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

JAMES G. STRATHMAN AND JANET R. HOPPER

Whereas automatic passenger counters (APCs) offer the potential for cost-effective data recovery, they introduce new complications in the data recovery process. Three issues associated with the use of APCs are addressed on the basis of the experience of the Tri-County Metropolitan Transportation District of Oregon. The first issue is validation, which concerns both recovery and accuracy of APC passenger data. The second concerns the design of a sampling methodology for APCs compatible with UMTA's Section 15 reporting requirements. Third is inferring system-level ridership from sample data in the presence of selective APC failures. APCs provided systematically accurate passenger counts. Given that APCs recover operating data for all bus trips making up a vehicle schedule, a cluster sampling method was developed. Selective data recovery failures can bias estimates of system-level ridership. When data recovery rates vary by bus type, route type, or time of day, inferences may over- or underrepresent total system ridership. In these circumstances, post hoc stratification of the sample is recommended. Several alternative corrections based on a priori knowledge of the mix of bus types and schedule characteristics in the system are presented.

Automatic passenger counters (APCs) offer potential benefits to transit operators in data acquisition, management, and utilization. Compared with manual collection, APCs are cost-effective for larger transit systems, and they provide better data turnaround and improved accuracy (*1*). They can also recover the 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 selected buses in the fleet—usually about 10 percent (*2*)—are equipped with APCs, 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, whereas APCs generally return more accurate data than manual counters, many of the data are screened out because of 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 nonresponse type of bias might undermine the sample ridership statistics and, consequently, inferences of systemwide operating performance. Third, with manual data collection, surveyors are typically assigned to randomly selected bus trips. With APCs the unit of observation is the "train" or "block," which consists of all the scheduled service performed by a bus during an operating day. The bus trips of a train are not independent, and thus the sampling framework recommended by UMTA (*3*) cannot be used. As a result, an alternative methodology must be designed consistent with the APCs' operating features.

These issues are addressed in the coming sections. On the basis of information drawn from the recent performance of APCs used by the Tri-County Metropolitan Transportation District of Oregon (Tri-Met), data recovery is considered by analyzing the accuracy of the data generated by APCs and the sources of data recovery failures. Whether the set of trains from which data have been successfully recovered represents an equal probability sample is then determined. A sampling methodology is developed ensuring that the selection of bus trips (through the selection of trains) is both random and of sufficient size to comply with UMTA's Section 15 reporting requirements. Finally, a remedy for correcting sample statistics subject to bias from nonrandom data recovery failures is suggested.

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 the day the recovered data are 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 APCs 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.

J. G. Strathman, Center for Urban Studies, School of Urban and Public Affairs, Portland State University, P.O. Box 751, Portland, Oreg. 97207-0751. J. R. Hopper, Tri-County Metropolitan Transportation District, 4012 SE 17th Avenue, Portland, Oreg. 97202.

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 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 percent, (*b*) time between pullout and pull in differs by more than 30 min from the service schedule, or (*c*) total boardings and alightings differ by more than 10 percent.

Validation standards covering distances and pullout and pull-in times at the bus trip level are also applied. If several 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 for an entire train.

The sampling plan used by Tri-Met is organized around the five sign-up periods making up annual scheduled service. The objective of the plan is to sample the scheduled trips in each sign-up uniformly. Execution of the sampling plan requires the involvement of several divisions. The scheduling division, using a selection program that assigns higher selection priorities to trains that have been undersampled previously, draws a sample of trains daily. 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 that 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 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 percent) were assigned APC buses. Another 325 trains that were not requested were assigned APC buses. Valid data were recovered from 286 of the trains that had been requested and from 82 unrequested trains. Thus data were recovered from 26 percent of all assignments.

Losses of data resulted from various causes, including exceeding time tolerances (7 percent of the total failures), distance tolerances (5 percent), discrepancies between boardings and alightings (7 percent), incorrect or missing assignment information in the train records (11 percent), recovered data that were unusable (8 percent), and failures due to bus or equipment malfunction (62 percent). The last category represents cases for which no data were returned by the APCs. Failures in this category include instances in which the APC unit accidentally reset, buses did not pull close enough to the transmitter to allow transfer of the data, the microprocessor's memory was filled and could not record more data, and data were not recorded because of equipment breakdown.

Of the 1,414 train assignments, 368, or 26 percent, returned valid data. This rate is considerably lower than what has been reported in other studies of APC performance (*1,4*). Generally, about 80 percent of all train assignments have been reported to return valid data. The reasons for this difference cannot be further explored because of 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 screening tolerances used in validating the data, differences due to the mix of APC-equipped bus types in Tri-Met's fleet, and differences in APC technology. Given both the relatively small data recovery rate and the inclusion of unrequested trains, the question of nonresponse or sampling bias, or both, arises. It is therefore necessary to determine if the data losses were random or were systematically related to train-specific characteristics.

TABLE 1 BREAKDOWN OF APC DATA RECOVERY, APRIL 1989 SIGN-UP

	No.	%[1]
1. Trains Requested	1,589	100
2. Trains Assigned		
a. As requested	1,089	77
b. Unrequested	325	23
c. Total assignments	1,414	100
3. Data Recovered		
a. From requested trains	286	78
b. From unrequested trains	82	22
c. From all assigned trains	368	100
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	100

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

Determinants of Successful Data Recovery

The September–November 1988 sign-up was selected for a regression analysis of factors related to successful data recovery. Tri-Met staff considered this sign-up typical in regard to APC performance and other operating and ridership characteristics. During the sign-up 588 trains provided daily weekday service. Valid data were recovered from 1,552 assignments (about 3.4 percent of total scheduled weekday service). The following model was specified to examine the effects of train-specific characteristics on successful data recovery:

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

where

SAMP = the number of assignments in each train that recovered valid data;
APC = the number of available APC buses of the requested type at the garage from which each train assignment was made;
REQ = the number of times each train was requested;
ASG = the number of times each train was assigned;
AM = 1 if the train provided only a.m. peak service, 0 otherwise;
PM = 1 if the train provided only p.m. peak service, 0 otherwise;
G_1 = 1 if the train was dispatched from Garage 1, 0 otherwise;
G_2 = 1 if the train was dispatched from Garage 2, 0 otherwise;
ARTIC = 1 if the train was an articulated bus model (Crown-Ikarus), 0 otherwise;
ADB = 1 if the train was an ADB bus model (40-ft GMC RTS-II), 0 otherwise;
B500 = 1 if the train was a B500 bus model (40-ft Flexible "Metro"), 0 otherwise; and
B300 = 1 if the train was a B300 bus model (35-ft Flexible "New Look"), 0 otherwise.

The APC variable was included in the specification to account for differences in the number of APC buses of each relevant type at each garage. Data recovery is expected to improve when more buses are available for assignment. The number of requests was included to control for trains that were not successfully assigned because of operational or mechanical problems. Tri-Met's sampling software places a higher subsequent selection priority on trains that are requested but not assigned. A greater frequency of requests would thus be associated with trains that are not successfully recovering data. The number of assignments controls for variations in data recovery attributable to the relative frequency of train assignments; in other words, some trains may recover valid data more frequently because they are assigned more frequently. AM and PM were included because these trains are in service for a shorter time and should be more reliable in 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 overrepresented. The garage variables were included to check for differences in data recovery attributable 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 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 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 returned 0.3 more observations per train than day trains, whereas the net increase for PM peak trains was about 0.8. Both are statistically significant and represent increases of approximately 10 and 30 percent over the data recovery rate for day trains.

TABLE 2 REGRESSION ESTIMATES OF THE DETERMINANTS OF TRAIN LEVEL DATA RECOVERY, SEPTEMBER 1988 SIGN-UP

Variable	Mean	St. Dev.	Coefficient	t-ratio
Constant	n.a.	n.a.	.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.

Among the various bus types, the ADB and B300 models recovered 2.1 and 1.3 more sample observations per train than the "reference" bus type (B100/1000, which includes 40-ft AMGeneral and 40-ft Flexible "New Look" models). Garage 1 produced 0.67 fewer and Garage 2 produced 0.86 more observations per train in relation to the central garage. These differences are most likely due to breakdowns of the fixed-station transmitters at the garages, because assignments are proportionately distributed among the three garages. The transmitter at Garage 2, by implication, experienced fewer problems than the transmitters at the other two garages. Alternatively, some routes may be more likely to return valid data than others; a variation in the composition of route types by garage could affect relative data recovery rates.

Besides isolating various determinants of successful data recovery, the regression results point to possible sources of over- and underrepresentation of trains in the effective sampling scheme. Of particular concern are the AM and PM peak trains and two of the bus types. Significant differences in ridership characteristics among the trains in question 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 are successfully recovered by the APCs, another concern is the accuracy of the passenger counts. Automatic counters have been described as more accurate than manual data recovery, particularly for high-volume routes and routes with peak-period standing loads (*1*). The errors that have been observed with APCs indicate a tendency to undercount rather than overcount passenger activity, whereas boardings tend to be counted more accurately than alightings.

In a demonstration study of APCs equipped with infrared beams, the Washington Metropolitan Area Transit Authority conducted an accuracy test on a sample of more than 400 bus trips involving about 18,000 boardings and alightings (*5*). Total boardings recorded by the APCs equaled 99.7 percent of the manual counts, and recorded alightings equaled 98.4 percent 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 using APCs (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 (*1*).

Previous research has thus consistently demonstrated that APC and manual passenger counts tend to correspond. The APC systems evaluated were relatively new, however. Tri-Met's APCs have been in service for nearly 7 years, and their low data recovery rate indicates that they have not been performing at the levels observed elsewhere. As a result, a statistical comparison of APC and manual passenger counts for Tri-Met's system was undertaken.

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 totaled 3,768 across all observations. A test of the mean difference between APC and manual counts of boardings and alightings per stop was conducted for each bus as well as for the overall sample. Table 3 gives the findings for the overall analysis and for those buses having significant differences between APC and manual counts. Across all buses and all stops, the average boardings per stop counted by the APC were 0.01 passenger higher than the manual count, and the number of alightings counted by the APCs averaged 0.01 passenger lower. Neither difference was statistically significant at the .05 level. Of the six instances in which the APC and manual boarding counts differed significantly, three involved overcounting and three involved undercounting. Of the five instances in which the APC and manual alighting counts differed, two involved overcounting by the APC. Three specific buses were associated with significant differences in both boardings and alightings.

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

Because significant differences between APC and manual counts were found in only a few cases, and because there was no pattern of divergence, the APCs appear to provide systematically accurate counts. With 92 applications of the hypothesis test at the 95 percent confidence level, about five rejections of the null hypothesis due to Type I error (i.e., rejecting the null hypothesis of no difference when it should have been accepted) are expected. Moreover, an underlying assumption is that the manual counts are free of error, and this is likely to be violated in some cases. Finally, the data recovered by the APCs were not subjected to the normal screening process, which would have purged substantial portions of the data recovered from several buses (i.e., Buses 347 and 731).

SAMPLING WITH APCs

Two issues concerning sampling with APCs must be addressed. The first concerns the low data recovery rate when APCs are used and the fact that observations on some bus trips were assigned but not requested in the sampling methodology. This raises questions about the representativeness of the sample, which could fail from assignment or response bias.

The second issue concerns the sampling methodology itself. The sampling procedure recommended by UMTA (*3*) was essentially designed with manual data collection in mind, because it provides solely for independent random selection of bus trips. With APCs, bus trips are necessarily selected in blocks composing trains. Whereas trains can be selected in an independent and random fashion, the individual bus trips cannot. As a result, a specific methodology for APCs must be developed that ensures satisfaction of the UMTA precision standards and minimizes the 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, 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 by evaluating 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, attributing the cause to problems associated with the request or the assignment process would be difficult.

A chi-square test was used to determine whether 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 compared with total assignments, which correspondingly reduces the comparative intertrain 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).

An APC Sampling Methodology

The objective in designing a sampling methodology for APCs 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 percent at the 95 percent level of confidence. The methodology must account for correlation among bus trips within trains, and it should set the sample size large enough to reflect the anticipated data recovery rate.

The special features associated with the APC data recovery process are compatible with a multistage cluster sampling method (*6*). The first stage in this methodology would consist of a random selection of trains, and the second stage would then be defined by the 100 percent "clusters" of bus trips composing the selected trains. Variations in cluster sizes would also be accommodated, because the number of bus trips can vary by train. The methodology would be designed for implementation at the train level, consistent with data recovery using APCs, yet ensuring that the sample statistics satisfy trip level precision requirements.

Cluster sampling has also been proposed for data collection by ride checkers (*7*). For many transit systems, run pieces (usually about 4 hr of service) represent a more convenient sampling unit than bus trips. Thus, whereas cluster sampling may be a necessity for data collection with APCs, it may also be a more cost-effective approach for other modes of data collection.

The determination of the required sample size for cluster sampling 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 on the basis of the observed average number of bus trips per train. In the presentation below, the sample size is determined for estimating passenger miles, because the relative variance of passenger miles tends to be larger than that of other operating data. The minimum number of bus trips to be sampled, in conformance with the UMTA Section 15 standards, is

$$n_c = [(1.96S_c)/(0.1M)]^2 \quad (1)$$

where

n_c = the number of bus trips required in a multistage cluster sample,

S_c = the standard deviation of passenger miles per bus trip for a multistage cluster sample,

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

1.96 = the critical z value at the .025 level, and
M = the mean passenger miles per bus trip.

Equation 1 is equivalent to the arrangement used to determine the required number of observations for a simple random sample, except for the cluster sample standard deviation term, which accounts for the interdependence of bus trips within trains and the variation in the number of bus trips per train. The standard deviation for a simple random sample need not be elaborated, but its counterpart for a multistage cluster sample warrants presentation. This standard deviation is defined as follows:

$$S_c = [1/(n - 1) \cdot \sum_i n_i \cdot (M_i - M)^2]^{0.5} \tag{2}$$

where

n_i = the number of bus trips in Train i,
M_i = the mean passenger miles per bus trip for Train i, and
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 Tri-Met's September–November 1988 sign-up as an example, the overall mean passenger miles per bus trip is 8,481 and the multistage cluster sample standard deviation is 19,159. The minimum required sample size for the sign-up in the example is thus

$$n_c = [(1.96 \cdot 19{,}159)/(0.1 \cdot 8{,}481)]^2 \tag{3}$$

or 1,961 bus trips.

The sample size derived above represents 14 percent of the 13,955 trip observations actually recovered during the September–November 1988 sign-up. By using the cluster-sampling framework, it was found that the sample produced precision of ±3.7 percent at the 95 percent level of confidence.

To achieve the required sample size, the data recovery rate should also be taken into account. Table 1 indicates that 26 percent of all assignments return usable data. This suggests that to achieve the necessary number of valid observations, 7,542 trip assignments (2.3 percent of all scheduled trips) would have to be made. This number of assignments is probably excessive, because an improved data recovery rate from smaller-sized samples (as indicated by the APC coefficient in the regression model) is expected.

Because trains are the unit of assignment with APCs, it is necessary to translate sample size requirements from bus trips to this unit. From the sign-up in the example, an average of 8.98 bus trips per train is found. 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 of the sign-up-level example presented above, with the key parameters in the sample size equation drawn from annual statistics.

Finally, 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 should be taken into account in assessing its relative merits. For example, under the assumption of simple random sampling, the minimum sample size for the September–November 1988 sign-up was determined to be 456 bus trips. The "design effect" (6,p.103) on the sample size resulting from recovering data with APCs rather than manually is 4.30. In other words, an APC sample would need to be more than four times as large as a simple random sample to achieve the same level of precision.

SAMPLE INFERENCES

The low data recovery rate experienced by Tri-Met with its APCs and the results of the statistical analysis of the determinants of successful data recovery indicate that the threat of sampling bias should be a concern for transit operators who use this technology. In Tri-Met's experience, the threats to randomness in sampling have been multifaceted and 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 APCs, attributable to the APC equipment itself or traceable to the buses, pose additional complications not found in manual data collection. Accounting for these factors in the sampling methodology would hardly be worthwhile because of their complexity and the likelihood that their effects are not constant over time. This suggests an alternative involving poststratification of the sample data as insurance against generating biased estimates of system performance.

The choice of stratification factors is the primary issue in reconciling APC data subject to sampling bias. The choice is essentially dictated by two considerations. First, over- and underrepresentation of various basic operating characteristics in the recovered sample should be accounted for. Second, among those operating factors identified as being over- or underrepresented, the subset exhibiting significant differences in ridership and representing nontrivial shares of the underlying population should be retained as stratification factors.

Several candidates for poststratification factors can be identified from the regression results reported earlier. 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:

$$R' = \sum_i t_i \cdot M_i \tag{4}$$

where

R' = the corrected total ridership estimate,
t_i = the total number of scheduled bus trips associated with Stratification Category i, and
M_i = the mean ridership value in Stratification Category i calculated from the sample observations.

Equation 4 pertains to an individual stratification factor. An extension to the joint application of two factors would be obtained as follows:

$$R' = \sum_{i} \sum_{j} t_{ij} \cdot M_{ij} \qquad (5)$$

Poststratification corrections involving time-of-day and bus type factors were applied to the sample data from the September–November 1988 sign-up (see 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 is 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, poststratification by bus type resulted in an estimate of 158,199 boarding riders per weekday (1.1 percent lower), and poststratification by time of day produced an estimate of 157,864 (1.3 percent lower). Thus stratification by bus type and time of day had virtually no effect on the ridership estimate. Table 5 indicates that the bus types that were oversampled 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 undersampled, the difference in estimated ridership would have been more noticeable. With the AM and PM peak corrections it is seen that because of their relatively higher ridership, the benchmark ridership estimate was overstated owing to the overrepresentation of these trips. The magnitude of the overestimate was muted, however, by the small ridership differential between peak and off-peak periods.

The application of poststratification corrections to the example above did not yield remarkable differences in estimated ridership. Because it had been previously established that the underlying data represented an equal probability sample, these results should not be surprising. Rather, the corrections offer a way to ensure that estimates of ridership are unbiased when the underlying sample data are not representative.

The relatively low data recovery rate for APCs, among other threats to randomness, indicates that a poststratification procedure ought to be included in the system software package and applied to inferencing as a matter of course. The specifics of stratification factors will be determined by the experience of transit operators in implementing APC sampling plans. 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

TABLE 5 POSTSTRATIFICATION 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.

previously recovered sample data along the lines reported can identify the types of operating characteristics associated with differential data recovery rates.

CONCLUSIONS

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

Another area of concern is the low data recovery rate. Besides being a potential source of sampling bias, the low recovery rate necessitates more train assignments to achieve the required sample size. More than 45 percent of the assigned trains returned with no data, indicating a need for further evaluating the design, installation, and maintenance of the APCs. Contributions toward improvement in the recovery rate from the remaining sources of data failure, which collectively affect 28 percent of all train assignments, are probably not as likely as are improvements in the basic operation of the APC units. Thus Tri-Met's attention has been directed toward the latter objective.

Whether the costs and complications associated with APCs are outweighed by the estimated benefits of the technology has not been considered. The analysis has not been extended to the route level, where APCs provide the only practical means of comprehensive data recovery and thus offer substantial potential benefits. The scope of the evaluation would have to be extended to these elements, along with data management issues, to achieve a comprehensive assessment of the relative merits of APCs.

APCs have been found to be cost-effective compared with manual data recovery (*1*), although 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 APCs are difficult to quantify, but on the basis of Tri-Met's experience the gains have not been substantial. This is due to Tri-Met's use of APCs primarily for UMTA Section 15 reporting, for which rapid data turnaround is not necessary.

Tri-Met also uses the data recovered by APCs to construct route performance reports for each of the five sign-up periods making up 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. In an analysis of 32 routes (representing about 20 percent of Tri-Met's system), an average route level precision for mean boarding riders of ±58 percent at the 95 percent level of confidence was found (*8*). This range is clearly too wide for route planning. To achieve route level precision comparable with what is required by UMTA at the system level would entail more than a 40-fold increase in sample size. Samples of this size can conceivably be recovered with APCs (which can be regarded as one of their potential benefits), but 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 at 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. The data recovery rate would be expected to 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 APCs. 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 faces a number of challenges. As an alternative to comprehensive data recovery, Tri-Met has been considering targeted applications of APCs. 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 to convert the sample data to an annualized estimate of ridership. It was thought that fewer problems would be encountered for this alternative if large-scale sampling were undertaken in a single sign-up as opposed to an ongoing basis.

After nearly 7 years of operating experience, Tri-Met has yet to fully capitalize on the reported merits of APC technology. Application has been essentially limited to data collection for Section 15 reporting. Whereas the APCs may still be cost-effective for this purpose, their potential is greater.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the efforts of Jon Lutterman, who provided insights on the sampling methodology presently used by Tri-Met as well as statistics on APC data recovery. Ken Dueker and Rick Gerhart provided helpful comments on an earlier version of this paper. Research assistance was also provided by Chall Kyoung Sung.

Support for this project was provided by the U.S. Department of Transportation and Tri-Met.

REFERENCES

1. *An Assessment of Automatic Passenger Counters.* DOT-I-82-43. Multisystems, Inc., 1982.
2. L. E. Deibel and B. Zumwalt. *NCTRP Report 9: Modular Approach to On-Board Automatic Data Collection Systems.* TRB, National Research Council, Washington, D.C., 1984.
3. *Sampling Procedures for Obtaining Fixed Route Bus Operating Data Required Under the Section 15 Reporting System.* Circular 2710.1. Urban Mass Transportation Administration, U.S. Department of Transportation, 1978.
4. C. C. Hodges. *Automatic Passenger Counter Systems: The State*

of the Practice. DOT-I-87-36. Urban Mass Transportation Administration, U.S. Department of Transportation, 1985.
5. *Washington Metropolitan Area Transit Authority 1986 Automatic Passenger Counter Project Accuracy Evaluation*. UMTA-URT-20-86-1. Urban Transportation Associates, Inc., 1986.
6. G. A. Moser and G. Kalton. *Survey Methods in Social Investigation* (2nd ed). Basic Books, New York, 1972.
7. P. G. Furth, K. L. Killough, and G. F. Ruprecht. Cluster Sampling Techniques for Estimating Transit Patronage. In *Transportation Research Record 1165*, TRB, National Research Council, Washington, D.C., 1988, pp. 105-114.
8. J. G. Strathman. *Sampling Bus Ridership at the Route Level: Initial Results from Efforts To Improve the Precision of Sample Estimates*. Final report. Center for Urban Studies, Portland State University, Portland, Oreg., 1989.

Publication of this paper sponsored by Committee on Bus Transit Systems.