Using Early Performance to Project Transit Route Ridership: Comparison of Methods

JAMES F. FOERSTER AND NEILA IMLAY

The performance of eight models for predicting ridership levels on new transit routes by using early performance data is summarized. Seven of the models are based on least-squares estimates of linear and nonlinear functions; the eighth model is a manual method based on quarterly ridership statistics. Comparisons are based on *r*-square statistics, leverage estimates, and ability to predict ridership levels for the second year of operation. The results of these comparisons indicate that (a) forecasts based on less than 6 months of data are unreliable for all conventional statistical models, (b) a simple manual method based on prior experience with other local routes is more effective than least-squares models if ridership forecasts must be produced on the basis of limited amounts of data, and (c) probit-, logit-, and power-function and linear-log models perform acceptably if more than 6 months of data are used.

Previously published reports have presented different approaches for predicting ultimate ridership levels on newly introduced bus routes. In particular, Cherwony and Polin demonstrated that a logit curve can be fit to early ridership figures to predict performance in later periods (*1*). Subsequently, Foerster et al. presented a manual method based on experience with other transit routes for this same purpose (*2*). These approaches differ from methods for predicting ridership for new routes on an a priori basis because they are based on actual route performance data (*3*).

This research was conducted to compare the performance of the previously described methods and to investigate the performance of alternative model forms. Results for all methods are presented and compared. Seven of the models are based on least-squares estimates of linear and nonlinear functions; the eighth model is a manual method based on quarterly ridership statistics. Comparisons are based on *r*-square statistics, leverage estimates, and ability to predict ridership levels for the second year of route operation.

The results of these comparisons indicate that (a) forecasts based on less than 6 months of data are unreliable for most of the models, (b) a simple manual method based on prior experience with other local routes is more effective than least-squares models if estimates must be produced on the basis of one or two quarters of ridership data, and (c) logit-, probit-, and power-function, and linear-log models perform acceptably if more than 6 months of data are available.

School of Urban Planning and Policy, University of Illinois at Chicago, P.O. Box 4348, M/C 351, Chicago, Ill. 60680.

MODEL SPECIFICATION

Seven models were calibrated using standard statistical techniques. All used time (t) as an independent variable to predict ridership (y). Five of the functions have upper limits. The first two of these functions are the logit function:

$$y = B0/(1 + \exp[B1 + B2 * t]), \tag{1}$$

which is of particular interest because it was found to produce acceptable results in previous research (*1*), and the probit function:

$$y = B0/\Phi[B1 + B2 * t]. \tag{2}$$

The probit function is similar to the logit function, but it is somewhat easier to estimate using numerical methods. It is slightly less sensitive than Function 1 to the values of initial and final observations. This function is also of considerable interest because it represents an accepted model of the rate at which new products and services are adopted (*4*).

Three other functional forms were included in the research design because they are typically used to model asymptotic growth processes. These are the negative exponential:

$$y = B0 * (1 - \exp[-B1 * t]), \tag{3}$$

a linear model with a reciprocal transformation of t:

$$y = B0 + B1/t, \tag{4}$$

and an exponential model with a reciprocal transformation of t:

$$y = \exp[B0 - B1 * 1/t]. \tag{5}$$

Two other functional forms were included in the evaluation. These were originally chosen to serve as a baseline for comparison of the asymptotic models. As will be seen, they also provide useful forecasts. These models are the power function:

$$y = B0 * t ** B1 \tag{6}$$

and a linear model with a logarithmic transformation of t:

$$y = B0 + B1 * \ln(t). \tag{7}$$

An eighth model was calibrated with data from other transit routes instead of actual ridership for the subject route. The model yields a set of indexes that are used as multipliers to factor early ridership levels up to an expected ultimate ridership estimate. This procedure is described in a previous paper by Foerster et al. (*2*).

DATA

The data used for model calibration were provided by Pace, the Suburban Transit Division of the Chicago Regional Transportation Authority. The data consisted of weekday ridership counts for the first 361 days of operation of Route 354. Service on this route was initiated in 1987. The data were divided into two sets. The first 262 observations were used for model calibration. An additional 99 records from the second year of operation were used as a hold-out sample to test the accuracy of model forecasts.

PROCEDURES

Models 1 through 7 were calibrated using the nonlinear modeling (NLIN) procedure of the Statistical Analysis System (*5*). A variety of starting values and search techniques were used to ensure that the solutions obtained were not local minimums. Twelve calibration runs were made for each of Equations 1 through 7. Each calibration run was based on $22 * n$ data points for $n = [1,2, \ldots ,12]$. Ridership forecasts, *r*-square statistics, parameter estimates, and leverage function values were obtained for each calibration run. This, in effect, simulates the results that an analyst would obtain if ridership data were analyzed with each of the methods at the end of each month of operation.

The manual estimation technique described by Foerster et al. (*2*) was used to obtain ultimate ridership estimates. The ratios of index values for the 5th and 6th quarters to ultimate ridership index values were used to produce ridership estimates for the forecast period.

COMPARISON OF CALIBRATION RESULTS

The *r*-square measures of fit obtained for each of the 12 calibration runs of each model are shown in Figure 1. These values show that none of statistically calibrated Models 1 through 7 fit the observed data very well for the first 6 months of data. For the following 6 months, the logit-, probit-, and power-function and linear-log models appear to fit better than the other models. It is clear that the negative exponential, linear-reciprocal, and reciprocal-exponential mod•ls do not fit the observed data well for any calibration period.

The influence statistic $S(\mathbf{h})$ was computed for all of the calibration runs. This statistic is the standard deviation of the leverage function

$$h(i) = \mathbf{x}(i)(X'X)\,\mathbf{x}'(i) \quad (8)$$

where $X = dF/d\mathbf{B}$ and $\mathbf{x}(i) = \text{row}(i)$ of $\mathbf{X}$. Since $h(i)$ is a measure of the influence of data point i on the parameter vector $\mathbf{B}$, $S(\mathbf{h})$ is a measure of the variation of the influence each data point has on the parameters that are estimated. In general, models that have a lower value of $S(\mathbf{h})$ should be preferred because such models do not give excessive weight to any one data point. Large variances in influence values were noted for all model calibrations based on data from Months 1 through 6. In addition, it was found that the parameters of the linear-reciprocal model were more strongly influenced by a small number of data points than were those of the other models.

FORECASTING RESULTS

The forecasting ability of the models was analyzed by examining forecasted and observed ridership levels. Months 13 through 17 of route operation were used as the forecast period, and data from these months were used to test forecast accuracy. (The data for these months were not used in model calibration.)

Figure 2 shows the root mean square (rms) error that would occur if this forecast was used in an applied setting. The values indicate that the manual method produced more accurate forecasts on the basis of first- and second-quarter perfomance than any of Models 1 through 7. It can be seen that the log-linear and power-function models outperform the previously recommended logit function and the associated probit function if limited amounts of data are available. Furthermore, it

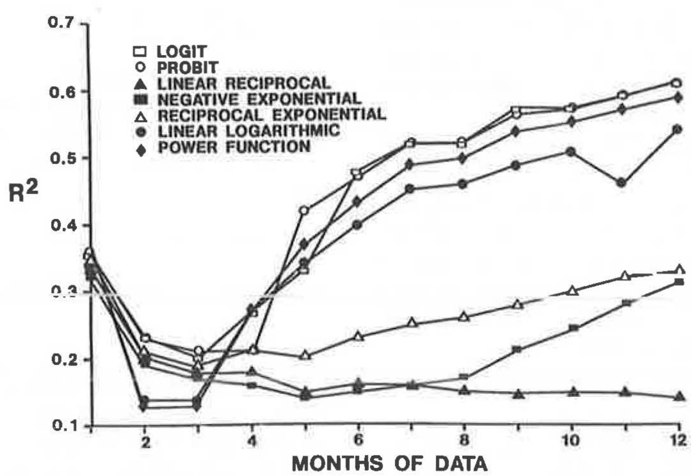

FIGURE 1 Model fit statistics.

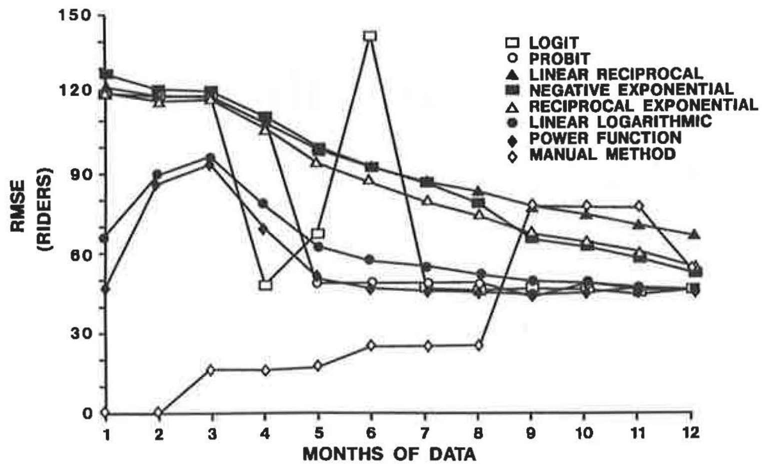

FIGURE 2 **RMS forecast error.**

appears that use of the manual method is preferable if forecasts must be generated during the first 6 months of route operation because of high error rates for all of the statistically calibrated models.

SUMMARY

These results are by no means conclusive. Although over 300 data points, 8 models, and 12 calibration runs per model were used, our sample has consisted of only one route. Different results may be obtained by other analysts using data from other locations.

However, we have shown that it may be misleading to develop trend forecasts for new routes, regardless of the functional forms used, if only a few months of data are available. We have also shown that simpler manual methods that take advantage of local experience should be given every consideration in spite of their simplicity. In fact, these methods may even be preferable because they can be applied in 5 to 10 min with the aid of a hand calculator, in contrast to the time and expense associated with developing calibrations for more-sophisticated, yet less-accurate statistical models.

Continuation of this research is clearly warranted. We will continue to observe the performance of the route in question and update our comparison of the forecast performance of the models considered. The results we have reported should be validated by subjecting additional data to similar analyses. We would gladly do this for any property willing to submit daily ridership data. In addition, we see a need to refine all of the methods to account for seasonality; this would be most easily accomplished by applying correction factors based on system-level trends.

REFERENCES

1. W. Cherwony and L. Polin. Forecasting Patronage on New Transit Routes. In *Traffic Quarterly*, Vol. 31, No. 2, April 1977, pp. 287–295.
2. J. F. Foerster, A. Jones, and D. Fisher. Monitoring Performance of New Bus Routes. In *Transportation Research Record 1165*, TRB, National Research Council, Washington, D.C., 1988, pp. 69–74.
3. Route-Level Demand Models: A Review. Report DOT-I-82-6. UMTA, U.S. Department of Transportation, 1982.
4. M. E. Rodgers. *Diffusion of Innovations*. Free Press, New York, 1962.
5. *SAS Users Guide: Statistics, Version 5 Edition*, SAS Institute, Inc., Cary, N.C., 1985.

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