Ridership Sampling for Barrier-Free Light Rail

PETER G. FURTH AND ASHOK KUMAR

The challenges and current practice in ridership estimation on light rail lines, particularly barrier-free lines, are reviewed. Two-stage sampling is an efficient plan because of the high level of accuracy demanded and the small number of scheduled trips. The theory of two-stage sampling is reviewed, and modifications are derived for times when the second-stage sample size varies between primary units. Sampling plans for light rail lines in greater Los Angeles and San Jose are offered as examples. Necessary samples sizes are as low as 25 round trips for 10 percent annual precision at the 95 percent confidence level, 80 round trips for 5 percent precision, and 400 round trips for 2 percent precision.

Responsible management demands reliable ridership estimates in order to monitor system performance, to track and forecast ridership and revenue trends, and to fulfill FTA Section 15 reporting requirements. However, estimating transit ridership poses a challenge for nearly every North American system because in most systems all passengers are not routinely counted. Ridership estimation is done by sampling, for which many techniques have been advanced. This paper describes a statistically based sampling technique that is appropriate to barrier-free light rail lines (although it certainly can be used in other situations as well) and its application to two new West Coast systems.

Statistical sampling is an established discipline, covered, for example, in the work by Cochran (*1*). Stopher and Meyburg's text reviews sampling in the context of transportation planning (*2*). Application of statistical sampling techniques to transit began with the *Bus Transit Monitoring Manual* (BTMM) (*3*) and its update, the *Transit Data Collection Design Manual* (TDCDM) (*4*), both of which were written for bus systems. In part, these manuals were driven by UMTA Section 15 requirements that systems receiving operating assistance report, among other things, systemwide annual estimates by mode of boardings and passenger miles that are based on a statistically sound methodology and satisfy specific accuracy requirements (10 percent precision at the 95 percent confidence level). UMTA published circulars describing two approved methodologies for bus systems (*5,6*) and one for demand-responsive systems (*7*), but no standardized method for light rail or other modes has been established. Applications of statistical methodologies seeking greater efficiency than the standard techniques have been published, but again they are designed for bus systems. These applications include cluster sampling (*8*), ratio-based estimation (*9*), and regression-based estimates (*10*).

Ridership sampling for barrier-free light rail systems demands special attention for several reasons. First, accurate ridership estimates are sought for a single line or a very small number of lines; for bus systems, however, total ridership must be estimated over a large number of lines. Light rail systems can therefore be expected to exhibit less variability than bus systems, and techniques that reduce between-line variability such as stratifying by line type or line length become moot. Second, greater accuracy is often sought for light rail lines, particularly new lines whose ridership levels are being watched carefully by public officials and the press. The intensive sampling needed to achieve high levels of accuracy, when applied to a single light rail line, means that there is a high likelihood that a particular scheduled trip will be sampled more than once, and it is even possible that every scheduled trip will be sampled once or more. Statistical techniques that assume an infinite population of scheduled trips can safely be used for a bus system in which a few trips are sampled each week from a sampling frame of thousands of daily trips, but not for a light rail system in which a large fraction of the scheduled trips is sampled each month.

Third, barrier-free systems with off-vehicle revenue collection do not easily lend themselves to revenue-based sampling, a very efficient technique in many bus systems. Revenue-based estimation typically requires jointly measuring ridership and revenue for a sample of trips, and when there is no on-board revenue collection it is virtually impossible to assign revenue to a specific trip. Revenue-based estimation can be done on a basis other than vehicle trips, for example, by correlating boardings during a sample of time intervals at a sample of stations to revenue collected at ticket-vending machines at those stations during those intervals; but this approach raises a host of issues of its own. For example, it provides no measurement of passenger miles; in addition, there can be a huge variance in revenue per passenger at different stations and different times of day due to transfer patterns, availability of round-trip tickets, and pass use patterns.

Finally, the greater accuracy demands of new light rail systems and their larger vehicles make load-based estimation, a technique used in several bus systems, impractical. One unpublished study of the accuracy of wayside bus load measurements suggests a mean absolute error of 13 percent and a bias of 3 to 6 percent. A load-based estimate has three sources of error. The first two are sampling errors, first in the load data and second in ride check data (load, boardings, and passenger miles data used to derive the ratio-to-load fac-

P. G. Furth, Department of Civil Engineering, Northeastern University, 360 Huntington Avenue, Room 420 SN, Boston, Mass. 02115. A. Kumar, Scheduling and Operations Planning Department, Los Angeles County Metropolitan Transportation Authority, 425 South Main Street, Los Angeles, Calif. 90013.

tors). These errors can be reduced by making more point checks and more ride checks. However, the third source of error, bias in the load measurements, cannot be diminished by increasing sample size, making high accuracy unattainable using wayside counts. Even if measurement error were eliminated (e.g., by having checkers count on board the vehicle, introducing a new set of problems), the compounded errors from the point check sample and the ride check sample limit the value of the technique to desired precision levels of 10 percent or greater.

ESTIMATION METHODS IN USE

An informal survey of light rail operators was conducted, supplementing an earlier survey performed by Kumar and Parry (*11*), to see what techniques are used to estimate ridership on light rail lines in the U.S. and Canada. Three relatively new barrier-free systems use a technique described later in this paper, a special case of two-stage sampling in which every scheduled trip is sampled several days a year. In another barrier-free system, ticketed boardings are counted directly from ticket-vending machines, and nonticketed (i.e., pass, transfer, and free) boardings are estimated from a random sample of trips expanded in proportion to ticketed boardings. A fifth barrier-free system expands a random sample of trip portions by service minute rather than simply by number of trips. Among other light rail systems, a variety of methods are used, including the methods of Circular 2710.1 (*5*) (direct expansion of a sample of about 550 trips), Circular 2710.4 (*6*) (revenue-based expansion of a sample of 208 trips), and others (*4*) (expansion of point load data based on ratio-to-load factors that are updated every few years). At least one system uses a sampling method for Section 15 reporting while using electronic farebox and turnstile counts without sampling for internal purposes. Another uses revenue-based expansion in which the ratio-to-revenue factors are updated every few years and there is no sampling during the intervening years. The Canadian systems, which are not subject to Section 15 requirements, do not sample for passenger miles and are less systematic in estimating boardings than the U.S. systems, preferring to concentrate on peak load.

TWO-STAGE SAMPLING METHODOLOGY DEVELOPED FOR SOUTHERN CALIFORNIA RAPID TRANSIT DISTRICT METRO BLUE LINE

The Metro Blue Line is a 23-mi-long light rail facility connecting downtown Los Angeles to downtown Long Beach. This 22-station line, operated by the Southern California Rapid Transit District (SCRTD), was opened to the public on July 14, 1990. Like recently completed light rail facilities in San Diego, Sacramento, San Jose, and Portland, the fare collection systems on the Metro Blue Line are barrier-free, meaning that patrons are required to neither go through any turnstiles or fare gates nor show their fare media to an on-board operator or conductor. Under a contract from SCRTD, fare payment is enforced by a team of roving county deputy sheriffs who, besides patrolling the line for security purposes, are authorized to randomly inspect passengers for valid fare media.

Being the first new rail project in the Los Angeles area in more than 30 years and the first segment of a 150-mi rail rapid transit system currently under development, the Metro Blue Line drew considerable attention from the local news media and elected officials at the time of its opening.

For about 2 months, ridership was tracked on a daily basis and shared with the news media. The passenger counting program is conducted by unionized schedule checkers employed by SCRTD. As operations stabilized, attention was focused on finding an efficient, statistically sound methodology for making reliable quarterly and annual patronage estimates for both internal management and external reporting purposes.

Several considerations dictated the statistical methodology chosen for estimating ridership. First, because there was no auxiliary variable such as revenue suitable for ratio estimation, boardings would be estimated and expanded directly. Second, both sampling efficiency and variance reduction suggest using the round trip as the sampling unit rather than the one-way vehicle trip. Third, variance reduction and economy suggest estimating weekday, Saturday, and Sunday ridership in separate strata. Finally, the population of transit trips has a natural two-dimensional structure—that is, the fundamental sampling element is trip (i,j), the ith scheduled trip on the jth day. Because there is one pattern of variation between scheduled trips and another pattern of variation between days, the appropriate technique is two-stage sampling.

Review of Two-Stage Sampling

Two-stage sampling means sampling a number of primary units, and then, for each selected primary unit, sampling a number of elements or subunits within that primary unit. An example of two-stage sampling is the methodology of UMTA Circular 2710.1 (*5*), in which a number of days (the primary unit) is selected, and for each selected day a number of trips is selected. Because for a given day type each trip is run every day, we have two-stage sampling with primary units of equal size. The discussion will assume simple random sampling at each stage unless noted otherwise. Notation and definitions, following Cochran's text (*1*), are as follows:

N = number of primary (Stage 1) units in population

M = number of subunits (Stage 2) in population of each primary unit

$\overline{Y}_i$ = population mean per subunit in primary unit i

$\overline{\overline{Y}} = $ population mean per subunit overall $= \frac{1}{N} \sum \overline{Y}_i$

n = number of primary units sampled

m = number of subunits sampled in each primary unit

y_{ij} = value for jth subunit in ith primary unit

$\overline{y}_i = \frac{1}{m} \sum_j y_{ij}$ = sample mean per subunit in ith primary unit

$\overline{\overline{y}} = \frac{1}{n} \sum \overline{y}_i$ = overall sample mean per subunit

$S_1^2 = \frac{1}{N-1} \sum (\overline{Y}_i - \overline{\overline{Y}})^2$ = variance among primary unit means

$S_2^2 = \frac{1}{N(M-1)} \sum_i \sum_j (y_{ij} - \overline{Y}_i)^2$ = variance among subunits within primary units

As Cochran demonstrates, $\overline{\overline{y}}$ is an unbiased estimate of $\overline{\overline{Y}}$, and its variance, accounting for the finite population correction, is

$$V(\overline{\overline{y}}) = \left(\frac{N-n}{N}\right)\frac{S_1^2}{n} + \left(\frac{M-m}{M}\right)\frac{S_2^2}{mn}$$
$$= \frac{1-f_1}{n}S_1^2 + \frac{1-f_2}{mn}S_2^2 \quad (1)$$

where $f_1 = n/N$ and $f_2 = m/M$ are the sampling fractions in the first and second stages, and $(1 - f_1)$ and $(1 - f_2)$ are the finite population corrections.

Of course, S_1^2 and S_2^2 are unknown and must be estimated from data. Define the sample Stage 1 variance and the sample Stage 2 variance as

$$s_1^2 = \frac{1}{n-1} \sum (\overline{y}_i - \overline{\overline{y}})^2 \quad (2)$$

$$s_2^2 = \frac{1}{n(m-1)} \sum_i \sum_j (y_{ij} - \overline{y}_i)^2 \quad (3)$$

As Cochran points out, s_2^2 is an unbiased estimator of S_2^2. However, s_1^2 overestimates S_1^2 because s_1^2 is calculated from sample means rather than true means, introducing additional variance that is proportional to the variance of these sample means. Correcting for this additional variance yields the following unbiased estimate of S_1^2:

$$s_1^2 - \frac{s_2^2(1 - f_2)}{m} \quad (4)$$

Two-Stage Sampling Schemes for Transit

As mentioned earlier, UMTA Circular 2710.1, as well as the BTMM, use two-stage sampling for transit ridership in which the day is the primary stage and the trip is the subunit. (Their sampling plans involve a degree of systematic sampling in the selection of days, for the same number of days is selected each week. This departure from simple random sampling is expected to have a small but beneficial effect on precision and a negligible effect on bias, and can therefore be ignored.) The other possible two-stage scheme is to have the trip be the primary unit. In this scheme a number of trips would be selected from the daily schedule, and each trip would then be sampled on a fixed number of randomly selected days.

The more efficient scheme is simply the one with the lowest total sample size (i.e., total trip-day combinations), since neither scheme offers any significant efficiencies in collecting the data. (In addition, the first scheme can concentrate the checking on a smaller number of days, which may have a small impact on manpower availability.) Under either scheme, it will usually be efficient to define "trip" as a round trip, since checkers must usually make round trips anyway in order to return to their starting point. Efficiency therefore depends on the Stage 1 and 2 variances and the finite population corrections, as indicated by Equation 1.

For estimating ridership of a bus system composed of many bus lines, the first scheme appears appropriate. There are only 250 or so weekdays per year as opposed to thousands of trips in the daily schedule, so if the day is the primary unit it is certainly plausible that the finite population correction for Stage 1 will be considerably below unity, or even zero if every day is sampled, substantially reducing or eliminating the Stage 1 variance's contribution to standard error and leaving as the main source of variability the variance between trips on a given day, which, while large, is divided in Equation 1 by mn, the total number of subunits in the sample. However, for a simple light rail line the second scheme, with the trip as the primary unit, appears more natural. The weekday schedule is likely to have about 100 round trips, fewer than the number of weekdays in a year, so the Stage 1 finite population correction factor can more easily become zero under Scheme 2. If the sample size is at least 250, Stage 1 variance vanishes for either scheme, and efficiency depends entirely on second-stage variance, which in Scheme 1 depends primarily on peaking of demand and in Scheme 2 depends on day-to-day variation in demand and operations. In Figure 1, second-stage variability is illustrated for both schemes on the basis of Metro Blue Line data. It was expected—and the data in Figure 1 confirm the expectation—that the second-stage variance would be smaller under Scheme 2. As it turned out, the difference between the two schemes was not as great as expected.

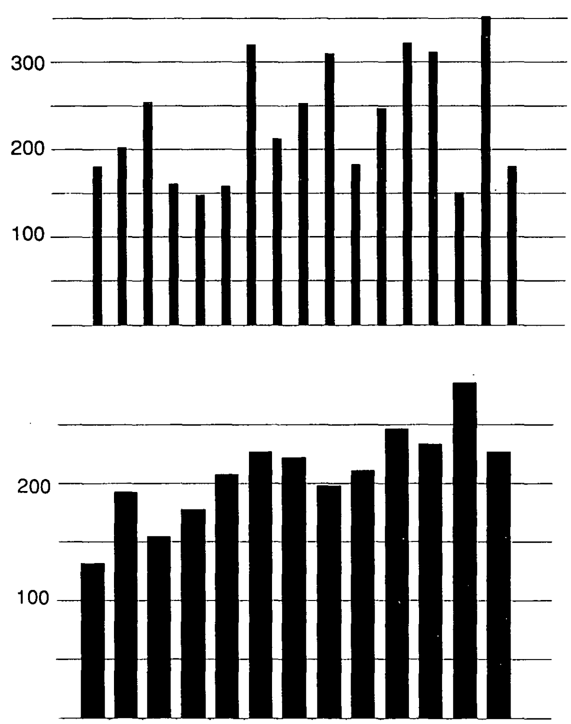

FIGURE 1 Second-stage variability for two sampling schemes: *top*, Scheme 1, within a day; *bottom*, Scheme 2, within a trip.

Varying Second-Stage Sample Sizes

The available Metro Blue Line data had been sampled following the first scheme—that is, days were selected at random each week and trips were selected at random for each selected day. Matters were complicated by the fact that the sampling rate had changed overtime—sampling had been more intensive soon after the line opened because both management and public interest were great at that time. Therefore, although each primary unit (the day) is still equal in size (same number of trips), the second-stage sample size was not the same across all sampled days. Letting

m_i = number of subunits sampled within primary unit i

the sample within unit variance should be defined as follows:

$$s_2^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\sum_{j=1}^{mi}(y_{ij}-\bar{y}_i)^2}{m_i-1}\right] \tag{3a}$$

$$s_2^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\sum_{j=1}^{mi}(y_{ij}-\bar{y}_i)^2}{m_i-1}\right] = \frac{1}{n}\sum_{i=1}^{n}s_{2i}^2 \tag{3b}$$

which is then an unbiased estimator of S_2^2. An unbiased estimator of S_1^2 is then

$$s_1^2 - \frac{s_2^2(1-f_2')}{\overline{m}'} \tag{4a}$$

where $\overline{m}'$ is the harmonic mean of the second-stage sample size given by

$$\overline{m}' = \frac{1}{\frac{1}{n}\sum\frac{1}{m_i}} \tag{5}$$

and

$$f_2' = \frac{\overline{m}'}{M} \tag{6}$$

Furthermore, the variance of the final estimate, given in Equation 1, should be modified by replacing m with $\overline{m}'$ and f_2 with f_2'.

To find the expected value of an estimate, first average over all possible second-stage samples for a fixed set of n Stage 1 units. Then average this result over all possible Stage 1 samples of n units. This logic can be represented as $E(\bullet) = E_1[E_2(\bullet)]$. Applying this logic to Equation 3a,

$$E_2\left[\frac{\sum_{i=1}^{mi}(y_{ij}-\bar{y}_i)^2}{m_i-1}\right] = S_{2i}^2 \tag{7}$$

where S_{2i}^2 is the variance among subunits within the ith primary unit. Therefore

$$E_2(s_2^2) = \frac{1}{n}\sum S_{2i}^2 \tag{8}$$

Then averaging over all possible first stage samples of size n,

$$E(s_2^2) = E_1[E_2(s_2^2)] = S_2^2 \tag{9}$$

To obtain the expected value of s_1^2 (still given by Equation 2), two results should first be established. First, given that unit i has been selected in Stage 1 and asigned to have m_i randomly selected subunits sampled, the variance of the Stage 2 sample mean of unit i is, from the theory of simple random sampling without replacement,

$$\text{var}_2(\bar{y}_i) = S_{2i}^2\left(\frac{M-m_i}{Mm_i}\right) \tag{10}$$

Likewise, for a given Stage 1 selection and a given assignment of Stage 2 sample sizes m_i,

$$\text{var}_2(\bar{\bar{y}}) = \text{var}_2\left(\frac{1}{n}\sum\bar{y}_i\right) = \frac{1}{n^2}\sum S_{2i}^2\left(\frac{M-m_i}{Mm_i}\right) \tag{11}$$

Considering now s_1^2, we know that

$$\sum(\bar{y}_i-\bar{\bar{y}})^2 = \sum\bar{y}_i^2 - n\bar{\bar{y}}^2 \tag{12}$$

so that

$$\begin{aligned}(n-1)E_2(s_1^2) &= E_2\left(\sum\bar{y}_i^2\right) - nE_2(\bar{\bar{y}}^2)\\ &= \sum\bar{Y}_i^2 + \sum S_{2i}^2\left(\frac{M-m_i}{Mm_i}\right)\\ &\quad - n\bar{\bar{Y}}_n^2 - \frac{1}{n}\sum S_{2i}^2\left(\frac{M-m_i}{Mm_i}\right)\\ &= \sum\left(\bar{Y}_i-\bar{\bar{Y}}_n\right)^2\\ &\quad + \left(\frac{n-1}{n}\right)\sum S_{2i}^2\left(\frac{1}{m_i}-\frac{1}{M}\right)\end{aligned} \tag{13}$$

where $\bar{Y}_n$ is the population mean per subunit for the selected n primary units.

When averaging over all possible Stage 1 samples of n units with assigned Stage 2 sample sizes m_i, random sampling and random assignment implies that

$$E_1\left[S_{2i}^2\left(\frac{1}{m_i}\right)\right] = E_1(S_{2i}^2)E_1\left(\frac{1}{m_i}\right) = S_2^2\frac{1}{\overline{m}'} \tag{14}$$

Therefore, dividing Equation 13 by $(n-1)$ and taking the expectation over Stage 1,

$$\begin{aligned}E(s_1^2) &= S_1^2 + S_2^2\left(\frac{1}{\overline{m}'}-\frac{1}{M}\right)\\ &= S_1^2 + \frac{S_2^2}{\overline{m}'}(1-f_2')\end{aligned} \tag{15}$$

from which Equation 4a follows. The proof for the modification to Equation 1 follows the same line of reasoning.

Metro Blue Line Results

Results for Scheme 1 are presented in Table 1, Column a on the basis of weekday data from July 1991 through February 1992, during which time the rail line schedule and the sampling schedule were relatively stable. Variance information is presented in the form of coefficient of variation (relative standard deviation) or cv = (square root of variance)/mean, since the cv is more readily visualized and is closely related to statistical precision. As indicated in the table, the Stage 1 sample cv (variation between the sample means of the sampled days) is 21 percent, on the basis of a sample of 135 days with a harmonic mean = 1.3 trips per day. The Stage 2 sample cv (variation between trips within a day) is 22 percent, on the basis of data from 56 days on which multiple trips were sampled (harmonic mean of 2.3 trips per day). This figure conforms with expectations based on occasional 100 percent ride checks (checking every trip on a given day). The estimated Stage 2 cv is the same as the sample Stage 2 cv, 22 percent, but the estimated Stage 1 cv, 9 percent, is far smaller than the sample Stage 1 cv. This is a good example of the bias inherent in the sample Stage 1 variance—it incorporates variance from both the true Stage 1 variance and from the sampling variance inherent in a sample of only (on average) 1.3 trips per day. Applying Equation 4a with cv's and ignoring the insignificant Stage 2 finite population correction,

$$cv_1^2 = CV_1^2 + CV_2^2/\overline{m}' \cong CV_1^2 + cv_2^2/\overline{m}'$$

$$0.21^2 = 0.09^2 + 0.22^2/1.3$$

where CV_i is population stage i coefficient of variation and cv_i is sample stage i coefficient of variation.

With these results, first- and second-stage sample sizes necessary to achieve a given precision level can be determined by trial and error. For the Metro Blue Line, and probably for most light rail lines, efficient plans for increasing levels of precision involve sampling one trip per day for n days until every day is covered once, then adding a second trip, and so on. (Under such a plan, the Stage 2 finite population correction can be neglected.) The sample size necessary to estimate annual weekday boardings to a precision of 10 percent was found to be one round trip per day for 22 days, which compares favorably with the more than 540 randomly selected one-way trips required by UMTA Circular 2710.1. For 5 percent precision, the necessary sample size is 1 round trip per day for 82 days, and for 2 percent precision it is 2 round trips per day, every weekday of the year, for a total sample size

TABLE 1 Sampling for Weekday Boardings

	(a)	(b)	(c)	(d)
sampling element	round trip	round trip	round trip	one-way trip
stage 1	date	trip	trip	trip
stage 2	trip	date	date	date
N	255	112	224	206
M	112	255	128	255
ANALYSIS DATASET				
n	135	112	same	206
m	1.3	1.3	as	2.3
total sampled trips	219	219	(b)	757
mean	395	385		46.5
cv	0.21	0.26		0.58
cv	0.22	0.19		0.32
Estimated CV	0.09	0.19	0.19	0.54
Estimated CV	0.22	0.19	0.19	0.32
SAMPLING PLAN FOR 10% PRECISION (m=1)				
n	22	26	27	98
SAMPLING PLAN FOR 5% PRECISION (m=1)				
n	82	76	91	190
SAMPLING PLAN FOR 2% PRECISION				
n	255	112	224	206
m	2	3.25	1.75	5
m	2	3.2	1.6	5
nm = sample size	510	364	392	1030

(a) Metro Blue Line, scheme 1
(b) Metro Blue Line, scheme 2
(c) Metro Blue Line, scheme 2 with thorough midyear timetable change
(d) Guadalupe Line, scheme 2

of 510 round trips. (Here and throughout the paper the 95 percent confidence level is used, for which the precision is 1.96 times the relative standard error.)

The same data set was used to test the second scheme for two-stage sampling, namely, selecting *n* trips at the first stage and then selecting *m* days on which to sample each selected trip. The results of this analysis are approximate since they are based on a sample that was selected after the first scheme, so that the selection of days is not entirely independent from one trip to another. The results, presented in Table 1, Column b, are that the Stage 1 (variation between the means of different round trips) *cv* is estimated to be 19 percent, and the Stage 2 (variation between days for a given round trip) *cv* is also estimated to be 19 percent. Consequently, the sample sizes needed to achieve 10, 5, and 2 percent precision, respectively, are 26, 76, and 364 round trips. The latter case involves sampling every trip 3.25 times on average, meaning 75 percent of the round trips are sampled three times a year and the rest are sampled four times. The harmonic mean of the second-stage sample size is 3.2, so whereas the sampling cost is proportional to 3.25 days per trip, the precision is calculated as if there were only 3.2 days per trip.

Comparing the two schemes, there appears to be little difference except when the sample size is large enough that nearly every Stage 1 unit is covered, in which case the effect of the Stage 1 variance is greatly diminished by the finite population correction and the critical factor is the Stage 2 variance. The prior expectation was that the Stage 2 variance in Scheme 2 (variation between days for a given trip) would be considerably smaller than the Stage 2 variance in Scheme 1 (variation between trips for a given day), making Scheme 2 more efficient. However, the available data show only a small advantage for Scheme 2. Given the limited scope and imperfections of the data set, further data collection and analysis will be needed before anything definitive can be concluded about the relative efficiencies of the two schemes. Further data collection and analysis will also be needed to see that the levels of variance observed are maintained as the systems matures.

Practical Considerations

Transforming these results into actual sampling plans for the Metro Blue Line required some modifications to account for several complications.

1. Data were available for 2 fiscal years, 1990–1991 and 1991–1992. Both data sets were analyzed and *cv*'s were taken as an average for the 2 years, with a greater weight placed on the more recent year.

2. The same analysis was applied to Saturday and Sunday data. In the final sampling plan each day type is a separate stratum. For the final annual estimate the strata are combined using standard formulas for stratified sampling.

3. The trip schedule does not consist only of simple round trips. There are pull-outs and pull-ins as well from the depot that is located along the line about 3 mi from the Long Beach terminal. Therefore, the sampling unit is a generalized round trip, which may be a simple round trip, a round trip plus a pull-in or pull-out between the depot and Long Beach, or a depot–Los Angeles–Long Beach trip, or any other combination whose ridership is expected to be comparable to that of a simple round trip. The daily schedule should be divided a priori into sampling units, and each sampling unit selected with equal probability (within a day type stratum).

4. What if the timetable changes during the year? For Scheme 1, this change could be ignored, assuming the underlying between-day and between-trip variations do not change much. For Scheme 2, a thorough timetable change can be treated like a doubling of the number of scheduled trips and a halving of the number of days each trip runs. This modification can affect significantly the magnitude of the finite population correction (since the population of trips doubles), reducing the efficiency of the method. If the timetable change is partial (as are most timetable changes), the effect is less severe. A sampling plan for the Metro Blue Line that assumes there will be one thorough timetable change in the year is shown in Table 1, Column c. Sample sizes are only a little higher than with no timetable change, assuming no change in the Stage 1 and 2 variances. Of course, if the timetable change significantly affects the underlying variances (e.g., changing departure times to smooth out vehicle loads should reduce the between-trip variance), the underlying variances should be reestimated, and the before and after parts of the year treated as separate strata.

5. As mentioned earlier, with either scheme the most efficient sampling plan for moderate to high levels of precision calls for only one second-stage sample per primary unit sample. The data generated by this kind of sampling are insufficient to reestimate S_1^2 and S_2^2 in the future. Therefore, it is recommended that a sampling plan involving 2 days per trip be followed every third year or so to permit reestimation of the underlying variances. To illustrate the loss of efficiency from increasing the Stage 2 sample size, an analysis was done for Scheme 2, considering a stratified sample with three day types, a mid-year schedule change, and a given number (124) of round trips. Sampling 62 round trips on 2 days each resulted in a precision of 7.2 percent, and sampling 124 round trips on 1 day each yielded a precision of 5.5 percent.

SAMPLING METHODOLOGY FOR SANTA CLARA COUNTY TRANSIT DISTRICT GUADALUPE LINE

The Guadalupe Light Rail Line is operated by the Santa Clara County Transit District in a north-south corridor through downtown San Jose, California. Like the SCRTD Metro Blue Line, the Guadalupe Line is barrier-free, with ticket-vending machines and multiticket canceling machines on station platforms. The ride check data available for this study came from 1990, when only the downtown and northern portions of the line were open.

Like the Metro Blue Line, the Guadalupe Light Rail line is the first light rail line in its county in several decades. Its ridership is changing rapidly as people accommodate their travel patterns to the line's availability and as new segments of the line open. Both management and the general public have a keen interest in tracking ridership. Therefore, ridership estimates must be more accurate than the annual 10 percent precision mandated by Section 15. After consulting with management, it was concluded that the desired precision was 1 to 2 percent for annual estimates, corresponding to a precision of 2 to 4 percent for quarterly estimates. At this level it was clear that data collection would have to be sufficiently intensive to cover every trip in the schedule once a month for weekday trips and once a quarter for Saturday and Sunday trips. The sampling method then is a special case of two-stage sampling following Scheme 2 (first select trips, then select days for each trip) in which every trip in the schedule is selected; in effect, it becomes stratified sampling with each trip representing a stratum. Because every Stage 1 unit is sampled, Stage 1 variance is inconsequential; the only contribution to the variance of the estimate comes from Stage 2 variance, the variance between days for a given trip.

Presented in Table 1, Column d are the results from analysis of Guadalupe Line weekday data. The Stage 2 *cv* is 32 percent. This value is considerably higher than the corresponding variance for the Metro Blue Line (22 percent). Recognizing that mean boardings were about 50 per one-way trip on the Guadalupe Line and about 400 per round trip on the Blue Line, this result is in keeping with the common finding that the *cv* diminishes as the ridership level increases and suggests that the variance on the Guadalupe Line will diminish as the line expands and ridership grows. The Saturday and Sunday Stage 2 *cv*'s are still greater: 64 and 45 percent, respectively. Estimating weekday ridership with a 2 percent annual precision requires sampling all 206 weekday trips five times a year, for a total of 1,030 one-way trips. For 5 and 10 percent precision estimates, the high between-trip *cv* necessitates sampling 190 and 98 one-way trips, respectively, once a year.

Alternative sampling plans were recommended for the time after the southern portion of the line opened. By way of illustration, one plan was to sample each weekday trip once every 2 months and each weekend trip once a quarter. Assuming that the *cv*'s will be the same after the line expands, an assumption that the authors consider conservative, this plan is expected to achieve 1.5 percent precision in the estimate of total annual boardings and 3 percent precision for quarterly estimates. Calculations for this stratified sampling plan are shown in Table 2. The following formulas are used in that table (the Stage 2 finite population correction is ignored):

h = stratum identifier

w_h = stratum weight = (total trips)$_h$/(total trips)

$$\text{variance} = \sum_h [w_h(\text{mean per trip})_h(\text{within trip } cv)_h]^2/n_h m_h$$

$$\text{mean} = \sum_h w_h(\text{mean per trip})_h$$

$$\text{precision} = 1.96\sqrt{\text{variance}}/\text{mean}$$

It is interesting to note that if each sampling unit (a date-trip combination) is treated as an independent unit without recognizing its two-stage structure, the sample size necessary to achieve a given level of precision would be a little more than three times larger than needed when the two-stage structure is exploited. If the sampling units are stratified into 10 strata by period (Saturday, Sunday, and four weekday periods) and by direction (for weekday only), the necessary sample size would still be almost twice as large as required by the two-stage scheme, which is effectively stratified sampling with about 500 strata (one each for 231 weekday, 135 Saturday, and 135 Sunday trips).

RELATIONSHIP TO OTHER PUBLISHED SAMPLING PLANS

It seems appropriate to contrast the methodology described in this paper with those in the four major publications dealing with statistical sampling for transit ridership: the BTMM (*3*); its update, the TDCDM (*4*); and UMTA Circulars 2710.1 (*5*) and 2710.4 (*6*), all of which were written primarily for bus systems.

The BTMM assumes a two-stage sampling plan. Circular 2710.1 is based on its methodology and established necessary sample sizes by using default values for the Stage 1 and 2 variances and ignoring the finite population correction. The BTMM formulas differ from those used in this paper in two

TABLE 2 Stratified Sampling Plan for Guadalupe Line

	Wkday	Sat	Sun	Total
N (trips)	231	135	135	
M (days)	255	52	58	365
total trips = MN	58905	7020	7830	73755
stratum weight	0.799	0.095	0.106	
mean per trip	46.5	37.7	35.2	
mean contribution	37.1	3.6	3.7	44.5
within trip cv	0.32	0.64	0.45	
SAMPLING PLAN				
n	231	135	135	
m	6	4	4	
sampled trips = mn	1386	540	540	2466
variance contribution	0.0976	0.0090	0.0049	0.1116
standard error				0.334
precision @ 95% conf.				1.5%

ways, and it seems proper to explain the differences for the sake of avoiding confusion.

1. The BTMM uses s_1^2, the variance between Stage 1 sample means, as an estimator of S_1^2, the variance between Stage 1 true means. As we have shown, s_1^2 is always greater than S_1^2 because the former incorporates a degree of Stage 2 sampling variation inherent in the sample means. Omitting the correction from the BTMM makes its results more conservative than necessary.

2. The BTMM formula for Stage 1 or between-day variance (*3*,p.118) differs from Equation 2 in that each primary unit is weighted by the number of subunits sampled, k_i (called m_i in this paper). The BTMM formula for Stage 2 or within-day variance (3,p.119) differs from Equation 3a likewise. Weighting is appropriate only in the case of a stratified sample (e.g., a sample that includes weekday and weekend trips). However, even in such a case the weights should be the population weights, not the sample weights. Therefore, these formulas are valid only when the sample weights are the same as the population weights (e.g., when every trip, regardless of stratum, has an equal probability of selection). The examples presented in the BTMM (*3*,pp.134–135) have equal sampling rates, and are therefore correct. However, if the sampling rate is not the same in all strata, or if some weekdays are sampled more often than others (say, due to invalid samples being discarded), the BTMM formula will bias the results in favor of the days that, for whatever reasons, are overrepresented in the sample.

Circular 2710.4 and the TDCDM model the sampling process as simpler one-stage sampling, assuming that trip-day combinations are simply selected at random. (In fact, as in Circular 2710.1, selections in Circular 2710.4 are not purely random but contain a systematic element in that an equal number of selections are made for every week.) These documents retreat from the two-stage sampling of their antecedents because when there are thousands of trips in the daily schedule, as is the case in large bus systems, and the required precision demands a sample of fewer than one trip a day, the possibility of a trip being repeated in a one-stage sample is so remote that a one-stage sample is essentially indistinguishable from a two-stage sample. It can easily be shown that (a) when $m = 1$ (i.e., each sampled trip is sampled only once) and (b) the finite population corrections at both stages are negligible, the two-stage estimates of the overall mean per trip and of its variance become identical to single stage estimates from an infinite population, which is the assumption of the TDCDM and Circular 2710.4. Why then does this paper revert to two-stage sampling? Because the vastly smaller number of scheduled trips and the greater demands for accuracy on a light rail line make it almost certain that a one-stage sample will include some scheduled trips many times while other trips are excluded, whereas a two-stage sample can be designed to provide for even coverage of all trips. Even coverage is more efficient statistically and benefits the transit agency more in meeting other data needs.

REFERENCES

1. W. G. Cochran. *Sampling Techniques*, 3rd ed. John Wiley & Sons, New York, 1977.
2. P. R. Stopher and A. H. Meyburg. *Survey Sampling and Multivariate Analysis for Social Scientists and Engineers*. D. C. Heath, Lexington, Mass., 1979.
3. J. P. Attanucci, I. Burns, and N. H. Wilson. *Bus Transit Monitoring Manual*. Report UMTA-IT-09-9008-811. UMTA, U.S. Department of Transportation, 1981.
4. P. G. Furth, J. P. Attanucci, I. Burns, and N. H. Wilson. *Transit Data Collection Design Manual*. Report DOT-I-85-38. U.S. Department of Transportation, 1985.
5. *Sampling Procedures for Obtaining Fixed Route Bus Operating Data Required Under the Section 15 Reporting System*. UMTA Circular 2710.1. UMTA, U.S. Department of Transportation, 1978.
6. *Revenue-Based Sampling Procedures for Obtaining Fixed Route Bus Operating Data Required Under the Section 15 Reporting System*. UMTA Circular 2710.4. UMTA, U.S. Department of Transportation, 1985.
7. *Sampling Procedures for Obtaining Demand Responsive Bus System Operating Data Required Under the Section 15 Reporting System*. UMTA Circular 2710.2. UMTA, U.S. Department of Transportation, 1978.
8. P. G. Furth, K. L. Killough, and G. F. Ruprecht. Cluster Sampling Techniques for Estimating Transit System Patronage. In *Transportation Research Record 1165*, TRB, National Research Council, Washington, D.C., 1988, pp. 105–114.
9. P. G. Furth and B. McCollom. Using Conversion Factors To Lower Transit Data Collection Costs. In *Transportation Research Record 1144*, TRB, National Research Council, Washington, D.C., 1987, pp. 1–6.
10. S. P. Phifer. Use of Sampling in Bus-Line Data Collection. In *Transportation Research Record 877*, TRB, National Research Council, Washington, D.C., 1982, pp. 41–44.
11. A. Kumar and S. T. Parry. Section 15 Patronage Reporting for Barrier-Free Transit Service. *Journal of Transportation Engineering*, ASCE, Vol. 119, May 1993.

Publication of this paper sponsored by Committee on Transit Management and Performance.