Empirical Modeling and Forecasting of Monthly Transit Revenue for Financial Planning: A Case Study of SCRTD in Los Angeles

DAVID SKINNER, ROBERT WAKSMAN, AND GEORGE H. WANG

Time series revenue data from the Southern California Rapid Transit District (SCRTD) in Los Angeles are used as a case study to develop empirical models and forecasts of monthly transit revenue for financial planning. Seasonal time series models for the five major types of transit revenues collected by SCRTD are specified and estimated. For all five types, the observed variation in revenue during the estimation period fits well with the values obtained from the models, as is demonstrated by the relevant regression statistics and the behavioral characteristics of the model. A data split technique is used to determine the prediction capabilities of the model. A comparison of the forecasts with the actual data shows that the models perform well for forecasting purposes. Finally the models are used in a simulation mode to estimate the impact of the SCRTD June 1982 fare rollback on revenues for the next year and a half.

The purpose of this analysis is to (a) construct and estimate seasonal models for the five major types of transit revenues collected by the Southern California Rapid Transit District (SCRTD); (b) examine the forecasting capabilities of these revenue models; and (c) use these models to forecast revenues for 1.5 years ahead with and without the July 1982 fare rollback. The five types of transit revenues are farebox (or cash), regular pass, senior citizen and handicapped pass, express stamp pass, and student pass.

Forecasts of the sale of transit services are the principal data needed for predicting cash flow. They reveal to financial planners when and how much cash income can be expected during the period being studied. This information along with a forecast of expenses assists planners in determining future cash needs, planning for financing these needs, and exercising control over the cash flow of the transit authority (1).

A survey of recent financial forecasting techniques in the transit industry indicates that various types of fare elasticity and simple trend extrapolation methods are the most popular devices used by transit planners as tools for forecasting transit revenue (2). The use of fare elasticities calculated from annual time series data or by shrinkage ratios tends to produce poor forecasts of monthly transit revenue because these methods are unable to predict the seasonal and working day variation in the revenue series. Furthermore, these methods do not take into account (or do so in a highly judgmental manner) the effects of changes in variables other than fares, e.g., gasoline prices and overall growth trends (3).

To improve the forecasts of monthly transit revenues, empirical modeling using the regression time series approach is adopted in this paper. Within the framework of a model using empirical monthly data, the impact on transit revenues of changes in real fares and real gasoline prices can be examined. Furthermore, alternative forecasts (simulations) can be generated from these models for alternative assumptions about the behavior of independent variables in the model.

This paper presents (a) a description and recent history of SCRTD transit fares and revenues, (b) the construction of monthly models for transit revenue, (c) empirical results, and (d) model assessment and forecasting performance of the models.

DESCRIPTION AND HISTORY OF SCRTD FARES AND REVENUES

The SCRTD fare structure has many different fare categories and pass types. There is a base adult cash fare for local bus service and additional incremental fares for express bus service for each of up to five express zones. Transfers between buses are priced at a small percentage of the base fare price. The three separate fare structures for handicapped and senior citizens, elementary and high school students, and college students are cash fares, incremental express zone fares, and transfer charges. For each of these user groups there is a monthly pass for unlimited rides on local bus service, and monthly express zone stamps are added to the monthly pass for unlimited rides on express service up to the indicated number of zones.

As shown in Table 1, from July 1976 to June 1982 there was a steady nominal increase in cash fares and pass prices for all user groups. Effective in July 1982, there was a sizable rollback in fare and pass prices. The rollback was mandated by Proposition A, which allocated the revenues from a one-half percent increase in the sales tax in Los Angeles County for SCRTD use. One purpose of this paper is to forecast the revenue changes that will result from the fare rollback.

This analysis makes use of monthly cash fare and pass revenue data collected since July 1976. Monthly farebox revenue data are aggregated from daily records of total cash fares kept by SCRTD. Every day, after each bus pulls into the garage following its last run of the day, the farebox is removed and is taken to a central counting facility where it is unlocked and the contents emptied into an automatic coin counter. The sum of the farebox revenues from all buses represents the total farebox revenue collected that day. This revenue includes cash fares of all types and for all user groups, i.e., base fares, zonal increments, and transfers for adults, handicapped and seniors, and students because the fareboxes cannot differentiate between fare categories.

Monthly pass revenue by type is obtained directly from SCRTD records. Passes for a given month are sold at the end of the previous month and at the beginning of the pass month. Records are kept of the number of passes of each type sold each month and of the revenue taken in for each pass type. In this analysis, for the purposes of simplification and of modeling, only pass revenues that contribute substantially to the SCRTD total revenue are included. Pass revenues are disaggregated into the following four categories: regular pass revenue, express stamp revenue, senior citizen and handicapped pass revenue, and student pass revenue.

Figure 1 shows trends in SCRTD total transit revenue and in the relative proportion derived from pass revenues since July 1976. Several interesting points can be summarized from the figure.

1. SCRTD total transit revenue has risen steadily from 1976 to 1981 at a compounded rate of approximately 20 percent.

2. Seasonal fluctuations exist in both farebox revenue and total pass revenue.
3. Pass revenue as a percentage of total revenue remained fairly constant between July 1976 and July 1980 at about 25 percent (perhaps increasing slightly during that time) and then jumped to about 35 percent after the July 1980 fare increase.

CONSTRUCTION OF MODELS FOR MONTHLY TRANSIT REVENUE

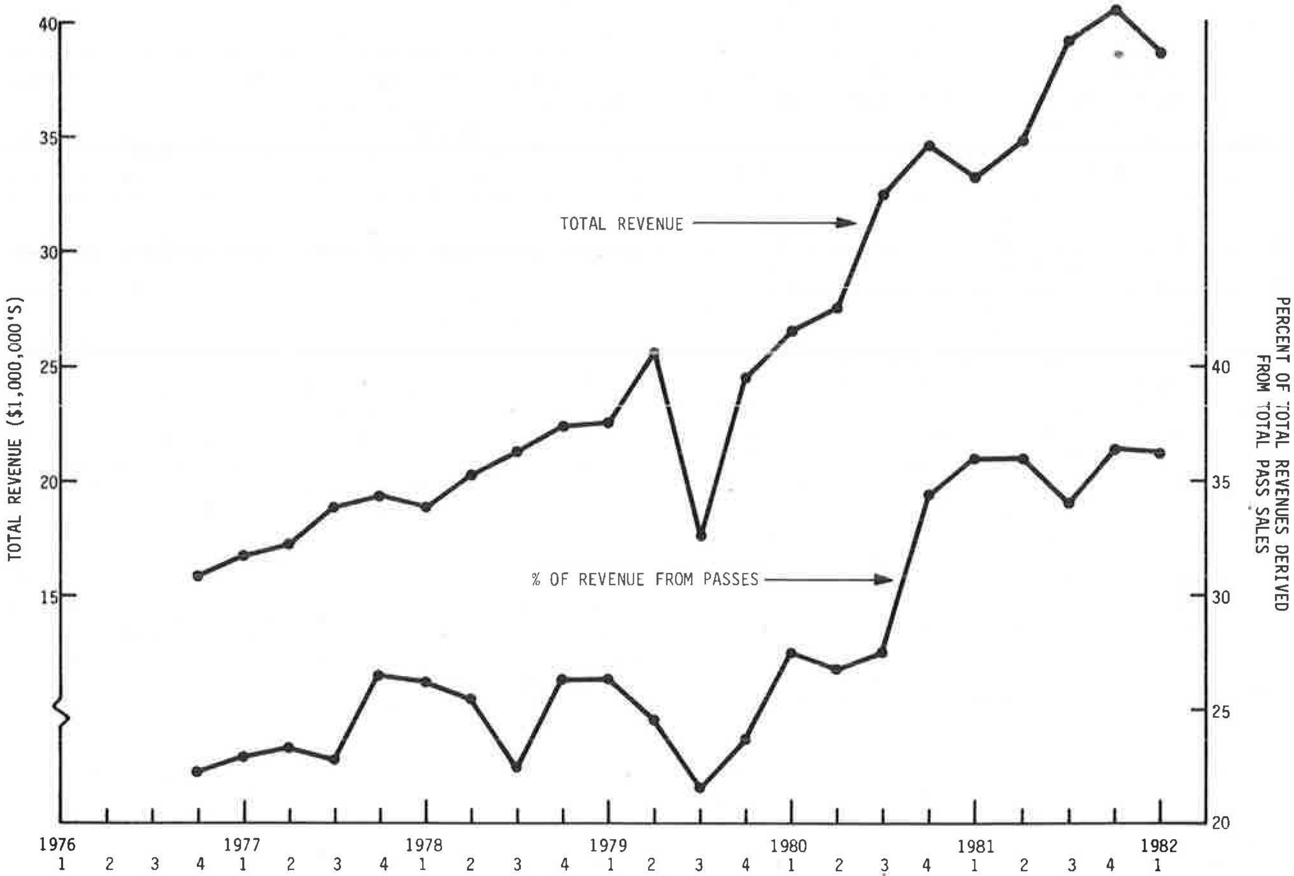
In general there are two approaches to modeling the level of transit revenue: (a) the direct approach, which uses revenue variables; and (b) the indirect approach, which first estimates ridership and then multiplies that ridership by an appropriate fare or average fare variable to obtain revenue. The choice of which approach to use depends on the relative availability--and quality--of the revenue and ridership data. The indirect approach was used by Wang, Ward, and Hassler (4) for modeling and forecasting New York City transit revenue. Ridership was regressed on a set of independent variables including (a) appropriate transit fares, (b) Consumer Price Index variables, (c) transit service variables, and

Table 1. SCRTD fare levels between July 1976 and July 1982 (dollars).

	July 1976-June 1977		July 1977-June 1978		July 1978-Oct. 1979		Nov. 1979-June 1980		July 1980-June 1981		July 1981-June 1982		July 1982-Present	
Fare Category	Cash	Pass	Cash	Pass	Cash	Pass	Cash	Pass	Cash	Pass	Cash	Pass	Cash	Pass
Adult base	0.35	14	0.40	18	0.45	20	0.55	20	0.65	26	0.85	34	0.50	20
Adult express zone increment	—[a]	—[a]	0.20	6	0.20	6	0.20	6	0.30	8	0.40	12	0.25	7
Adult with 1 transfer	0.45	14	0.50	18	0.55	20	0.60	20	0.85	26	1.00	34	0.60	20
Adult with 2 transfers	0.45	14	0.50	18	0.55	20	0.60	20	1.05	26	1.15	34	0.60	20
Adult with 3 transfers	0.45	14	0.50	18	0.55	20	0.60	20	1.70	26	2.00	34	0.60	20
Senior and handicapped	0.10	4	0.15	4	0.20	4	0.20	4	0.30	6	0.40	7.50	0.20	4
Senior and handicapped with 1 transfer	0.20	4	0.25	4	0.30	4	0.25	4	0.40	6	0.45	7.50	0.25	4
Student (thru high school)	0.25	—[b]	0.40	12[b]	0.45	14	0.45	14	0.50	16	0.65	22	0.20	4
Student (thru high school) with 1 transfer	0.35	—[b]	0.50	12[b]	0.55	14	0.50	14	0.60	16	0.70	22	0.25	4
Student (college)	0.25	—[b]	0.40	12[b]	0.45	—[c]	0.45[c]	—[c]	0.65	20	0.85	26	0.20	4
Student (college) with 1 transfer	0.35	—[b]	0.50	12[b]	0.55	—[c]	0.50[c]	—[c]	0.85	20	1.00	26	0.25	4

[a]Before July 1977 an additional 35 cents was charged for travel between two overlapping zones.
[b]Student pass not available until September 1977.
[c]Student pass not available to college students between July 1978 and June 1980.

Figure 1. Total revenue and percent of total revenue derived from pass sales by yearly quarter.



(d) variables representing seasonal and working day variation. Then, the estimated values for ridership were multiplied by a fare variable to give an estimate of revenue. This two-step approach, however, is not applicable to modeling SCRTD transit revenue because historical monthly pass ridership and transfer ridership data are not directly recorded, and the fare structure is quite complex as indicated in Table 1. Also, SCRTD cash fare revenue is not recorded by type. The only data available are total farebox revenue and pass revenue by pass type. For these reasons the direct approach of modeling transit revenue is adopted in this paper.

The direct approach is based entirely on empirical estimation of models using time series data. The initial statistical model for monthly farebox revenue is postulated as follows:

$$Y_t = B_0 + B_1X_{1t} + B_2X_{2t} + B_3X_{3t} + B_4X_{4t} + B_5X_{5t} + \sum_{j=1}^{m} \delta_j Y_{t-j} + \sum_{k=1}^{5} SS_{kt} + \sum_{k=1}^{6} SC_{kt} + WD_t + U_t \qquad (1)$$

where

- Y_t = monthly farebox revenue (in $1,000s).
- X_{1t} = adult cash fare deflated by the Consumer Price Index for Los Angeles-Long Beach (5). This is a proxy variable for the general level of fares and serves to improve the predictive ability of the model and to evaluate the effects of changes in fares on farebox revenue.
- X_{2t} = monthly gasoline price (obtained from Oil and Gas Journal) (6) deflated by the Consumer Price Index for Los Angeles-Long Beach.
- X_{3t} = price of alternative methods of payment. This variable tests the effects of substitution; e.g., the price of a regular monthly pass would be a possible substitute for cash payment.
- X_{4t} = a linear trend variable representing change occurring over a period of years.
- X_{5t} = total monthly nonagricultural employment in Los Angeles-Long Beach (7).
- SS_{kt}, SC_{kt} = seasonal variation component of transit revenue (see Equation 2 for explanation).
- WD_t = working day variation component of transit revenue.
- Y_{t-j} = lagged dependent variables where j = 1, 2,. . ., m. These variables are used to capture the systematic effects in the data that cannot be explained by the X_t variables.
- U_t = stationary time series error term.

Ideally variables that measure the underlying supply of, and demand for, transit service over time should be included in the preceding model. Unfortunately some of these variables are not readily obtainable on a monthly basis. For instance, SCRTD has changed its routes and schedules considerably during the period of the revenue model estimation (1976 to 1982); yet, the resulting changes in daily scheduled vehicle hours and miles have varied by only 14 percent during the 6-year period. With such a small variability, neither scheduled vehicle hours or miles was considered to be a significant variable. Developing monthly values for alternative variables that reflect changes in the level of service provided by SCRTD would be a formidable task involving considerable subjective judgment and, hence, was not undertaken.

Total monthly nonagricultural employment in Los Angeles-Long Beach was considered an important proxy variable for demand of transit service because work trips and shopping trips by employed workers are a large percentage of all trips. Data on employment for this area of California are readily available and, importantly, are forecast by several econometric services.

Seasonal variation in the revenue time series is specified as follows:

$$SS_{kt} = \alpha_k \operatorname{Sin}(w_k \cdot X_{4t}) \quad \text{for } k=1,\ldots,5$$
$$SC_{kt} = \delta_k \operatorname{Cos}(w_k \cdot X_{4t}) \quad \text{for } k=1,\ldots,6 \qquad (2)$$

where $w_k = 2\pi K/12$, and X_{4t} is defined as a trend variable.

Strong month-to-month variation in the revenue data is the result of working day variations. The reasons for the existence of these working day variations in the transit revenue data are as follows:

1. The level of ridership on weekdays is higher than that of weekends and holidays;
2. Monthly revenue data are an aggregate of daily data; and
3. Each calendar month has a different number of working days and holidays.

The specification of working day variation variables for transit series suggested by Wang (8) was adopted in this analysis. The set of working day variables are specified as

$$WD_t = \psi_1 WW2_t + \psi_6 WW6_t \qquad (3)$$

where

$WW2_t = \overline{TD}_t - 5*(TD_{7t}$ + number of holidays in the t^{th} month), and

$WW6_t = TD_{6t} - (TD_{7t}$ + number of holidays in the t^{th} month).

ψ_1 is the coefficient of the weekday effect, and ψ_6 is the coefficient of the Saturday effect. The coefficient for the Sunday effect can be derived as $\psi_{sun} = -5\psi_1 - \psi_6$. TD_{it} denotes the number of occurrences of the i^{th} day of the week in month t. TD_{it} can take only two values, 4 and 5. $\overline{TD}_t$ represents the total number of working days occurring in month t. The number of holidays occurring on weekdays of month t has been subtracted. The inclusion of working day variables in the model permits us to perform a direct test on the existence of working day variation in the data and to estimate this variation directly from data.

The models for pass revenues are similar in specification to the farebox revenue model described previously except that variables representing working day variation are excluded. Separate models for each of the four types of pass payment methods were constructed to allow for differences in the behavior of the parameter values and functional forms of each revenue time series.

EMPIRICAL RESULTS

Cash Fare Revenue

The cash fare revenue equation was estimated using monthly data from July 1976 to April 1982. The estimated equation for cash fare revenue is presented in Figure 2 as are the estimated equations for the four types of pass revenue.

Figure 3 shows that the observed variation in cash fare revenue during the estimation period fits well with the values obtained from the equation. The coefficients for the real cash fare and the real gasoline price are both positive and significant at the 1-percent level. Thus, as the cash fare and gasoline prices in the Los Angeles area increased (decreased) in real terms during the estimation period, revenue increased (decreased). The positive coefficients for real cash fare and real gasoline price and their significance at the 1-percent level confirm the expectation that increasing fares (in the range of fares being examined) will lead to increased revenues and that there is some switch to transit as real gasoline prices increase.

Working day variation is present in the cash revenue data and the estimated equation accounts for it with the variables WW2 and WW6. This variation indicates that revenue alone will increase (decrease) as the number of working days within a month increases (decreases).

Figure 2. Estimated equations.

● $$\mathrm{REV.CASH} = -\underset{(-5.97)}{8880.92} + \underset{(8.31)}{18712.60}*\mathrm{CASHFARE} + \underset{(4.69)}{7357.79}*\mathrm{LAGR} + \underset{(4.06)}{2.22}*\mathrm{LAEMP} + \underset{(4.88)}{29.41}*\mathrm{WW2} - \underset{(-0.87)}{32.43}*\mathrm{WW6} + \underset{(2.25)}{155.57}*\mathrm{SS1}$$
$$+ \underset{(1.49)}{49.44}*\mathrm{SS3} + \underset{(7.16)}{187.20}*\mathrm{SS4} - \underset{(-1.38)}{31.16}*\mathrm{SS5} - \underset{(-2.17)}{152.32}*\mathrm{SC1} + \underset{(2.11)}{55.33}*\mathrm{SC4} + \underset{(2.10)}{51.31}*\mathrm{SC5} + \underset{(2.63)}{44.63}*\mathrm{SC6}$$

$\bar{R}^2 = .864$ F(13/56) = 34.67 D-W = 2.11 GLS RHO1 = 0.70 PERIOD: 7/76 TO 4/82

● $$\mathrm{REV.REGPASS} = -\underset{(-2.17)}{390.19} + \underset{(3.54)}{52.02}*\mathrm{REGPASS} + \underset{(4.95)}{1396.23}*\mathrm{LAGR} + \underset{(7.66)}{533.75}*\mathrm{D80.7} + \underset{(8.97)}{0.50}*\mathrm{REV.REGPASS}(-1) + \underset{(2.29)}{34.67}*\mathrm{SS1}$$
$$- \underset{(-1.45)}{21.45}*\mathrm{SS3} + \underset{(1.54)}{22.68}*\mathrm{SS4} + \underset{(4.22)}{62.48}*\mathrm{SC3} - \underset{(-2.18)}{32.50}*\mathrm{SC4}$$

$\bar{R}^2 = .986$ F(9/60) = 544.19 D-h = 0.80 OLS PERIOD: 8/76 TO 5/82

● $$\mathrm{REV.S{+}H} = \underset{(0.84)}{33.43} + \underset{(7.05)}{65.09}*\mathrm{SHPASS} + \underset{(0.81)}{61.68}*\mathrm{LAGR} + \underset{(6.01)}{999.22}*\mathrm{SHFARE} + \underset{(11.97)}{3.91}*\mathrm{TTD}$$

$\bar{R}^2 = .959$ F(4/66) = 406.23 D-W = 2.10 GLS RHO1 = 0.68 PERIOD: 7/76 TO 5/82

● $$\mathrm{LOG(REV.EX)} = \underset{(12.84)}{7.13} + \underset{(1.52)}{0.27}*\mathrm{LOG(EXPASS)} + \underset{(6.64)}{1.28}*\mathrm{LOG(LAGR)} + \underset{(4.52)}{0.01}*\mathrm{TTD} + \underset{(3.08)}{0.54}*\mathrm{LOG(EXFARE)} - \underset{(-1.66)}{0.02}*\mathrm{SS3} + \underset{(1.89)}{0.02}*\mathrm{SS4}$$
$$+ \underset{(2.07)}{0.03}*\mathrm{SC3}$$

$\bar{R}^2 = .943$ F(7/51) = 140.65 D-W = 1.98 GLS RHO1 = 0.35 PERIOD: 7/77 TO 5/82

● $$\mathrm{LOG(REV.STUDENT)} = \underset{(11.07)}{4.65} + \underset{(8.02)}{1.65}*\mathrm{LOG(STUDPASS)} + \underset{(12.00)}{1.73}*\mathrm{LOG(LAGR)} - \underset{(-2.23)}{0.12}*\mathrm{DNC} + \underset{(3.22)}{0.10}*\mathrm{SS2} - \underset{(-3.25)}{0.09}*\mathrm{SS3} + \underset{(2.53)}{0.07}*\mathrm{SS4}$$
$$- \underset{(-2.94)}{0.07}*\mathrm{SS5} + \underset{(10.41)}{0.33}*\mathrm{SC1} - \underset{(-5.60)}{0.17}*\mathrm{SC2} + \underset{(7.12)}{0.20}*\mathrm{SC3} - \underset{(-5.21)}{0.14}*\mathrm{SC4}$$

$\bar{R}^2 = .900$ F(11/43) = 45.29 D-W = 1.95 GLS RHO1 = 0.15 PERIOD: 11/77 TO 5/82

LEGEND

REV.CASH:	Cash revenue
REV.REGPASS:	Regular pass revenue
REV.S+H:	Senior and handicapped pass revenue
REV.EX:	Express stamp pass revenue
REV.STUDENT:	Student pass revenue
CASHFARE:	Real cash fare
REGPASS:	Real regular pass cost
SHPASS:	Real senior and handicapped pass cost
SHFARE:	Real senior and handicapped cash fare
EXPASS:	Real express stamp pass cost
EXFARE:	Real express stamp cash cost
STUDPASS:	Real student pass cost
LAGR:	Real gasoline price
WW2:	Weekday effect
WW6:	Saturday effect
LAEMP:	Non-Agricultural employment for Los Angeles-Long Beach
D80.7:	July 1980 fare increase dummy variable
TTD:	Trend
DNC:	Dummy variable for no college study provision: August 1978 to June 1980
SS1,...,SC6:	Seasonal effects

Figure 3. Actual and estimated cash fare revenue by month.

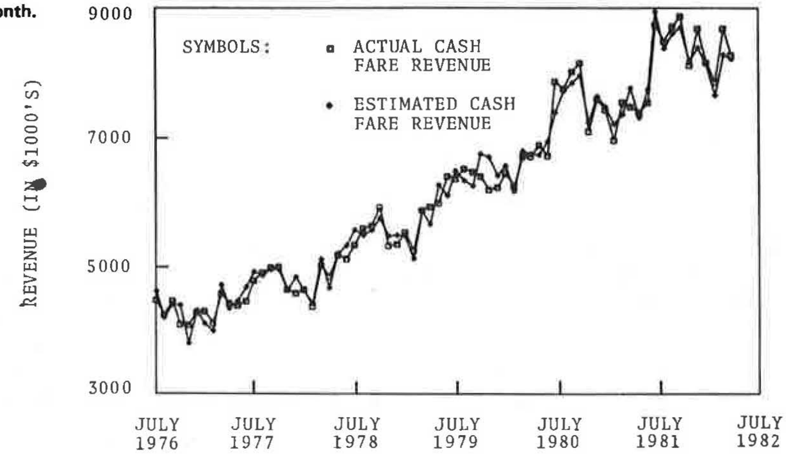
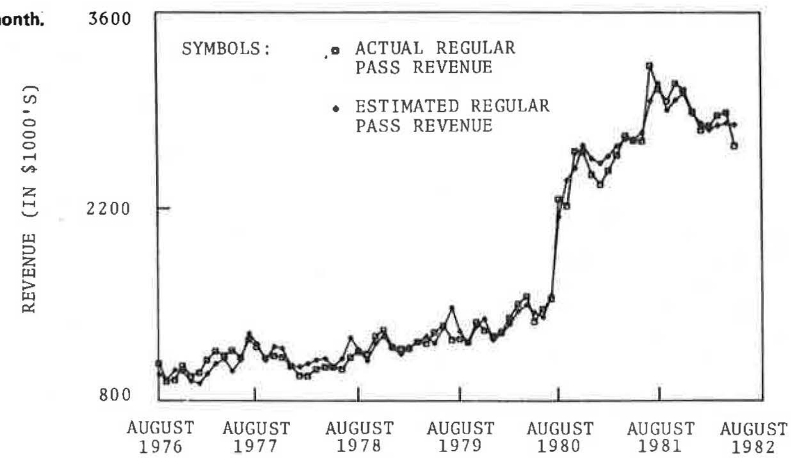


Figure 4. Actual and estimated regular pass revenue by month.



Regular Pass Revenue

The regular pass revenue equation was estimated using monthly data from August 1976 to May 1982 (see Figure 2). Figure 4 shows that the observed variation in regular pass revenue during the estimation period fits well with the values obtained from the model. The coefficients for the real price of the monthly regular pass and the real gasoline price are both positive and significant at the 1-percent level.

A dummy variable (D80.7) was included in the equation to measure the effects of the July 1980 fare change in which significant changes in the transfer provisions of cash-paying passengers made passes more attractive. The dummy variable has a positive coefficient and is significant at the 1-percent level.

Ordinary least-squares (OLS) is an appropriate estimation procedure for the regular pass revenue model even though the model includes a lagged dependent variable. The Durbin-h statistic (used instead of the Durbin-Watson statistic when the equation contains a lagged dependent variable) (9) of 0.803 indicates that the null hypothesis of no first-order autocorrelation cannot be rejected (critical value is 1.645). If such autocorrelation were present, OLS would not provide a consistent estimator and other consistent and asymptotic efficient estimation procedures would have to be used (10).

Senior Citizen and Handicapped Pass Revenue

The senior citizen and handicapped pass revenue equation was estimated using monthly data from July 1976 to May 1982 (see Figure 2). The coefficient for the real price of the senior citizen and handicapped pass is positive and significant at the 1-percent level. The coefficient of the real gasoline price is not significant even at the 10-percent level, possibly because this group is not as likely as others to have automobiles as an alternative mode of transportation. Senior and handicapped persons can pay for transit service in cash, if they desire, at a special cash fare. A variable for real senior and handicapped cash fare is included in the equation. It has a positive coefficient that is significant at the 1-percent level, indicating that revenues from senior and handicapped monthly passes increase as the senior and handicapped cash fare increases. This suggests a substitution effect between pass use and cash payment for this group.

Express Stamp Pass Revenue

The express stamp pass revenue equation was estimated using monthly data from July 1977 to May 1982, a shorter period than that of the previous models, because comparable express stamp passes were not available before July 1977 (see Figure 2).

The coefficient for the real price of the express stamp pass is positive and significant at the 10-percent level. The coefficient for the real gasoline price is also positive and significant at the 1-percent level. SCRTD passengers can pay for the express service in cash rather than purchasing the monthly pass. A variable for the real cash price of the express stamp is included in the equation (Figure 2). It has a positive coefficient and is significant at the 1-percent level, indicating that express stamp pass revenue will increase as the cash price of the express stamp increases. A substitution effect between express stamp pass use and cash payment is suggested by this relationship.

Student Pass Revenue

The student pass revenue equation was estimated using monthly data from October 1977 to May 1982, a shorter time period than that of the first four models because the student pass was not made available until October 1977 (see Figure 2).

The coefficients for the real monthly student pass price and the real gasoline price are both positive and significant at the 1-percent level. A dummy variable (DNC) was included in the equation (Figure 2) because student passes were not available to college students during the period July 1978 to June 1980. The coefficient for this dummy variable is negative and significant at the 1-percent level, indicating lower pass revenue during the period that college students could not buy student passes.

General Results

Seasonal variation is present in the data for four of the five types of revenue. Senior citizen and handicapped pass revenue is excepted because it does not exhibit much seasonal fluctuation from month to month. It should be noted that the seasonal revenue patterns are different for each of the variables; therefore, the disaggregated approach to modeling adopted in this paper is appropriate.

Before running any regressions, the dependent variables were interpolated for missing and low values due to transit strikes and service interruptions by using a seasonal time series regression model. Values for these periods with incomplete data were adjusted to reflect a complete set of data for that period. Adjusted values for cash revenue were estimated for August and September 1976 and for May, August, and September 1979. For regular pass revenue and senior citizen and handicapped pass revenue, adjusted values were estimated for September 1976 and September and October 1979. Adjusted values for express stamp revenue were estimated for September and October 1979, and student pass revenue was estimated for September 1979.

MODEL ASSESSMENT AND FORECASTING

It is well known that a high R^2 for a model estimation in the sample period does not necessarily imply that the model will predict well in the period outside the sample. For this reason, a data split technique is used to test whether the estimated model equations will be satisfactory for forecasting purposes. This is our model validation procedure.

The data are split into two parts: the first part is used for estimating the model equation. The second part, the last 6 months of data, is reserved to test the forecasting performance of the model. In this case, the values of the independent variables used to forecast revenue are known. A forecast so obtained is known as an ex-post forecast. Therefore, the forecasting errors can be attributed to the specification of the models. Following this rule, all five model equations were reestimated and the results are shown in Figure 5.

Forecasting accuracy may be evaluated by two standards: absolute accuracy and relative accuracy. The average absolute percentage error (AAPE) is used as a statistical measure of absolute accuracy in this analysis and is defined as

$$AAPE = 1/N \sum_{t=1}^{N} |(A_t - P_t)/A_t| \times 100 \quad (4)$$

where A_t and P_t are the actual values and the forecasts, respectively, in period t (where P_t are ex-post forecasts), and N is the number of forecasting periods.

It should be emphasized that comparing a forecast with actual data is the equivalent of making the comparison with an entirely error-free forecast. This is a strong criterion and may be more stringent than needed in practice.

It is unrealistic to expect a forecast to be entirely error free. Thus, a more relevant criterion is how a forecast compares with other available forecasts. To do this, Theil's U-statistic for seasonal time series was used as a standard for relative accuracy. This statistic is defined as

$$U_{12} = \sqrt{\Sigma(A_t - P_t)^2/\Sigma(A_t - A_{t-12})^2} \quad (5)$$

where A_t and P_t are the actual values and the forecasts, respectively, in period t (where P_t are ex-post forecasts), and A_{t-12} are actual values lagged 12 months; and U_{12} represents a comparison of the sum of squares of the forecast errors with the sum of squares of a seasonal random walk model.

It is interesting to observe that when the forecasts are free of error, the value of U_{12} is zero. A value of U_{12} that is less than one implies that the nonnaive model performs better than the naive, random walk model.

Table 2 gives the forecast evaluations for both the absolute and relative accuracy standards. From this table, it is concluded that these models perform quite well for forecasting purposes. Therefore, the model equations estimated in the previous section (using the complete set of data from about July 1976 to May 1982) will be used for forecasting transit revenue by type of payment for the period May or June 1982 to December 1983 under the alternative assumptions of future fare policies.

The forecasting procedure consists of two steps: selecting forecasted values for independent variables under various assumptions and substituting these forecasted variables into the models. The forecasts rest on the assumption that the basic structural relationships among the variables will remain unchanged during the forecasting period. The forecasts for the monthly Consumer Price Index and labor employment were obtained from Chase Econometrics (11) and the gasoline prices from Wharton (12). The fare structure (fare rollback) implemented by SCRTD in July 1982 was input to the model. Two sets of forecasts were generated: the first was based on the fare structure actually initiated in July 1982, and the second was based on the fare structure that prevailed during the period July 1981 to June 1982.

The forecast for student pass revenue, given that the pass price dropped from $24 to $4, some 83 percent, was not judged to be reasonable. Hence, an alternative forecasting methodology was developed. The historical average number of passes sold (student pass revenue divided by nominal pass price) per month was regressed on corresponding real gasoline prices, seasonality, and a dummy variable (DNC) to take into account a period (July 1978 to June 1980) when student passes were available only to high school students. Using the estimated parameters, forecasted values of the average number of passes sold were obtained and multiplied by nominal pass

Figure 5. Equations for model validation.

$$\bullet\ \mathrm{REV.CASH} = -\underset{(-7.02)}{8575.58} + \underset{(6.60)}{15565.30}*\mathrm{CASHFARE} + \underset{(6.04)}{8110.48}*\mathrm{LAGR} + \underset{(4.11)}{25.61}*\mathrm{WW2} - \underset{(-0.64)}{24.44}*\mathrm{WW6} + \underset{(4.92)}{2.24}*\mathrm{LAEMP} + \underset{(1.94)}{126.87}*\mathrm{SS1}$$

$$+ \underset{(1.35)}{45.49}*\mathrm{SS3} + \underset{(6.56)}{178.92}*\mathrm{SS4} - \underset{(-1.23)}{28.62}*\mathrm{SS5} - \underset{(-3.35)}{230.24}*\mathrm{SC1} + \underset{(1.26)}{33.70}*\mathrm{SC4} + \underset{(1.93)}{48.08}*\mathrm{SC5} + \underset{(2.80)}{48.46}*\mathrm{SC6}$$

$\bar{R}^2 = .899$ F(13/50) = 44.31 D-W = 2.05 GLS RHO1 = 0.62 PERIOD: 7/76 TO 10/81

$$\bullet\ \mathrm{REV.REGPASS} = -\underset{(-2.28)}{416.25} + \underset{(3.60)}{53.51}*\mathrm{REGPASS} + \underset{(4.74)}{1431.99}*\mathrm{LAGR} + \underset{(7.41)}{527.14}*\mathrm{D80.7} + \underset{(8.31)}{0.50}*\mathrm{REV.REGPASS(-1)} + \underset{(2.31)}{36.96}*\mathrm{SS1}$$

$$- \underset{(-1.54)}{23.94}*\mathrm{SS3} + \underset{(1.16)}{18.05}*\mathrm{SS4} + \underset{(3.44)}{53.60}*\mathrm{SC3} - \underset{(-2.36)}{36.99}*\mathrm{SC4}$$

$\bar{R}^2 = .983$ F(9/54) = 417.84 D-h = 0.77 OLS PERIOD: 8/76 TO 11/81

$$\bullet\ \mathrm{REV.S\&H} = -\underset{(-0.30)}{11.90} + \underset{(7.71)}{70.90}*\mathrm{SHPASS} + \underset{(1.71)}{137.61}*\mathrm{LAGR} + \underset{(5.69)}{978.46}*\mathrm{SHFARE} + \underset{(10.21)}{3.64}*\mathrm{TTD}$$

$\bar{R}^2 = .966$ F(4/60) = 448.66 D-W = 2.01 GLS RHO1 = 0.58 PERIOD: 7/76 TO 11/81

$$\bullet\ \mathrm{LOG(REV.EX)} = \underset{(12.41)}{7.34} + \underset{(1.46)}{0.27}*\mathrm{LOG(EXPASS)} + \underset{(5.67)}{1.44}*\mathrm{LOG(LAGR)} + \underset{(2.52)}{0.01}*\mathrm{TTD} + \underset{(3.20)}{0.57}*\mathrm{LOG(EXFARE)} - \underset{(-1.65)}{0.02}*\mathrm{SS3}$$

$$+ \underset{(1.45)}{0.02}*\mathrm{SS4} + \underset{(1.59)}{0.02}*\mathrm{SC3}$$

$\bar{R}^2 = .941$ F(7/45) = 119.64 D-W = 2.00 GLS RHO1 = 0.31 PERIOD: 7/77 TO 11/81

$$\bullet\ \mathrm{LOG(REV.STUDENT)} = \underset{(8.53)}{4.68} + \underset{(5.80)}{1.64}*\mathrm{LOG(STUDPASS)} + \underset{(11.05)}{1.72}*\mathrm{(LOG\ LAGR)} - \underset{(-2.08)}{0.12}*\mathrm{DNC} + \underset{(2.90)}{0.10}*\mathrm{SS2} - \underset{(-2.80)}{0.09}*\mathrm{SS3} + \underset{(2.22)}{0.07}*\mathrm{SS4}$$

$$- \underset{(-2.60)}{0.07}*\mathrm{SS5} + \underset{(8.94)}{0.33}*\mathrm{SC1} - \underset{(-4.94)}{0.17}*\mathrm{SC2} + \underset{(6.17)}{0.20}*\mathrm{SC3} - \underset{(-4.76)}{0.14}*\mathrm{SC4}$$

$\bar{R}^2 = .871$ F(11/37) = 30.57 D-W = 1.94 GLS RHO1 = 0.15 PERIOD: 11/77 TO 11/81

LEGEND

REV.CASH:	Cash revenue	STUDPASS:	Real student pass cost
REV.REGPASS:	Regular pass revenue	LAGR:	Real gasoline price
REV.S+H:	Senior and handicapped pass revenue	WW2:	Weekday effect
REV.EX:	Express stamp pass revenue	WW6:	Saturday effect
REV.STUDENT:	Student pass revenue	LAEMP:	Non-Agricultural employment for Los Angeles-Long Beach
CASHFARE:	Real cash fare		
REGPASS:	Real regular pass cost	D80.	July 1980 fare increase dummy variable
SHPASS:	Real senior and handicapped pass cost	TTD:	Trend
SHFARE:	Real senior and handicapped cash fare	DNC:	Dummy variable for no college study provision: August 1978 to June 1980
EXPASS:	Real express stamp pass cost		
EXFARE:	Real express stamp cash cost	SS1,...,SC6:	Seasonal effects

Table 2. Model validation: Values for percentage error, average absolute percentage error (AAPE), and relative errors (U_{12}).

Item	Cash Fare Revenue	Regular Pass Revenue	Senior Citizens and Handicapped Pass Revenue	Express Stamp Revenue	Student Pass Revenue
Absolute Percentage Error and AAPE					
Nov. 1981	-0.70	–	–	–	–
Dec. 1981	4.73	0.97	1.98	-4.02	-0.50
Jan. 1982	4.26	-1.3	1.40	-0.60	2.95
Feb. 1982	7.56	0.24	4.29	4.47	1.02
March 1982	11.46	3.22	4.17	8.82	10.18
April 1982	10.25	4.52	3.35	12.34	-6.28
May 1982	–	-3.87	1.10	-0.12	-4.56
AAPE	6.49	2.19	2.72	5.06	4.25
Relative Errors (time period: November 1981 to May 1982)					
Values for U_{12}	0.633	0.026	0.144	0.593	0.285

price to produce the desired revenue estimates. All of the forecasts are shown in Table 3. From this table the effects of the fare rollback on aggregate and disaggregate transit revenue can be seen.

Finally, it should be mentioned that no single forecasting method is uniformly better than all other methods at all times. Different approaches follow different philosophies to extract information from the available data. Therefore, the practical approach is to examine alternative forecasts by alternative approaches such as judgmental forecasts, time series forecasts, and causal model forecasts. Then, these forecasts can be combined into a single forecast by applying to each type of forecast weights based on different specific criteria. The flexibility of this approach lends itself to the incorporation of topical future events into the forecasts.

Table 3. Revenue forecasts with and without the July 1982 fare rollback ($1,000s).

Item	Cash Fare Revenue		Regular Pass Revenue		Senior and Handicapped Pass Revenue		Express Stamp Revenue		Student Pass Revenue	
	Without	With	Without	With	Without	With	Without	With	Without	With
July 1982	8,547	6,261	2,661	2,407	481	331	309	208	962	160
Aug. 1982	8,235	5,961	2,484	2,105	483	334	286	192	628	105
Sept. 1982	8,335	6,072	2,303	1,863	485	337	289	194	452	75
Oct. 1982	8,233	5,982	2,341	1,872	487	340	307	207	871	145
Nov. 1982	7,840	5,601	2,343	1,861	489	343	298	200	1,083	181
Dec. 1982	8,054	5,826	2,220	1,733	492	346	289	194	870	145
Jan. 1983	7,881	5,665	2,188	1,699	494	349	294	198	921	154
Feb. 1983	7,656	5,451	2,225	1,738	496	352	299	201	1,093	182
March 1983	8,273	6,080	2,270	1,784	499	355	307	206	1,099	183
April 1983	8,083	5,900	2,293	1,809	501	359	301	202	1,122	187
May 1983	8,133	5,960	2,236	1,753	504	362	292	196	1,064	177
June 1983	8,388	6,226	2,265	1,785	506	365	314	211	988	165
July 1983	8,374	6,224	2,383	1,905	509	368	323	217	947	158
Aug. 1983	8,386	6,248	2,302	1,827	511	371	300	201	618	103
Sept. 1983	8,361	6,236	2,173	1,700	513	374	306	205	449	75
Oct. 1983	8,289	6,174	2,241	1,771	516	378	328	220	873	145
Nov. 1983	7,920	5,814	2,262	1,794	518	381	320	215	1,090	182
Dec. 1983	8,087	5,991	2,153	1,687	521	384	313	211	881	147

CONCLUDING REMARKS

In this paper it has been demonstrated that an empirical regression model using the time series approach is appropriate for modeling seasonal farebox revenue and the four types of pass revenue generated in the SCRTD bus system. All empirical results confirm the a priori expectation about the empirical relationship between revenues and the real fare variables and real gasoline price variables.

The regression model approach is superior to the elasticity approach because it rigorously takes into account changes in other variables, e.g., gasoline prices, seasonal effects, overall revenue trends, substitute payments, as well as the change in the fare variables. The models provide empirical relationships among the variables which test a priori expectations. Furthermore these empirical results perform quite well for predicting purposes; and these models can provide conditional forecasts of various types of transit revenues under varying assumptions about the economy, energy conditions, and fare structure. In practice, forecasts generated from this approach should be combined with judgmental forecasts because forecasts will be more accurate if based on a wide range of information.

REFERENCES

1. J.C. Van Horne. Financial Management and Policy. Prentice-Hall, Englewood Cliffs, N.J., 1980.
2. A. Anagnostopoulous, D. Damm, T. Dooley, and W. Maling. Financial Forecasting Techniques in the Transit Industry, A Summary of Current Practice. U.S. Department of Transportation, Rept. No. UMTA-MA-06-0039-82-1, 1982.
3. D.E. Agthe and R.B. Billings. The Impact of Gasoline Price on Urban Ridership, In Annals of Regional Science, 1978, pp. 90-96.
4. H.G. Wang, D. Ward, and F. Hassler. Possible Effects of Fare Increases on New York Ridership and Revenue. Transportation Systems Center Staff Study SS-24-U3-199 and U.S. Senate Hearings before the Committee on Appropriations, Department of Transportation and Related Agencies Appropriations, Fiscal year 1983, 97th Congress, Second Session, part 2, 1981, pp. 334-358.
5. CPI Detailed Report. U.S. Department of Labor, Bureau of Statistics, Consumer Price Index, monthly.
6. Oil and Gas Journal. The Petroleum Publishing Company, Tulsa, Okla., from Vol. 74, No. 36 to Vol. 80, No. 25.
7. Employment and Earnings. U.S. Department of Labor, Bureau of Labor Statistics, from Vol. 22 No. 9 to Vol. 29 No. 11, monthly.
8. H.G. Wang. An Intervention Analysis of Interrupted Urban Transit Time Series: Two Case Studies. Proc. of Business and Economic Statistics Section, American Statistical Association, 1981, pp. 424-429.
9. J. Johnston. Econometric Methods. McGraw-Hill, New York, 1980, pp. 313.
10. H.G. Wang, M. Hidiroglou, and W.A. Fuller. Estimation of Seemingly Unrelated Regression with Lagged Dependent Variables and Autocorrelated Errors. Journal of Statistical Computation and Simulation, Vol. 10, 1980, pp. 133-146.
11. Chase Econometrics Associates, Inc. U.S. Macroeconomic Forecasts and Analysis. Table 1.2, May 24, 1982.
12. Wharton Econometric Forecasting Associates. Quarterly Model Outlook. Vol. 1, No. 2, Table 8.2, Energy, March 1982.

Publication of this paper sponsored by Committee on Public Transportation Planning and Development.