estimating PMT:

- Option 2 – estimate APTL through traditional sampling with manual ridechecks
- Option 4 – estimate APTL through intentional sampling with APCs

- Option 6 – estimate PMT directly with the methodology in this guidebook through extrapolating all valid APC data on both UPT and PMT
- Option 7 – estimate PMT directly with the methodology in this guidebook through extrapolating all valid APC data on PMT only

Traditional sampling is largely irrelevant for this choice. After obtaining FTA certification of its APC system for NTD reporting, an agency will not continue using traditional sampling with manual ridechecks due to its labor cost.

This Methodology vs. Intentional Sampling

This choice is between Options 4 and 6. It is clear that this methodology (Option 6) will always lead to better estimates than intentional sampling with APCs (Option 4):

- These two options suffer the same problem of missing data to the same degree.
- For both options, the problem of missing data can lead to two undesirable outcomes for certain segments of the service. These service segments either are no longer represented or are disproportionately represented in the valid APC data. If not properly mitigated, both outcomes potentially can result in bias in the estimates.
- The potential bias is likely to be realized with intentional sampling. With intentional sampling, the problem of missing data is significantly more likely to lead to certain service segments not being represented or being disproportionally represented. Any bias for a service segment resulting from not being represented in the valid data cannot be mitigated. Even in the case of disproportionate representation, estimation for such service segments is highly unreliable due to the small remaining sample size for each segment.
- The potential bias can be largely mitigated by using all valid APC data through this methodology. The problem of missing data still can lead to certain service segments being disproportionally represented, but is highly unlikely to have certain service segments not represented. When a service segment is disproportionately represented, the significantly large amount of valid data allows relatively reliable estimation for this segment.
- The large amount of all valid APC data can mitigate potential bias from the problem of non-random missing data and also will result in estimates with much better precision than those from intentional sampling.

APTL vs. Average PMT

This choice is between two different ways of using the methodology in this guidebook. The methodology can be used to estimate both UPT and PMT and the APTL as the ratio of these two estimates for Option 6. In other words, Option 6 expands the APTL from this methodology by the 100% UPT from another source. On the other hand, the methodology can be used to estimate only PMT for Option 7. That is, Option 7 expands the average PMT per train trip from this methodology by the total number of one-way train trips actually operated.

If the 100% UPT from another source was the true value, the agency should choose Option 6. That is, it should report the annual total PMT that is estimated as the product of the 100% UPT from another source and the APTL from this methodology. The precision reached by this APTL-based estimate most likely will be better than the average-based estimate. This is simply because APTL varies much less than does PMT.

If the 100% UPT count understates the true value to some degree, the choice is not immediately clear. If the APC data have similar magnitudes of error in APTL versus in average PMT, the agency should choose Option 7. This is because the expansion factor for Option 7 is the number of trips operated and is error-free. The expansion factor for Option 6, on the other hand, is the 100% UPT and understates the true value. As discussed in testing the statistical equivalence between APC data and paired manual data, the magnitude of error can be estimated with paired APC and manual data for both APTL and average PMT. If the magnitude of error is smaller in APTL than in average PMT, the choice is a difficult one.

Not Reporting 100% UPT

With 100% UPT from another source, an agency can still choose to use the methodology in this guidebook and get an estimate of annual total UPT. The choice in this case is whether it should report the 100% count or the APC-based estimate to the NTD:

- FTA requires that agencies report the 100% UPT when available and reliable. It should report the APC-based estimate from this methodology if the 100% count is considered not reliable. However, “reliability” is not defined; it is up to the agency to determine whether the 100% count is reliable. If the 100% UPT understates the true value by more than 10%, for example, the agency can still consider it as being available and reliable and, therefore, report it. To meet the 10% precision and 95% confidence levels in this case, however, the agency should report the APC-based estimate.
- From the perspective of the agency, another consideration is that the APC-based estimate is likely to be larger than the 100% count if the latter understates the true value to some degree. In this case, the agency has an incentive to report the APC-based estimate.

3.04.2 No 100% UPT

If an agency does not have a 100% UPT count, it has three potential options from which to choose for estimating both UPT and PMT:

- Option 3 – Estimate through traditional sampling with manual ridechecks
- Option 5 – Estimate through intentional sampling with APCs
- Option 8 – Estimate through using the methodology in this guidebook

Similar to the choice among Options 2, 4, and 6 discussed earlier, the real choice is between Options 5 and 8. In addition, Choice 8 will always give better estimates than Option 5.

4.00 Certification Conditions

4.01. Introduction

To use the methodology in this guidance as a certified alternative sampling technique, an agency must meet all of the following conditions:

1. The agency operates at least 1,000 train trips annually (e.g., 4 trips daily).
2. APCs must cover all doors of all vehicles of the fleet.
3. The agency should have a continuous monitoring and maintenance program.
4. The agency must recover valid APC data from at least 50% of all vehicle trips operated.
5. The agency must pass a test of statistical equivalence for its APC data.
6. The agency must use all valid APC data for estimation.

4.02 Coverage, Operations, and Data Recovery

The agency must have APC units installed at every door of every passenger car in the fleet. To ensure a minimum of 500 train trips with valid APC data during a year, the agency is also required to operate at least 1,000 train trips annually and to recover valid APC data from at least 50% of all train trips operated.

4.03 Continuous Monitoring

The agency must have a process in place to monitor and maintain individual APC units and the APC system continuously. An effective process, for example, would immediately notify the agency when the APC units of a door failed to acquire any raw data on the on-off activities through that door and would indicate the source of the equipment failure problem. Once equipment failure occurs, the same process would have an ongoing partnership with vendors for quick maintenance and repair. Reducing train trips with missing data due to equipment failure benefits not only NTD reporting but also service and other planning efforts within the agency.

4.04 Use of All Valid APC Data

The agency must be able to use a significant amount of APC data at a disaggregated level. This should not be real challenge for most LR, SR, or YR systems due to their simple network. One effective approach to an agency having this ability is that it should have an annual database of train-level information for all one-way train trips actually operated during a full year. This database would accumulate such information as individual train trips are being operated and as the raw APC data are being processed from the first day to the end of the year.

4.05 Test of Statistical Equivalence

4.05.1 Introduction

To use the methodology in this guidebook as a certified alternative sampling procedure for estimating annual total UPT, annual total PMT, or both, a transit agency must do a test of statistical equivalence for its APC data and must pass the test. It must collect paired APC and manual data from the same sample of train trips selected at random during a period that is at least one week long.

- To pass the test of statistical equivalence, the agency must demonstrate that its APC data are statistically equivalent to the paired manual data within ±7.5% at the 95% confidence level.
- This condition of statistical equivalence is to ensure that the estimate of annual total transit usage meets the 10% precision and 95% confidence levels. The equivalence bounds are set at ±7.5% to allow up to ±2.5% of other types of errors in the estimates.
- Requiring the paired APC and manual data coming from a random sample is necessary to access the level of confidence achieved for the test of statistical equivalence.
- The minimum duration of one week for the testing period is specified to minimize the burden on transit agencies while non-seasonable variation is mostly being captured.

If an agency plans to use the pre-certified alternative sampling procedure only for estimating annual total UPT, it should focus the test on the car-level UPT data from the testing sample. Similarly, if it plans to use the pre-certified alternative sampling procedure only for estimating annual total PMT, it should focus the test on the car-level PMT data from the testing sample. If the agency wants to use the pre-certified alternative sampling procedure for estimating both annual total UPT and PMT, it should test statistical equivalence for both UPT and PMT.

During the procurement process for an APC system, a transit agency may want to consider if the RFP should include the requirement of having the new APC units tested for statistical equivalence and if its acceptance of the APC system is conditional on passing the test of statistical equivalence. If the agency plans to use this methodology as a pre-certified alternative sampling procedure, it should include testing statistical equivalence and passing the test as system acceptance requirements.

This section provides detailed guidance for an agency to test statistical equivalence on its own in an Excel environment.

4.05.2 Selection of Trips for Testing

The selection of trips for testing statistical equivalence consists of several aspects, including:

- **Duration** – The shortest duration for selecting trips is a full week to capture most non-seasonal variations. It is better, for course, to cover a much longer duration so both seasonal and non-seasonal variations are accounted for.
- **Unit of Trips** – The proper unit of trips would be one-way train trips. It would be difficult to balance on-off activities for individual car trips. Onboard passengers can walk across cars any time before they alight the train. It is almost impossible to keep track of these passenger activities between stations.
- **Number of Trips** – Refer to Section 4.05.6, Sample Size for Equivalence Testing, for guidance on the appropriate number of trips to be selected.

- **Method of Selection** – The trips must be selected at random from all scheduled trips during the chosen period. One commonly-used approach to selecting a random sample from a sampling frame of train trips is to use the random number function in Excel. Assuming a testing sample of 50 train trips, this approach involves the following steps:

    1. Starting at the 2nd row and column B in a blank Excel worksheet, list all one-way train trips to be operated during the selected data period. The train trips on a weekday schedule should be repeated for each weekday in the period. Similarly, the train trips on the Saturday schedule should be repeated for each Saturday in the period. This list should reflect any case in which the Sunday schedule is operated on a weekday or Saturday holiday.
    2. In cell A2, type "=rand()" and press “Enter.” This will produce a random number (decimal) between 0 and 1.
    3. Copy and paste the formula in cell A2 down column A to the bottom of the list. One way to do this is to hover the mouse pointer over the lower right corner of cell A2 until it changes to a thin black cross and then double-click. This will automatically copy the formula to the bottom of the list.
    4. Copy all data in column A. Perform a “Paste Special” to paste values over the random formula.
    5. Sort by column A. Highlight entire table of data, select data, and sort from the menu. In the “Sort by” window, select column A or the heading for that column, if applicable. Select “Ascending” and click on “OK.”
    6. The train trips in the first 50 rows of the sorted list are the random sample.

4.05.3 Processing and Extraction of APC Data

Once the train trips have been selected at random during the testing period, the agency should make sure the APC units for each selected train trip are in working conditions before each train trip takes place. This will increase the data recovery rate for the testing sample.

It is beyond the scope of the current guidance to cover the collection and processing of raw APC data in general. But the agency should make sure that the extracted APC data are processed from the raw APC data with the identical set of procedures as those used for processing raw APC data in general.

The processed APC data for each of these train trips must be extracted from the database of all processed APC data. At a minimum, the extracted APC data must have the train-level UPT and PMT for each train trip in the sample. The agency should make sure that the extracted APC data are for the correct trips. The result is train-level UPT and PMT data for each of the train trips in the sample to be paired with train-level UPT and PMT from the manual data.

4.05.4 Collection and Processing of Manual Data

Collection

The manual data for testing equivalence are commonly collected by sending human ridecheckers to ride with passengers on the selected train trips. It is also possible to collect the manual data through watching recorded videos if enough cameras are properly positioned to capture passenger activities through all doors of every car on a train. For either method of data collection, it is critical that manual data collected for the selected train trips are correctly identified so they can be correctly matched with the APC data extracted for these train trips.

To ensure that the manual data used in testing represent the true value:

- When using ridecheckers, agencies are highly encouraged to use one ridechecker per door, especially for selected trips that are expected to be heavily used. When necessary, two checkers for each door should be considered, with one checking boardings and the other checking alightings. When passenger volumes are high, simultaneous boarding and alighting of multiple passengers through the same door at the same time can occur frequently.

  Watching recorded videos could be more cost-effective, especially during peak periods. Train trips typically also have more cars during peak periods. A train trip with 4 cars would require 16 ridecheckers if two are used for each door.

- If using recorded videos, the agency should make sure that enough cameras are adequately positioned to capture all passenger activities throughout a train. In addition, recorded videos may be kept for a limited time, and the agency should make sure that data collection from the recorded videos is done before they are deleted.

- When using either method, at least one extra column for recording the passenger loading between each consecutive pair of stations should be included. Each ridechecker in a passenger car should try to observe and record this additional information. For the typical simple network of light rail systems, it is unlikely to have the following:
  - Inherited passengers from the previous trip at the first station
  - Passengers continuing to the next trip at the last station

  If these passenger activities do occur for a particular agency, it should require the ridecheckers to observe and record them as well.

Although proper processing can correct minor errors in the raw manual data for many cases, the additional data collected can help for other cases with bigger errors in the raw manual data.

It is in the best interest of an agency to minimize errors in the manual data, as it is likely that errors in the manual data would increase variation in the difference between APC and manual data. This increase in variation will mean larger sample sizes for equivalence testing in the future and will reduce the chance of the APC data passing the test of statistical equivalence.

Processing

For either method of collecting manual data, train-level UPT and PMT should be calculated for each selected trip only after the boarding and alighting activities have been properly balanced for each train trip. When on-off activities are properly balanced for a given train trip, train-level boardings and train-level alightings are equal and passenger load at the last station is zero.

If using the station-level on-off activities fails to balance a train trip, the agency should compare the calculated passenger load from these passenger activities with the passenger load recorded in the field. Having the additional data on passenger loads has three advantages:

- A high degree of accuracy in the processed manual data through using the observed loads in balancing passenger on-off activities and calculating passenger loads.
- Less time and effort to balance passenger on-off activities and calculate passenger loads with the additional data than without them.
- Fewer sampled train trips whose passenger on-off activities cannot be balanced; unbalanced trips must be excluded for testing statistical equivalence.

The result is train-level UPT and PMT data for each of the train trips in the testing sample to be paired with the corresponding APC data.

4.05.5 Statistical Testing in Excel

Once the train-level UPT and PMT are ready for both APC and manual data, the agency can easily test the statistical equivalence in Excel. An example is used to illustrate the detailed steps for the test. The example assumes that only PMT data are being tested.

Testing Equivalence in UPT or PMT

Example

The steps for such a test are discussed in the context of hypothetical paired APC and manual data on PMT. The screen capture in Figure III-1 is used for illustration.

The paired APC and manual data are entered into columns A and B. Column C has the number of cars on each train trip. Their difference between APC and manual data is in column D. For subsequent calculations and testing, Column G lists the relevant statistical measures, Column H shows the Excel formulas and functions, and Column I presents the calculated values for the example data used. For easy identification, the functions are in bold. The reference to cells in the formulas assumes that the agency will perform the computations in column H. The steps are numbered from 1) through 20). The 95% confidence level is reflected through Step 15), and the equivalence bounds of ±7.5% are reflected in Steps 18) and 19).

	A	B	C	D	E	F	G	H	I
1	1) APC	2) Manual	3) Cars	4) Diff	5) Diff pe Car		Measures	Excel Formulas and Fuctions	Values
2	65.3	145.2	2	-79.9	-39.9				
3	55.9	92.3	1	-36.5	-36.5	6)	Train trips in sample	=**COUNT**(E2:E34)	33
4	136.6	125.1	2	11.5	5.7				
5	107.4	98.0	1	9.4	9.4	7)	Car trips in sample	=**SUM**(C2:C34)	89
6	377.7	367.0	4	10.7	2.7				
7	400.7	330.4	4	70.3	17.6	8)	Mean of manual data	=**SUM**(B2:B34)/H5	118.1
8	131.4	109.9	2	21.5	10.8				
9	124.6	120.0	3	4.6	1.5	9)	Mean difference	=**AVERAGE**(E2:E34)	0.29
10	350.2	360.5	4	-10.3	-2.6				
11	506.1	429.9	4	76.2	19.0	10)	Standard deviation	=**STDEV**(E2:E34)	15.27
12	214.8	241.2	2	-26.4	-13.2				
13	339.3	337.7	3	1.6	0.5	11)	Standard error	=H11/**SQRT(H3)**	2.66
14	566.3	567.6	4	-1.3	-0.3				
15	132.1	142.1	1	-10.0	-10.0	12)	Relative mean difference	=H9/H7	0.25%
16	231.2	262.9	2	-31.7	-15.9				
17	417.2	364.1	3	53.1	17.7	13)	Relative standard deviation	=H11/H7	12.93%
18	297.1	278.2	2	18.9	9.4				
19	342.2	308.9	3	33.3	11.1	14)	Relative standard error	=H13/H7	2.25%
20	477.0	409.0	3	68.1	22.7				
21	4.2	4.2	1	0.0	0.0	15)	t_critical value	=**TINV**(2*0.05,H3-1)	1.69
22	368.1	351.2	2	16.9	8.5				
23	363.0	325.4	2	37.6	18.8	16)	CI lower bound	=H15-H21*H19	-3.57%
24	404.7	404.0	3	0.7	0.2				
25	493.3	425.5	3	67.8	22.6	17)	CI upper bound	=H15+H21*H19	4.06%
26	352.2	352.2	2	0.0	0.0				
27	409.4	439.6	4	-30.2	-7.5	18)	CI Lower bound > -7.5%?	=**IF**(H23>-7.5%,"Yes","No")	Yes
28	586.6	585.9	4	0.7	0.2				
29	490.1	565.7	4	-75.6	-18.9	19)	CI Upper bound < 7.5%?	=**IF**(H25<7.5%,"Yes","No")	Yes
30	274.0	339.7	3	-65.7	-21.9				
31	470.0	512.6	4	-42.6	-10.6	20)	Equivalent?	=**IF(AND(H27**="Yes",H29="Yes"),"Yes","No")	Yes
32	361.3	341.2	2	20.0	10.0				
33	393.3	402.1	3	-8.8	-2.9				
34	375.0	372.0	2	3.0	1.5				

Figure III-1. Steps for Equivalence Testing in Excel

Testing Procedures

Transit agencies can do equivalence testing using the following steps:

1) Enter the processed APC data in column A. In the example, the APC data are in range A2:A34.
2) Enter the processed manual data in column B. For the example, the APC data are in range B2:B34. It is critical that the APC and manual data in the same row come from the same train trip.
3) Enter the number of cars on the train trip in column C.
4) Subtract the manual data from the APC data for each train trip in column D.
5) Calculate the APC and manual difference per car trip in column E by dividing column D by column C.
6) Determine the number of train trips in the sample in H3 by counting the number of rows in range E2:E34.

7) Determine the number of car trips in the sample in H5 by summing up range C2:C34 with the **SUM** function.

8) Calculate the mean of the manual data in cell H7 as the ratio of the sum of range B2:B34 using the **SUM** function over the number of car trips in cell H5.

9) Calculate the mean of APC and manual differences per car trip in cell H9 by using the **AVERAGE** function.

10) Calculate the standard deviation of the APC and manual differences per car in cell H11 with the **STDEV** function.

11) Calculate the standard error of the per-car difference in cell H13 as the ratio of the standard deviation in cell H11 over the square root of the number of train trips in cell H3 with the **SQRT** function.

12) Calculate the relative mean difference in cell H15 as the ratio of the mean difference in cell H9 over the mean of the manual data in cell H7.

13) Calculate the relative standard deviation in cell H17 as the ratio of the standard deviation in cell H11 over the mean of the manual data in H7.

14) Calculate the relative standard error in cell H19 as the ratio of the standard error in cell H13 over the mean of the manual data in cell H7.

15) Determine the critical value at the 95% confidence level in cell H21 with the **TINV** function and the number of train trips in cell H3.

16) Determine the lower bound of the confidence interval for equivalence testing in cell H23.

17) Determine the upper bound of the confidence interval for equivalence testing in cell H25.

18) Determine if the lower bound is equal or greater than -7.5% in cell H27.

19) Determine if the upper bound is equal or smaller than 7.5% in cell H29.

20) The agency passes the test of statistical equivalence if “Yes” appears in cell H31. It fails to pass it if “No” is in cell H31.

Testing Equivalence in APTL

To do an equivalence test for APTL, the agency needs to conduct the test for both UPT and PMT first. If they pass the test for both UPT and PMT, they pass the test for APTL as well. If they fail the test for UPT, for PMT, or for both UPT and PMT, they do not pass the test for APTL.

4.05.6 Sample Size for Equivalence Testing

The proper sample size for testing the statistical equivalence of paired APC and manual data depends on the relative standard deviation in the difference of APC and manual data and the expected recovery rate of the APC system. Table III-1 shows the minimum sample size for each of five ranges of relative standard deviation and 5% increments of the expected recovery rate.

Table III-1. Sample Size for Equivalence Testing

Expected Recovery Rate of Valid APC Data	Minimum Sample Size in One-Way Train Trips by Range of Relative Standard Deviation				
	≤5%	**>5% & ≤10%**	**>10% & ≤15%**	**>15% & ≤20%**	**>20%**
≥95%	7	25	55	98	152
≥90% & <95%	7	26	58	103	160
≥85% & <90%	7	28	61	109	170
≥80% & <85%	8	29	65	116	180
≥75% & <80%	8	31	70	123	192
≥70% & <75%	9	33	75	132	206
≥65% & <70%	9	36	80	142	222
≥60% & <65%	10	39	87	154	240
≥55% & <60%	11	42	95	168	262
≥50% & <55%	12	47	104	185	288

If the agency has paired APC and manual data previously collected, it should compute the relative standard deviation as illustrated in Figure III-1, determine the range into which its value falls, determine the expected data recovery rate, and pick the proper sample size accordingly from Table III-1. For the example in Figure III-1, the relative standard deviation for PMT is 12.93%, which falls between 10% and 15%. Therefore, the proper sample size for this case would be 55 train trips if the recovery rate is at least 95% but would increase to at least 70 if the recovery rate does not exceed 78%.

A potential source of paired APC and manual data is that collected during the procurement process or from a previous effort of equivalence testing.

An agency that has failed a previous equivalence test may want to select a sample size that is bigger than what its relative standard deviation suggests from Table III-1. The larger the sample size, the greater the chance of a successful equivalence test.

The relative standard deviation in general differs between UPT and PMT. In fact, it generally is larger for PMT than for UPT. They should do the following to select the sample size from Table III-1:

- For testing UPT only, use the relative standard deviation for UPT.
- For testing PMT only, use the relative standard deviation for PMT.
- For testing both UPT and PMT, use the relative standard deviation for PMT.

For many cases, however, an agency may not have its own previously-paired APC and manual data to determine the relative standard deviation. Limited evidence indicates that the relative standard deviation typically does not exceed 15% and the expected recovery rate typically is at least 70%. For these cases, a reasonable sample size to use would be 75.

4.05.7 Future Testing of Statistical Equivalence

Once the agency has passed this test, it stays valid until it fails the benchmarking test in FTA's certification policy for an APC system. The benchmarking test is required in every fiscal year that is evenly divisible by three (i.e., 2019, 2022, 2025, etc.). If that happens, the agency should use the paired APC and manual data from the last test of statistical equivalence to determine the relative standard deviation and use Table III-1 for the proper sample size to use for the new test of equivalence.

5.00 Defining Service Segments

5.01 Introduction

A critical element of this methodology is to divide the whole service during an entire year into individual pieces referred to as service segments. Each service segment consists of a set of individual one-way train trips actually operated during the year. The individual one-way train trips in all defined service segments add up to all one-way train trips actually operated during the year. The following are several examples of segmentation meeting this requirement:

1. Individual train trips on weekday, Saturday, and Sunday schedules
2. Individual train trips on weekday schedules, individual Saturday routes, and individual Sunday routes
3. Weekday peaks, weekday midday, weekday night, Saturday, Sunday

To facilitate the estimation of annual average daily UPT and PMT by schedule type, schedule type should be one attribute of the service segments defined. All three of the above examples meet this requirement.

5.02 Data Variation

For the purpose of this methodology, one objective in selecting a level of segmentation and defining service segments is to minimize variation within each service segment. To better meet this objective in defining service segments, it is important to understand the nature of variation. This involves several aspects:

- Train-level UPT varies significantly. Train-level UPT and PMT for a given line typically are higher for trips in the peak direction, during peak periods, and on weekdays. Train-level UPT and PMT also can vary significantly across lines. In addition, train-level PMT typically is higher for longer lines than for shorter lines. For lines of similar length, train-level PMT typically is higher for lines whose passengers are more likely to travel from one end to the other than for other lines. For single-line systems, there is no cross-line variation.
- The number of cars on a train also varies, adding another source of variation in train-level UPT and PMT.
- Train-level PMT varies more than train-level UPT in most cases, which results from the variation in the distance traveled by individual passengers.
- APTL varies much less than both UPT and PMT across individual train trips.
- APTL typically does not vary significantly across individual one-way train trips for a given line; however, it can vary significantly across lines of different length.

5.03 Adequate Sample Size

Another objective in selecting the level of segmentation and defining service segments is to make sure that each defined service segment has an adequate number of train trips with valid APC data.

The agency should understand the importance of the data recovery rate of an APC system in defining service segments. For a service segment consisting of 40 train trips actually operated during a year, the number of train trips with valid APC data would be only 20 if the rate is 50%, 28 trips if the rate is 70%, and 36 trips if the rate is 90%. Although an agency can never be sure in advance of the actual recovery rate for the whole APC system or for individual segment segments, it can keep track of missing data for individual train trips actually operated throughout the year. At a minimum, it can learn from previous years.

Also important to understand is how the number of one-way train trips in individual service segments depends on the level of segmentation.

- A train trip on the weekday schedule for an entire year may be operated as many as 260 (=52*5) times during the year.
- A train trip on the Saturday schedule may be operated 52 times, and a trip on the Sunday schedule can be operated 52 times plus non-Sundays with Sunday service.
- These numbers drop quickly, however, if an agency changes its schedule during a year. In general, rail schedules are changed less frequently than bus schedules.
  - A train trip on a weekday schedule may be operated only 65 times during the year with quarterly service changes and only 43 times with bi-monthly service changes.
  - A train trip on a Saturday schedule may be operated only 13 times a year with quarterly service changes.

A determining factor is that each service segment defined should have at least 10 one-way train trips with valid APC data. To ensure this target sample size and to account for the uncertainty in the actual recovery rate for each service segment, **the minimum number of train trips actually operated should be at least 20 for any service segment defined for using this methodology.** For example, service segments defined by train trips on the weekday schedule would meet this requirement even if service changes occur bi-monthly. Service segments defined by train trips on weekend schedules with quarterly or bi-monthly service changes, however, would not meet this requirement. Service segments for weekend services, as a result, would need to be more aggregated than individual train trips on weekend schedules.

5.04 Convenience

The third objective in selecting the level of segmentation and defining service segments is the level of convenience in using this methodology. Service segments defined with fewer attributes generally would be better than those defined with more attributes. Individual routes and individual trips on schedules are examples that would be considered to meet this convenience requirement. Service segments defined as individual routes by direction, time of day, and day of week would be complex to use for this methodology. However, having an adequate sample size and minimizing variation are more important than convenience.

5.05 Frequency of Segmentation

It is possible to modify the service segments defined and used for the previous year using data from the current year. Estimation for a given year with this methodology is done after the year has ended and all raw APC data have been collected and processed. An agency can examine the pattern of train trips with missing data, evaluate the service segments defined previously, and modify if needed so each new service segment has an adequate number of train trips.

Having a new set of service segments every year is not suggested. Instead, the agency should study the pattern of train trips with missing data, define a set of service segments accordingly, and use them for this methodology for a number of years. If the monitoring process indicates that the pattern of train trips with missing data has changed significantly from the time when these service segments were defined, service segments should be modified according to the new pattern.

5.06 Suggested Segmentation

An agency may choose the level of segmentation and define the service segments accordingly. For agencies that are new to the methodology, a specific set of service segments is suggested. This segmentation consists of the following service segments:

- Individual train trips on a weekday schedule
- Individual lines for Saturdays
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight
- Individual lines Sundays
  - Midnight to noon
  - Noon to 6 PM
  - 6 PM to midnight

This suggested segmentation meets the requirement of being mutually-exclusive and adding up to the whole service, uses schedule type as one attribute, and balances all three objectives for defining service segments: minimizing variation, adequate sample size, and relative convenience.

Column A in Figure III-2 lists all service segments for a hypothetical service as an illustrative example. In this case, there are 26 service segments, including 3 Saturday segments, 3 Sunday segments, and 20 trips on the weekday schedule. No service changes occur during the year in this example. The service segments in this example follow the suggested segmentation approach.

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Define Service Segments**	**Schedule Type**	**Service Segment**	**Weekday Time Period**	**Count Annual Total Car Trips Operated**	**Count Annual Car Trips with Valid APC Data**	**Aggregate Annual UPT from Valid APC Data**	**Aggregate Annual PMT from Valid APC Data**	**Calculate Average UPT per Car Trip**	**Calculate Average PMT per Car Trip**	**Estimate Annual UPT**	**Estimate Annual PMT**
2					1	2	3	3	4	4	5	5
3	1	Saturday	Morning	N/A	300	185	15,267	43,159	82.5	233.3	24,758	69,988
4	2	Saturday	Afternoon	N/A	350	285	20,349	57,525	71.4	201.8	24,990	70,645
5	3	Saturday	Night	N/A	300	271	8,163	23,077	30.1	85.2	9,037	25,547
6	4	Sunday	Morning	N/A	300	277	5,128	15,279	18.5	55.2	5,554	16,548
7	5	Sunday	Afternoon	N/A	350	254	11,792	35,135	46.4	138.3	16,249	48,415
8	6	Sunday	Night	N/A	300	278	7,080	21,094	25.5	75.9	7,640	22,763
9	7	Weekday	Trip 1	AM peak	255	234	6,211	15,447	26.5	66.0	6,769	16,833
10	8	Weekday	Trip 2	AM peak	255	167	4,587	16,625	27.5	99.6	7,004	25,386
11	9	Weekday	Trip 3	AM peak	255	220	39,216	90,354	178.3	410.7	45,455	104,729
12	10	Weekday	Trip 4	AM peak	255	214	13,618	32,005	63.6	149.6	16,227	38,137
13	11	Weekday	Trip5	Midday	255	232	4,191	15,294	18.1	65.9	4,607	16,810
14	12	Weekday	Trip 6	Midday	255	194	7,653	36,413	39.4	187.7	10,059	47,862
15	13	Weekday	Trip 7	Midday	255	163	6,212	17,566	38.1	107.8	9,718	27,480
16	14	Weekday	Trip 8	Midday	255	178	20,465	55,558	115.0	312.1	29,318	79,592
17	15	Weekday	Trip 9	Midday	255	215	9,722	32,168	45.2	149.6	11,531	38,152
18	16	Weekday	Trip 10	Midday	255	154	15,306	51,286	99.4	333.0	25,345	84,921
19	17	Weekday	Trip 11	Midday	255	157	13,668	34,377	87.1	219.0	22,200	55,835
20	18	Weekday	Trip 12	Midday	255	231	9,302	31,291	40.3	135.5	10,269	34,542
21	19	Weekday	Trip 13	PM peak	255	177	7,423	25,577	41.9	144.5	10,694	36,849
22	20	Weekday	Trip 14	PM peak	255	207	16,529	39,861	79.8	192.6	20,362	49,104
23	21	Weekday	Trip 15	PM peak	255	173	1,908	7,634	11.0	44.1	2,813	11,252
24	22	Weekday	Trip 16	PM peak	255	165	7,194	30,460	43.6	184.6	11,118	47,075
25	23	Weekday	Trip 17	PM peak	255	186	16,760	61,758	90.1	332.0	22,977	84,668
26	24	Weekday	Trip 18	Other	255	190	11,489	22,112	60.5	116.4	15,420	29,676
27	25	Weekday	Trip 19	Other	255	233	3,987	49,367	17.1	211.9	4,364	54,028
28	26	Weekday	Trip 20	Other	255	162	4,489	77,983	27.7	481.4	7,065	122,751
29												
30	**Estimate Annual Total UPT and PMT for the Whole Service** (6)										381,542	1,259,588
31												
32	**Weekday Total UPT and PMT** (7)										293,315	1,005,682
33	**Saturday Total UPT and PMT** (7)										58,784	166,180
34	**Sunday Total UPT and PMT** (7)										29,443	87,726
35	**Total UPT for weekday AM peak** (8)										75,455	N/A
36	**Total UPT for weekday Midday** (8)										123,047	N/A
37	**Total UPT for weekday PM peak** (8)										67,964	N/A
38	**Total UPT for weeday other** (8)										26,849	N/A

Figure III-2. Example of Calculations for Estimation Procedure

6.00 Data Requirements and Estimation Procedure

6.01 Data Requirements

Ideally, an agency should have an annual database of train-level information for all one-way train trips actually operated during a full year. This database would accumulate such information as individual train trips are being operated and as the raw APC data are being processed from the first day to the end of the year. At a minimum, this database should have the following data items for each one-way train trip actually operated:

- Date on which the trip was operated
- Whether it was an atypical day
- Route
- Schedule type
- Number of cars
- Start time
- Service segment to which the trip belongs
- Whether raw APC data were returned for all APC units on the train
- Whether train-level valid APC data are available
- Train-level UPT if valid APC data are available
- Train-level PMT if valid APC data are available

Often, a transit agency operates its Sunday schedule on holidays that fall on Monday through Saturday. Agencies should include the data for these holidays under the day for the schedule that they operate (e.g., if operating on a Sunday schedule for a holiday on a Tuesday, the data would be included under Sunday).

These minimum data items are based on the suggested definition of service segments and the need for estimating monthly total UPT, annual average daily UPT and PMT by schedule type, and annual average weekday UPT by time period. Additional data items may be included for alternative ways of defining service segments. Trip direction would need to be added, for example, if an agency uses it in defining its service segments.

6.02 Continuous Monitoring

To use the methodology in this guidance as a certified alternative sampling technique, the agency must have a process in place to monitor and maintain individual APC units and the APC system continuously. An effective process, for example, would immediately notify the agency when an APC unit on a train trip just operated failed to get any raw APC data on the on-off activities for the trip and would indicate the source of the equipment failure problem. Once equipment failure occurs, the same effective process would have an ongoing partnership with vendors for quick maintenance and repair. Reducing the number of train trips with missing data due to equipment failure benefits NTD reporting and internal service and other planning efforts within an agency.

6.03 Estimation Procedure for Annual Totals

Once service segments are defined and the annual database is ready, the agency is ready for estimation with this methodology. As detailed below, the procedure for estimating annual totals with this methodology consists of six steps. The first three steps extract data from the database of train trips actually operated. The next two steps involve calculations for each service segment separately. The last step aggregates the results across all service segments. These steps are illustrated with an illustrative example in Figure III-2.

1. **Aggregate annual car trips operated** – This step obtains the annual number of one-way car trips actually operated during an entire year for each service segment. This would simply sum the number of cars across all train trips operated in the database for each service segment during the whole year. Column E of Figure III-2 shows these numbers for the example.
2. **Aggregate annual car trips with valid APC data** – This step obtains the annual number of one-way car trips actually operated for which valid train-level data on UPT and PMT are available for each service segment during an entire year. This would simply sum the number of cars across all train trips in the database with valid APC data for each service segment for the whole year. Column F of Figure III-2 adds these trips with valid APC data for each service segment.
3. **Aggregate annual UPT and PMT from valid APC data** – This step aggregates the train-level UPT and PMT across all one-way train trips actually operated with valid APC data during an entire year. This requires summing the train-level UPT and PMT across all rows in the database for each service segment. Columns G and H of Figure III-2 show the aggregated annual UPT and PMT from valid APC data, respectively, for each service segment.
4. **Calculate average UPT and PMT** – This step calculates the average UPT per one-way car trip and average PMT per one-way car trip by dividing the annual aggregated UPT and PMT from Step 3 by the annual number of one-way car trips operated with valid APC data from Step 2. Columns I and J of Figure III-2 calculate the average UPT and PMT per car trip, respectively, for each service segment. In the case of no trips with valid APC data for any specific service segment, use the averages for a similar service segment.
5. **Estimate annual UPT and PMT by service segment** – This step estimates annual UPT and PMT for each service segment by multiplying the average UPT and average PMT per car trip from Step 4 by the annual total number of one-way car trips actually operated from Step 1. Columns K and L of Figure III-2 have the estimated annual UPT and PMT, respectively, for each service segment.
6. **Estimate annual total UPT and PMT for the whole service** – In this step, the agency estimates the annual total UPT and PMT for the entire service by summing across all service segments the already-estimated annual UPT and PMT for each service segment. Row 30 of Figure III-2 presents the estimated annual totals. Shown in cell K30, the annual total UPT for the whole service is estimated at 381,542. Shown in cell L30, the annual total PMT for the whole service is estimated at 1,259,588.

6.04 Estimating Annual Total UPT and PMT by Schedule Type

6.04.1 No Atypical Days

If an agency had no atypical days during an entire year, it can easily estimate annual total UPT and PMT by schedule type from using the summary data at the level of individual service segments used in estimating annual totals—simply sum the estimated annual UPT and PMT across all service segments for each schedule type. In the example, the annual total PMT is 1,005,682 for weekday schedule, 166,180 for Saturday schedule, and 87,726 for Sunday schedule.

Once the annual totals by schedule type are estimated, it is straightforward to estimate annual average daily UPT and PMT by schedule type—simply divide the annual totals for each schedule type estimated above by the number of days operated for the same schedule type. The number of days operated by schedule type used here should be consistent with those reported to the NTD.

6.04.2 Having Atypical Days

If an agency did have atypical days during a year, it cannot use the segment-level summary data from estimating annual totals for the whole service. Instead, it needs to follow the six steps again without including train trips operated on atypical days to estimate the annual totals needed for this purpose by schedule type. With atypical days being identified in the database, excluding them for Steps 1 through 3 is straightforward.

Annual average daily trips by schedule type are similarly estimated. In this case, however, the number of days operated must exclude atypical days as well.

6.05 Estimating Average Weekday UPT by Time Period

6.05.1 No Atypical Days

If an agency had no atypical days during an entire year, it can easily estimate annual total UPT for each weekday time period from using the summary data at the level of individual service segments used in estimating annual totals—simply sum the estimated annual UPT across all weekday service segments for each weekday service period. In the example, the annual total UPT is 75,455 for weekday AM peak, 123,047 for weekday midday, 67,964 for PM peak, and 26,849 for weekday other period.

Once the annual total UPT for each weekday time period is estimated, it is straightforward to estimate annual average weekday UPT by time period—simply divide the annual total UPT for each time period estimated above by the number of days operated for the weekday schedule. The number of days operated for the weekday schedule used here should be consistent with those reported to the NTD.

6.05.2 Having Atypical Days

If an agency did have atypical days during a year, it cannot use the segment-level summary data from estimating annual totals for the whole service. Instead, it needs to follow the six steps again without including train trips operated on atypical days to estimate the annual totals needed for this purpose by weekday time period. With atypical days being identified in the database, excluding them for Steps 1 through 3 is straightforward.

Annual average weekday UPT by time period are similarly estimated. In this case, however, the number of days operated for the weekday schedule must exclude atypical days as well.

6.06 Estimating Monthly UPT

Although the estimation procedure is stated for estimating annual total UPT and PMT, it is just as applicable to monthly estimation once the period of data is changed from one year to one month. Only total UPT is needed for monthly reporting.

To maintain consistency between the estimated annual total UPT from this methodology for annual reporting with the sum of the monthly UPT values for monthly reporting, it is important that the reported monthly UPT comes from the same methodology.

A potential complication with monthly estimation with this methodology is that the chance is higher for some service segments to have no train trips operated with valid APC data during a month. This complication is not an inherent shortcoming of this methodology; in fact, it can happen with intentional sampling as well if the monthly portion of the annual sample is used in monthly reporting. In those cases, the average UPT per car trip from a similar service segment from Step 4 should be applied for Step 5.

Agencies are discouraged from using the sum of monthly UPT estimated with this methodology as the annual total UPT for annual reporting. Potentially, it is possible to estimate monthly UPT and PMT every month and to add these monthly estimates as the annual total UPT and PMT. The advantage of doing this is that the annual total UPT would be equal to the sum of the reported monthly UPT values. The shortcoming is that the monthly number of train trips with valid APC data is likely too small for some service segments. In addition, the annual total UPT from adding the monthly total UPT would be inconsistent with the annual total PMT estimated annually.

6.07 Estimating APTL

Once an agency has estimated annual total UPT and PMT for the whole service and annual total UPT and PMT for each schedule type, it can easily estimate the corresponding APTL. If the agency has a reliable 100% UPT count from a non-APC source and choose to report it to the NTD, it may want to estimate APTL from the estimated UPT and PMT data with the methodology in this guidebook.

7.00 Appendix A – Certification Document

7.01 Introduction

This appendix serves as the document of certification for the methodology in this guidebook for an agency to estimate annual total unlinked passenger trips (UPT), annual total passenger miles traveled (PMT), or average passenger trip length (APTL) as the ratio of annual total PMT over annual total UPT through using data from its automatic passenger counter (APC) system. It certifies that this methodology meets the 10% precision and 95% confidence levels as long as the agency meets the necessary conditions and follows the guidance in this guidebook. This document lists the certification conditions, describes the methodology, justifies the certification, and summarizes my qualifications as a qualified statistician.

7.02 Conditions

To use the methodology as a pre-certified alternative sampling technique, the agency must meet all of the following conditions:

- The agency operates at least 1,000 train trips annually (e.g., 4 trips daily).
- APCs must cover all doors of all passenger cars of the fleet.
- The agency should have a continuous monitoring and maintenance program.
- The agency must recover valid APC data from at least 50% of all train trips operated.
- The agency must pass a test of statistical equivalence for its APC data.
- The agency must use all valid APC data for estimation.

7.03 Methodology

This methodology consists of a sampling component and an estimation component. The sampling component is based on full-population sampling (i.e., every train trip is sampled) with some degree of nonresponse (i.e., valid APC data are not recovered from some train trips operated). To minimize any potential bias from missing APC data, the estimation component requires post-stratification as follows:

- Divides the whole service into a set of mutually-exclusive service segments (e.g., a trip on the weekday schedule, a route, etc.).
- Estimates total transit usage for each service segment by extrapolating all valid APC data for that service segment. For each segment, this is to multiply the average transit usage per trip among all trips with valid APC data by the annual total number of trips actually operated for that service segment.
- Sums the estimates of segment-level total transit usage across all service segments to get total transit usage for the whole service.

7.04 Justification

Since this methodology uses full-population sampling, estimates from this methodology have no sampling error. However, these estimates can still have errors from other sources. As summarized in Table III-2, these errors combined are within ±9.0% at the 95% confidence level.

Table III-2. Precision at 95% Confidence

Error Type	Magnitude
Random sampling error	0%
Random measurement error	≤ ±1.12% with at least 4 daily trips operated on 250 days a year & data recovery rate ≥ 50%
Systematic measurement error	≤ ±7.5%
Systematic error from missing data	< ±0.2%
Total @95% confidence	< ±9.0%

The valid APC data for any train trip can differ from the unknown true value. Systematic measurement errors exist because the average of differences between the valid APC data and the true value cross all trips typically is not zero. Random measurement errors exist because the difference varies randomly across train trips. Systematic errors can result not only from the valid APC data being different from the true value but also from the trips with valid APC data not being representative of all trips operated. The systematic error due to missing data may be referred to as the missing-data bias.

Both types of measurement errors can be assessed through paired APC and manual data from a random sample of trips with an adequate sample size. This assessment is part of conducting and passing the test of statistical equivalence. Random measurement errors become negligible from using all valid APC data. Random measurement errors would be within ±1.12% for the extreme case of operating only 4 trips daily on weekdays only with a 50% recovery rate of valid APC data. Random measurement errors are smaller for all other cases and, in fact, are well within ±0.5% for agencies with at least 20 train trips daily. Systematic measurement errors do not become smaller from using more data but will be controlled to be within ±7.5% by the requirement of passing the test of statistical equivalence. As a result, the two measurement errors combined would not exceed ±8.62% for all cases and would be within ±8.0% for most cases.

The missing-data bias is not directly measurable. As part of the effort in developing the guidance for this methodology, a comprehensive simulation was carried out to measure a base level of the missing-data bias when estimation is done without using this methodology and to examine how much of this base bias would disappear once estimation is done with this methodology. This assessment was done for a range of possibilities on how train-level usage varies, how the data recovery rate varies, etc. The results show that the base level of bias can be as high as 7% for some circumstances. One example of such circumstances is when train-level PMT varies significantly across service segments and data recovery rates are low and vary significantly across these service segments. The same simulation, however, also showed that the large base bias largely disappeared (i.e. within ±0.2%) once estimation is done through this methodology.

7.05 Qualification

I have a Ph.D. in Economics with a concentration in Transportation Economics from the University of California, Irvine. I have developed and certified alternative sampling techniques for many transit agencies. The Federal Transit Administration has adopted the *National Transit*

Database Sampling Manual that I developed to give detailed guidance to NTD reporters for determining annual service-consumed data for all modes.

8.00 Appendix B – Glossary and Formulas

8.01 Glossary

The following definitions are largely based on the *NTD Glossary* and the *2015 NTD Policy Manual*.

Atypical Days: Atypical days occur when a transit agency does not operate its normal, regular schedule. Instead, the agency:

- Provides extra service to meet demands for special events, such as conventions, parades, or public celebrations, or
- Operates significantly reduced service because of unusually bad weather (e.g., snowstorms, hurricanes, tornadoes, earthquakes) or major public disruptions (e.g., terrorism)

The concept of atypical days is relevant only for scheduled, fixed-route services, such as motorbus (MB), commuter bus (CB), bus rapid transit (RB), trolley bus (TB), and rail modes.

Average Passenger Trip Length (APTL): The average distance ridden for an unlinked passenger trip (UPT) computed as passenger miles traveled (PMT) divided by UPT; may be determined by sampling, or calculated based on actual data.

Days Operated: Number of days that service was actually operated according to the schedule of service.

Hybrid Rail (YR): Rail system primarily operating routes on the National system of railroads, but not operating with the characteristics of commuter rail. This service typically operates light rail-type vehicles as diesel multiple-unit trains. These trains do not meet Federal Railroad Administration standards, and so must operate with temporal separation from freight rail traffic.

Light Rail (LR): A transit mode that typically is an electric railway with a light volume traffic capacity compared to heavy rail (HR). It is characterized by:

- Passenger rail cars operating singly (or in short, usually two car, trains) on fixed rails in shared or exclusive right-of-way;
- Low or high platform loading; and
- Vehicle power drawn from an overhead electric line via a trolley or a pantograph.

Passenger Miles Traveled (PMT): The cumulative sum of the distances ridden by all passengers.

Streetcar Rail (SR): This mode is for rail transit systems operating entire routes predominantly on streets in mixed-traffic. This service typically operates with single-car trains powered by overhead catenaries and with frequent stops.

Unlinked Passenger Trips (UPT): Number of passengers who board public transportation vehicles. Passengers are counted each time they board vehicles no matter how many vehicles they use to travel from their origin to their destination.

Vehicles Operated in Annual Maximum Service (VOMS): The number of revenue vehicles operated to meet the annual maximum service requirement. This is the revenue vehicle count during the peak season of the year on the week and day that maximum service is provided. VOMS excludes atypical days or one-time special events.

8.02 Formulas for Equivalence Testing

Those interested in the statistics involved in this test equivalence may refer to the following symbols, statistical concepts, related formulas, hypotheses, and testing criteria:

- *n* paired observations for a given certification metric (e.g., PMT per trip)
- y_{ai} (APC) and y_{mi} (manual)
- Mean of manual data: $\overline{y_a} = \frac{\sum y_{ai}}{n}$
- Difference for each pair: $d_i = y_{ai} - y_{mi}$
- Mean difference: $\bar{d} = \frac{\sum d_i}{n}$ (or $\frac{\bar{d}}{\overline{y_a}}$ in percent terms)
- Standard deviation of difference: $s_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1}$
- Standard error of mean difference: $s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$ (or $\frac{s_{\bar{d}}}{\overline{y_a}}$ in percent terms)
- Significance level for statistical testing: α (e.g., 5%)
- Critical value in the t distribution: $t_{n-1,1-\alpha}$
- Population mean difference: μ_d (the true value but not observed)
- Equivalence bounds: $\pm\theta$

The null hypothesis is H_0: $\mu_d \le -\theta$ or $\mu_d \ge \theta$. The alternative hypothesis is H_1: $|\mu_d| < \theta$.

Not Equivalent **Equivalent**

The equivalence of APC and manual data is accepted if the null hypothesis is rejected at the 1-α confidence level. More specifically, equivalence is established if the following two one-sided t-tests both are rejected:

$$t_1 = \frac{\bar{d}-(-\theta)}{s_{\bar{d}}} > t_{n-1,1-\alpha} \text{ AND } t_2 = \frac{\bar{d}-\theta}{s_{\bar{d}}} < -t_{n-1,1-\alpha}$$

Alternatively, equivalence is established if the (1-2α)% confidence interval falls within (-θ, θ).