9-1989

An Evaluation of Automatic Passenger Counters: Validation, Sampling, and Statistical Inference

James G. Strathman
Portland State University

Citation Details

Strathman, James G., "An Evaluation of Automatic Passenger Counters: Validation, Sampling, and Statistical Inference" (1989). *Center for Urban Studies Publications and Reports*. 109.
https://pdxscholar.library.pdx.edu/cus_pubs/109

An Evaluation of Automatic Passenger Counters: Validation, Sampling and Statistical Inference

James G. Strathman

Final Report
September 1989

Center for Urban Studies
School of Urban and Public Affairs
Portland State University
Portland, OR 97207-0751

Dr. Strathman is an Associate Professor of Urban Studies and Planning and Research Associate in the Center for Urban Studies, Portland State University.

Support for this project was provided by the U.S. Department of Transportation and the Tri-County Metropolitan Transportation District of Oregon (Tri-Met).

Abstract

While automatic passenger counters (APC's) offer the potential for cost effective data recovery and management, they also introduce new complications in the data recovery process. This report addresses three issues associated with the implementation of APC's, based on an evaluation of the recent experiences of the Tri-County Metropolitan Transportation District of Oregon (Tri-Met). First is the issue of validation, which is concerned with both the recovery and accuracy of APC passenger data. The second issue concerns the development of a sampling methodology for APC's compatible with UMTA's Section 15 reporting requirements. Third is the issue of inferring system-level ridership from sample data in the presence of selective APC failures.

We find that the APC's are providing systematically accurate passenger counts. Analysis of the data recovered from September to November 1988 also shows that sampling was representative, based on the set of "trains" from which ridership data were successfully recovered. The initial selection/assignment of trains, however, was not representative.

Given that APC's record operating data for all bus trips comprising a train assignment, a cluster sampling method is formulated that ensures an overall random selection of bus trips via a random first stage selection of trains.

Selective data recovery failures can hamper the process of inferring system-level ridership from the sample estimates. For example, when failure rates vary by bus type or time of day, inferences drawn from the sample of recovered data may over or under-represent total system ridership. In such circumstances, post hoc stratification of the sample data may be required. We outline several alternative corrections based on a-priori knowledge of the mix of bus types and schedule characteristics in the system.

Acknowledgements

The author gratefully acknowledges the efforts of Janet Hopper, who undertook a large part of the statistical analysis presented in this report, and whose knowledge of transit operations proved vital in framing the research problems encountered in this project. Jon Lutterman provided information on the sampling methodology presently used by Tri-Met, as well as statistics on APC data recovery. Rick Gerhart provided helpful comments on the draft of the report. Research assistance was also provided by Chall Kyoung Sung.

Introduction

Automatic passenger counters (APC's) offer a number of potential benefits to transit operators in the areas of data acquisition, management and utilization. Compared with manual collection, APC's have been found to be cost effective for larger transit systems, while also providing better data turnaround and improved accuracy [5]. They are also technically capable of recovering the very large quantities of information required in analyzing transit performance at the disaggregate level, thus permitting greater sensitivity in service scheduling and planning. Along with these potential gains, however, come several complications not found with manual data collection. First, only a selected number of buses in the fleet - usually around 10 per cent [2] - are APC-equipped, and this results in a dependence on bus-specific assignments to selected routes rather than random assignment of surveyors. Even under the best of circumstances - where the requests for and actual assignments of APC buses are well coordinated - less flexibility exists in the data recovery process. Second, while APC's generally return more accurate data than manual counters, much of the data that is recovered is screened out due to functional inconsistencies. Apart from the resulting need for larger sample sizes is the question of whether, following the screening of unusable data, the remaining information still constitutes a representative sample of bus trips for the system. If "failure" rates are systematically related to route or other operational-specific characteristics, a non-response type of bias might undermine the sample ridership statistics and, consequently, inferences of system-wide operating performance. Third, with manual data collection, surveyors are assigned to randomly selected bus trips. In contrast, with APC's the unit of observation is the train, which consists of all the scheduled service performed by an APC bus during an operating day. The bus trips comprising a train cannot be assumed to be independent, and thus the sampling framework recommended by UMTA [7] cannot be employed. Short of gross over-sampling, an alternative methodology must be designed consistent with the APC's operating features.

These issues are addressed in the coming sections. Utilizing information drawn from the recent performance of APC's employed by the Tri-County Metropolitan Transportation District of

Oregon (Tri-Met), we first take up the matter of data recovery by analyzing the accuracy of the data generated by APC's as well as the sources of data recovery failures. We then proceed to determine whether the set of trains from which data has been successfully recovered represents an equal probability sample. Following this, a sampling methodology is developed which ensures that the selection of bus trips (via the selection of trains) is both random and of sufficient size to comply with UMTA's Section 15 reporting requirements. Finally, we suggest a remedy for correcting sample statistics subject to bias from non-random data recovery failures.

Evaluation of APC Performance

Data Recovery

Tri-Met's APC system uses infrared sensors located about waist-high at the stairwells of the front and rear bus doors. An on-board microprocessor records passenger boardings and alightings, times and distances. At the end of a run the recovered data is transferred to a microcomputer using an automated infrared transmitter that scans the buses from fixed stations at each of the agency's three garages. The system was manufactured by Red Pine Instruments of Denbigh, Ontario, and is installed on 50 of Tri-Met's 567 buses. Implementation of the APC's was initiated in 1982, and Tri-Met has relied on the system to provide data for UMTA Section 15 reporting since the 1986 fiscal year. APC generated data are also used internally for route performance reporting, and contribute to a lesser extent to scheduling and analysis.

Software for validating and managing the APC data was developed in-house. Incoming data are assigned route and bus identification codes, and are then aggregated to the bus trip level. A program then checks the data for compatibility with various validation standards. Train or trip level data that fail to meet these standards are purged. At the train level, observations are deleted for the following reasons:

a. Recorded distance differs from actual by more than 15 per cent;

b. Pull out to pull in time differs by more than 30 minutes from the service schedule;

c. Total boardings and alightings differ by more than 10 per cent.

Validation standards covering distances and pull-out and pull-in times at the bus trip level are also applied. If a number of the trips in a train are deleted, the remaining trips in that train are more thoroughly evaluated manually, which may result in purging the data from an entire train.

The sampling plan used by Tri-Met is organized around the five sign-up periods comprising annual scheduled service. The objective of the plan is to uniformly sample the scheduled trips in each sign-up. The execution of the sampling plan requires the involvement of several divisions. The scheduling division is responsible for drawing a daily sample of trains, using a selection program that assigns higher selection priorities to trains which have been previously under-sampled. The trains selected for sampling by the scheduling division are called "requests." Daily lists of requests are provided to the operations division, which is responsible for assigning an appropriate APC-equipped bus model to each of the trains requested. In practice, not all the trains from the daily list of requests are successfully assigned an APC bus, and sometimes APC buses are assigned to trains which were not requested. Thus the daily tally of "assignments" consists of a group of trains for which APC buses were both requested and assigned, and a group of trains for which APC buses were assigned but were not requested. Finally, the train assignments (both requested and unrequested) that return valid data are defined to represent the set of "successfully sampled" trains.

Information on the degree of success recently encountered by Tri-Met in recovering data with the APC system is presented in Table 1. Records from the first half of the April-June 1989 sign-up identify 1,589 requests, of which 1,089 (69 per cent) were assigned APC buses. Another 325 trains that were not requested were assigned APC buses. Valid data was successfully recovered from 286 of the trains that had been requested and from 82 unrequested trains. Thus data was recovered from 26 per cent of all assignments.

Data losses resulted from various causes, including exceeding the time tolerances (7 per cent of the total failures), distance tolerances (5 per cent), boardings/alightings discrepancies (7 per cent), incorrect or missing assignment information in the train records (11 per cent), recovered data that was unusable (8 per cent) and failures due to bus or equipment malfunction (62 per cent).

The latter category represents cases where no data was returned by the APC's. The reasons for failures in this category would include instances where the APC unit accidently resets, where buses did not pull close enough to the transmitter to allow transfer of the data, where the microprocessor's memory was filled and could not record more data, and where data was simply not recorded due to equipment breakdown.

Table 1
Breakdown of APC Data Recovery, April 1989 Sign-up

	No.	%[1]
1. Trains Requested	1,589	—
2. Trains Assigned		
a. As requested	1,089	77
b. Unrequested	325	23
c. Total assignments	1,414	—
3. Data Recovered		
a. From requested trains	286	78
b. From unrequested trains	82	22
c. From all assigned trains	368	—
4. Data Recovery Failures, due to		
a. Time tolerances	71	7
b. Distance tolerances	56	5
c. On/off tolerances	71	7
d. Incorrect/missing assignment information	113	11
e. Unusable data	85	8
f. No data	650	62
g. All sources (a-f)	1,046	—

[1] The percentage figures pertain to the breakdowns within each numbered category.

Of the 1,414 train assignments 368, or 26 per cent, returned valid data. This rate of successful data recovery is considerably lower than what has been reported in other studies of APC performance [3,5]. Generally, about 80 per cent of all train assignments have been reported to return valid data. The reasons for this difference cannot be further explored given the lack of more detailed information about the performance of other APC systems. Among the factors contributing to Tri-Met's low data recovery rate could be differences in the screening tolerances employed in validating the data, differences due to the mix of APC-equipped bus types in Tri-Met's fleet, and differences in the APC technology. Given both the relatively small data recovery rate and the inclusion of unrequested trains, the question of non-response and/or sampling bias also arises. As a result, it is necessary to determine if the data losses were a random phenomenon or if they were systematically related to train-specific characteristics.

Determinants of Successful Data Recovery

We selected the September-November 1988 sign-up for a statistical analysis of factors related to successful data recovery. This sign-up was considered to be typical by Tri-Met staff in regard to APC performance and other operating and ridership characteristics. The sign-up consisted of 588 weekday trains and 1,552 assignments that returned valid data. Trains were defined to represent the unit of analysis, and the following model was specified to examine the effects of train-specific characteristics on successful data recovery:

$$\text{SAMP} = f(\text{APC, REQ, ASG, AM, PM}, G_1, G_2, \text{ARTIC, ADB, B500, B300}), \text{ where}$$

SAMP = the number of assignments in each given train that recovered valid data;
APC = the number of available APC buses of the requested type at the garage from which each given train assignment was made;
REQ = the number of times each given train was requested;
ASG = the number of times each given train was assigned;
AM = a dummy variable equalling 1 if the train provided only AM peak service and 0 otherwise;
PM = a dummy variable equalling 1 if the train provided only PM peak service and 0 otherwise;
G_1 = a dummy variable equalling 1 if the train was dispatched from Garage #1 and 0 otherwise;

G_2 = a dummy variable equalling 1 if the train was dispatched from Garage # 2 and 0 otherwise;

ARTIC = a dummy variable equalling 1 if the train was an articulated bus model (Crown-Ikarus) and 0 otherwise;

ADB = a dummy variable equalling 1 if the train was an ADB bus model (40 foot GMC RTS-II) and 0 otherwise;

B500 = a dummy variable equalling 1 if the train was a B500 bus model (40 foot Flexible "Metro") and 0 otherwise;

B300 = a dummy variable equalling 1 if the train was a B300 bus model (35 foot Flexible "New Look") and 0 otherwise.

The APC variable is included in the specification to account for differences in the number of APC buses of each relevant type at each garage. Controlling for the variation in assignments, we would expect that data recovery would improve when more buses are available for assignment. The number of requests is included to control for trains whose assignments were deterred for a variety of functional and mechanical reasons. Tri-Met's sampling software places a higher subsequent selection priority on trains that are requested but not assigned. The number of assignments controls for variations in data recovery attributable to the relative frequency of train assignments; in other words, some trains may have recovered data more frequently because they were assigned more frequently. The AM and PM peak dummy variables are included because these trains are in service for a shorter time period, and should be expected to be more reliable in terms of returning data successfully. They are also likely to have higher ridership per bus trip than "day" trains, and thus could shift the sample statistics upward if they are over-represented. The garage dummy variables are included to check for differences in data recovery that could be attributed to factors that could be traced to the performance of the system among Tri-Met's three garages. The variables G_1 and G_2 represent the operator's two satellite facilities. The four fleet type dummy variables are included to determine whether variations in data recovery can be linked to the mix of bus types in the system.

Table 2 presents descriptive statistics and the parameter estimates for the data recovery model. The R^2 of .62 and overall F value of 86.14 indicate that the model provides a moderately strong fit of the data. The parameter estimates for APC, REQ and ASG have the expected signs and are highly significant. AM peak trains are found to return .3 more observations per train than day trains, while the net increase for PM peak trains is about .8. Both are statistically significant and represent increases of approximately 10 and 30 per cent over the data recovery rate for day trains. Among the various bus types, the ADB and B300 models were found to recover 2.1 and 1.3 more observations per train than the "reference" bus type (B100/1000, which includes 40 foot AMGeneral and 40 foot Flexible "New Look" models) during the sign-up. Garage 1 generated .67 fewer observations per train, while Garage 2 produced .86 more in relation to the central garage. These differences are most likely due to breakdowns of the fixed-station transmitters at the garages, given that assignments are proportionately distributed among the three garages. The transmitter at Garage 2, by implication, experienced fewer breakdowns than the transmitters at the other two garages. Alternatively, it may be that some routes are more likely to return valid data than others, and if the composition of route types varies by garage this could also affect their relative data recovery rates.

Table 2

Regression Estimates of the Determinants of Train Level Data Recovery, September 1988 Sign-up

Variable	Mean	St. Dev.	Coefficient	t-ratio
Constant	----	----	.26	1.76
APC	6.86	5.01	.169	4.67**
REQ	4.46	5.03	-.252	-10.81**
ASG	5.52	3.66	.235	7.09**
AM	.29	.45	.311	2.43*
PM	.30	.46	.782	6.10**
G_1	.31	.46	-.666	-4.20**
G_2	.32	.47	.861	6.16**
ARTIC	.15	.35	-.136	-.76
ADB	.15	.36	2.090	10.62**
B500	.09	.28	-.061	-.29
B300	.30	.46	1.285	8.48**

$R^2 = .62$
$F = 86.14$
$n = 588$

---

* Significant at the .01 level.
** Significant at the .0001 level.

Apart from isolating various determinants of successful data recovery, the regression results generally point to possible sources of over and under-representation of trains in the effective sampling scheme. Of particular concern in this regard are the AM and PM peak trains and two of the bus types. To the extent that the trains in question differ significantly in their ridership characteristics, these differences can represent a source of bias in the overall sample estimates of ridership and other operating characteristics. This issue is addressed further in the section on sample inferences.

Measurement Accuracy

For the data that is successfully recovered by the APC's, another concern regards the accuracy of the passenger counts generated. Automatic counters have been described as more accurate than manual data recovery, particularly for high volume routes and routes with peak period standing loads [6]. The errors that have been observed with APC's indicate a tendency to undercount rather than over-count passenger activity, while boardings tend to be counted more accurately than alightings.

In a demonstration study of APC's equipped with infrared beams, the Washington Metropolitan Area Transit Authority conducted an accuracy test on a sample of over 400 bus trips involving about 18,000 boardings and alightings [8]. It was found that the total boardings recorded by the APC's equalled 99.7 per cent of the manual counts, while recorded alightings equalled 98.4 per cent of the manual counts. However, the circumstances of this evaluation were quite controlled, with a limited number of routes included in the survey. A field test in 1982 of five properties employing APC's (Minneapolis/St. Paul, Columbus, Kalamazoo, Seattle and Los Angeles) found slightly larger discrepancies between APC counts and recordings by manual checkers, although the differences were not statistically significant [5].

Previous research on the issue of accuracy has thus consistently demonstrated that APC and manual passenger counts tend to correspond. The APC systems evaluated were relatively new, however. Tri-Met's APC's have been in service for nearly seven years and, given their low data recovery rate, have not been performing to the levels observed elsewhere. A statistical comparison of APC and manual passenger counts for Tri-Met's system was undertaken as a result.

Forty-six APC buses were selected for the evaluation. The buses were assigned to a representative set of routes, and both manual and automatic counts of boardings and alightings were recovered for each stop. The number of stops per bus ranged from 44 to 148, and totalled 3,768 across all observations. A test of the mean difference in APC versus manually recorded boardings and alightings per stop was conducted for each bus, as well as for the overall sample. Table 3 reports the findings for the overall analysis and for those buses where significant differences between APC and manual counts were found. Across all buses and all stops the average boardings per stop counted by the APC were .01 passenger higher than the manual count, while the number of alightings counted by the APC's averaged .01 passenger lower. Neither of these differences was statistically significant at the .05 level. Of the six instances where the APC and manual boarding counts differed significantly, three involved over-counting and three involved undercounting. Of the five instances where the APC and manual alighting counts differed, two involved over-counting by the APC. Three specific buses were associated with significant differences of both boardings and alightings.

Even when significant differences between APC and manual counts are found, as a result, no consistent pattern of divergence is evident. Given 92 observations, we would expect nearly five instances of Type I error in the analysis. Moreover, an underlying assumption is that the manual counts themselves are measured without error, and this is most likely to be violated in some cases. Finally, the data recovered by the APC's was not subjected to the normal screening process, which would have purged substantial portions of the data recovered from several buses (i. e., #347 and #731).

Table 3

Tests of Differences Between APC and Manual Counts: Overall Results and Cases Involving Significant Differences

Boardings

Bus #	No. of Stops	APC - Manual	t - ratio
347	80	.25	2.78
350	142	.13	2.71
901	81	-.11	-2.58
731	62	-.35	-2.50
119	82	.09	2.16
1040	81	-.10	-2.04
All Buses	3,768	.01	.68

Alightings

731	62	-.52	-3.12
347	80	.15	2.80
119	82	-.12	-2.43
526	85	.09	2.19
900	138	-.07	-2.07
All Buses	3,768	-.01	-1.38

Sampling With APC's

There are two primary issues that need to be addressed in regard to sampling with APC's. The first concerns the fact that the data recovery rate with the APC's is relatively low, and includes observations on some bus trips that were assigned but not requested in the sampling methodology. This raises questions about the representativeness of the sample, which could be found to be failing from assignment and/or response bias. The second issue concerns the sampling methodology itself. The sampling procedure recommended by UMTA [7] is essentially designed with manual data collection in mind, given that it provides solely for independent random selection of bus trips. With APC's, bus trips are necessarily selected in blocks comprising trains. Thus while trains can be selected in an independent and random fashion, the individual bus trips cannot. As a result, a specific methodology for APC's must be developed to ensure that the UMTA precision standards are satisfied while minimizing the necessary number of bus trips required to be sampled.

Evaluation of the Recovered Sample

There are three possible threats to representativeness in the sampling of APC-equipped trains. First (and of least concern here), the initial requests for train assignments may not be representative. Second, the actual assignments may not be representative if they do not fully correspond with the requests. Third, the trains from which data are ultimately recovered may not be representative, given the previously identified association between selected train characteristics and successful data recovery. The latter two possibilities are addressed in this section in an evaluation of the September-November 1988 sign-up. Train requests are not evaluated because the selection procedure used by Tri-Met assigns a higher priority to trains that were previously requested but not assigned. Thus if requests were found to be unrepresentative, it would be difficult to distinguish if this were due to problems associated with the request or the assignment process.

A chi-square test was employed to determine if the systematic patterns of trains that were requested and assigned, assigned, and successfully sampled represented an equal probability sample. The results of the tests are given in Table 4. The null hypothesis that the observations constituted an equal probability sample is rejected at the .05 level for trains that were requested and assigned, and for total assignments. It could not be rejected, however, for the trains that successfully generated data. This finding is in part attributable to the smaller number of successful assignments in comparison with the total assignments, which correspondingly reduces the comparative inter-train variance and the calculated chi-square value. It also indicates why the chi-square is considered to be a relatively weak test statistic (i.e., it is sensitive to the scale of measurement).

Table 4

Chi-Square Results for Trains in the September 1988 Sign-up

	Requested/Assigned	All Assignments	Recovered Data
Mean observations per train	3.1	5.5	2.6
Calculated chi-square value	2,236.0	1,147.0	710.0
Critical value, .05 level	720.0	720.0	720.0
Number of trains	588.0	588.0	588.0

An APC Sampling Methodology

The objective in designing a sampling methodology for the APC's is to identify the minimum number of randomly selected trains required to generate passenger information at the bus trip level that will satisfy UMTA's precision standard of +/- 10 per cent at the 95 per cent level of confidence. The methodology must account for correlation among bus trips within trains, and it should set the sample size sufficiently large enough to reflect the anticipated data recovery rate.

The special features associated with the APC data recovery process are compatible with a multi-stage cluster sampling method [4]. The first stage in this methodology would be defined to consist of a random selection of trains, and the second stage would then be defined by the 100 percent "clusters" of bus trips comprising the selected trains. Variations in cluster sizes would also be accommodated, recognizing that the number of bus trips can vary by train. The methodology would be designed for implementation at the train level, consistent with the manner of data recovery using APC's, while yet ensuring that the sample statistics satisfy trip level precision requirements.

The determination of the required sample size for the cluster sampling method follows from the convention for simple random sampling, with modification to account for the trip clustering effect. The sample size is first determined at the bus trip level and then converted to the train level based on the average number of bus trips observed per train. In the presentation that follows, the sample size is determined on the basis of recorded passenger miles, given that the relative variance of passenger miles tends to be larger. We can thus be confident that the sample size will be more than sufficient for the other operating data to be collected. The minimum number of bus trips to be sampled, in conformance with the UMTA Section 15 standards, is then defined as follows:

$$n_c = [(1.96 \cdot S_c) / (.1 \cdot M)]^2 ,\ \text{where}$$

n_c = the number of bus trips required in a multi-stage cluster sample;
S_c = the standard deviation of passenger miles per bus trip for a multi-stage cluster sample;
1.96 = the critical z value at the .025 level;
M = the mean passenger miles per bus trip.

The sample size equation presented above is equivalent to the arrangement used to determine the required number of observations for a simple random sample, with the exception of the cluster sample standard deviation term, which accounts for the interdependence of bus trips within trains and variation in the number of bus trips per train. While the standard deviation for a simple random sample need not be elaborated, its counterpart for a multi-stage cluster sample warrants presentation. This standard deviation is defined as follows:

$$S_c = [(1/n - 1) \cdot \sum_i n_i \cdot (M_i - M)^2]^{.5}, \text{ where}$$

n_i = the number of bus trips in train i;
M_i = the mean passenger miles per bus trip for train i;
M = the mean passenger miles per bus trip across all bus trips.

Sample statistics from previously collected data can be used to derive the required sample size. Using the September-November 1988 sign-up as an example, the overall mean passenger miles per bus trip is 8,481 and the multi-stage cluster sample standard deviation is 19,159. The minimum required sample size for the sign-up in the example is thus derived as follows:

$$\begin{aligned} n_c &= [(1.96 \cdot 19{,}159) / (.1 \cdot 8{,}481)]^2 \\ &= 1{,}961 \text{ bus trips.} \end{aligned}$$

The sample size derived above represents 14 per cent of the 13,955 trip observations actually recovered during the September-November 1988 sign-up. Using the cluster sampling framework we found that sample produced precision of +/- 3.7 percent at the 95 percent level of confidence. The new cluster sample size represents .6 per cent of the 332,154 scheduled trips during that sign-up.

To achieve the required sample size we should also take the data recovery rate into account. From Table 1 we see that 26 per cent of all assignments return usable data. This would suggest that to achieve the necessary number of valid observations, a total of 7,542 bus trip assignments would have to be made. In other words, about 2.3 per cent of all scheduled trips would need to be assigned APC buses to generate a sufficient number of validated sample observations. This number of assignments is probably excessive, given that we should expect to observe an improved data recovery rate from smaller sized samples (as indicated by the APC coefficient in the regression model).

Given that trains are the unit of assignment with APC's, it is also necessary to translate sample size requirements from bus trips to this unit. From the sign-up in the example, we find an average of 8.98 bus trips per train. Thus a minimum sample size of 218 trains is needed for the sign-up, which translates to 838 train assignments when the data recovery rate is accounted for.

The determination of the required sample size on an annual basis is a straightforward extension from the sign-up level example presented above, with the key parameters in the sample size equation being drawn from annual statistics.

Finally, it should be noted that because of the influence of the clustering effect on the required sample size, economic evaluation of APC performance in relation to manual data recovery should not be based on straightforward comparisons of costs-per-observation. The APC approach requires more observations to achieve the same level of precision as the manual approach, and this difference should be taken into account in assessing its relative merits. For example, under the assumption of simple random sampling we determined that the minimum sample size for the September-November 1988 sign-up would be 456 bus trips. The "design effect" [4, p. 103] on the sample size resulting from recovering data with APC's rather than manually is 4.30. In other words, an APC sample would need to be more than four times larger than a simple random sample to achieve the same level of precision.

Sample Inferences

Considering both the low data recovery rate experienced by Tri-Met with its APC's and the results of the statistical analysis of the determinants of successful data recovery, the threat of sampling bias should be a concern for transit operators who are using this technology. In Tri-Met's experience, the threats to randomness in sampling have been multi-faceted and have been associated with both technical and procedural factors. In regard to procedural aspects of sampling, successful APC implementation mainly requires effective coordination among "schedulers," bus dispatchers and drivers. Hardware malfunctions involving APC's, attributable to the APC equipment itself or traceable to the buses, pose additional complications not found with a reliance on manual data collection. Accounting for these factors in the sampling methodology would hardly be worthwhile, considering their complexity and the likelihood that their effects are not constant over time. This suggests an alternative involving post-stratification of the sample data as insurance against generating biased estimates of system performance.

The choice of stratification factors represents the primary issue in reconciling APC data subject to sampling bias. This choice is essentially dictated by two considerations. First, we should account for over and under-representation in the recovered sample with respect to various basic operating characteristics. Second, among those operating factors identified as being over and under-represented, the subset of factors exhibiting significant differences in ridership and representing non-trivial shares of the underlying population should be retained as stratification factors.

From the regression results reported earlier we can identify several candidates to serve as post-stratification factors. They include the AM and PM peak variables (or, more generally, time-of-service stratification), which were associated with higher data recovery rates, and the bus type variables, which showed higher data recovery rates for two bus models. By stratifying these variables a correction of the system ridership estimate, accounting for sampling bias, is obtained as follows:

$$R' = \sum_i t_i \cdot M_i \text{, where}$$

R' = the corrected total ridership estimate;

t_i = the total number of scheduled bus trips associated with the stratification category i;

M_i = the mean ridership value in stratification category i calculated from the sample observations.

The correction presented above pertains to an individual stratification factor. An extension to the joint application of two factors would be obtained as follows:

$$R' = \sum_i \sum_j t_{ij} \cdot M_{ij} .$$

Post-stratification corrections involving time-of-day and bus type factors were applied to the sample data from the September-November 1988 sign-up, with the outcome presented in Table 5. A benchmark value of 159,937 average weekday boarding rides was obtained by multiplying the overall sample mean by the total number of scheduled trips. The benchmark total represents the estimate that would be obtained using the procedure recommended in the UMTA guidelines, which assumes that the underlying sample of bus trips is random. In contrast with this value, post stratification by bus type resulted in an estimate of 158,999 boarding riders per weekday (.6 per cent lower), and post-stratification by time-of-day produced an estimate of 157,864 (1.3 per cent lower). Thus stratification by bus type had virtually no effect on the ridership estimate, while the effect of the time-of-day correction produced a marginally greater change. We see from the table that the bus types which were over-sampled in the sign-up are little different from the overall sample in terms of the average boarding rides per trip. Had the articulated buses been over or under-sampled, the difference in estimated ridership would have been more noticeable. With the AM and PM peak corrections we see that because of their relatively higher ridership, the benchmark ridership estimate was overstated due to the over-representation of these trips. The magnitude of the overestimate was muted, however, by the small ridership differential between peak and off-peak periods.

Table 5

Post-Stratification Estimates of Average Weekday Boarding Riders: September-November 1988 Sign-up

Stratified by Bus Type

Bus Type	Average "ons"/trip	Scheduled Trips	Estimated Boardings
B100/1000	28	1,883	52,724
B300	19	1,802	34,238
ARTIC	40	608	24,320
ADB	24	1,083	25,992
B500	27	775	20,925
			158,199

Stratified by Time-of-Day

AM Peak[*]	27	911	24,597
Midday	28	3,146	88,088
PM Peak[**]	22	1,967	43,274
Other	15	127	1,905
			157,864

---

[*] The AM Peak period includes all trips initiated between 6:00 and 8:00 AM.

[**] The PM Peak period includes all trips initiated between 4:00 and 6:00 PM.

The application of post-stratification corrections to the example above did not yield remarkable differences in estimated ridership. Given that we had previously established that the underlying data represented an equal probability sample, these results should not be surprising. Rather, the corrections provide an illustration of a means for insuring that estimates of ridership are unbiased in instances where the underlying sample data are not representative.

The relatively low data recovery rate for APC's, among other threats to randomness, indicates that a post-stratification procedure ought to be included in the system software package and applied to inferencing as a matter of course. The specifics involving stratification factors will be essentially determined by the experience of transit operators in implementing APC sampling plans, recognizing that variations in APC hardware and software, fleet mix and type, general ridership and scheduling characteristics, and coordination among personnel preclude the development of standardized correction procedures. For those operators who have already implemented APC systems, an analysis of previously recovered sample data along the lines pursued in this report can serve to identify the types of operating characteristics associated with differential data recovery rates.

Conclusions

Tri-Met's reliance on APC's to provide transit operating data has introduced both procedural complexities and a certain rigidity not found with manual data collection. Among the concerns that arose as a result of their experience with APC's were the underlying precision, accuracy and representativeness of the sample data. In light of those concerns, we have developed methodologies covering the areas of sampling and inference that provide a determination of the sample size required to meet a given precision standard, as well a means of reconciling unrepresentative sample data. We have also verified the accuracy of APC's with respect to passenger counts.

Another area of concern regards the low data recovery rate. Apart from representing a potential source of sampling bias, the low recovery rate results in the need for considerably more train assignments to achieve the necessary sample size. Over 45 per cent of the assigned trains returned with no data, indicating a need for further evaluation of the APC's in regard to their design, installation and maintenance. The prospects for improvements in the recovery rate as related to the remaining sources of data failure, which collectively affect 28 per cent of all train assignments, are probably not as good as they are for improvements in the basic operation of the APC units. Thus Tri-Met's attention in the area of data recovery has been directed toward the latter objective.

We have not evaluated whether the costs and various complications associated with APC's are outweighed by the estimated benefits of the technology. We have also not extended the evaluation to the route level, where APC's provide the only practical means of comprehensive data recovery and thus offer substantial potential benefits. Clearly, the scope of the evaluation would have to be extended to include these elements, along with data management issues, to achieve a comprehensive assessment of the relative merits of APC's.

APC's have been found to be cost effective in comparison with manual data recovery [5], although it should be stressed that such analysis should account for differentials in sample sizes required to meet a given level of precision. The benefits of more rapid data turnaround with APC's would be difficult to quantify, but based on Tri-Met's experience the gains have not been substantial. This is due to their use of APC's primarily for UMTA Section 15 reporting, for which rapid data turnaround is not necessary.

Tri-Met also uses the data recovered by APC's to construct route performance reports for each of the five sign-up periods comprising annual service, but questions about the underlying precision of ridership estimates at the route level have precluded a more prominent contribution of APC data to route analysis and scheduling. For example, based on a 20 per cent sample of routes, we found the average estimated route level precision to be +/- 58 per cent at the 95 per cent level of confidence. This range is too broad to provide an acceptable basis for transit planning, and to achieve route level precision comparable to what is required by UMTA at the system level would entail more than a forty-fold increase in sample size. While samples of this size can conceivably be recovered with APC's (which can be regarded as one of their potential benefits), one can also expect that problems associated with coordination in executing the associated sampling plan would be considerable.

Assuming that difficulties associated with sampling and data recovery at the route level can be overcome, a more refined set of validation standards - targeted to the stop or route segment rather than the trip level - would be needed. This would require the development of detailed base level information on times and distances for the route network, which presently does not exist, against which the APC data could be validated. One would also expect that the data recovery rate would decline with more strictly defined validation standards applied to the present data recovery process. As a result, Tri-Met has considered acquiring an automatic vehicle locating system to supplement the APC's. The accuracy of the recorded APC data on times and distance would also need to be verified, in a manner consistent with the approach used to test the validity of passenger counts.

Implementation of a comprehensive route level data recovery program thus appears to face a number of challenges. As an alternative to comprehensive data recovery, Tri-Met has been considering targeted applications of APC's. For example, one possible targeting strategy would be to reserve those APC buses not assigned to recover Section 15 data for intensive data recovery from routes where service changes are being considered. Another would be to select one of the five annual sign-ups for comprehensive sampling (i.e., combining Section 15 sampling efforts with route level sampling), and converting the sample data to an annualized estimate of ridership. Regarding this alternative, it was thought that fewer problems would be encountered if large scale sampling were undertaken in a single sign-up as opposed to an ongoing basis.

After nearly seven years of operating experience, Tri-Met has yet to fully capitalize on the reported merits of the APC technology. Application has instead been essentially limited to data collection for Section 15 reporting. While the APC's may still be cost effective for this purpose, their conceivable potential is much greater. At this point, however, it is not clear whether the various impediments to full application discussed above will be effectively overcome.

References

1. Balaram, A, G. Gruver and H. Thomas, Evaluation of Passenger Counter System for an AVM Experiment, Volume I: Technical Report, Gould Information Identification, Inc., Fort Worth, Texas, February 1979.
2. Deibel, Lawrence E. and Barbara Zumwalt, Modular Approach to On-Board Automatic Data Collection Systems, Report 9, National Cooperative Transit Research & Development Program, Transportation Research Board, December 1984.
3. Hodges, Carol C., Automatic Passenger Counter Systems: The State of the Practice, Final Report for the Urban Mass Transportation Administration, No. DOT-I-87-36, June 1985.
4. Moser, G.A. and G. Kalton, Survey Methods in Social Investigation, 2nd Edition, New York, Basic Books, Inc., 1972.
5. Multisystems, Inc., An Assessment of Automatic Passenger Counters, Interim Report to the Urban Mass Transportation Administration, No. DOT-I-82-43, 1982.
6. Multisystems, Inc., Transit Data Collection Design Manual, Final Report for the Urban Mass Transportation Administration, No. DOT-I-85-38, June 1985.
7. Urban Mass Transportation Administration, U.S. Department of Transportation, "Sampling Procedures for Obtaining Fixed Route Bus Operating Data Required Under the Section 15 Reporting System," Circular No. 2710.1, February 1978.
8. Urban Transportation Associates, Inc., Washington Metropolitan Area Transit Authority 1986 Automatic Passenger Counter Project Accuracy Evaluation, Report to the Urban Mass Transportation Administration, No. UMTA-URT-20-86-1, 1986.