Temporal Analysis of Handicapped Ridership in Specialized Transportation Service: Lexington/ Fayette County Experience

MANOUCHEHR VAZIRI

This paper focuses on modeling of Kentucky's Lexington/Fayette County specialized transportation (WHEELS) monthly ridership. The 1979 through 1985 time-series data suggest an intervention model to replicate monthly ridership. The identified model successfully incorporated the lag structure and functional forms that constitute the relationships between monthly ridership and service changes, such as service area expansion and fare increase. The selected model satisfies all estimation and diagnostic requirements. Model predictions for 1985 were quite reasonable when compared with actual ridership: cyclical patterns were correctly replicated. The supremacy of intervention modeling when compared with multiple linear regression analysis was found to be in capturing ridership seasonality, properly reflecting the impact of changes in service attributes, and displaying uncorrelated residuals.

Specialized transportation services are often provided to persons who do not have the physical or mental ability to use alternative means of transportation. In the last two decades, specialized transportation services have seen a tremendous growth, mostly in the form of paratransit services (*1*,*2*). The development of paratransit systems has been accompanied by the development of mathematical modeling for better planning and management, particularly in routing and scheduling. These models are often designed to determine the delicate balance between supply, demand, and cost of a paratransit system (*2*,*3*). The demand is considered as a whole range of levels of ridership that would eventuate from a variety of different fare levels, different area coverage levels, different times of operation, different months and years, different policies of passenger eligibility, and so forth. The methods used for measuring and forecasting the demand for specialized transportation systems have not adequately taken into account all its major dimensions (*4–6*). Nonetheless, in recent years, implementation of microcomputer and data base management software has alleviated many previous problems of trip information gathering and recording (*7*,*8*).

In planning and management of specialized transportation services, major characteristics of ridership that should be considered include spatial and temporal variations. Information about time variation of ridership is essential to determine the level of service that is most appropriate for different points in time. Most of the existing demand models predict trip density, trips per square mile per day or hour (*3*). Although these models are useful in ridership forecasting because of structural changes in users' condition and density, they become problematic and insensitive when generating short-run predictions. A class of models proven to be particularly well suited to short-term forecasting is that often referred to as ARIMA, autoregressive integrated moving average (*9*,*10*). These models replicate past behavior of a univariate time-series rather than determine direct multivariate structural relationships. Such models are particularly useful for short-term forecasting when it is expected that underlying factors determining the level of the variable of interest in the past, herein specialized transportation monthly ridership, will behave the same in the near future. An extension of univariate ARIMA models to the multivariate domain can be presented by intervention modeling. This type of model is specially structured to deal with intervening events affecting the time-series process, herein changes in fare, fleet size, and coverage area affecting ridership. In recent years, there have been several published works related to time-series demand modeling for regular transit systems (*11–17*). However, its application has not been sufficiently addressed in paratransit and specialized transportation ridership modeling.

This paper presents an intervention model for modeling and forecasting of Lexington/Fayette County (Kentucky) specialized transportation ridership. Managers and planners of specialized transportation can use the methodology and findings of this study to enhance ridership forecasts and assess the impact of service policy changes.

LEXINGTON/FAYETTE COUNTY SPECIALIZED TRANSPORTATION

The Lexington/Fayette County specialized transportation service, WHEELS, was established in 1978. WHEELS is designed to meet the needs of handicapped persons by overcoming the lack of economical and accessible transportation. The disability of individuals using WHEELS must be documented and a person must fill out an application and be registered in order to become eligible for service. Trip reservations, usually by telephone, are made at least 1 day in advance.

Department of Civil Engineering, University of Kentucky, Lexington, Ky. 40506-0046. Current address: 889 East High Street, Lexington, Ky. 40502.

WHEELS provides service Monday through Friday from 7 a.m. to 6 p.m. and Saturday from 10 a.m. to 4 p.m. Since January 1979, when WHEELS began operation, there have been several modifications of this service. Among these, there were four major events that could be predicted to most profoundly impact system ridership:

- In January 1981, coverage (service) area was expanded from a pilot area in the north end to the whole urban area; simultaneously, fleet size was increased from four to eight vehicles.
- In September 1983, Saturday service was initiated.
- In July 1984, fare was increased from \$0.50 to \$0.75.
- In July 1985, fare was increased from \$0.75 to \$1.00.

Monthly ridership during the study period is shown in Figure 1. Figure 1 suggests a general secular increase after expansion of the service area in 1981, seasonal variation involving periodicity over 12-month cycles with a minimum most often occurring during midwinter, and the possible negative impacts of 1984 and 1985 fare increases. The combination of distinct seasonality, secular increase, and four intervening events suggests that the time-series of monthly ridership is a good candidate for intervention modeling.

MODEL STRUCTURE

The intervention model consisted of a mathematical relationship known as a transfer function, which expresses the degree to which intervening events affect the time-series. The model and the method for assessing its parameters as presented in the following section are known as Box-Tiao Intervention Analysis (*19*). This technique is a generalization of the multiple linear regression model with k independent variables:

$$Y_t = b_0 + \sum_{i=1}^{k} b_i X_{it} + e_t \tag{1}$$

where Y_t is the dependent variable at time t, b_0 is a constant, b_i is the coefficient of the ith independent variable X_{it}, k is the number of independent variables, and e_t is the error term. The basic assumption of Equation 1 is that covariance $(e_t, e_{t'}) = 0$ for $t \neq t'$. This presents a serious constraint for application to monthly ridership because of factors such as seasonality. Such a problem does not exist in intervention modeling. The intervention model applied to the 1979 through 1985 monthly time-series for specialized transportation ridership had the following general functional form:

$$Y_t = \sum_{i=1}^{k} \frac{\omega_i(B)}{\delta_i(B)} X_{it} + N_t \tag{2}$$

where Y_t is the time-series dependent variable; B is a backshift operator pertinent to the time index of variables, that is, $BY_t = Y_{t-1}$ and $B^2Y_t = Y_{t-2}$; ω_i and δ_i are polynomial operators for ith intervention variable, that is, $\omega_i(B) = \omega_{0i} - \omega_{1i}B - \ldots - \omega_{gi}B^{gi}$ for polynomial operator of order gi, where $\omega_{0i}, \ldots, \omega_{gi}$ are coefficients and $\delta_i(B) = 1 - \delta_{1i}B - \ldots - \delta_{hi}B^{hi}$ for polynomial operator of order hi, where $\delta_{1i}, \ldots, \delta_{hi}$ are coefficients; X_{it} is the ith intervention variable and $X_{it} = 1$ for month t, wherein the ith intervening event is taking place and $X_{it} = 0$ otherwise; and N_t is the noise that can be presented by an ARIMA process such as $N_t = (\theta(B)/\phi(B))a_t$, where θ and ϕ are polynomial operators and a_t is the white noise variable for month t, independent and normally distributed with mean of zero and variance of σ^2. The advan-

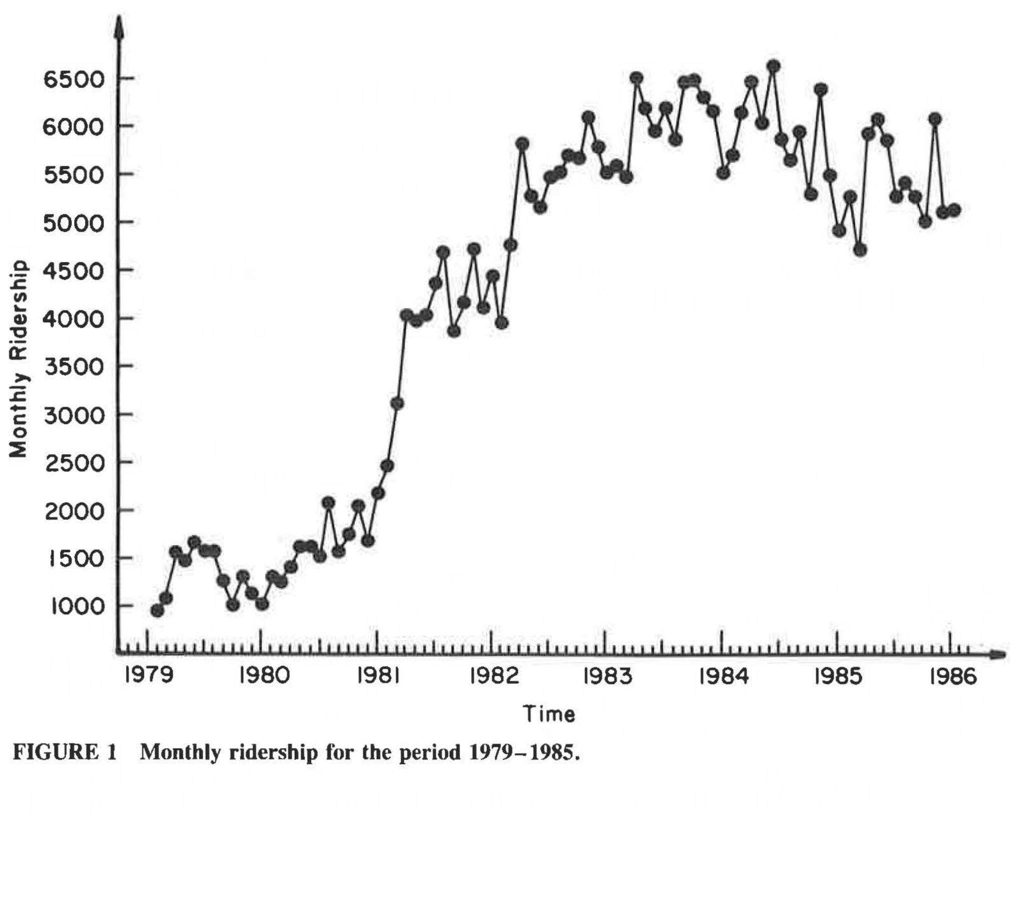

FIGURE 1 Monthly ridership for the period 1979–1985.

tage of Equation 2 is that it allows estimates of Y_t to reflect prior levels of Y, prior levels of X_i's, and prior levels of white noise. This is not feasible for Equation 1 when Y_t is only dependent on current levels of the X_i's.

MODEL BUILDING

The selection of a model for any time-series data from the family of intervention models presented by Equation 2 is in large part a matter of judgment. Nevertheless, a generally accepted model building strategy includes iterative identification, estimation, and diagnosis stages (*9,10,20*). The identification stage is often accomplished on the basis of (a) prior knowledge of the data pattern, (b) evaluation of the plotted time-series, (c) evaluation of the sample autocorrelation coefficients (ACFs), and (d) evaluation of the sample partial autocorrelation coefficients (PACFs). The identification often starts with initial noise modeling, the ARIMA modeling for N_t, based on (a) the portion of the data containing no unusual events or (b) all the data by using robust estimation to reduce the effect of unusual events. Once components of the ARIMA model are identified, the information is used to identify the transfer function components. Once a tentative model is identified, its parameters are estimated and tested for statistical significance. In addition, parameter estimates must meet stationarity-invertability requirements (*9,10,20*). If either criterion is not met, a new model must be identified and its parameters estimated and tested. After successful estimation and testing, the model is finally diagnosed. To pass diagnosis, the autocorrelation of the residuals (RACFs) from the estimated model should be sufficiently small and should resemble white noise. If the residuals remain significantly correlated among themselves, a new model should be identified.

After several trials, following basically the aforesaid stages of modeling and using the Time Series Program of the Biomedical Package (School of Public Health, UCLA), the selected intervention model with the smallest residual mean square (RMS) was found to have the following form (*21*):

$$Y_t = \frac{\omega_{01}}{1 - \delta_{11}B} X_{1t} + \omega_{02}X_{2t} + \omega_{03}X_{3t} + \omega_{04}X_{4t} + \frac{(1 - \theta_1 B - \theta_4 B^4 - \theta_5 B^5)(1 - \theta_{12}B^{12})}{(1 - B)(1 - B^{12})} a_t \quad (3)$$

where Y_t is monthly ridership for month t; X_{1t} is a dummy variable reflecting the intervening event of fleet and service area expansion, $X_{1t} = 0$ for months of 1979 through 1980 and $X_{1t} = 1$ thereafter; X_{2t} is a dummy variable reflecting the intervening event of Saturday service, $X_{2t} = 0$ for months before September 1983 and $X_{2t} = 1$ thereafter; X_{3t} is a dummy variable reflecting the fare increase from \$0.50 to \$0.75, $X_{3t} = 0$ for months before July 1984 and $X_{3t} = 1$ thereafter; X_{4t} is a dummy variable reflecting the intervening event of fare increase from \$0.75 to \$1.00, $X_{4t} = 0$ for months before July 1985 and $X_{4t} = 1$ thereafter; a_t is the white noise variable for month t, independent and normally distributed with mean of zero and variance of σ^2; ω_{01}, δ_{11}, ω_{02}, ω_{03}, ω_{04}, θ_1, θ_4, θ_5, and θ_{12} are parameters and B is the backshift operator.

Based on Equation 3, the estimated intervention model for 1979 through 1985 monthly ridership has the following form:

$$Y_t = \frac{601.8}{1 - 0.712B} X_{1t} + 82.23X_{2t} - 512.9X_{3t} - 152.5X_{4t} + \frac{(1 - 0.584B - 0.290B^4 + 0.506B^5)(1 - 0.811B^{12})}{(1 - B)(1 - B^{12})} a_t \quad (4)$$

where notations are the same as in Equation 3, t statistics for parameter estimates are all greater than 2 except 0.35 for X_{2t} and 0.63 for X_{4t}, and RMS is 97,579. The autocorrelations of the residuals—shown in Figure 2—are inside the range of 95 percent confidence interval and therefore not significant. To check whether the entire residual autocorrelation is different from what could be expected of white noise, the Portmanteau test was performed (*9*). Following is a summary of the test results.

K	*Q*	*Degree of Freedom*	*Level of Significance*
6	4.33	2	0.123
12	8.20	8	0.425
18	15.31	14	0.370
24	19.98	20	0.464

The Q statistic is the sum of the first K residual autocorrelations multiplied by the number of observations minus the maximum back order for the period of time-series study. The Q values are distributed approximately chi-square with the degree of freedom equal to K minus the number of estimated parameters. The data show that the results are not significant at the 0.05 level. For Equation 4, the roots of $\theta(B)$ lie outside and those of $\phi(B)$ lie on the unit circle, thus meeting stationarity and invertability requirements (*9*).

The parameter estimates for X_{1t} suggest that monthly ridership increased by roughly 2,090 because of the service expansion of January 1981. Nevertheless, the response was not immediate but rather a first-order dynamic response like that in Figure 3. The parameter estimate for X_{2t} suggests that Saturday service increased monthly ridership by roughly 82. The parameter estimate for X_{3t} suggests that the fare increase

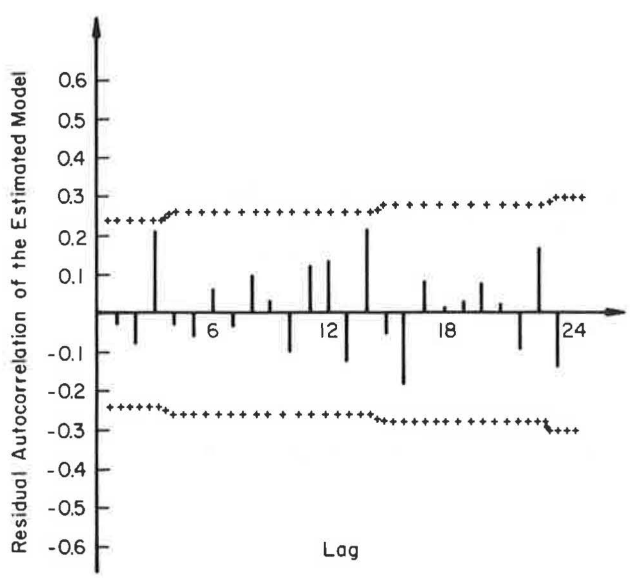

FIGURE 2 Residual autocorrelation function of the estimated intervention model.

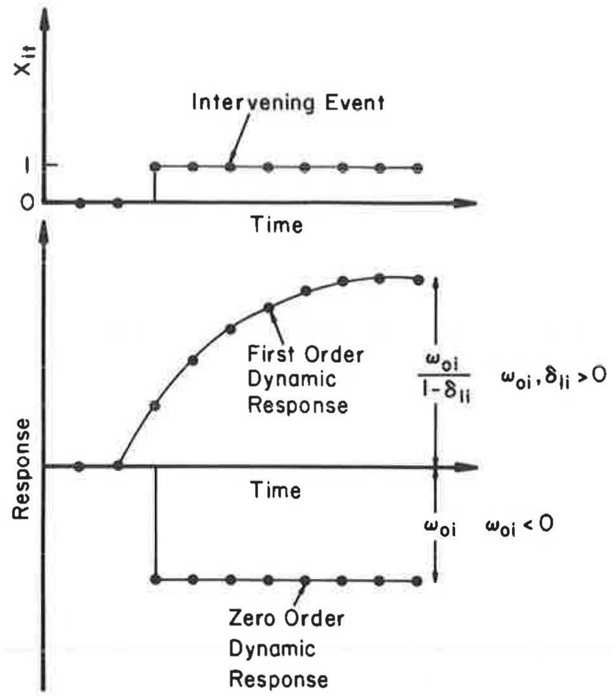

FIGURE 3 Response to a step-intervening event.

from \$0.50 to \$0.75 resulted in a monthly ridership decrease of roughly 513. The parameter estimate for X_{4t} suggests that the fare increase from \$0.75 to \$1.00 resulted in a monthly ridership decline of roughly 152. The intervention events, X_{2t}, X_{3t}, and X_{4t} produced immediate responses, the zero order dynamic response type of Figure 3 (*19*). In view of statistical significance at a level of 0.05, X_{2t} and X_{4t} should be excluded from the model because their parameter estimates have t statistics smaller than 2. Exclusion of X_{2t} and X_{4t} from the model resulted in the following intervention model:

$$Y_t = \frac{582.6}{1 - 0.725B} X_{1t} - 464.1X_{3t} + \frac{(1 - 0.563B - 0.296B^4 + 0.514B^5)(1 - 0.806B^{12})}{(1 - B)(1 - B^{12})} a_t \qquad (5)$$

where the notations are the same as in Equation 4, t statistics for parameter estimates are greater than 2, and RMS is 96,018. Equation 5 meets stationarity and invertability requirements and passes all checks of diagnosis. The parameter estimates for X_{1t} suggest that monthly ridership increased by roughly 2,119 because of the service expansion of January 1981. The parameter estimate for X_{3t} suggests that the fare increase from 50 cents to 75 cents in July 1984 resulted in a monthly ridership decrease of roughly 464.

MODEL EVALUATION AND PREDICTION

To demonstrate the advantage of intervention modeling, the same time-series data were used to calibrate two regression models. The first is the simpler version that assumes time as the only independent variable:

$$Y_t = 1392.1 + 67.073t + e_t \qquad (6)$$

where Y_t is monthly ridership for month t ($t = 1, \ldots, 84$). The coefficient of t has a t statistic greater than 2. The RMS of Equation 6 is 991,151, which is 10 times larger than the RMS of Equations 4 and 5. The parameter estimate for time variable t suggests that monthly ridership increased by roughly 67 per month. Introduction of intervention variables as further independent variables resulted in the second regression model:

$$Y_t = 589.4 + 71.874t + 1525.8X_{1t} - 414.3X_{2t} - 1354.7X_{3t} - 881.4X_{4t} + e_t \qquad (7)$$

where the notations are as defined before and the t statistics for parameter estimates are all greater than 2 except 1.8 for X_{2t}. The RMS of Equation 7 is 222,479, which is 2.3 times larger than the RMS of Equation 5. Although Equation 7 is superior to Equation 6 because of its smaller RMS, the negative coefficient of X_{2t}, introduction of Saturday service, is not logical. One should expect an increase in total monthly ridership as a result of Saturday service, as Equation 4 correctly predicted. Nevertheless, the parameter estimate for the time variable suggests a monthly ridership increase of roughly 72 per month. The parameter estimate for X_{1t} suggests that monthly ridership increased by roughly 1,526 because of service expansion. The parameter estimate for X_{2t} suggests that the Saturday service decreased monthly ridership by roughly 414. The parameter estimate for X_{3t} suggests that the fare increase from \$0.50 to \$0.75 resulted in a monthly ridership decrease of roughly 1,355. The parameter estimate for X_{4t} suggests that the fare increase from \$0.75 to \$1.00 resulted in a monthly ridership decrease of roughly 881. In view of statistical significance at the 0.05 level, X_{2t} should be excluded from the model. Exclusion of X_{2t} from the model resulted in the following equation:

$$Y_t = 700.2 + 63.014t + 1719.6X_{1t} - 1431.1X_{3t} - 801.7X_{4t} + e_t \qquad (8)$$

where the notations are the same as in Equation 7, t statistics for parameter estimates are greater than 2, and RMS is 228,780. The parameter estimate for the time variable suggests that monthly ridership increased by roughly 63. The parameter estimate for X_{1t} suggests that monthly ridership increased by roughly 1,720 because of service expansion. The parameter estimate for X_{3t} suggests that monthly ridership decreased by roughly 1,431 because of the fare increase from \$0.50 to \$0.75. The parameter estimate for X_{4t} suggests that the fare increase from \$0.75 to \$1.00 resulted in a monthly ridership decline of roughly 802.

The major drawback of Equations 6, 7, and 8 is the assumption of residual independency. Indeed, residual autocorrelations of Equations 6, 7, and 8 showed several statistical significances, especially for Lag 1 and Lag 12. However, such a problem does not exist for the intervention models, Equations 4 and 5.

The calibrated regression and intervention models were used to predict the monthly ridership of WHEELS for the 12-month period beginning in January 1985. Figure 4 presents the 12-month predictions for Equations 5, 6, and 8. It also shows the actual values and 95 percent confidence interval for the predicted values from the intervention model, Equation 5. The intervention model extends the seasonality throughout 1985, whereas the regression models are insen-

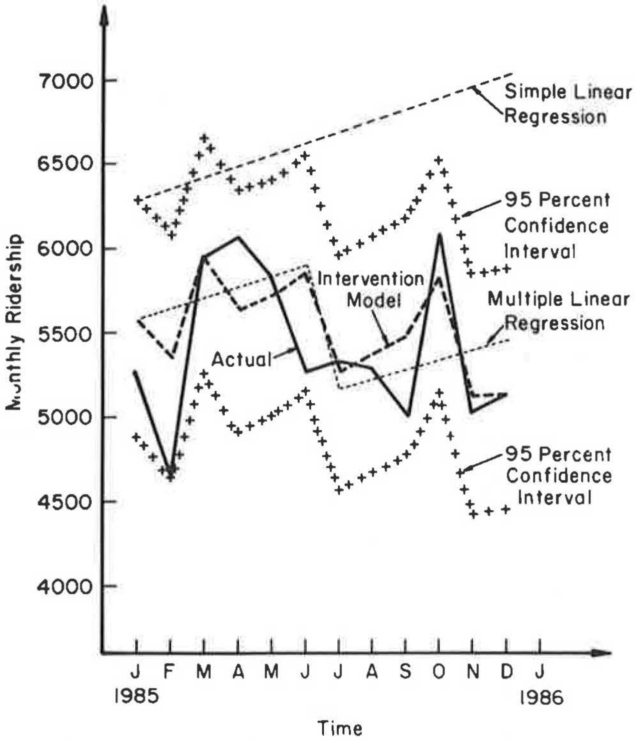

FIGURE 4 Comparison of actual and predicted monthly ridership.

sitive to such seasonal behavior. Parameter estimates for the time variable t from the developed regression models are positive. It is noteworthy that in the least square method of regression modeling with time as a monotonically increasing independent variable, the first and the last observations of the time-series usually make the greatest contribution to the sum of squares (*20*). Thus, the coefficients are derived so that the trend line passes close to the first and last data points. This would suggest that predictions by such regression models are inferior for mid-periods. Because of the statistical significance of the coefficient of X_{4t}, Equation 8 has apparently reflected the impact of the second fare increase. However, this to a great extent could have been because of the aforesaid characteristic of the least square method. This has not been the case for Equation 5, which was developed excluding X_{4t}. Availability of 1986 data for the time-series will eventually allow clarification of the assumption with respect to X_{4t} for the intervention model. Implementation of the second fare increase was in July 1985 and the impact was reflected in only six monthly ridership data points. This is usually considered to be an insufficient number of data points for an intervention model to show the impact of an intervening event. Indeed, it would have been desirable to develop models for the 1979 through 1984 period, and then for model evaluation, predict 1985 monthly ridership. Unfortunately, because of the second fare increase in July 1985, this was an unsuitable basis for evaluation.

Ridership predictions for 1985 are summarized in Table 1. The superiority of the intervention model is evident because of smaller RMS and more accurate prediction of both average monthly ridership and yearly ridership.

CONCLUSIONS

Intervention modeling applied herein to time-series monthly ridership was based on Box-Tiao Intervention Analysis. The applied intervention modeling for Lexington/Fayette County specialized transportation (WHEELS) monthly ridership consisted of iterative stages of identification, estimation, and diagnosis. The selected model for 1979 through 1985 time-series data showed that monthly ridership has a seasonality of 12 months and depends on the past month's ridership as well as Lag 1, Lag 4, Lag 5, and Lag 12 white noises. Furthermore, the monthly ridership was found to be affected by changes in service attributes, such as fare increase and service expansion. With 0.05 as a level of significance criterion for parameter estimates in the intervention model, it was found that service expansion in January 1981 resulted in a monthly ridership increase of roughly 2,119 and the fare increase from 50 cents to 75 cents in July 1984 resulted in a monthly ridership decline of roughly 464.

For specialized transportation service of the type analyzed, intervention modeling is more appropriate and powerful than traditional multiple linear regression in evaluating and predicting time-series data with intervening events. Although requiring somewhat more historical data points, intervention analysis successfully treated time lag structure and interrelations of the time-series data. The superiority of intervention modeling lies in the ability to capture seasonality in the time-series and properly reflect the impact of changes in service attributes. Unlike traditional multiple linear regression models, the residual autocorrelations of the estimated intervention models were found to be uncorrelated. Although the selected intervention model is dependent on 1979 through 1985 WHEELS ridership data, the same methodology can be applied to study any specialized transportation or paratransit system with time-series nonstationary behavior characteristics that have been affected by service policy changes or other intervening events.

TABLE 1 COMPARISON OF PREDICTED AND ACTUAL 1985 RIDERSHIP

		Regression		
Variable	Actual	Simple	Multiple	Intervention Model
Monthly avg.	5,415	6,657	5,535	5,524
Yearly total	64,981	79,883	66,408	66,291
RMS	0	1,809,664	208,925	111,450

ACKNOWLEDGMENT

The author gratefully acknowledges the cooperation by WHEELS personnel, especially N. Barker, in provision of data.

REFERENCES

1. W. G. Bell and J. S. Revis. *Transportation for Older Americans, Issues and Options for the Decade of the 1980s.* Report DOT-I-83-42. Office of Technology and Planning Assistance, U.S. Department of Transportation, 1983.
2. J. W. Billheimer, G. R. Lucas, and R. W. Wilmuth. Review and Assessment of Paratransit Models. In *Transportation Research Record 724*, TRB, National Research Council, Washington, D.C., 1979, pp. 29–35.
3. S. Kikuchi. Dispatching of Specialized Transportation Vehicles: Problems and Methods. Presented at 10th National Conference on Specialized Transportation, Florida State University, Sarasota, 1985.
4. *NCHRP Synthesis of Highway Practice 39: Transportation Requirements for the Handicapped, Elderly, and Economically Disadvantaged.* TRB, National Research Council, Washington, D.C., 1976, 54 pp.
5. *NCHRP Synthesis of Highway Practice 209: Market Opportunity Analysis for Short-Range Public Transportation Planning, Transportation Services for the Transportation Disadvantaged.* TRB, National Research Council, Washington, D.C., 1979, 52 pp.
6. *NCHRP Synthesis of Highway Practice 262: Planning Transportation Services for Handicapped Persons, User's Guide.* TRB, National Research Council, Washington, D.C., 1983, 74 pp.
7. D. Dornan and D. Middendorf. *Planning Services for Transportation Handicapped People, Data Collection Manual.* Report DOT-I-83-40R. Technology Sharing Program, U.S. Department of Transportation, 1985.
8. N. Hamer. Implementation of Computer Systems for Specialized Transit: Three Case Studies. Presented at 10th National Conference on Specialized Transportation, Florida State University, Sarasota, 1985.
9. G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control.* Holden-Day, Inc., Oakland, Calif., 1976.
10. J. P. Cleary and H. Levenbach. *The Professional Forecaster: The Forecasting Process Through Data Analysis.* Lifetime Learning Publications, Belmont, Calif., 1982.
11. M. Gaudry. An Aggregate Time Series Analysis of Urban Transit Demand: The Montreal Case. *Transportation Research*, Vol. 9, 1975, pp. 249–258.
12. M. Gaudry. Seemingly Unrelated Static and Dynamic Urban Travel Demand Models. *Transportation Research*, Vol. 12, 1978, pp. 195–212.
13. M. Kemp. *Bus Service in San Diego: A Study of Patronage Growth in the Mid-1970s.* Working Paper 1470-1. The Urban Institute, Washington, D.C., 1981.
14. M. Kemp. *A Simultaneous Equations Analysis of Route Demand and Supply, and Its Application to the San Diego Bus System.* Working Paper 1470-2. The Urban Institute, Washington, D.C., 1981.
15. M. Kyte, J. Stoner, and J. Cryer. Development and Application of Time-Series Transit Ridership Models for Portland, Oregon. In *Transportation Research Record 1036*, TRB, National Research Council, Washington, D.C., 1985, pp. 9–18.
16. G. H. Wang. An Intervention Analysis of Interrupted Urban Transit Time Series Data: Two Case Studies. *Proc., American Statistical Association, Business and Economic Section*, Washington, D.C., 1981, pp. 424–429.
17. G. H. Wang, W. Maling, and D. Skinner. *Modeling Monthly Ridership for Seven U.S. Transit Authorities.* Transportation System Center, U.S. Department of Transportation, Cambridge, Mass., 1982.
18. N. Barker. *WHEELS Monthly Report.* WHEELS, Lexington, Ky., Jan. 1979 through Dec. 1985.
19. G. E. P. Box and G. C. Tiao. Intervention Analysis with Application to Economic and Environmental Problems. *Journal of the American Statistical Association*, Vol. 70, 1975, pp. 70–79.
20. R. McCleary and R. A. Hay. *Applied Time Series Analysis for the Social Sciences.* Sage Publications, Inc., Newbury Park, Calif., 1980.
21. *BMDP Statistical Software.* (W. J. Dixon, ed.) University of California Press, Berkeley, 1983.

Publication of this paper sponsored by Committee on Transportation for the Transportation Disadvantaged.