U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-265 326

An Econometric Analysis of Public Transportation Planning Norms

Peat, Marwick, Mitchell and Co, Washington, D C

Prepared for

Office of the Assistant Secretary for Policy, Plans and International Affairs (DOT), Washington, D C

Jan 75

TECHNICAL MEMORANDUM

An Econometric Analysis of Public Transportation Planning Norms

prepared for
Office of the Secretary
U.S. Department of Transportation

prepared by
Peat, Marwick, Mitchell & Co.

January 1975

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-TPI-77-10-05		

4. Title and Subtitle	5. Report Date
An Econometric Analysis of Public Transportation Planning Norms	January 1975
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
9. Performing Organization Name and Address Peat, Marwick, Mitchell and Co. 1025 Connecticut Avenue, N.W. Washington, D. C. 20036	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DOT-OS-40064
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Secretary 400 Seventh Street, S.W. Washington, D. C. 20590	13. Type of Report and Period Covered Technical Memorandum
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The purpose of this working paper is to extend the analysis of urbanized areas to explore the hypothesis that it is possible to develop some preliminary public transportation planning norms from the NTS data. The analysis proceeds on three sections. The first section provides a description of the procedures and results of the econometric analysis. The second section uses the econometric results in an analysis of transit planning norms. The first part of this section discusses and interprets the econometric analysis, and the second part uses the econometric analysis to identify urbanized areas which are exceptions to the typical values for the group of urbanized areas being analyzed. The final section suggests some conclusion about the possibilities for using the 1974 NTS data to develop norms for evaluating public transportation plans.

17. Key Words	18. Distribution Statement
econometric; urbanized areas; physical, financial, and operating characteristics; similarity of urbanized areas	For sale to the public National Technical Information Service U.S. Department of Commerce Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
		59	PC A04-MF A01

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

In Section II of the Appendix to the 1974 National Transportation Report (1974 NTR) selected individual urbanized area public transportation statistics were examined to determine whether transit systems serving different urbanized areas have similar physical, financial, and operating characteristics. Although similarities in the statistics of the individual urbanized areas were identified, significant variations were also apparent. Examination of these variations suggests that frequently the dispersion of individual urbanized area values about a mean value can be explained by related causal variables.

The purpose of this working paper is to extend the analysis of urbanized areas to explore the hypothesis that it is possible to develop some preliminary public transportation planning norms from the NTS data. The analysis proceeds in three sections. Section II provides a description of the procedures and results of the econometric analyses. Section III uses the econometric results in an analysis of transit planning norms. The first part of this section discusses and interprets the econometric analyses, and the second part uses the econometric analyses to identify urbanized areas which are exceptions to the typical values for the group of urbanized areas being analyzed. Finally, Section IV suggests some conclusions about the possibilities for using the 1974 NTS data to develop norms for evaluating public transportation plans.

II. ECONOMETRIC ANALYSES

This section contains a technical description of the econometric analyses on which the results of Section III are based. These analyses test the hypothesis, stated in the introduction, that it is possible to develop norms from the 1974 NTS data for the development and operation of transit systems.

OBJECTIVE AND APPROACH

The general purpose of the econometric analyses in this section is to explore the notion that differences in variables which characterize the performance of transit systems can be explained by means of some economic relations, including supply, demand, production, operating cost, and operating revenue functions. If it is possible to specify and estimate equations which provide statistically good fits to the data, there is some presumption that these equations could then be used as norms against which evaluations could be made of the transit plans of individual urbanized areas. In order to accomplish this purpose, regression equations have been specified for supply, demand, production, revenue, and cost. Proposed explanatory variables in these equations are of two kinds: variables representing socio-economic characteristics of urbanized areas; and variables representing performance characteristics of transit systems.

The general approach of these analyses has been to use conventional econometric techniques to estimate the postulated equations. The equations estimated include both linear and logarithmic specifications. Analyses have been made by individual mode (bus and rail)[1] and for all modes (including bus, rail, personal rapid transit and demand actuated systems). In a few cases, too, it has been possible to stratify the analyses by system size to make rough estimates of economies of scale.

INVENTORY OF EQUATIONS

Table 1 is an inventory of the regression equations used in the econometric analyses. The equations are organized into the economic categories used in the appendix and some production functions have been added. Each of the economic categories includes alternative model specifications (estimated in both linear and logarithmic forms) with alternative variable specifications. There is a wide variation in the explanatory capability of the individual equations. The equations will be discussed in detail in subsequent paragraphs of this supplement.

Specification of the Regression Equations

Each of the equations listed in Table 1 was specified from a subset of the set of variables in the 1974 NTS data. Since the NTS socioeconomic data contained only population and square mile statistics, specifying the equations to be estimated was more limited than would have been desired. In the interest of consistency, it was decided not to introduce additional

1/ In these analyses, rail and commuter rail data were combined, except in cases where a dummy variable was specified to distinguish between the two.

TABLE 1
INVENTORY OF REGRESSION EQUATIONS

Category	Description	General Hypothesis	General Form	Code	Year	Mode
Supply	linear	Seat miles per capita is a function of land concentration per capita and the capability of the transit system to provide service (line miles per capita, vehicles per capita).	$\frac{stmi}{pop} = f\left\{\frac{sqmi}{pop}, \frac{limi}{pop}, \frac{v}{pop}\right\}$	S1 S2 S3 S4 S5 S6	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
	linear	Change in seat miles per capita is a function of current utilization (passenger miles per seat mile) and changes in those factors hypothesized to affect seat miles (line miles per capita, vehicles per capita, square miles per capita).	$\Delta\frac{stmi}{pop} = f\left\{\frac{pmi}{stmi}, \Delta\frac{sqmi}{pop}, \Delta\frac{limi}{pop}, \Delta\frac{v}{pop}\right\}$	S7 S8 S9 S10	1980 1990 1980 1990	bus bus rail rail
	logarithmic	Passenger miles per capita is a function of price and vehicles per capita. Passenger miles per capita is a function of price and line miles per capita (since the vehicles measure was unavailable for all modes).	$\ln\left(\frac{pmi}{pop}\right) = f\left\{\ln\left(\frac{oprev}{p}\right), \ln\left(\frac{v}{pop}\right)\right\}$ $\ln\left(\frac{pmi}{pop}\right) = f\left\{\ln\left(\frac{oprev}{p}\right), \ln\left(\frac{limi}{pop}\right)\right\}$	S11 S12 S13 S14 S15 S16	1972 1990 1972 1990 1972 1990	bus bus rail rail all modes all modes
	logarithmic: estimated for subgroups of urbanized areas, distinguished by system line miles.	Seat miles per capita is a function of price and vehicles per capita.	$\ln\left(\frac{stmi}{pop}\right) = f\left\{\ln\left(\frac{oprev}{p}\right), \ln\left(\frac{v}{pop}\right)\right\}$	S17 S18 S19 S20 S21 S22 S23 S24	1972 1972 1972 1972 1990 1990 1990 1990	bus bus bus bus bus bus bus bus
Production	linear	Change in seat miles per capita is a function of investment per capita.	$\Delta\frac{stmi}{pop} = f\left\{\frac{k}{pop}\right\}$	S25 S26 S27 S28 S29 S30	1980 1990 1980 1990 1980 1990	bus bus rail rail all modes all modes

TABLE 1 (cont.)
INVENTORY OF REGRESSION EQUATIONS

Category	Description	General Hypothesis	General Form	Code	Year	Mode
	logarithmic	Seat miles per capita is a function of capital and labor inputs, both per capita.	$\ln\left(\frac{\text{stmi}}{\text{pop}}\right) = f\left\{\ln\left(\frac{\text{avank cost}}{\text{pop}}\right), \ln\left(\frac{\text{opcost}}{\text{pop}}\right)\right\}$	S31 S32 S33	1990 1990 1990	bus rail all modes
	logarithmic	Seat miles per capita is a function of capital and labor inputs, both per capita and in combination.	$\ln\left(\frac{\text{stmi}}{\text{pop}}\right) = f\left\{\ln\left(\frac{\text{avank cost}}{\text{pop}}\right), \ln\left(\frac{\text{opcost}}{\text{pop}}\right), \left[\ln\left(\frac{\text{avank cost}}{\text{pop}}\right) - \ln\left(\frac{\text{opcost}}{\text{pop}}\right)\right]^2\right\}$	S34 S35 S36	1990 1990 1990	bus rail all modes
	logarithmic	Seat miles per capita is a function of capital and labor inputs and highway investment, all per capita.	$\ln\left(\frac{\text{stmi}}{\text{pop}}\right) = f\left\{\ln\left(\frac{\text{avank cost}}{\text{pop}}\right), \ln\left(\frac{\text{opcost}}{\text{pop}}\right), \ln\left(\frac{\text{hywk cost}}{\text{pop}}\right)\right\}$	S37	1990	all modes
	logarithmic	Seat miles per capita is a function of bus capital and labor inputs and rail capital and labor inputs, all per capita.	$\ln\left(\frac{\text{stmi}}{\text{pop}}\right)_{\text{all modes}} = f\left\{\ln\left(\frac{\text{avank cost}}{\text{pop}}\right)_{\text{bus}}, \ln\left(\frac{\text{opcost}}{\text{pop}}\right)_{\text{bus}}, \ln\left(\frac{\text{avank cost}}{\text{pop}}\right)_{\text{rail}}, \ln\left(\frac{\text{opcost}}{\text{pop}}\right)_{\text{rail}}\right\}$	S38	1990	all modes
	logarithmic: estimated for subgroups of urbanized areas, distinguished by system line miles.	Seat miles per capita is a function of capital and labor inputs, both per capita.	$\ln\left(\frac{\text{stmi}}{\text{pop}}\right) = f\left\{\ln\left(\frac{\text{avank cost}}{\text{pop}}\right), \ln\left(\frac{\text{opcost}}{\text{pop}}\right)\right\}$	S39 S40 S41 S42 S43 S44	1980 1980 1980 1990 1990 1990	bus bus bus bus bus bus

TABLE 1 (cont.)
INVENTORY OF REGRESSION EQUATIONS

Category	Description	General Hypothesis	General Form	Code	Year	Mode
Demand	linear: demand per capita	Passenger miles per capita is a function of land concentration per capita and quality of service.	$\frac{pmi}{pop} = f\left(\frac{sqmi}{pop}, pctpop, pctemp, avhdwy\right)$	D1 D2 D3	1972 1990 1972	bus bus rail
		Passenger miles per capita is a function of land concentration per capita, quality of service, and price.	$\frac{pmi}{pop} = f\left(\frac{sqmi}{pop}, pctpop, pctemp, avhdwy, \frac{oprev}{p}\right)$	D4 D5 D6 D7 D8 D9	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
	linear: change in demand per capita	Change in passenger miles per capita is a function of changes in land concentration and in quality of service factors.	$\Delta\left(\frac{pmi}{pop}\right) = f\left(\Delta\frac{sqmi}{pop}, \Delta pctpop, \Delta pctemp, \Delta hdwy\right)$	D10 D11 D12 D13	1980 1990 1980 1990	bus bus rail rail
	logarithmic: demand per capita	Passenger miles per capita is a function of price and quality of service: a) (peak-hour speed = quality of service); and	$\ln\left(\frac{pmi}{pop}\right) = f\left\{\ln\left(\frac{oprev}{p}\right) \ln\left(pkhrsp\right)\right\}$	D14 D15 D16 D17	1972 1990 1972 1990	bus bus rail rail
		b) (capital input is a proxy for quality of service).	$\ln\left(\frac{pmi}{pop}\right) = f\left\{\ln\left(\frac{oprev}{p}\right) \ln\left(\frac{avank\ cost}{pop}\right)\right\}$	D18 D19	1972 1990	all modes all modes
	linear: demand utilization	Passenger miles per seat mile is a function of land concentration per capita, price, and quality of service.	$\left(\frac{pmi}{atmi}\right) = f\left(\frac{sqmi}{pop}, avhdwy, pctpop, pctemp, \frac{oprev}{p}\right)$	D20 D21 D22 D23 D24 D25	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
	logarithmic: demand utilization	Passenger miles per seat mile is a function of price and quality of service a) (peak-hour speed = quality of service); and	$\ln\left(\frac{pmi}{atmi}\right) = f\left\{\ln\left(\frac{oprev}{p}\right) \ln\left(pkhrsp\right)\right\}$	D26 D27 D28 D29 D30 D31	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
		b) (capital input is a proxy for quality of service.	$\ln\left(\frac{pmi}{atmi}\right) = f\left\{\ln\left(\frac{oprev}{p}\right) \ln\left(\frac{avank\ cost}{pop}\right)\right\}$	D32 D33	1972 1990	all modes all modes

TABLE 1 (cont.)
INVENTORY OF REGRESSION EQUATIONS

Category	Description	General Hypothesis	General Form	Code	Year	Mode
Operating Costs	linear	Costs per capita is a function of seat miles per capita.	$\frac{opcosts}{pop} = f\left(\frac{stmi}{pop}\right)$	C1 C2 C3 C4 C5 C6	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
	linear	Costs per capita is a function of the variables hypothesized to relate to seat miles.	$\frac{opcosts}{pop} = f\left(\frac{sqmi}{pop}, \frac{limi}{pop}, \frac{v}{pop}\right)$	C7 C8	1972 1972	bus rail
	linear	Costs per capita is a function of seat miles per capita and utilization.	$\frac{opcosts}{pop} = f\left(\frac{stmi}{pop}, \frac{pmi}{stmi}\right)$	C9 C10 C11 C12 C13 C14 C15 C16 C17	1972 1980 1990 1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail all modes all modes all modes
	linear: estimated for subgroups of urbanized areas distinguished by size of line mile system.	Costs per capita is a function of seat miles per capita and utilization.	$\frac{opcosts}{pop} = f\left(\frac{stmi}{pop}, \frac{pmi}{stmi}\right)$	C18 C19 C20 C21 C22 C23 C24 C25	1972 1972 1972 1972 1990 1990 1990 1990	all modes all modes all modes all modes all modes all modes all modes all modes
Operating Revenue	linear	Revenue per capita is a function of passenger miles per capita.	$\frac{oprev}{pop} = f\left(\frac{pmi}{pop}\right)$	R1 R2 R3 R4 R5 R6	1972 1980 1990 1972 1980 1990	bus bus bus rail rail rail
		Revenue per capita is a function of the variables believed to be related to passenger miles.	$\frac{oprev}{pop} = f\left(\frac{sqmi}{pop}, avhdwy, pctpop, pctemp\right)$	R7 R8	1972 1972	bus rail

KEY TO VARIABLES

Name	Description
$\frac{stmi}{pop}$	seat miles per capita (thousands)
$\frac{sqmi}{pop}$	square miles per capita (square miles per thousand inhabitants)
$\frac{limi}{pop}$	line miles per capita (miles per thousand inhabitants)
$\frac{v}{pop}$	vehicles per capita (vehicles per thousand inhabitants)
$\frac{pmi}{stmi}$	passenger miles per seat mile at peak hours
$\frac{pmi}{pop}$	passenger miles per capita (thousands)
$\frac{oprev}{p}$	operating revenue per passenger (dollars per thousand passengers)
$\frac{k}{pop}$	capital investment per capita over the period beginning in 1972 (dollars)
$\frac{avank\ cost}{pop}$	average annual capital costs per capita (dollars)
$\frac{opcost}{pop}$	operating costs per capita (dollars)
$\frac{hvwk\ cost}{pop}$	highway capital costs per capita (dollars)
avhdwy	average weekday headway (minutes)
pctpop	percent population in specified band width around transit system
pctemp	percent employment in specified band width around transit system
pkhrsp	transit speed at peak hour (miles per hour)
Δ	change in value of the variable over the designated time period
C	data describing either a rail system or a commuter rail system (i.e., C equals '0' for rail data, and C equals '1' for commuter rail data)

data items into the data base. The resultant specification bias in the estimates must be recognized, although it is difficult to appraise how serious this bias may be. Variables for land use, income, and economic activity were not included in the NTS data base; consequently, they are not included here.

All of the equation specifications given in Table 1 have been estimated to determine which provide the best explanations of the dependent variables and which explanatory variables are significant. Many of the equations are statistically significant, although some are not.

Supply and Production Equations

Seat miles has been chosen as the best measure of the supply of transit service. To obtain comparable measures for different size urbanized areas, the dependent and the independent variables have been normalized by dividing by population for each urbanized area. Variations in seat miles per capita have been analyzed through three hypotheses:

- Seat miles per capita is related to both the intensity of land use (square miles divided by population) and the variables which define the ability of the system to supply transit services (line miles and number of vehicles each divided by population).
- Changes in seat miles per capita from 1972 to 1980 and from 1972 to 1990 are related to current utilization (passenger miles per seat miles), changes in square miles per capita in the urbanized area, changes in line miles per capita, and changes in vehicles per capita.

. The number of seat miles is a function of capital and labor inputs in a conventional production relation.

Table 2 contains the first hypothesis, as reflected in linear equations S1, S2, S3 (for bus for 1972, 1980, and 1990), and S6 (for rail for 1990). The R statistic indicates that the three independent variables explain a relatively high proportion (more than one half) of the variation in seat miles per capita. Equations S4 and S5 (for rail for 1972 and 1980 as shown in Table 1), not included in Table 2, have R^2 statistics of .17 and .27, respectively. Number of seat miles was expected to relate positively to number of line miles and number of vehicles, and to relate negatively to intensity of land use because transit is used more often and more effectively in more densely populated urbanized areas. The number of vehicles per capita is a significant variable in all of the equations, and the number of line miles per capita is significant in equations S1 and S2. Intensity of land use (square miles per capita) is never significant.

The second hypothesis is reflected in linear equations S7 through S10 (Table 3) which estimate the change from 1972-1980 and from 1972-1990 in seat miles per capita for bus and rail as a function of the utilization factor of the ratio of peak-hour passenger miles to peak-hour seat miles and the change in the per capita values of the intensity of land use, the number of line miles, and the number of vehicles. The independent variables successfully explain a high proportion of the variation in supply for bus and

TABLE 2

COEFFICIENT ESTIMATES FOR SEAT MILES PER CAPITA AS A FUNCTION OF SQUARE MILES PER CAPITA, LINE MILES PER CAPITA, AND VEHICLES PER CAPITA #

Dependent Variable: $\frac{\text{Seat Miles}}{\text{Population}}$ (thousands)

		Variables				Regression Statistics	
Equation	Constant	Square Miles per Capita (square miles per thousand inhabitants)	Line Miles per Capita (miles per thousand inhabitants)	Vehicles per Capita (vehicles per thousand inhabitants)	C* Dummy Variable	R^2	σ_y
S1 Bus, 1972						.76	.105
Coefficient	-.0147	.0319	.224	1.10			
σ	.0568	.0525	.092	.11			
β		.046	.191	.790			
S2 Bus, 1980						.70	.203
Coefficient	-.123	.00436	.358	1.43			
σ	.104	.11461	.131	.19			
β		.003	.248	.697			
S3 Bus, 1990						.54	.433
Coefficient	-.0337	-.266	.116	1.95			
σ	.2490	.405	.249	.38			
β		-.072	.059	.679			
S6 Rail, 1990						.57	.359
Coefficient	.452	-.830	4.21	2.74	-.210		
σ	.244	.475	4.49	.83	.170		
β		-.222	.127	.553	-.203		

*C is a variable which denotes whether the data describe a rail system or a commuter rail system. C equals '0' for rail data, and '1' for commuter rail data.

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
β = standardized regression coefficients;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

TABLE 3

COEFFICIENT OF ESTIMATES OF CHANGE IN SEAT MILES PER CAPITA AS A FUNCTION OF CURRENT PASSENGER MILES PER SEAT MILE AND CHANGES IN SQUARE MILES PER CAPITA, LINE MILES PER CAPITA, AND VEHICLES PER CAPITA #

Dependent Variable: $\Delta \frac{\text{Seat Miles}}{\text{Population}}$ (thousands)

Equation	Constant	Variables					Regression Statistics	
		1972 Passenger Miles per Seat Mile, peak hour	Change in Square Miles per Capita (square miles per thousand inhabitants)	Change in Line Miles per Capita (line miles per thousand inhabitants)	Change in Vehicles per Capita (vehicles per thousand inhabitants)	C* Dummy Variable	R^2	σ_y
S7 Bus, 1972-1980							.63	.173
Coefficient	.0503	-.0357	.174	.256	1.63			
σ	.0657	.1043	.352	.166	.27			
β		-.031	.045	.173	.680			
S8 Bus, 1972-1990							.62	.377
Coefficient	.0260	.113	.0858	-.323	2.51			
σ	.1324	.232	.3630	.256	.37			
β		.015	.022	-.163	.884			
S9 Rail, 1972-1980							.40	.298
Coefficient	.0810	-.00214	-.189	6.04	3.13	-.0590		
σ	.2362	.34106	1.326	5.61	1.34	.1250		
β		-.001	-.026	.205	.455	-.054		
S10 Rail, 1972-1990							.68	.315
Coefficient	.156	.0695	.531	1.22	3.90	-.0977		
σ	.218	.3090	.404	3.45	.74	.1287		
β		.023	.143	.049	.780	-.093		

*C is a variable which denotes whether the data describe a rail system or a commuter rail system. C equals '0' for rail data and '1' for commuter rail data.

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
β = standardized regression coefficients;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

rail. Change in the number of vehicles per capita is the only significant variable related to change in seat miles per capita for both bus and rail. The strong collinearity between line miles and number of vehicles (determined from the correlation matrix) could be masking the relation between seat miles and line miles. Changes in the intensity of land use show no significant relation to changes in seat miles per capita for either mode or either time period.

An analysis was also made to estimate the relation between changes in seat miles per capita and capital cost per capita. The regression equation specified a linear relation between changes in seat miles per capita as the dependent variable and changes in capital cost per capita and a dummy variable that could distinguish between rail and commuter rail ('0' and '1', respectively) as independent variables. The results of the regression analysis are shown in Table 4. Although the coefficient for capital cost per capita is statistically significant in all of the equations, the relatively low R^2 statistics show that little of the variance in seat miles per capita is explained by differences in capital cost per capita. The dummy variable is not statistically significant.

The third hypothesis is reflected in equations S31 through S33, which specify Cobb-Douglas production functions relating transit system output to inputs of capital and labor. The production function can be written as: $Y = A\ K^{b_1}\ L^{b_2}$

TABLE 4

COEFFICIENT ESTIMATES OF CHANGE IN SEAT MILES PER CAPITA AS A FUNCTION OF CAPITAL INVESTMENT PER CAPITA#

Dependent Variable: $\Delta \frac{\text{Seat Miles}}{\text{Population}}$ (thousands)

		Variables		Regression Statistics	
Equation	Constant	Capital Cost per Capita	C* Dummy Variable	R^2	σ_y
S25 Bus, 1972-1980				.37	.221
Coefficient	.048	.00570			
σ	.045	.00107			
β		.605			
S26 Bus, 1972-1990				.13	.549
Coefficient	.134	.00305			
σ	.105	.00114			
β		.361			
S27 Rail, 1972-1980				.31	.300
Coefficient	.081	.000385	-.083		
σ	.098	.000312	.125		
β		.503	-.118		
S28 Rail, 1972-1990				.47	.390
Coefficient	.198	.000956	-.195		
σ	.140	.000256	.158		
β		.563	-.185		
S29 All Modes, 1972-1980				.28	.323
Coefficient	.184	.00117			
σ	.055	.00027			
β		.533			
S30 All Modes, 1972-1990				.16	1.738
Coefficient	.169	.00234			
σ	.368	.00078			
β		.401			

*C is a variable which denotes whether the data describe a rail system or a commuter rail system. C equals '0' for rail data and '1' for commuter rail data.

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
β = standardized regression coefficients;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

where: Y = seat miles per capita

A = coefficient of efficiency

K = capital

L = labor

For this analysis, operating cost per capita has been used as a proxy for labor inputs and capital cost per capita[1] as the measure of capital inputs. The parameters b_1 and b_2 are elasticities of output with respect to capital and labor, respectively. Also, $b_1 + b_2$ is a measure of economies of scale,

where: $b_1 + b_2 = 1$ indicates constant returns to scale,

$b_1 + b_2 > 1$ indicates increasing returns to scale, and

$b_1 + b_2 < 1$ indicates decreasing returns to scale.

The coefficient of efficiency, A, indicates that two functions which have identical elasticities may still have different levels of output if they have different values of A. The degree of factor intensity can be assessed by the ratio, $b_2 : b_1$. A production function with a higher $b_2 : b_1$ ratio represents a more labor intensive activity than a function with a low $b_2 : b_1$ ratio.

The estimated coefficients of the Cobb-Douglas production function are given for bus, rail, and all modes for 1990 in Table 5. The elasticities

---

1/Capital cost is the total capital expenditure from 1972 to 1990 and disregards capital in place as of 1972. This exclusion is important only for the larger, existing rail systems, such as New York, Chicago, and Boston. Possible distortions of the results are mitigated, however, by the use of capital cost as a proxy for physical capital.

TABLE 5

COEFFICIENT ESTIMATES FOR COBB-DOUGLAS PRODUCTION FUNCTIONS FOR BUS, RAIL, AND ALL MODES

Dependent Variable: $\frac{\text{Seat Miles}}{\text{Population}}$ (thousands)

Equation	Constant	Variables		Regression Statistics	
		b_1 (capital)	b_2 (labor)	R^2	σ_y
S31 Bus, 1990				.50	.466
Coefficient	.130	.109	.552		
σ	1.330	.081	.105		
S32 Rail, 1990				.67	.982
Coefficient	.046	.094	.828		
σ	1.321	.099	.130		
S33 All Modes, 1990				.62	.544
Coefficient	.0768	.030	.805		
σ	1.4096	.079	.130		

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

for labor are substantially higher than those for capital for bus, rail, and all modes. The relative size of the bus and rail coefficients is consistent with the notion of factor intensity discussed above; a higher ratio of the labor coefficient to the capital coefficient (higher ratio of $b_2 : b_1$) indicates a more labor intensive activity. That is the case in equations S31 for bus and S32 for rail, which have $b_2 : b_1$ ratios of 5.06 and 8.81, respectively. The estimated coefficients for the labor input (operating cost) are statistically significant in all of the equations. The statistical significance of the capital input (capital cost) coefficients is low in equations S31 and S32, and nonexistent in equation S33.

A limited analysis was also made of returns to scale in bus operations, as represented in the 1974 NTS data. The results are shown in Table 6. For 1972 and 1990, bus operations were stratified into large, medium, and small systems on the basis of the number of line miles operated. The estimated parameters of the regressions conform to the expected pattern for the large and small systems, but for medium-size systems, the capital coefficient has the wrong sign. The large systems have almost the same pattern of capital and labor elasticities for both 1972 and 1990 and show almost constant returns to scale for both years. The R^2 statistics are reasonably good for both 1972 and 1990 for the large and small systems, indicating that these four equations achieve about the same fit to the data.

TABLE 6

PRODUCTION FUNCTIONS FOR BUS SYSTEMS STRATIFIED BY NUMBER OF LINE MILES OPERATED

System Size (Line Miles)	Parameters of a Cobb-Douglas Production Function (bus mode)						
	A	b_1 (capital)	b_2 (labor)	R^2	Returns to Scale	Factor Intensity for Labor	n
1972							
Large >1000 miles	.088	.324 (.111)	.652 (.170)	.66	.976	2.01	16
Medium 500-1000 miles	.163	-.104 (.202)	.618 (.336)	.34	--	--	12
Small <500 miles	.102	.0922 (.1031)	.616 (.095)	.75	.708	6.68	23
1990							
Large >1000 miles	.079	.364 (.137)	.638 (.245)	.60	1.002	1.75	16
Medium 500-1000 miles	.082	-.137 (.108)	.829 (.239)	.51	--	--	15
Small <500 miles	.134	.278 (.151)	.453 (.139)	.70	.731	1.63	17

#/ R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

n = number of observations.

The small systems show approximately the same elasticity of output to labor input as the large systems in 1972, but a significantly smaller elasticity in 1990. The elasticities of output to labor inputs are similar for all three system stratifications in 1972, and for the large system in 1990, but the elasticity is higher for the medium system and lower for the small systems in 1990. The small systems for both 1972 and 1990 have decreasing returns to scale (implying increasing long-run average cost). The returns to scale are similar for 1972 and 1990, although the relative elasticities of labor and capital are quite different in the two years. The factor intensity of the production function is given by the ratio of the coefficients of labor and capital. The labor intensity of production for small systems decreases from 6.68 in 1968 to 1.63 in 1990, and capital intensity increases accordingly. This shift may, however, be due to the use of dollar values rather than real measures of inputs.

A supply relation was also postulated which incorporated a proxy for price. A logarithmic function was specified, using seat miles per capita as the dependent variable. The full specification was:

$$S = P^{c_1} V^{c_2}$$

where:

S = seat miles per capita

P = operating revenue per passenger

V = vehicles per capita

c_1 = supply elasticity with respect to operating revenue per passenger

c_2 = supply elasticity with respect to vehicles per capita

This specification was estimated for bus systems stratified by size for 1972 and 1990. The results appear in Table 7. The price elasticities of supply are not significant for the large- and medium-size systems in 1972 but are significant in all other cases.

The R^2 statistics are high for all of the 1972 equations and quite high for the 1990 equations. In general, the elasticity of supply is less associated with changes in price than with changes in the number of vehicles per capita.

Operating Revenue

Operating revenue was hypothesized to be related to two sets of variables--passenger miles per capita and those variables closely related to passenger miles (namely, square miles per capita, average headway, and percent of population and percent of employment within a specified band width around the transit system). The dependent variable specified was operating revenue per capita. Operating revenue equations were estimated for bus and rail in each of the three time periods, 1972, 1980, and 1990. As shown in Table 8, the only variable which was consistently significant was passenger miles per capita. The R^2 statistic was greater than .4 in only three of the six equations estimated.

Operating Cost

Operating cost equations were specified with operating costs per capita as a function of seat miles per capita, and also with operating costs per capita as a function of land area per capita, line miles per capita, and

TABLE 7

COEFFICIENT ESTIMATES IN LOGARITHMIC SUPPLY FUNCTIONS FOR BUS SEAT MILES PER CAPITA #

System Size (Line Miles)	Variables		R^2	n
	Operating Revenue per Passenger	Vehicles per Capita		
Total, 1972			.95	50
Coefficient	.04	.91		
σ	(.01)	(.06)		
Large, 1972 (>1,000 miles)			.93	9
Coefficient	.01	.75		
σ	(.02)	(.17)		
Medium, 1972 (500-1,000 miles)			.91	11
Coefficient	.002	.57		
σ	(.05)	(.3)		
Small, 1972 (<500 miles)			.96	30
Coefficient	.03	.9		
σ	(.02)	(.08)		
Total, 1990			.95	50
Coefficient	.04	.91		
σ	(.01)	(.06)		
Large, 1990 (>1,000 miles)			.66	16
Coefficient	.1	1.2		
σ	(.03)	(.2)		
Medium, 1990 (500-1,000 miles)			.65	15
Coefficient	.11	1.11		
σ	(.04)	(.2)		
Small, 1990 (<500 miles)			.88	17
Coefficient	.07	1.08		
σ	(.08)	(.17)		

#/ σ = standard errors of estimate for the regression coefficients;
R = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.
n^2 = number of observations.

TABLE 8

COEFFICIENT ESTIMATES FOR OPERATING REVENUE AS A FUNCTION OF PASSENGER MILES PER CAPITA AND A RAIL DUMMY VARIABLE#

Dependent Variable: $\frac{\text{Operating Revenue}}{\text{Population}}$ (dollars)

		Variable		Regression Statistics	
Equation	Constant	Passenger Miles Per Capita	C* Dummy Variable	R^2	σ_y
R1 Bus, 1972				.49	3.6
Coefficient	2.54	49.3			
σ	.98	7.2			
β		.70			
R3 Bus, 1990				.43	7.8
Coefficient	6.72	36.6			
σ	1.71	6.1			
β		.66			
R6 Rail, 1990				.44	9.3
Coefficient	6.37	39.2	-4.4		
σ	3.46	10.6	3.6		
β		.56	.18		

*C is a variable which denotes whether the data describe a rail system or a commuter rail system. C equals '0' for rail data and '1' for commuter rail data.

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
β = standardized regression coefficients;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

vehicles per capita. In addition, an attempt was made to determine if utilization of the transit system affected operating costs per capita. The relation between the dependent and independent variables was expected to be positive. The following relations were hypothesized:

- Operating costs per capita is positively related to seat miles per capita, i.e., as seat miles per capita increases, operating cost per capita will increase.
- Operating costs per capita is positively related to square miles per capita, line miles per capita, and vehicles per capita.
- Operating costs per capita is positively related to seat miles per capita and to the utilization of the system.

The results of estimating the operating cost equations are given in Table 9 for all equations with R^2 statistics of .4 or more. A strong relation was found between operating costs per capita and seat miles per capita, suggesting that per capita costs increase proportionately with supply as measured by seat miles per capita. This relation was estimated for bus and rail systems for 1972, 1980, and 1990. The results also showed a significant relation between seat miles and operating costs for bus for all three periods and for rail for both projected periods (1980 and 1990). Only for bus did the seat mile variable explain at least 40 per cent of the variation (i.e., an $R^2 > .4$, see equations C1, C2, and C3 in Table 9).

TABLE 9

COEFFICIENT ESTIMATES FOR OPERATING COSTS AS A FUNCTION OF SEAT MILES PER CAPITA, SQUARE MILES PER CAPITA, LINE MILES PER CAPITA, VEHICLES PER CAPITA, AND PASSENGER MILES PER CAPITA#

Dependent Variable: $\frac{\text{Operating Cost}}{\text{Population}}$ (dollars)

		Variables					Regression Statistics	
Equation	Constant	Seat Miles per Capita	Square Miles per Capita	Line Miles per Capita	Vehicles Per Capita	Passenger Miles per Seat Mile	R^2	σ_y
C1 Bus, 1972							.63	3.7
Coefficient	-1.06	23.4						
σ	1.29	2.6						
β		.79						
C2 Bus, 1980							.50	6.8
Coefficient	3.34	18.6						
σ	2.06	2.7						
β		.70						
C3 Bus, 1990							.42	10.5
Coefficient	9.29	14.1						
σ	2.44	2.4						
β		.65						
C7 Bus, 1972							.76	3.1
Coefficient	1.70		-3.16	-3.79	34.96			
σ	1.67		1.54	2.71	3.35			
β			- .16	- .11	.85			
C15 All Modes, 1972							.80	5.92
Coefficient	-5.16	24.2				6.59		
σ	1.52	2.6				1.93		
β		.72				.27		
C16 All Modes, 1980							.75	8.9
Coefficient	-4.32	21.7				10.49		
σ	2.46	2.6				3.23		
β		.70				.27		
C17 All Modes, 1990							.43	15.0
Coefficient	12.15	4.59				13.87		
σ	3.82	1.14				3.28		
β		.45				.47		

#/ σ = standard errors of estimate for the regression coefficients;
σ_y = standard errors of estimate for the regression equations;
β = standardized regression coefficients;
R^2 = coefficients of multiple determination, which show the proportion of total variance explained by the regression equation.

For 1972, bus operating costs were also estimated as a function of square miles per capita, line miles per capita, and vehicles per capita. The results (equation C7) showed that this set of variables explained more variation in bus operating costs than did seat miles (equation C1). An R^2 of only .17 (not shown in Table 9) was obtained for rail costs with this specification. In the bus equation, vehicles per capita and square

For 1972, bus operating costs were also estimated as a function of

miles per capita showed a significant relation to costs.

It was hypothesized, in addition, that operating costs per capita were related to utilization (defined as passenger miles per seat mile) in addition to seat miles per capita. This relation was specified for bus, rail, and all modes data for 1972, 1980, and 1990. Adding the utilization variable did not improve the equation fits for bus and rail modes (the equations are not shown in the table), but it did produce good fits for the all mode equations (see equations C15, C16, and C17 in Table 9.)

Although the coefficient estimates are not reproduced herein, equations C15 and C17 were also estimated for stratification by system size. As before, three alternative size classifications were used. In the eight equations estimated, only in 1972 for medium size systems was the utilization factor not significant. The equation lends itself to a straightforward economic interpretation. The equation is:

$$TC = AX_1 + B\frac{X_2}{X_1}$$

where: TC = total operating cost per capita

X_1 = seat miles per capita

X_2 = passenger miles per capita

$$\text{Average Cost} = \frac{TC}{X_1} = A + B\frac{X_2}{X_1^2}$$

$$\text{Marginal Cost} = \frac{dTC}{dX_1} = A - B\frac{X_2}{X_1^2}$$

Substituting the size class mean values for seat miles for passenger miles per seat miles yields the average and marginal costs shown in Table 10.

For both 1972 and 1990, marginal cost lies below average cost for all size systems, indicating that average cost is decreasing. In three instances in 1972, marginal cost is declining; in the fourth case for 1972 and in all cases for 1990, marginal cost is rising.

The average cost pattern for 1972 indicates average costs are rising as system size increases; that is, the data exhibit decreasing returns to scale. The data for 1972 show decreasing short-run average costs and decreasing returns to scale (increasing long-run average cost). The data for 1990 show decreasing short-run average costs. Returns to scale and, therefore, long-run average cost are approximately constant.

The present results are not comparable with the conclusions on returns to scale from the production function analysis. The present analysis is for all modes, whereas the production function analysis was for bus only.

TABLE 10

AVERAGE AND MARGINAL COST
FOR ALL MODES FOR 1972 AND 1990
(1971 dollars)

System Size (Line Miles)	1972 Marginal Cost	1972 Average Cost	1972 n	1990 Marginal Cost	1990 Average Cost	1990 n
Total	-11.68	19.41	49	8.51	22.22	43
Large (than 1,000 line miles)	18.11	28.50	8	9.12	21.82	13
Medium (500 to 1,000 line miles)	-14.30	24.36	12	13.94	23.85	15
Small (than 500 line miles)	-5.69	12.54	29	14.77	23.41	15

Demand

Passenger miles per capita was selected as the best measure of transit service demanded. Using this variable, a number of hypotheses can be formulated which suggest specifications of simple transit demand models. Several variables hypothesized as related to transit demand include:

- Price. It is expected that the price of transit service (defined here as operating revenue per passenger) will be an important determinant of transit usage, with higher prices causing fewer passenger miles per capita.
- Population density of the urbanized area. It is expected that the dispersion of inhabitants in an urbanized area (measured here as square miles per capita) will influence demand, but the direction of influence is uncertain. It could be expected that denser areas (i.e., those with fewer square miles per capita) will have a greater demand since transit is more convenient in these areas and the alternative, automobile travel, is more bothersome. In denser urbanized areas, however, passengers make shorter trips. Therefore, it is not clear whether demand, as measured by passenger miles, would increase or decrease with a greater number of square miles per capita.
- Coverage of the transit system. The percent of population and percent of employment served by the transit system implies

more complete service offered to the passenger. Accordingly, higher values of these variables are expected to be associated with more passenger miles per capita.

- Average headway. This variable describes an aspect of the quality of service that a passenger can expect. Shorter headways mean better service and should generate more passenger demand.

The relations between these explanatory variables, in different combinations, and passenger miles per capita were estimated individually for bus and rail systems for 1972 and 1990. As shown in Table 1, linear and logarithmic demand equations were specified. The principal result was that, for both bus and rail systems, the explanatory variables of square miles per capita, price, and headway were not sufficient to explain very much of the variation among urbanized areas in demands for transit services. For bus, about one quarter of the variation in demand was explained by the variables (for 1972 data), and for rail, 41 percent was the maximum explained (in 1990). None of these was considered worth reproducing here. Thus, it can be concluded that either no consistent pattern exists to explain the variation in transit demand among urbanized areas or that additional variables will explain the variation. Possible variables, not in included in the NTS data, are demographic characteristics of the residents of the urbanized areas.

III. ANALYSIS OF TRANSIT PLANNING NORMS

This section uses the results of the econometric analysis in Section II to determine whether it is possible to develop transit planning norms from the 1974 NTS data. The first part of this section discusses and interprets the econometric analyses, and the second part uses the econometric analyses to identify urbanized areas which are exceptions to the typical values for the group of urbanized areas being analyzed.

ECONOMETRIC ANALYSES

The main purpose of econometrics is to give empirical content to *a priori* reasoning about relations among economic variables. In this memorandum, econometric analysis has been used to define and estimate relations among transit variables in an attempt to refine the search for similarities as well as to account for differences in transit planning among the largest urbanized areas. Briefly, the approach used in the analyses was the:

- identification of a framework which simplified the characterization of individual transit operations;
- specification of the relations between those explanatory and explained variables that best accounted for individual variations in the data; and
- application of linear regression analysis, using the 1974 NTS data, to the 52 urbanized areas studied.

The framework chosen to simplify the characterization of individual transit operations is based on the treatment of transit service as an economic service. Accordingly, urban transit operations may be viewed from the perspective of:

- supply and production (expressed as seat miles per capita);
- demand (expressed as passenger miles per capita of service); and
- financial operations (expressed by the operating revenues and operating costs resulting from the supply of a given level of transit service).

The analyses in this section specify relations between the foregoing variables and sets of explanatory variables. The explanatory variables are of two kinds: variables representing socioeconomic characteristics of urbanized areas; and variables representing performance characteristics of transit systems.

General Description of the Econometric Analyses

The purpose of the econometric analyses is to determine whether differences in the variables characterizing transit operations may be explained by the suggested explanatory variables. Each of the analyses has three separate parts.

The first part is the identification of specific relations between the explanatory variables and the variable to be explained. This includes the selection of which variables are to be explained, which variables are to be suggested as explanatory, and what form is to be used for the descriptive equations to best explain the relation among variables. The technical term for this identification process is specification.

The second part of the statistical analysis is essentially a measurement of the variations encountered. The process uses available data (i.e., the 1974 NTS data) to measure the degree of variation in the variable to be explained (the dependent variable) which results from variation in one or more suggested explanatory variables (the independent variables). The technical term for this measurement process is estimation. As this term implies, the mathematical relation between the dependent and independent variables is not known with certainty; rather, it is estimated from the data using the appropriate statistical techniques.

The third part of the statistical analysis is the evaluation and interpretation of the estimated mathematical relations between the dependent and independent variables. Although the estimated relations are not exact, it is possible to determine from the statistical analysis the degree of confidence that one can place in them.

Specification of Equations

The framework for the econometric analyses is a set of regression equations which have potential for describing the physical, financial, and operational characteristics of a group of individual transit systems. The analysis of the supply characteristics includes equations which appear to be possible causal or descriptive determinants of the level of supply of transit service and of the magnitude of the operating costs which are incurred as a result of providing those services. The analysis of demand characteristics includes equations which appear to be possible causal or descriptive determinants of the level of demand for transit service and of the magnitude of the corresponding operating revenues resulting from that demand.

The framework for the statistical analyses was applied separately to bus and to rail modes because the technical differences between these two modes suggest that the explanatory variables may differ or at least may exert different influences on the variables to be explained. The rail and commuter rail data have been combined here for cases where a dummy variable was used to distinguish between the two modes.

The individual equations are discussed here in their order of appearance in Section II. Because they are shown in detail in Section II (Table 1), the equations will not be repeated here. The purpose of this section is, rather, to describe the findings and implications of the econometric analyses.

Supply and Production

Seat miles per capita is the best measure of the supply of transit services, when the effects of different size urbanized areas are taken into account. The supply and production equations are specified in both linear and logarithmic form, and they use both seat miles per capita and change in seat miles per capita as the variables to be explained (i.e., as dependent variables) in the regression equations. In one set of equations (S1 through S6), the supply of transit services is postulated as being explained by the variables:

- square miles per capita;
- line miles per capita; and
- vehicles per capita.

In a second set of equations (S7 though S10), the change from 1972 to 1980 and from 1972 to 1990 in the supply of seat miles per capita is postulated as being dependent on changes in the foregoing variables and on a utilization factor measured by the variable, passenger miles per capita/seat miles per capita.

In general, the specified variables provide relatively good explanations of differences in bus seat miles per capita for 1972 and 1980 but a poor explanation for bus seat miles per capita for 1990. The variables provide a partial explanation for rail seat miles per capita for 1990 but are poor for rail for other years. In addition, a dummy variable was used in the rail equations to distinguish between rail and commuter rail, but it did not prove statistically significant. The estimated coefficients of these equations are given in Table 2.

The most consistently important explanatory variable for rail seat miles per capita is vehicles per capita; for bus it is line miles per capita. Square miles per capita is a significant explanatory variable only for rail systems in 1990. There is an inverse relation between seat miles per capita and square miles per capita, which means a direct relation between seat miles per capita and population density (inhabitants per square mile); the higher the density, the more seat miles per capita. Again, when change in seat miles per capita is regressed against the specified set of variables, only change in vehicles per capita is a statistically significant exploratory variable.

Some equations (S25 through S30) which related changes in seat miles per capita to capital investment per capita and to a rail/commuter rail dummy variable were also estimated. None of these equations (Table 4) accounts for very much of the variation among urbanized areas in planned changes in seat miles per capita. Although the capital investment coefficient is always statistically significant, the small proportion of the total variance explained by the regression equation indicates only an imprecise relation between planned output and planned capital investment.

A set of production functions was also specified using seat miles per capita as the output (or dependent) variable with operating cost as a proxy for labor input and capital cost as a proxy for capital input. The labor coefficient was consistently significant, but the capital coefficient was never significant, even for rail systems (Table 5).

Equations were estimated only for 1990. None of the three equations explained as much as two-thirds of the total variance in seat miles per capita.

A limited analysis of the returns to scale was possible for bus systems. (Returns to scale indicates whether output increases in proportion with, more than, or less than proportional increases in all inputs. Constant returns to scale means constant long-run average total cost as the planned output of the system increases; increasing returns to scale means decreasing long-run average total cost as the planned output of the system increases; and decreasing returns to scale means increasing long-run average total cost as the planned output of the system increases.) The systems were stratified by size (i.e., number of line miles) into large (more than 1,000 miles), medium (between 500 and 1,000 miles), and small (less than 500 miles). Large systems show approximately constant returns to scale in both 1972 and 1990 (Table 6); but small systems showed decreasing returns to scale. These results imply that average costs of bus systems increase as the planned output of the system increases, until the system becomes quite large. At that point, further increases in planned output can be accomplished at approximately constant average total cost.

One additional, and more conventional, supply function was postulated which specified bus seat miles per capita as a function of operating revenue per passenger (used as a proxy for price) and vehicles per capita (used as a measure of output capability). The analyses were again stratified by bus system size as determined by number of line miles comprising the system. In general, the regression equations explain a high proportion of the total variance in seat miles per capita. For 1972, the regression equations explain more than 90 percent of the total variation among urbanized areas in the supply of bus seat miles per capita (Table 7). Vehicles per capita is a significant variable in all of the equations, and price is significant in most of the equations. The ability of the equations to explain variation in the supply of seat miles per capita among the urbanized areas in 1990 is not quite as good as it is for 1972. Nonetheless, the amount of explained variation is also quite high in this latter set of equations.

The results of these analyses of supply and production must be interpreted with a good deal of caution. Although a number of the equations estimated provide relatively good explanations of variations in the supply of seat miles per capita among the larger urbanized areas, a number of the equations did not yield any usable results. The failure of some of the equations and the success of others may be a phenomenon inherent in the data, or it may reflect the inadequacy of the equation specification. In a number of instances, there are some a priori presumtions--additional socioeconomic variables would help obtain equations which better fit the data. But, as is noted in Section II, more inclusive specifications are not possible so long as the NTS data set is the sole source of data. Although the equations discussed above have, at least in part, been capable of explaining variation in the supply of seat miles per capita for the transit system variables such as line miles per capita or vehicles per capita, they have not succeeded in explaining variations in terms of different socioeconomic characteristics of the urbanized areas to the degree that had been hoped.

Operating Revenue

Two separate equation specifications were made for operating revenue. The first was used to estimate the relation between operating revenue per capita and passenger miles per capita. If a close relation could be found, it would mean that urbanized areas are generally setting about the same transit fares per passenger mile. In fact, however, only relatively little of the variation among urbanized areas in operating revenue per capita is explained by differences in passenger miles per capita. The equations were estimated for bus and rail in all three time periods (i.e., six equations). In only three of the six equations (Table 8) did passenger miles per capita explain at least 40 percent of the variation in passenger revenue per capita.

An expanded specification, using square miles per capita, average headway, and percent population and percent of employment within a specified band width around the transit system provided no better results. It must be concluded, therefore, that either there is a good deal of randomness in operating revenue per capita among the larger urbanized areas, or it has not been possible to determine what variable or variables would explain the observed variation. Certainly the variables specified do not provide the explanation expected.

Operating Cost

Two separate equation specifications were made for operating cost. The first was used to estimate the relation between operating cost per capita and seat miles per capita. Seat miles per capita was found to be a highly significant explanatory variable, but it did not explain a high proportion of the total variation in operating costs per capita among the urbanized areas.

In the second of the all mode equations, system utilization (i.e., passenger miles per seat mile) was specified along with seat miles per capita and was a significant variable. These two variables explained 80 percent and 75 percent of the total variation in operating costs per capita in 1972 and 1980, but only 43 percent in 1990.

A third equation specified square miles per capita, line miles per capita, and vehicles per capita as explanatory variables. The equation explained 76 percent of the variation in operating costs per capita for bus data for 1972, with vehicles per capita and square miles per capita as significant variables. None of the other equations with this specification explained more than 40 percent of the variation among the urbanized areas in operating costs per capita.

Although the econometric analysis was able to produce some interesting equations explaining operating costs per capita, many of the equations that used the same set of variables and different data (i.e., for different years or different modes) did not produce meaningful results. Much of the difference in the operating cost data appears to be the result of random variation.

Demand

Passenger demand was measured in passenger miles per capita. In general, it was expected that the level of transit usage would vary inversely with price (measured as operating revenue per passenger), inversely with headway, and directly with the percent of employment and population conveniently accessible to the transit system. A number of equations were specified using these variables, and none explained more than 41 percent of the total variation in passenger miles per capita. It seems likely that a change in the equation specification to include more socioeconomic variables would improve the fit of the equations to the data. The poor equation fits precluded any realistic possibility of identifying demand norms from the NTS data.

Prediction of Rail Systems

An attempt was made to specify some variables that would predict the planning for a rail or commuter rail system. In this analysis the dummy variable, K, takes a value of '1' for a planned rail or commuter rail system and '0' otherwise. The explanatory variables include population, population density, and planned highway capital outlay. Only some 21 percent of the total variance was explained by the equation. The population variable is statistically significant and has the expected sign. Population density is not very significant, but it also has the expected sign. Perhaps the most interesting result occurs for highway capital investment. The variable is statistically significant and indicates that investment plans for rail or commuter rail are positively associated with higher highway investment plans, holding constant for population and population density. This result strongly suggests that urbanized areas are not planning rail or commuter rail and highway systems as substitutes for one another, as might be expected.

ANALYSIS OF EXCEPTIONS

In the previous part of this section, several relations were suggested and analyzed, through the use of econometric techniques, in order to refine the search for similarities in transit planning of the urbanized areas studied. Although econometric analysis provides an estimate of the relation between one or more explanatory variables and a dependent variable, this estimate is not exact; actual values of individual urbanized areas will deviate from those values which are predicted by the estimated equations. Individual urbanized values which deviate by a significant amount from the predicted values may be investigated as exceptions to the pattern established by that group of variables representing an urbanized area.

Method of Analysis

The standard error of estimate (σ_y) is considered a measure of the typical deviation which may be expected between the actual value of a variable and the value predicted by the regression equation. For the purposes of this analysis, therefore, the standard error of estimate constitutes a criterion by which exceptions to the urbanized area aggregate may be evaluated. In particular, the process for determining individual urbanized area 'exceptions' may be summarized as follows:

- specify and estimate regression equations;
- determine, for each urbanized area and each equation, the difference between the predicted values and the actual values which were reported in the 1974 NTS; and
- identify those urbanized areas for which the difference between the reported value and the predicted value exceeds the standard error of the estimate.

Results

Tables 11 through 13 display the results of the foregoing process for all modes, bus, and rail. Each equation that explains at least 40 percent

TABLE 11

TABLE OF EXCEPTIONS: ALL MODES

Urbanized Areas	Variables Analyzed*: Seat Miles / Population 1990 (S33)#	Operating Costs / Population 1972 (C15)#	Operating Costs / Population 1980 (C16)#	Operating Costs / Population 1990 (C17)#
	R^2=.62	R^2=.80	R^2=.75	R^2=.43
New York		+	+	
Boston		+		
Cleveland				+
Pittsburgh			+	
Baltimore				+
Seattle		-	-	
Cincinnati	+			
San Diego	+			
Milwaukee				+
Dallas		-		
Portland			-	
San Juan		+		+
Providence		-		
Norfolk-Portsmouth		-		-
Memphis		-	+	
Akron	-			
Birmingham	+	-	-	-
Tampa	-			
Toledo	+			-
Orlando				-

*/ + indicates that the urbanized area reported statistic was at least one standard deviation higher than the value computed from the corresponding regression equation.

\- indicates that the urbanized area reported statistic was at least one standard deviation lower than the value computed from the corresponding regression equation.

#/ See Table 1.

TABLE 12

TABLE OF EXCEPTIONS: BUS

	Variables Analyzed[a]											
	(S1)# Seat Miles / Population 1972	(S2)# Seat Miles / Population 1980	(S3)# Seat Miles / Population 1990	(S7)# Δ Seat Miles / Δ Population 1972-1980	(S8)# Δ Seat Miles / Δ Population 1972-1990	(S31)# Seat Miles / Population (production) 1990	(C1)# Operating Costs / Population 1972	(C2)# Operating Costs / Population 1980	(C3)# Operating Costs / Population 1990	(C7)# Operating Costs / Population 1972	(R1)# Operating Revenue / Population 1972	(R3)# Operating Revenue / Population 1990
Urbanized Areas	R^2=.76	R^2=.70	R^2=.54	R^2=.63	R^2=.62	R^2=.50	R^2=.63	R^2=.50	R^2=.42	R^2=.76	R^2=.49	R^2=.43
New York						-	+	+	+		+	
Los Angeles		+										
Chicago		-	-	-	-		+	+		+	+	
Philadelphia							+		+			
San Francisco			-		-		+	+	+			
Boston						-	+	+	+	+		
Washington, D.C.	-			+			+	+			+	+
Cleveland									+		+	+
St. Louis	-	-				-						
Pittsburgh							+	+	+	+		
Baltimore			-	-	-							
Houston	+						-					
Seattle												-
Atlanta	+	+					-					
Miami		+		+							-	
Cincinnati		+	+	+		+			-			
San Diego		+	+	+	+	+			-			-
Milwaukee	+	+	+	+	+	+	-				+	+
Buffalo	-			+								
Dallas	+						-	-				
Portland						+						
San Juan	+	-		-			+	+		+		+
Columbus		+									+	
Norfolk-Portsmouth							-			-		
Memphis	+						-					
Sacramento								-				
San Antonio							-					
Rochester		+		+				-				
Oklahoma City						-						
St. Petersburg					-							
Akron						-						
Birmingham	+					+	-	-	-		-	-
Springfield, Mass.	-	-	-							-		
Hartford	-											
Albany	-			+								
Tampa				-								
Toledo		-		-								
Omaha				-								
Orlando										+		

a/+ indicates that the urbanized area reported statistic was at least one standard deviation higher than the value computed from the corresponding regression equation.

\- indicates that the urbanized area reported statistic was at least one standard deviation lower than the value computed from the corresponding regression equation.

#/Equation numbers correspond to those in Table 1

TABLE 13

TABLE OF EXCEPTIONS: RAIL

	Variables Analyzed*				
	Seat Miles / Population 1990 (S6)#	Δ Seat Miles / Population 1972-1980 (S9)#	Δ Seat Miles / Population 1972-1990 (S10)#	Seat Miles / Population 1990 (S32)#	Operating Revenue / Population 1990 (R6)#
Urbanized Areas	R^2=. 57	R^2=. 40	R^2=. 68	R^2=. 67	R^2=. 44
Los Angeles	-		-	-	
Chicago	-				
Philadelphia	-				
Detroit				+	
Boston				-	
Washington, D.C.	+	+	+	+	+
Cleveland					+
Balt.more	+		+		
Atlanta					-
Kansas City					+
Portland	+		+	+	
San Juan	+		+		+
Fort Lauderdale			-		
Jacksonville			-		

*/+ indicates that the urbanized area reported statistic was at least one standard deviation higher than the value computed from the corresponding regression equation.

- indicates that the urbanized area reported statistic was at least one standard deviation lower than the value computed from the corresponding regression equation.

#/Equation numbers correspond to those in Table 1.

of the total variance is used to analyze extreme variation among urbanized areas in their transit plans or forecasts. The urbanized areas for which the reported values differ by more than one standard error from the values predicted by the equation have been marked with a (+) indicating a reported value one standard error or more above the predicted value, or a (-) indicating a reported value one standard error or more below the predicted value.

Observations of exceptions for all modes are shown in Table 11 for all modes. Each column in the table gives the results for one equation. The dependent variable, equation number, and proportion of total variance explained by the equation are indicated at the head of each column. Equations C15, C16, and C17 specify operating costs per capita as a function of seat miles per capita and passenger miles per seat mile. The equations for 1972 and 1980 explain a high proportion of the total variance (as indicated by the R^2 statistic). New York is an exception in 1972 and 1980, and Boston is an exception in 1972. The exceptions, in this case, mean that operating costs per capita are higher for those urbanized areas marked + and lower for those marked - than would be predicted within one standard error of estimate from the equation. Cleveland is high in 1990, Pittsburgh in 1980, and Baltimore in 1990. Seattle is low in both 1972 and 1980, but this probably reflects the effects of the limited monorail system. In general, the smaller urbanized areas tend to deviate above the predicted value, but there are some exceptions to this generalization.

A tabulation of exceptions in the bus analysis is given in Table 12. The tabulation includes six supply equations, four operating cost equations, and two operating revenue equations.

The first three equations (S1, S2, and S3) specify seat miles per capita as a function of square miles per capita, line miles per capita, and vehicles per capita. In these supply equations, the larger urbanized areas tend to deviate below the predicted values, and the medium size areas tend to deviate above the value predicted from the equations. Los Angeles plans a larger supply of bus service than the equation predicts for 1980. Washington, D.C., and St. Louis were low in 1972. Chicago is low in its planning for 1980 and 1990, and San Francisco is low in 1990.

Equations S7 and S8 specify changes in seat miles per capita, from 1972 to 1980 and from 1972 to 1990 respectively, as a function of the change in passenger miles per seat mile, change in square miles per capita, change in line miles per capita, and change in vehicles per capita. Chicago and Baltimore deviate below the values predicted by the equation for both periods, and Washington, D.C., deviates above the predicted value for the period 1972 to 1980.

Equation S31 is a production function which specifies seat miles per capita as a function of capital and labor inputs. New York, Boston, and St. Louis are the largest urbanized areas that are less productive than would be predicted from the equation. Some of the medium size cities expect to be more productive than the equation predicts.

Equations C1, C2, and C3 specify bus operating costs per capita for 1972, 1980, and 1990 as a function of seat miles per capita. In 1972, six of the first seven urbanized areas (except Los Angeles) deviate above the costs predicted by the equation; San Juan also deviates above the prediction. Deviations below the predicted operating cost per capita occur with the medium and smaller urbanized areas. Similar patterns occur for 1980 and 1990, except that there are relatively fewer positive deviations in 1980. There is a reduction in the proportion of variance explained by the equations in 1980 and 1990, which at least partly explains the reduction in the number of deviant urbanized areas.

Equation C7 specifies bus operating cost per capita for 1972 as a function of square miles per capita, line miles per capita, and vehicles per capita. Chicago, Boston, and Pittsburgh are the larger urbanized areas showing operating costs above that which would be predicted by the equation.

Equations R1 and R3 specify bus operating revenue per capita as a function of passenger miles per capita for 1972 and 1990. The larger urbanized areas with revenues above those that would be predicted by the equation for 1972 include New York, Chicago, Washington, D.C., and Cleveland. Conversely, Cincinnati reported 1972 revenues substantially below those that would be predicted by the equation. Of the larger urbanized areas, Washington, D.C., and Cleveland also forecast revenues above that which would be predicted by the equation. Of the smaller areas in the sample, Buffalo and San Juan forecast operating revenues per capita above those that would be forecast by the 1990 equation.

A tabulation of exceptions in the rail analysis is given in Table 13. The tabulation includes four equations for supply and one for operating revenues. Equation S6 specifies seat miles per capita in 1990 as a function of square miles per capita, line miles per capita, vehicles per capita, and a dummy variable distinguishing between rail and commuter rail data. Planned output of seat miles per capita in Los Angeles, Chicago, and Philadelphia is below that which would be predicted by the equation. Washington, D.C., Baltimore, Portland, and San Juan all are planning outputs of seat miles in 1990 at lower levels than those which would be predicted by the equation.

Equations S9 and S10 specify change in seat miles per capita for 1972 to 1980 and 1972 to 1990 as a function of passenger miles per seat mile, change in square miles per capita, change in line miles per capita, change in vehicles per capita, and a dummy variable distinguishing between rail and commuter rail. In equation S9, Washington, D.C., is the only urbanized area deviating more than one standard error from the value predicted by the equation. As indicated by the R^2 statistic, the equation explains only 40 percent of the variance in the data, so the equation is not a very good predictor of change in seat miles per capita for 1972 to 1980. Equation S10 explains more (68 percent) of the variance in the data for change in seat miles per capita for 1972 to 1990. The change in seat miles per capita over this period planned for Los Angeles is below that which would be predicted by the equation. As before, Washington, D.C., is above the predicted value, as are Baltimore, Portland, and San Juan.

Equation S31 is a rail production function for 1990 and specifies seat miles per capita as a function of capital and labor inputs. Detroit, Washington, D.C., and Portland planning reflects expectations for higher productivity than that which would be predicted by the equation; Boston predicts a lower productivity.

Equation R6 specifies rail operating revenue for 1990 as a function of passenger miles per capita and a dummy variable distinguishing between rail and commuter rail. The equation explains only some 44 percent of the variance among the urbanized areas in operating revenue per capita for 1990 and has a wide dispersion of the data around the regression line. Nonetheless, Washington, D.C., Cleveland, Kansas City, San Juan, and Atlanta forecast operating revenue per capita for 1990 below that predicted by the equation.

IV. CONCLUSIONS

There were two major objectives of this Appendix. The first, accomplished in Section II, was to display selected data for individual urbanized areas concerning their planning for urban mass transportation. The second objective, accomplished in Section III, was more analytical. This objective was to investigate whether or not norms for certain physical, financial, and operating characteristics of urban mass transportation planning could be developed from the 1974 NTS data and to determine which of the norms would be usable in evaluating the public transportation plans of individual urbanized areas.

METHOD OF ANALYSIS

The search for a set of norms was based on an analysis of the 1974 NTS data. (These data are composed of submissions from the 52 largest urbanized areas, i.e., those areas that reported 1972 population estimates in excess of 500,000. Together, these areas account for nearly 95 percent of all planned capital investment in mass transportation between 1972 and 1990.)

The analysis of the data proceeded in the following four stages:

- identification of a framework for analysis which simplified the review of NTS data and provided an appropriate representation of the physical, financial, and operating characteristics of the transit industry;

- investigation of the consistency among urbanized areas through analysis of the means, medians, and coefficients of variation of selected NTS statistics relative to local area or transit performance characteristics;
- identification of a set of multivariate relations which might be specified as causal or descriptive determinants of the physical, financial, and operating characteristics of transit operations;
- application of linear regression techniques to estimate the precise nature of these relations and the accuracy with which they might be considered to represent the characteristics of mass transit within the group of urbanized areas studies; and
- identification of the urbanized areas which diverge substantially more than one standard error from the values predicted by the regression equations.

FRAMEWORK OF THE ANALYSIS

The framework of this analysis treated transit as an economic service. Accordingly, the supply of seat miles per capita of various transit systems was associated with specific capital investment and operating expenditures. Similarly, the amount of transit patronage (i.e., passenger miles per capita) was associated with a set of transit services and a price (fare) at which those services are offered. Fare and patronage together, determine transit operating revenues. Thus, the 1974 NTS data were analyzed in terms of supply (using variables for both the level and the quality of supply), demand, acapital investment, operating cost, and operating revenue characteristics.

ECONOMETRIC ANALYSES

In Section III, econometric techniques were applied to an a priori set of relations among variables expressing transit characteristics, in order to extend the search for norms in those characteristics. The idea was to see if variations among urbanized areas in the characteristics of their transit systems could be explained either by variables representing socioeconomic characteristics of urbanized areas, or by other variables representing performance characteristics of transit systems. Only the latter set of variables proved to be statistically significant after transit characteristics were normalized for population. This result probably occurred for two reasons. First, it is reasonable to expect population to be the variable which would be the principal determinant of differences among urbanized areas in transit system characteristics. Consequently, when the data are normalized for population, much of the variance is removed. Secondly, the only other socioeconomic characteristic in the data set was the size, in square miles, of the urbanized area. As a result, the data did not really have enough socioeconomic characteristics of urbanized areas to determine how effective these might be in explaining variations in transit characteristics among urbanized areas.

Estimation of the regression equations demonstrated that some of the explanatory variables were significant determinants of the values of the corresponding transit performance measures. The regression analysis was, in many cases, an improvement over the previous analysis of means, medians, and standard deviations in explaining variations.

Although the econometric analyses yielded a number of interesting results, the overall result was not sufficiently consistent to allow an generalizations about the existence of transit performance norms in the data. The indications are hopeful, but not conclusive.

The results of the analyses lead to two specific recommendations about further research efforts in this direction:

- First, similar analyses should be conducted on a larger sample of actual operating data than was possible in the present analysis, which was limited to 1972 inventory data. A larger data set of operating experience would potentially yield better results.
- Secondly, as has been pointed out before, a larger set of socioeconomic data would potentially be helpful in explaining more of the variance among urbanized areas in transit performance characteristics.