

gained through a random sample of transit riders instead of a sample of employees who use transit.

Finally, as stated at the outset, the objective of this study was to develop sound explanatory econometric models of pass-purchasing behavior. The estimation results presented here provide significant insight into the transit payment method decision. Given an adequate data set, the models developed here could be readily adapted as tools for predicting demand for alternative payment options.

Abridgment

Analysis of Revenue-Ridership Relationship of Selected RTA Carriers

DILIP R. JHAVERI

Revenue and ridership are the two most important indicators of transit system performance. Many management decisions are based on them. How reliable and accurate are these data? How does one affect the other? How do they compare among carriers? Using percentage changes in the time-series of revenue and ridership and ordinary least squares, it is shown that the approach provides a valuable tool to examine the consistency of data and compare structural relationship of revenue and ridership of carriers without regard to size, location, or other attributes. It is noted that, with one exception, all Regional Transportation Authority carriers showed marginal revenue productivity of riders constant but less than one for the study period. Six of the 12 carriers, most small ones, showed poor to very poor revenue-ridership relationship.

Revenue and ridership are the two important indicators of transit operation. The two indexes, however, may not move in the same direction or at the same rate. Strikes, accidents, change in fares or composition of riders, and faulty and inconsistent reporting of revenue or ridership may account for discrepancies.

This study addresses the question of reliability of joint ridership-revenue data of 12 carriers of the Regional Transportation Authority (RTA) in Chicago. The carriers include the Chicago Transit Authority's bus and rail operations, five commuter railroads, and five suburban bus systems. The management objectives include understanding of (a) the expected change in revenue given the change in ridership and vice versa and (b) the evaluation of current and past ridership-revenue data.

It is noted that many carriers, including those in the RTA system, report significant increase in ridership without comparable growth in revenue. Could this be because of the increasing number of discount riders, such as elderly, handicapped, and monthly-pass users, that allow for unlimited rides at a far lower rate than the single fares? Or, are our data suspect?

STUDY APPROACH AND METHODOLOGY

This paper attempts to determine the association between revenue and ridership through a study of linear relationships between sequential changes in the time-series of revenue and ridership. The suggested statistical approach provides confidence intervals for accepting or rejecting the data based on past relationships. The study also helps to compare the structural relationship of revenue and ridership

ACKNOWLEDGMENT

The work in this paper was supported as part of the service and methods demonstration project of the Urban Mass Transportation Administration, U.S. Department of Transportation.

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

of RTA carriers irrespective of size, location, fare levels, or other agency-specific characteristics. The method allows the measurement of changes in relationships over time.

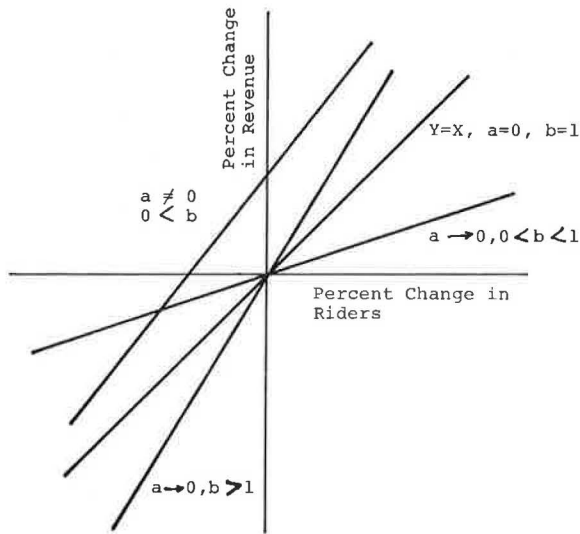
Simple revenue and ridership time-series are good descriptives of the transit systems but do not reveal their dynamic relationship. Further, seasonality and other fluctuations make them not very useful. They usually fail the statistical requirement of independence of observations. Further, because of size differences, comparison of carriers is not feasible. This is an important practical drawback that does not permit the establishment of norms against which performance of carriers may be measured.

The time-series method of percentage change in revenue and ridership overcomes these objections and provides readily interpretable criteria for inter-carrier comparisons and assessment of revenue-ridership data. Even when significant changes such as fare increase take place, the series are not affected except for the observation following the change. The study approach aids in the detection of shifts in the ridership-revenue relationship.

Regression analysis of percentage-change data has an appealing analytical and practical meaning. For instance, in the absence of subsidized and special fares, every percentage change in ridership results in an identical percentage change in revenue. Each carrier can be measured against this unit state in terms of percentage change in revenue associated with a percentage change in riders.

As illustrated in Figure 1, in unit state, the regression intercept is zero and the slope of the regression line is 1, that is, each percentage change in ridership is expected to result, on the average, in a percentage change in revenue. A consistent revenue-ridership reporting system will always suggest an intercept zero, or nearly so, while the slope may vary from zero to 1, but non-zero for all practical purposes because a zero value would indicate no relationship between revenue and ridership. The value of the slope is determined by the average fare level of new riders, if the system is growing, or riders leaving the system, in case of ridership decline, relative to the base riders and revenue. If new riders' average fare is greater than that of the base riders, the slope value ex-

Figure 1. Linear regression of percentage change in revenue (Y) and percentage change in ridership (X).



ceeds 1; if less, the slope is less than 1. Abnormal and sometimes infinite percentage changes in revenue or ridership resulting from strike, fare increase, or change in report formats are substituted by seasonally adjusted average values. For instance, if fares double, the revenue may nearly double and ridership may decline in the period following the increase. But no percentage increase in ridership or revenue may be noted in subsequent periods, all things being equal. This stability in the values of percentage changes in revenue and ridership is highly desirable for intercarrier comparative analysis of reporting performance. One must, however, allow for nonlinear revenue-ridership relationship over a short range when a system is evolving, or for kinks in the regression line as a result of sudden policy shifts.

REGRESSION ANALYSIS OF RTA CARRIER DATA

Table 1 presents important test results of regression of percentage change in revenue and ridership of 12 RTA carriers. An overall qualitative classification of carriers' data is made based on a combination of statistical test results.

Test of Parameters

All carriers indicate value of the constant a not

significantly different from zero at 95 percent confidence. For some carriers, however, this value was significantly larger than for others, such as G and F, and has affected their ratings.

The slope of regression lines of RTA carriers indicates that, on the average, percentage increase in revenue lags behind that in ridership. There is significant disparity among carriers in their expected increase in revenue from ridership growth. Carrier I had the smallest growth coefficient of 0.56, i.e., every percentage growth in riders is expected to increase, on the average, revenue by only 0.56 percent. Carrier L, an exception, has percentage revenue-growth rate exceeding ridership increase. Each percentage growth in ridership of L increases revenue by 1.056 percent, a possible indication that increasingly more new riders on L are full-fare passengers.

Test of Regression Models

The validity of regression models is largely determined by the test of residuals, i.e., the proportion of sum-of-squares of deviations around the mean unexplained, the mean-square error (MSE), and autocorrelation.

For four of the carriers, B,G,H, and I, the models explained less than one-half the sum of squares suggesting relative deviation from the regression line. For this reason, the rating of these carriers was adversely affected. For carrier B, however, the residuals are quite small in absolute value as suggested by the MSE. This helped its rating. For carriers with a higher rating, the R² values were mostly greater than 0.8.

Figure 2 presents the distributions of percentage changes in revenue and ridership of two carriers, D and G. Discounting the scale differences, it is apparent that the observations in D are significantly closer to the regression line than are in G. This explains the high R² value for D, 0.814, and low for G, 0.40. The MSE for G is 97.14 compared with only 2.00 for D. These and other characteristics of the data and regression tests account for the "excellent" rating for D but "very poor" for G.

Figure 3 contrasts the distributions of residuals of carriers D and G. Discounting the scale differences, it is observed that the residuals of D are more random and significantly smaller than that of G. Residuals of G also tend to increase with changes in ridership that indicate significant variance in the period-to-period growth rates of revenue and ridership of G. Examination of data-gathering and reporting practices of carrier G confirms our evaluation that the noticeable discrepancies in growth rates of G are primarily the

Table 1. Summary of least-square regression of percentage change in revenue and ridership of selected RTA carriers.

Carriers	Intercept		Parameter		R-Square	Mean Square Error	Durbin Watson-D	Percentage of Residuals Within ± 1%	Percentage of Residuals Outside ± 2%	Autocorrelations		Assessment of Revenue-Ridership Data
	A	t-Statistic	B	t-Statistic						1st Order	Annual	
A	0.5040	1.14	0.8177	7.43	0.7978	2.5573	3.0535	62.50	18.75	-0.6897	0.1273	Acceptable
B	0.2977	0.70	0.6543	3.74	0.4237	3.4034	2.1301	28.60	33.33	-0.1186	0.4807	Acceptable
C	0.4704	1.69	0.7292	4.10	0.5834	6.1729	1.4748	35.71	50.00	0.0768	0.0182	Poor
D	0.5763	1.66	0.7668	9.12	0.8140	2.0024	2.2599	47.62	4.76	-0.1766	0.0292	Excellent
E	0.4379	0.52	0.6572	4.40	0.5637	11.4554	2.4101	35.29	58.82	-0.2210	0.2332	Poor
F	-0.8884	-0.73	0.6010	9.06	0.6891	57.9349	2.1327	10.25	82.05	-0.1667	0.1205	Poor
G	1.4527	0.88	0.6881	4.83	0.4003	97.1385	2.6468	13.51	72.97	-0.3324	0.3181	Very poor
H	0.2532	0.25	0.7123	5.0	0.4100	37.7082	2.6184	13.16	71.05	-0.3205	-0.1221	Very poor
I	0.6563	0.27	0.5563	5.53	0.4593	216.3400	2.7366	5.25	81.58	-0.3761	-0.3074	Very poor
J	-0.2531	-0.39	0.8431	14.07	0.8389	18.7447	2.4574	30.00	62.50	-0.2929	0.1805	Acceptable
K	-0.1366	-0.52	0.9308	26.74	0.8584	8.3742	2.5576	47.56	24.39	-0.1505	-0.0356	Good
L	-0.0711	-0.11	1.0562	13.62	0.6988	31.2548	2.2999	65.52	20.69	-0.2819	0.2119	Good

Note: Data sets modified to remove known abnormal variations because of strike, fare increase, etc., and some extreme observations.

Figure 2. Percentage change in revenue and riders, carriers D and G.

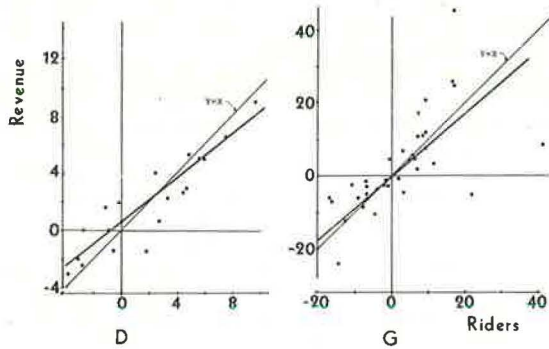
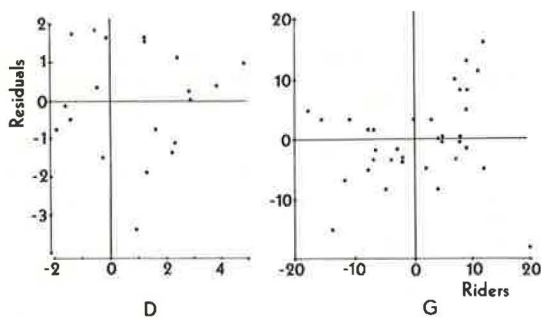


Figure 3. Plot of residuals and percentage change in riders, carriers D and G.



result of poor data quality.

The MSEs are valuable indicators of the average size of residuals even when the model may not explain a large proportion of the square deviations. A small size of deviations indicates a close grouping of observations around the regression line and small estimation errors. RTA carriers with "poor" or worse ratings exhibited MSEs in excess of 35. Carrier B with rather poor R^2 of only 0.42 but an MSE of only 3.40 was considered "acceptable" for its small deviations. Sources of its small but unexplained deviations may be found in its relatively minor error in data-gathering and reporting practices. Because MSE may be affected by a few large residuals, a visual check of residuals was made and

extreme residuals were removed when justified.

Test of Seasonality

Because seasonal, systematic patterns of changes in the data could affect regression equations, all models were examined for autocorrelation in residuals, that is, seasonality in errors after linear trend is taken into account. Only carriers B and I showed statistically significant annual seasonality. Seasonal parameters for these carriers, however, failed to decrease MSE or overall evaluation of their data.

Assessment of Data Quality

No single index of data quality has been suggested since each carrier must be considered in its own unique environment. Relative values of various statistics for the carriers must be compared with caution. The MSE is perhaps the single most important index to watch in any relative test, mindful that a single large residual may affect this value.

CONCLUSION

This study demonstrated an application of ordinary linear regression analysis to evaluate the consistency and reliability of revenue and ridership data of 12 RTA-Chicago carriers. Using percentage changes over sequential periods in revenue and ridership, the simple linear models provide a useful mechanism to examine the quality of data, compare carriers in their revenue-ridership relationships irrespective of their size or location, and enable the tracking of changes in the composition of riders that affects revenue.

ACKNOWLEDGMENT

This paper was prepared as part of my work assignment at the request of L. Michael Fultz, Treasurer, Northeastern Illinois Regional Transportation Authority. The original draft of this paper was prepared in December 1979.

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

Abridgment

Bus Costing Information in Short-Range Planning: Survey of Principles and Practice

MICHAEL A. KEMP, MICHAEL E. BEESLEY, AND ROBERT G. MCGILLIVRAY

This paper discusses the major principles involved in the use of bus transit cost information for planning and policymaking purposes. It is argued that there is no such thing as comprehensive costing, which is immediately and uniformly applicable to all kinds of decisions. Rather, the types and treatments of the costs that should properly be considered vary with the nature of the decision being contemplated. Particular emphasis is placed on the relevance, structure, and valuation of cost items. Three major categories of decisions that require different approaches are considered. These are characterized as service changes,

innovation, and the allocation of deficits and subsidies. The paper also provides a brief critical review of currently available procedures for employing cost information in short-range transit planning. A direct estimation of costs may be made by planning proposed service changes in full operational detail, but this is cumbersome and expensive. Short-cut techniques include computerized operational research models (so far, they are not well developed), simple average cost procedures, and more sophisticated causal factor allocation methods. Enhancements to the widely used causal factor approach attempt to take better