TRANSPORTATION RESEARCH RECORD 928

Improving the Quality and Efficiency of Transportation Data

TRANSPORTATION RESEARCH BOARD

NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1983

Transportation Research Record 928
Price $6.00
Edited for TRB by Susan Singer-Bart

modes
1 highway transportation
2 public transit

subject area
12 planning

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Improving the quality and efficiency of transportation data.

(Transportation research record; 928)
1. Transportation–United States–Congresses. I. National Research Council (U.S.). Transportation Research Board. II. Title. III. Series.
TE7.H5 no. 928 380.5 s [380.5'0973] 84-8322
[HE203]
ISBN 0-309-03618-6 ISSN 0361-1981

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Kenneth W. Heathington, University of Tennessee, chairman

Transportation Forecasting Section
George V. Wickstrom, Metropolitan Washington Council of Governments, chairman

Committee on Transportation Information Systems and Data Requirements
Alan E. Pisarski, Gellman Research Associates, Inc., chairman
Rolf R. Schmitt, Federal Highway Administration, secretary
George A. Bonina, H.W. Bruck, Donald M. Chapman, Kenneth J. Dueker, David J. Dunlap, Robert T. Dunphy, Robert A. Ferlis, James Finane, Alexander French, Philip N. Fulton, Larry R. Goode, William A. Lesansky, Ginger Levin, Joseph M. Manning, James J. McDonnell, Morton A. Meyer, Philip D. Patterson, John L. Skinner, Jr., Arthur B. Sosslau, Robert D. Studer, Robert N. Tap, Jack Williams

Transportation Systems and Planning Innovations Section
Joseph L. Schofer, Northwestern University, chairman

Committee on Statewide Multimodal Transportation Planning
Roger L. Creighton, Roger Creighton Associates, Inc., chairman
Harvey R. Atchison, Robert C. Blensly, W.F. Caddell, Jr., Wm. John Cameron, Robert E. David, Richard E. Esch, Peter A. Fausch, Clarence W. Friesen, Arne L. Gausmann, Philip I. Hazen, Carl A. Hennum, Jack L. Housworth, Thomas F. Humphrey, Wallace L. Larsen, Lowell T. Livingston, Colin Ian MacGillivray, Isaac Shafran, Nat Simons, Jr., Howard L. Slavin, Carl N. Swerdloff, William C. Taylor, Paul F. Wilkinson

James A. Scott, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units, officers, and members are as of December 31, 1982.

928

Contents

Authors of the Papers in This Record

Erlbaum, Nathan S., New York State Department of Transportation, Building 4, State Office Campus, Albany, N.Y. 12232
Fulton, Philip N., Population Division, U.S. Bureau of the Census, Washington, D.C. 20233
Goode, Larry R., North Carolina Department of Transportation, Planning and Research Branch, Raleigh, N.C. 27611
Hartgen, David T., New York State Department of Transportation, Building 4, State Office Campus, Albany, N.Y. 12232
Heimbach, Clinton L., North Carolina State University, Civil Engineering Department, Raleigh, N.C. 27607
Koeppel, K.W. Peter, New York State Department of Transportation, Building 4, State Office Campus, Albany, N.Y. 12232
Lemmerman, John H., New York State Department of Transportation, Building 4, State Office Campus, Albany, N.Y. 12232
McPherson, Larry W., North Carolina Department of Transportation, Planning and Research Branch, Raleigh, N.C. 27611
Yelich, Betty J., New York State Department of Transportation, Building 4, State Office Campus, Albany, N.Y. 12232

Public Transportation: Solving the Commuting Problem?

PHILIP N. FULTON

In this paper journey-to-work data from the 1980 census are used to provide a perspective on how well public transportation is coping with the increasing spatial complexity of metropolitan communities. The data show that about 6 percent of all workers in the United States used some form of public transportation to commute to work. Most commuter use of public transit occurred in the Northeast. Transit use among workers was lowest in the South. About 67 percent of all workers who used public transportation lived in the central city of a metropolitan area. The number of workers who use public transportation to commute to work in the United States dropped by about 487,000 between 1970 and 1980—a decline of approximately 7 percent. Significant declines in transit use occurred in each region except the West. The Northeast experienced the most drastic decline—about 596,000 workers or 17 percent. In contrast, about 378,000 more workers used public transportation in the West in 1980 than in 1970, an increase of 67 percent. The decline in commuter use of public transportation is closely associated with the movement of people and jobs to places where public transportation is not available or easily accessible. The continued shift of the population from the North to the South and West means that the public transportation market is moving from regions that have the most transit service to regions that have the least. Furthermore, the nonmetropolitan sector of the country, where public transit is virtually nonexistent, is growing faster than metropolitan areas. Finally, in recent years within metropolitan areas, the suburbs have far surpassed the central cities in population growth, Many large central cities, where transit is concentrated, have experienced losses of population. In addition, increasing suburbanization of employment and population has resulted in a predominance of lateral commuting in large metropolitan areas—intersuburban work trips for which public transportation is not well acclimated.

Advances in transportation technology have played an integral part in the growth of U.S. cities. During the last half of the 19th century, the horse-drawn street railway allowed dense industrial cities to expand their radius of urban settlement into previously inaccessible territory. The spatial growth of urban centers was further accelerated by the appearance of the electric street railway in the 1890s. The streetcar increased the commuting distance from the commercial and industrial core of the city dramatically. More and more new residences were built in outlying areas, and industries cramped for space in the central business district found relocation to the city's periphery to be an attractive alternative to the congestion that surrounded them.

The appearance of the automobile around the turn of the century marked the evolution from city growth to metropolitan development. In 1900 about 8,000 cars were registered in the United States. By 1925 that number had risen to more than 17 million. The speed and flexibility of the automobile, coupled with the construction of roads and highways to accommodate it, dramatically altered the spatial scale of urban regions through increased accessibility. Deconcentration, both in terms of suburban population growth and the relocation of industries outside the urban center, has typified most of this century, especially the period since World War II. Today the benefits of close urban association no longer depend on proximity but on the automobile, the telephone, and the computer. Advances in personal transportation and communication technology have given rise to the sprawling, postindustrial metropolitan community.

Data from the 1980 census are used to provide a perspective on how well mass transit is coping with the complexity of large metropolitan areas. Trends in commuting patterns are examined and analyzed in the context of changes in population distribution and the location of employment.

Data from the decennial census on means of transportation to work refer to the principal mode of travel that the respondent usually used to get from home to work during the week before enumeration. Persons who used different means of transportation on different days of the week were asked to specify the one they used most often. Persons who used more than one means of transportation to get to work each day were asked to report the one used for the longest distance during the work trip. Census data do not reflect total transit ridership or total trips.

COMMUTER USE OF PUBLIC TRANSPORTATION IN UNITED STATES

Final results from the decennial census indicate that 6 million people, (about 6 percent of all workers in the United States) used some form of public transportation to get to work in 1980 [Table 1 (1)]. The majority of Americans (about 84 percent) used a car, truck, or van. Almost two-thirds of all workers drove to work alone and approximately 20 percent rode in carpools.

Most commuter use of public transportation occurred in the Northeast, where transit is most widely available [Table 2 (1)]. About 14 percent of all workers in that region in 1980 used public transportation compared with about 5 percent of the workers in both the North Central region and the West and about 3 percent of the workers living in the South.

The extent to which public transportation is concentrated in the Northeast is further emphasized by examining the proportion of the nation's transit use that is attributable to each region. Table 2 demonstrates that nearly half of all commuter use of public transportation occurred in the Northeast. Actually the Northeast accounted for more than twice as much transit use as the North Central States and the South and about three times that of the West. Although the West accounted for a smaller proportion of the nation's public transportation users than the South, the rate of transit use was significantly higher in the West. The ratio of the South's share of all public transit riders to its share of the total work force (0.51) indicates that it had the lowest rate of transit use among the four regions.

Public trasportation is also used predominantly by workers who live in the central cities of metropolitan areas [Table 3 (1)]. About 67 percent of the workers who used public transit in 1980 lived in a central city of a standard metropolitan statistical area (SMSA), 30 percent of all transit users lived in the suburbs of an SMSA, and only about 3 percent of the workers who used public transporta-

Table 1. Principal means of transportation to work, 1980.

Means of Transportation	Number (000s)	Percent
All workers	96,617	100.0
Car, truck, or van	81,258	84.1
Drive alone	62,193	64.4
Carpool	19,065	19.7
Public transportation[a]	6,175	6.4
Walked only	5,413	5.6
Other means	1,591	1.6
Worked at home	2,180	2.3

Note: Column totals may differ due to rounding. Workers are 16 years and older.

[a]Category includes bus or streetcar, subway or elevated, railroad, and taxicab.

Table 2. Workers using public transportation by region, 1980.

Region	All Workers in United States Number (000s)	Percent	Workers Using Public Transportation[a] Number (000s)	Percentage of All Workers	Percentage of U.S. Total	Ratio of Percentage of U.S. Public Transportation Users to Percentage of U.S. Workers
Northeast	20,922	21.7	2,973	14.2	48.1	2.22
North Central	24,936	25.8	1,222	4.9	19.8	0.77
South	31,742	32.9	1,036	3.3	16.8	0.51
West	19,018	19.7	945	5.0	15.3	0.78
United States	96,618	100.0	6,176	6.4	100.0	1.00

Note: Column totals may differ due to rounding. Workers are 16 years and older.
[a]Category includes bus or streetcar, subway or elevated, railroad, and taxicab.

Table 3. Workers using public transportation by type of residence, 1980.

Region	All Workers Using Public Transportation[a]		Inside SMSAs						Outside SMSAs	
			Total		Inside Central Cities		Outside Central Cities			
	Number (000s)	Percent	Number (000s)	Percent	Number (000s)	Percent	Number (000s)	Percent	Number (000s)	Percent
Northeast	2,973	100.0	2,936	98.8	2,133	71.8	803	27.0	36	1.2
North Central	1,222	100.0	1,189	97.3	818	66.9	372	30.4	33	2.7
South	1,036	100.0	963	93.0	643	62.1	320	30.9	73	7.0
West	945	100.0	909	96.2	563	59.6	345	36.6	36	3.8
United States	6,175	100.0	5,997	97.1	4,157	67.3	1,840	29.8	178	2.9

Note: Column totals may differ due to rounding. Workers are 16 years and older.
[a]Category includes bus or streetcar, subway or elevated, railroad, and taxicab.

tion lived in nonmetropolitan territory. Transit use was notably higher than the national average in the central cities of the Northeast (72 percent), in the suburbs of the West (37 percent), and in the nonmetropolitan sector of the South (7 percent).

CHANGE IN COMMUTER USE OF PUBLIC TRANSPORTATION: 1970 TO 1980

Table 4 (1-4) presents changes in the use of public transportation that have taken place between 1970 and 1980 in the United States, the four regions, and in each SMSA that has a population of 1 million or more. The number of workers using public transportation to get to work dropped by about 487,000 between 1970 and 1980--a decline of approximately 7 percent. In 1970 about 9 percent of all workers used public transit compared with about 6 percent in 1980.

Significant declines in transit use occurred in each region except the West. The Northeast had the most drastic decline--about 596,000 workers or 17 percent. Use of public transportation declined by about 13 percent in the North Central region and by 7 percent in the South. In marked contrast with the other regions, commuter use of transit increased substantially in the West during the decade. About 378,000 more workers used public transportation in the West in 1980 than in 1970--an increase of 67 percent.

Public transit use in all of the large metropolitan areas in the Northeast declined between 1970 and 1980; in most cases quite substantially. The largest absolute decline (about 355,000 workers) occurred in the New York SMSA. Each of the large SMSAs in the North Central region also experienced an absolute decline, with the exception of Minneapolis-St. Paul. The number of workers using public transportation in the Twin Cities area grew by about 24,000--an increase of 36 percent. Furthermore, transit maintained its share of the region's commuter travel market at about 8.5 percent.

Transit use in the South increased in all but three of the large SMSAs (Baltimore, Tampa-St. Petersburg, and New Orleans). Washington, D.C., and Atlanta showed the largest gains of 38,000 and 17,000 riders, respectively. Yet, in every metropolitan area that experienced a numerical increase in public transportation use, the share of the labor force using transit actually decreased. Commuter use of public transit declined by 21,000 workers in the New Orleans SMSA and by about 13,000 workers in the Baltimore SMSA.

Dramatic increases in public transportation ridership occurred in all the large SMSAs in the West except San Diego. In San Diego the number of workers using transit increased by 21 percent, but transit's share of the market declined from 4.3 percent to 3.3 percent. The two most populous areas, Los Angeles-Long Beach and San Francisco-Oakland, both showed healthy increases in transit use, and in six other areas the number of workers using public transportation more than doubled during the decade.

The nation's large SMSAs, taken as a group, accounted for about 81 percent of all commuter use of public transportation in the United States in 1980 [Table 5 (1)]. New York alone accounted for about 28 percent of the national total; Chicago was a distant second at about 9 percent. Chicago, in turn, had more than twice as many commuters using public transit as the Philadelphia SMSA, the third-ranked area. The top six SMSAs, ranked on the basis of the number of workers that use mass transit to get to work, contained more than half of all transit commuters in the United States.

If New York is excluded from the analysis, the census results show that the number of workers using public transportation elsewhere in the country dropped by about 132,000 workers between 1970 and 1980--a decline of approximately 3 percent. Exclud-

Table 4. Workers using public transportation for regions and SMSAs of 1 million or more, 1980 and 1970.

SMSA by Region	Number of Workers Using Public Transportation[a] (000s) 1980	1970	Change, 1970 to 1980 Number (000s)	Percent	Percentage of Workers Using Public Transportation[a] 1980	1970
Northeast	2,973	3,569	-596	-16.7	14.2	19.1
New York, N.Y.-N.J.	1,711	2,067	-355	-17.2	45.1	52.5
Philadelphia, Pa.-N.J.	275	383	-108	-28.1	14.0	20.7
Boston, Mass.	204	224	-19	-8.7	15.6	19.7
Nassau-Suffolk, N.Y.	142	145	-3	-1.7	12.5	15.5
Pittsburgh, Pa.	105	123	-18	-14.9	11.5	14.6
Newark, N.J.	94	143	-49	-34.4	10.7	17.4
Buffalo, N.Y.	33	52	-19	-36.3	6.6	10.5
North Central	1,222	1,409	-187	-13.3	4.9	6.7
Chicago, Ill.	568	650	-82	-12.6	18.0	23.3
Detroit, Mich.	64	125	-61	-49.0	3.7	7.9
St. Louis, Mo.-Ill.	57	71	-14	-19.5	5.7	8.0
Minneapolis-St. Paul, Minn.-Wis.	91	67	24	35.8	8.7	8.5
Cleveland, Ohio	87	107	-20	-18.9	10.6	13.4
Cincinnati, Ohio-Ky.-Ind.	38	42	-4	-8.8	6.5	8.3
Milwaukee, Wis.	49	67	-18	-26.2	7.7	12.0
Kansas City, Mo.-Kans.	25	28	-3	-10.1	4.1	5.4
Indianapolis, Ind.	17	25	-8	-32.9	3.2	5.8
Columbus, Ohio	23	29	-6	-21.9	4.6	7.4
South	1,036	1,117	-82	-7.3	3.3	5.0
Washington, D.C.-Md.-Va.	241	203	38	18.5	15.5	16.3
Dallas-Fort Worth, Tex.	50	50	1	1.4	3.4	5.1
Houston, Tex.	43	42	1	2.4	3.0	5.4
Baltimore, Md.	99	113	-13	-11.8	10.3	13.8
Atlanta, Ga.	72	55	17	30.4	7.6	8.4
Miami, Fla.	48	46	3	5.9	6.6	9.1
Tampa-St. Petersburg, Fla.	11	11	-[b]	-1.9	1.8	3.0
New Orleans, La.	53	74	-21	-28.4	10.9	20.4
San Antonio, Tex.	21	18	2	11.6	4.6	5.6
Ft. Lauderdale-Hollywood, Fla.	8	5	4	84.7	2.0	2.1
West	945	567	378	66.6	5.0	4.6
Los Angeles-Long Beach, Calif.	235	154	81	52.5	7.0	5.6
San Francisco-Oakland, Calif.	256	194	63	32.3	16.4	15.5
Anaheim-Santa Ana-Garden Grove, Calif.	20	2	18	764.6	2.1	0.4
San Diego, Calif.	28	23	5	20.8	3.3	4.3
Denver-Boulder, Colo.	50	22	28	129.2	6.1	4.4
Seattle-Everett, Wash.	74	38	36	94.4	9.6	7.1
Riverside-San Bernardino-Ontario, Calif.	6	3	2	64.5	0.9	0.9
Phoenix, Ariz.	13	5	9	188.8	2.0	1.3
San Jose, Calif.	20	9	11	118.0	3.1	2.3
Portland, Oreg.-Wash.	48	23	24	104.3	8.4	6.0
Sacramento, Calif.	15	7	9	124.8	3.5	2.3
United States	6,175	6,662	-487	-7.3	6.4	9.0

Note: Column totals may differ due to rounding. Workers are 16 years and older.
[a]Category includes bus or streetcar, subway or elevated, railroad, and taxicab.
[b]Less than 500.

ing New York, Philadelphia, and Chicago (SMSAs that lost a combined total of 545,000 transit commuters during the decade), the number of workers using public transit elsewhere increased by about 57,000 workers from 1970 to 1980--an increase of only about 2 percent. Transit's overall share of the commuter market with these three SMSAs omitted declined from 5.4 percent in 1970 to 4.1 percent in 1980.

Analysis

The changes that occurred during the decade 1970 to 1980 in the use of public transportation for commuting to work appear to be closely associated with changes in population distribution, shifts in the location of employment, and the types of commuting patterns that result from these spatial modifications. People and jobs are moving to places where, historically, public transportation has either not been available or has not been accessible. Trends in transit use, in large measure, reflect the extent to which transit service has adapted to these changes.

Changes in Population Distribution

The continuing shift of the population from the North to the South and West means that, as a whole, the public transportation market is moving from regions of the country that have the most transit service to those that have the least [Table 6 (5)]. The population in the West and South grew between 1970 and 1980 by about 24 percent and 20 percent, respectively; however, in the North Central States the population grew by only 4 percent and in the Northeast the population grew by a mere 0.2 percent. Also, from 1970 to 1980 population in the nonmetropolitan sector of the nation, where public transit is virtually nonexistent, grew at the expense of metropolitan areas. The population outside SMSAs grew by about 15 percent, compared with a 10 percent growth rate within SMSAs. Moreover, large SMSAs in the South and West, which typically have less-developed transit systems, grew much faster than those in the Northeast and North Central regions, where many transit-oriented SMSAs actually suffered a decrease in their populations.

Changes in population distribution at the local level appear to have the greatest effect on public transportation ridership. Within SMSAs suburban population growth far surpassed that of the central cities. In the Northeast and North Central regions central cities lost population and the suburbs grew moderately. In the South and West rapid suburban population growth outpaced that of the central cities. Thus, the population balance moved increasingly away from the metropolitan core, where the most mass transit service is available, and toward the suburban fringe, where less public transportation is provided. Table 7 (1-4) compares the population changes that occurred between 1970 and 1980 with changes in commuter use of public transportation for each SMSA of 1 million or more.

In the Northeast SMSAs public transportation use declined in both the central cities and suburbs of every large metropolitan area. Central cities lost a greater proportion of their ridership than did the suburbs in each SMSA. These trends closely parallel the demographic shifts that occurred during the decade, with the central cities experiencing a substantial loss of population. Viewed in this context, the decline in transit ridership in these areas is not surprising.

Most of the large SMSAs in the North Central region exhibited the same pattern as those in the Northeast; i.e., a decline in public transportation use in both the central cities and the suburbs. Again, the decline in transit use in the central

Table 5. Rank of SMSAs of 1 million or more population by number of workers that use public transportation, 1980.

		Workers Using Public Transportation[a]			
SMSA	All Workers (000s)	Number (000s)	Percentage of All Workers	Percentage of U.S. Total	Rank
New York, N.Y.-N.J.	3,792	1,711	45.1	27.7	1
Chicago, Ill.	3,163	568	18.0	9.2	2
Philadelphia, Pa.-N.J.	1,959	275	14.0	4.5	3
San Francisco-Oakland, Calif.	1,562	256	16.4	4.2	4
Washington, D.C.-Md.-Va.	1,553	241	15.5	3.9	5
Los Angeles-Long Beach, Calif.	3,374	235	7.0	3.8	6
Boston, Mass.	1,308	204	15.6	3.3	7
Nassau-Suffolk, N.Y.	1,140	142	12.5	2.3	8
Pittsburgh, Pa.	912	105	11.5	1.7	9
Baltimore, Md.	967	99	10.3	1.6	10
Newark, N.J.	879	94	10.7	1.5	11
Minneapolis-St. Paul, Minn.-Wis.	1,046	91	8.7	1.5	12
Cleveland, Ohio	820	87	10.6	1.4	13
Seattle-Everett, Wash.	776	74	9.6	1.2	14
Atlanta, Ga.	948	72	7.6	1.2	15
Detroit, Mich.	1,710	64	3.7	1.0	16
St. Louis, Mo.-Ill.	1,004	57	5.7	0.9	17
New Orleans, La.	485	53	10.9	0.9	18
Dallas-Fort Worth, Tex.	1,465	50	3.4	0.8	19
Denver-Boulder, Colo.	809	50	6.1	0.8	20
Milwaukee, Wis.	642	49	7.7	0.8	21
Miami, Fla.	726	48	6.6	0.8	22
Portland, Oreg,-Wash.	568	48	8.4	0.8	23
Houston, Tex.	1,415	43	3.0	0.7	24
Cincinnati, Ohio-Ky.-Ind.	588	38	6.5	0.6	25
Buffalo, N.Y.	500	33	6.6	0.5	26
San Diego, Calif.	855	28	3.3	0.5	27
Kansas City, Mo.-Kans.	618	25	4.1	0.4	28
Columbus, Ohio	488	23	4.6	0.4	29
San Antonio, Tex.	450	21	4.6	0.3	30
San Jose, Calif.	650	20	3.1	0.3	31
Anaheim-Santa Ana-Garden Grove, Calif.	962	20	2.1	0.3	32
Indianapolis, Ind.	523	17	3.2	0.3	33
Sacramento, Calif.	435	15	3.5	0.2	34
Phoenix, Ariz.	659	13	2.0	0.2	35
Tampa-St. Petersburg, Fla.	606	11	1.8	0.2	36
Fort Lauderdale-Hollywood, Fla.	424	8	2.0	0.1	37
Riverside-San Bernardino-Ontario, Calif.	614	6	0.9	0.1	38
All SMSAs of 1 million or more	41,397	4,995	12.1	80.9	
Elsewhere	55,220	1,180	2.1	19.1	
U.S. Total	96,617	6,175	6.4	100.0	

Note: Column totals may differ due to rounding. Workers are 16 years and older.
[a]Category includes bus or streetcar, subway or elevated, railroad, and taxicab.

Table 6. Change in workers that use public transportation and change in population for regions by type of residence, 1970 to 1980.

		Percentage Change in Population, 1970 to 1980				
			Inside SMSAs			
Region	Percentage Change in Workers Using Public Transportation[a]	Total	Total	Inside Central Cities	Outside Central Cities	Outside SMSAs
Northeast	-16.7	0.2	-1.9	-10.5	4.3	13.5
North Central	-13.3	4.0	2.7	-9.2	12.0	7.4
South	-7.3	20.0	21.4	8.7	32.9	17.1
West	66.6	23.9	22.6	15.3	28.0	30.6
United States	-7.3	11.4	10.2	0.1	18.2	15.1

[a]Includes bus or streetcar, subway or elevated, railroad, and taxicab.

Table 7. Change in workers that use public transportation and change in population for SMSAs of 1 million or more by region and type of residence, 1970 to 1980.

SMSA	Number of Workers Using Public Transportation[a] (000s)		Change in Workers Using Public Transportation, 1970 to 1980		Percentage Change in Population, 1970 to 1980
	1980	1970	Number (000s)	Percent	
Northeast					
New York, N.Y.-N.J.	1,711	2,067	-355	-17.2	-8.6
Inside central city	1,597	1,915	-338	-17.6	-10.4
Outside central city	134	152	-18	-11.7	-1.4
Philadelphia, Pa.-N.J.	275	383	-108	-28.1	-2.2
Inside central city	183	274	-91	-33.0	-13.4
Outside central city	92	109	-17	-15.8	5.4
Boston, Mass.	204	224	-19	-8.7	-4.7
Inside central city	84	102	-17	-17.1	-12.2
Outside central city	120	122	-2	-1.7	-2.6
Nassau-Suffolk, N.Y.	142	145	-3	-1.7	2.0
Inside central city	...	...	...	...	...
Outside central city	142	145	-3	- 1.7	2.0
Pittsburgh, Pa.	105	123	-18	-14.9	-5.7
Inside central city	47	55	-9	-15.7	-18.5
Outside central city	58	68	-10	-14.3	-2.2
Newark, N.J.	94	143	-49	-34.4	-4.5
Inside central city	28	49	-22	-43.9	-13.8
Outside central city	66	94	-27	-29.4	-2.4
Buffalo, N.Y.	33	52	-19	-36.3	-7.9
Inside central city	22	36	-14	-40.0	-22.7
Outside central city	11	16	-5	-28.6	-0.2
North Central					
Chicago, Ill.	568	650	-82	-12.6	1.8
Inside central city	386	484	-98	-20.2	-10.8
Outside central city	183	167	-16	-9.7	13.6
Detroit, Mich.	64	125	-61	-49.0	-1.9
Inside central city	44	98	-54	-55.1	-20.5
Outside central city	20	27	-7	-26.3	7.8
St. Louis, Mo.-Ill.	57	71	-14	-19.5	-2.3
Inside central city	31	48	-16	-34.2	-27.2
Outside central city	26	24	2	10.0	6.3
Minneapolis-St. Paul, Minn.-Wis.	91	67	24	35.8	7.6
Inside central city	59	53	6	10.7	-13.9
Outside central city	32	13	18	134.7	20.6
Cleveland, Ohio	87	107	-20	-18.9	-8.0
Inside central city	40	61	-21	-34.3	-23.6
Outside central city	46	45	1	1.9	0.9
Cincinnati, Ohio-Ky.-Ind.	38	42	-4	-8.8	1.0
Inside central city	23	26	-3	-11.3	-15.0
Outside central city	15	16	-1	-4.8	8.8
Milwaukee, Wis.	49	67	-18	-26.2	-0.5
Inside central city	39	56	-17	-29.6	-11.3
Outside central city	10	11	-1	-8.9	10.8
Kansas City, Mo.-Kans.	25	28	-3	-10.1	4.2
Inside central city	19	22	-2	-11.2	-11.7
Outside central city	6	6	-[b]	-6.4	14.6
Indianapolis, Ind.	17	25	-8	-32.9	5.0
Inside central city	16	23	-8	-32.7	-4.9
Outside central city	1	2	-1	-36.1	24.5
Columbus, Ohio	23	29	-6	-21.9	7.4
Inside central city	18	25	-6	-25.7	4.6
Outside central city	4	4	-[b]	-0.4	10.6
South					
Washington, D.C.-Md.-Va.	241	203	38	18.5	5.2
Inside central city	112	126	-14	-11.1	-15.7
Outside central city	129	77	52	66.9	12.5
Dallas-Forth Worth, Tex.	50	50	1	1.4	25.1
Inside central city	44	45	-1	-2.9	4.1
Outside central city	6	4	2	50.2	47.9
South (continued)					
Houston, Tex.	43	42	1	2.4	45.3
Inside central city	38	39	-1	-2.4	29.2
Outside central city	5	3	2	72.5	71.2
Baltimore, Md.	99	113	-13	-11.8	5.0
Inside central city	75	92	-17	-18.5	-13.1
Outside central city	24	20	4	18.4	19.1
Atlanta, Ga.	72	55	17	30.4	27.2
Inside central city	42	43	-2	-3.7	-14.1
Outside central city	30	12	18	156.4	45.8
Miami, Fla.	48	46	3	5.9	28.3
Inside central city	21	25	-4	-14.4	3.6
Outside central city	27	21	6	29.8	37.1
Tampa-St. Petersburg, Fla.	11	11	-[b]	-1.9	44.2
Inside central city	7	9	-2	-20.6	2.9
Outside central city	4	2	2	75.5	78.4
New Orleans, La.	53	74	-21	-28.4	13.4
Inside central city	45	64	-19	-29.8	-6.1
Outside central city	8	10	-2	-20.0	38.9
San Antonio, Tex.	21	18	2	11.9	20.7
Inside central city	19	17	2	10.5	20.1
Outside central city	2	1	-[b]	27.9	22.4
Fort Lauderdale-Hollywood, Fla.	8	5	4	84.7	63.5
Inside central city	4	3	1	47.8	9.7
Outside central city	5	2	3	131.6	99.0
West					
Los Angeles-Long Beach, Calif.	235	154	81	52.5	6.2
Inside central city	156	111	45	40.0	5.0
Outside central city	79	43	36	85.0	7.2
San Francisco-Oakland, Calif.	256	194	63	32.3	4.6
Inside central city	160	137	23	16.8	-5.5
Outside central city	97	57	40	69.4	10.0
Anaheim-Santa Ana-Garden Grove, Calif.	20	2	18	764.6	35.9
Inside central city	7	1	6	738.9	23.8
Outside central city	13	1	12	778.6	41.4
San Diego, Calif	28	23	5	20.8	37.1
Inside central city	18	16	3	16.6	25.5
Outside central city	10	8	2	29.2	49.4
Denver-Boulder, Colo.	50	22	28	129.2	30.7
Inside central city	29	18	11	62.5	-2.3
Outside central city	21	4	17	430.3	59.9
Seattle-Everett, Wash.	74	38	36	94.4	12.8
Inside central city	49	33	15	45.8	-6.2
Outside central city	26	5	21	421.3	26.0
Riverside-San Bernardino-Ontario, Calif.	6	3	2	64.5	36.7
Inside central city	2	1	1	48.0	21.4
Outside central city	3	2	1	77.4	42.4
Phoenix, Ariz.	13	5	9	188.8	55.3
Inside central city	10	4	6	155.0	30.9
Outside central city	3	1	3	378.0	92.1
San Jose, Calif.	20	9	11	118.0	21.6
Inside central city	10	3	7	204.5	38.4
Outside central city	10	6	4	69.8	8.8
Portland, Oreg.-Wash.	48	23	24	104.3	23.3
Inside central city	27	17	10	59.1	-3.6
Outside central city	21	6	14	222.4	39.6
Sacramento, Calif.	15	7	9	124.8	26.2
Inside central city	6	5	1	14.4	7.2
Outside central city	9	2	8	464.1	35.0

Note: Column totals may differ due to rounding. Workers are 16 years and older.
[a]Includes bus or streetcar, subway or elevated, railroad, and taxicab.
[b]Less than 500.

city appears to be closely associated with a significant loss of population. The striking exception to the rule among large SMSAs in the North Central region (and in the entire northern United States) is Minneapolis-St. Paul, where transit use increased by 36 percent between 1970 and 1980. Most of the growth took place in the suburbs, where the number of commuters using public transportation more than doubled (an increase of about 18,000 workers). The Twin Cities area experienced a sizable growth in its suburban population, which partly accounts for this increase in the number of transit users. In addition, according to planners in the region, the suburbanization was accompanied by an aggressive public transportation program in that sector. Frequency and coverage of transit service within the suburbs were expanded significantly to give commuters an alternative to their automobiles. Intersuburban express-bus service was developed, and special transit programs for the elderly and handicapped were introduced. Fares were also kept constant throughout most of the decade to encourage transit ridership; however, the subsidy levels required to maintain low fares increased markedly during the 10-year period.

In the South, with the exception of the San Antonio, Ft. Lauderdale-Hollywood, and New Orleans SMSAs, every large metropolitan area experienced a decline in transit use in the central city and a corresponding increase in transit use in the suburbs. These areas also showed a decline in their central city populations and growth in the population in the suburbs. In the Washington, D.C., area, where the new Metro rail system began operating during the decade, a substantial increase of almost 52,000 workers using transit in the suburbs offset what appears to have been a market-related decrease in central city ridership. This gave the area an overall 19 percent rise in use of mass transit by commuters. Similarly, in the Atlanta area, where the Metropolitan Atlanta Rapid Transit Authority (MARTA) brought a large influx of additional suburban bus routes coupled with rail service, suburban transit use increased more than twofold and commuting by public transportation increased by 30 percent.

In San Antonio, where population increased rapidly in both the central city and the suburbs, and in Ft. Lauderdale-Hollywood, where suburban population nearly doubled during the decade, only small increases in transit use occurred. Other rapidly growing areas with limited mass transit service--Dallas-Ft. Worth, Houston, and Tampa-St. Petersburg--also showed little or no growth in transit use. In the New Orleans area the population shifted from the central city to the suburbs, and a decline in transit use was experienced in both geographic sectors.

Finally, in the West, where commuter use of public transportation increased substantially, each of the large SMSAs (with the exception of San Jose) experienced a proportionately greater increase in transit use in the suburbs than in the central cities. San Jose, with central city growth due to population increase and annexation, had a higher rate of transit growth in the central city than in the suburbs. The changes in the other SMSAs generally appear to be associated with large increases in suburban population. Additional explanations offered by planners in the various localities include expansion of bus service in the suburbs in response to population shifts, modification of routes to accommodate changing commuting patterns, higher vehicle productivity, and readily available funding for system development and fare subsidies.

Although they each lost central city population during the decade, the Denver-Boulder, Seattle-Everett, Portland, and San Francisco-Oakland SMSAs all showed substantial gains in transit use in the central city in addition to increases in the suburbs. Even the Los Angeles-Long Beach metropolitan area, so often maligned for its automobile orientation, had an increase of about 81,000 workers using public transportation, which represents an increase of 53 percent in regionwide transit use during the decade.

Changes in Location of Employment and Commuting Patterns

Another factor that is associated with the commuter use of public transportation is shifts in the location of employment. Deconcentration of industry into the suburbs has continued for most of this century, especially since World War II. Suburbanization of employment, coupled with an increasingly large proportion of the population residing outside central cities, is changing the nature of commuting patterns within the metropolitan areas of the United States.

The majority of workers are now lateral commuters; i.e., they both live and work in the suburbs. The Census Bureau's last national study of commuting (6) reported that about 18 million workers lived and worked in the suburbs, about 16 million lived and worked within central cities, and about 9 million lived in the suburbs and commuted to the city to work. Another 4 million workers were reverse commuters, living in the central city and commuting to a job in the suburbs. Why are these differentials important? Because most public transportation is not geared to intersuburban travel. The traditional function of public transportation has been to move people within the congestion of the city and to get suburban residents downtown and back. Table 8 (6) shows that in 1975 about 16 percent of all workers who lived and worked within an SMSA central city and about 10 percent of those who commuted from the suburbs to the city used public transportation to get to work. In contrast, only about 2 percent of the intersuburban commuters used transit.

Detailed place-of-work data from the 1980 census are not yet available, but findings from journey-to-work surveys conducted by the Census Bureau for the U.S. Department of Transportation during the 1970s provide evidence of the trends that have occurred in commuting patterns among the nation's large metropolitan areas (Tables 9-11). The first two columns of the tables show that the proportion of workers who live and work within the central city (those who have the highest rate of transit use) has decreased between 1970 and the most recent survey in every

Table 8. Workers in metropolitan commuting flows by public transportation use, 1975.

		Workers Using Public Transportation[a]	
Commuting Flow	All Workers (000s)	Number (000s)	Percent
Living in SMSA central cities			
Working in central city of same SMSA	16,338	2,674	16.4
Working in same SMSA, outside central cities	3,724	209	5.6
Living in SMSA outside central cities			
Working in central city of same SMSA	8,932	862	9.7
Working in same SMSA, outside central cities	18,001	318	1.8

Note: Column totals may differ due to rounding. Workers are 14 years and older.
[a]Includes bus or streetcar, subway or elevated, railroad, and taxicab.

Table 9. Percentage of resident workers in selected commuting flows and percentage of resident workers working outside central cities for selected SMSAs, 1977 and 1970.

Survey Group[a]	SMSA Resident Workers Who Live and Work Inside Central Cities (%)		SMSA Resident Workers Who Live and Work Outside Central Cities (%)		SMSA Resident Workers Who Work Outside Central Cities (%)	
	1977	1970	1977	1970	1977	1970
Los Angeles-Long Beach, Calif.	31.8	33.0	38.9	35.8	50.3	47.4
Detroit, Mich.	17.5	22.4	56.9	48.6	66.7	59.9
Dallas-Forth Worth, Tex.[b]	38.7	44.0	27.1	25.7	34.3	33.6
Boston, Mass.	16.1	17.1	56.3	54.3	61.4	59.1
Pittsburgh, Pa.	13.6	17.2	63.8	57.6	68.2	62.0
Minneapolis-St. Paul, Minn.	23.3	33.5	44.9	32.3	53.3	40.5
Newark, N.J.	6.6	9.7	59.1	55.8	63.0	60.6
Anaheim-Santa Ana-Garden Grove, Calif.	13.2	15.5	40.2	34.9	52.1	44.5
Phoenix, Ariz.	44.9	53.0	28.5	24.2	38.1	33.0

Notes: Workers 14 years and older.
SMSAs listed within survey group by population size.
SMSA definition is as of the 1970 census.
Data are from the Journey-to-Work Supplement to the Annual Housing Survey; various reports; the Census of Population and Housing, 1970; and Census Tracts series reports.

[a]Among the SMSAs of 1 million or more in 1980, Tampa-St. Petersburg, Ft. Lauderdale-Hollywood, and San Jose were not surveyed. Also, the Nassau-Suffolk SMSA was part of the New York SMSA in 1970, and other SMSAs have changed their definitions since 1970.
[b]Data for the 1970 Dallas and Fort Worth SMSAs are combined.

Table 10. Percentage of resident workers in selected commuting flows and percentage of resident workers working outside central cities for selected SMSAs, 1976 and 1970.

Survey Group[a]	SMSA Resident Workers Who Live and Work Inside Central Cities (%)		SMSA Resident Workers Who Live and Work Outside Central Cities (%)		SMSA Resident Workers Who Work Outside Central Cities (%)	
	1976	1970	1976	1970	1976	1970
New York, N.Y.	60.1	62.7	25.8	21.9	28.0	24.3
Houston, Tex.	52.2	57.4	22.9	18.7	29.5	24.4
St. Louis, Mo.-Ill.	14.4	19.2	60.7	52.0	65.6	57.1
Baltimore, Md.	23.8	30.5	42.1	37.3	51.8	46.9
Cleveland, Ohio	18.4	24.7	45.0	36.5	53.1	44.0
Denver, Colo.	25.5	35.7	42.9	27.8	49.9	39.4
Seattle-Everett, Wash.	32.9	37.9	36.3	29.2	41.9	34.7
Buffalo, N.Y.	18.6	24.1	54.5	45.9	61.6	54.7
Indianapolis, Ind.	57.5	53.7	18.9	17.2	23.7	29.4
Sacramento, Calif.	19.9	25.3	47.0	35.3	54.5	40.8

Notes: Workers 14 years and older.
SMSAs listed within survey group by population size.
SMSA definition is as of the 1970 census.
Data are from the Journey-to-Work Supplement to the Annual Housing Survey; various reports; the Census of Population and Housing, 1970; and Census Tracts series reports.

[a]Among the SMSAs of 1 million or more in 1980, Tampa-St. Petersburg, Ft. Lauderdale-Hollywood, and San Jose were not surveyed. Also, the Nassau-Suffolk SMSA was part of the New York SMSA in 1970, and other SMSAs have changed their definitions since 1970.
[b]Data for the 1970 Dallas and Fort Worth SMSAs are combined.

SMSA except Indianapolis. Conversely, the middle two columns show that the proportion of workers who live and work in the suburbs (those who have the lowest rate of transit use) has increased in each SMSA except San Bernardino-Riverside-Ontario. The last two columns of the tables show that in all the SMSAs except Indianapolis and San Bernardino the percentage of the area's commuting trips that ends in the suburbs has increased since 1970. The central cities in both of these SMSAs were active in annexation after 1970, which may affect the comparability of the census and survey data.

Availability Versus Consumer Choice

Preliminary results from the 1980 Annual Housing Survey add a further perspective on the decline in commuter use of public transportation. Respondents who used a car, truck, or van to get to work were asked to specify the main reason why they did not use public transportation instead. Their responses are given in Table 12 (7).

Almost half of those who used a car, truck, or van said that they did not use public transportation because it was not available. Only about 13 percent said that they simply prefer to use their own private vehicle. Another group of reasons, including "Available transit does not go to place of work," "Takes too long to get to work," "Time schedule is not convenient," and "Transit stop is too far from residence," provides an indication that the available public transportation systems are not meeting the commuters' needs. Taken together, they represent about 25 percent of the respondents. Thus, in sum, about 75 percent of all the workers who commute in a private vehicle did not use public transportation because it was either not available or it could not get them to work conveniently. Less than 1 percent of the respondents said that they did not use public transportation because it was too expensive.

POLICY IMPLICATIONS

Transportation policies should be determined by the problems and circumstances that are unique to each locality. The results of the 1980 census highlight

Table 11. Percentage of resident workers in selected commuting flows and percentage of resident workers working outside central cities for selected SMSAs, 1975 and 1970.

	SMSA Resident Workers Who Live and Work Inside Central Cities (%)		SMSA Resident Workers Who Live and Work Outside Central Cities (%)		SMSA Resident Workers Who Work Outside Central Cities (%)	
Survey Group[a]	1975	1970	1975	1970	1975	1970
Chicago, Ill.	34.8	38.2	43.0	38.1	50.0	46.0
Philadelphia, Pa.-N.J.	29.4	33.0	51.0	43.9	55.5	48.6
San Francisco-Oakland, Calif.	26.0	29.9	46.6	43.0	51.3	47.9
Atlanta, Ga.	19.2	26.3	45.7	36.9	52.1	43.9
San Diego, Calif.	41.5	43.7	32.2	30.9	40.0	38.7
Miami, Fla.	13.7	16.0	54.1	46.6	66.5	58.1
San Bernardino-Riverside-Ontario, Calif.	17.6	17.8	49.7	51.0	56.7	58.8
Cincinnati, Ohio-Ky.-Ind.	20.8	23.8	44.6	38.0	52.1	46.1
Milwaukee, Wis.	34.2	38.7	34.8	30.2	47.2	42.2
Kansas City, Mo.-Kans.	29.7	30.8	38.0	34.8	45.9	42.7
Portland, Oreg.-Wash.	27.2	30.5	39.1	34.9	45.7	42.2
New Orleans, La.	42.5	47.7	29.3	24.4	35.8	31.2
Columbus, Ohio	44.3	45.9	20.7	18.6	30.6	29.4
San Antonio, Tex.	54.5	62.6	15.1	10.5	30.5	18.3

Notes: Workers 14 years and older.
SMSAs listed within survey group by population size.
SMSA definition is as of the 1970 census.
Data are from the Journey-to-Work Supplement to the Annual Housing Survey; various reports; the Census of Population and Housing, 1970; and Census Tracts series reports.

[a]Among the SMSAs of 1 million or more in 1980, Tampa-St. Petersburg, Ft. Lauderdale-Hollywood, and San Jose were not surveyed. Also, the Nassau-Suffolk SMSA was part of the New York SMSA in 1970, and other SMSAs have changed their definitions since 1970.
[b]Data for the 1970 Dallas and Fort Worth SMSAs are combined.

Table 12. Workers who use car, truck, or van to get to work by main reason for not using public transportation and type of residence, 1980.

		Inside SMSAs			
Reasons for Not Using Public Transportation	All Workers Using Car, Truck, or Van[a] (N=75,525 (000s)	Total (N=51,900) (000s)	Inside Central Cities (N=18,200 (000s)	Outside Central Cities (N=33,702 (000s)	Outside SMSAs (N=23,623) (000s)
Rather use a car, truck, or van	12.7	16.4	24.8	11.9	4.4
Available transit does not go to place of work	10.4	13.3	13.3	13.4	4.0
Takes too long to get to work	5.1	7.2	10.9	5.2	0.6
Time schedule is not convenient	8.4	11.1	15.0	8.9	2.4
Public transportation is not available	49.4	34.7	15.9	44.9	81.7
Transit stop is too far from residence	1.2	1.6	1.6	1.6	0.3
Too expensive	0.6	0.6	0.9	0.5	0.4
Need car, truck, or van for work	8.9	10.9	11.9	10.3	4.5
Physical or mental impairment	0.2	0.2	0.4	0.1	-[b]
Other reason	3.2	4.0	5.3	3.2	1.5

Note: Workers are 14 years and older.
[a]Excludes workers who did not report their main reason for not using public transportation.
[b]Less than 0.1 percent.

differences in the demographic changes that mass transit must cope with in each large metropolitan area. At the same time, they also point to some striking similarities.

Areawide population losses during the past decade have resulted in an overall drop in the potential transit market in the older, traditionally more-transit-oriented SMSAs in the Northeast. These losses have been especially significant in the central cities, where most transit service is concentrated. Where the population declines have been smaller or, in some SMSAs, where population has increased slightly in the suburbs, the balance of the population is shifting outward into areas that have less transit service. This shift is even more distinct in large SMSAs of the North Central region, where more substantial suburban population growth has taken place.

Transit service in the South has not kept pace with the rapid population growth that has taken place in many of the metropolitan areas. Most of this growth has taken place in the suburbs. The suburbs in large metropolitan areas of the West also accounted for most of the dramatic growth in population that has occurred in those areas since 1970. In the West, however, mass transit has also made dramatic strides to serve the booming suburban labor force.

The continuing shift of the population toward areas that have limited transit service, both at the regional and local levels, implies a greater reliance on the automobile as an increasingly smaller proportion of the population has access to conventional public transportation. In order for transit to maintain or increase its share of the commuter market, the census results indicate that the geographic coverage of transit service must be increased, particularly in the rapidly growing metropolitan areas in the South and West and in the suburbs of all large SMSAs.

Given the low-density dispersion of residences and work places, policymakers need to maintain realistic expectations of what public transportation can accomplish in the urban environment in which it exists. Public transportation policies should be formulated in coordination with and in anticipation of demographic trends instead of depending on the alteration of such trends to achieve success.

The results of the census underscore the deterministic effect that demographic changes have on the use of public transportation for commuting to work in the United States. Planners have long made the fundamental assumption that mass transit would provide the ultimate remedy to urban transportation problems by reshaping urban form and by modifying consumer behavior. On the contrary, the principal lesson to be learned from the census is that, for transit to retain its public, it must better adapt itself to the changes in urban form and consumer preference that are taking place around it.

Discussion

Joel Markowitz*

Fulton has done a service to the transportation planning community by carefully amassing the comparable 1970 and 1980 U.S. Census commuting data. Only someone close to the data can make all the needed adjustments for travel mode categories, worker age definitions, and metropolitan area boundaries that changed between the two surveys. Mr. Fulton has done the job meticulously so that we may now compare apples with apples. His interpretation of the data, however, warrants a closer look.

I would have preferred that the public transportation category be restricted to conventional transit and not include taxis and that the transit user share be taken of only those who travel by vehicle and not include those who walk to work or work at home. Those two changes would result in a more policy-relevant definition of the public transit market and market share, although the resulting statistics would not be substantially different.

My major objections have to do with the general approach of the analysis. Data rarely speak for themselves. They must always be placed in a context for interpretation, and that interpretation invariably introduces a particular point of view.

In the first place, trends are presented on the basis of only two points in time. The least likely case is that a straight line connects the two points. The true curve may be U-shaped, with a low-point at middecade and an upward trend at the end, it may be an inverted U that peaked during the energy crisis and now is accelerating downward, or it may vary in some inexplicable or cyclical way. The decennial census is not frequent enough to reflect such variation.

A more fundamental problem is that the analysis was only a sorting procedure that used the standard census categories of four regions, 38 SMSAs, and whether the data were for locations inside or outside of central cities. The use of standard census geography may obscure actual patterns of transit use. The analysis implies that transit service is spread relatively uniformly across the nation, yet transit is, was, and will remain a big-city phenomenon. The question, What portion of all workers in the nation uses transit? has the same answer as the question, What portion of the total U.S. land mass is occupied by urbanized areas. Regional and national aggregate summary measures are misleading when the object of measurement is so spatially concentrated.

An illustration of this problem is the continued emphasis on the total decline in the number of transit commuters and the decrease in the overall transit share. Those numbers are indisputably correct, but not very helpful. Fortunately, Table 4 provides almost all the data needed to assemble a more useful picture. Table 13 was derived from Table 4 by dividing the number of workers who use public transportation by the percentage of workers who use public transportation to estimate the number of total workers in each year and each geographic division. The minor differences between Table 13 and other tables in Fulton's paper are due to rounding in my procedure.

The U.S. totals (bottom of Table 13) repeat the paper's main contention of general transit decline in the country. The top of Table 13 gives a high level of transit use in metropolitan areas that have greater than 1 million in population in 1980--81 percent of all transit users and double the national average of the percentage that use transit (12 versus 6 percent). More interesting is the breakdown of the greater than 1 million group in the center of Table 13. Five metropolitan areas (New York, Philadelphia, Newark, Chicago, and Detroit) accounted for more than 95 percent of the total losses in number of transit commuters nationwide from 1970 to 1980 (656,000 out of 691,000), which more than erases the 205,000 gain in transit commuters elsewhere in the country. Of the other 32 large metropolitan areas, 19 had increases in the number of transit commuters and 10 had increases in the share of transit commuters (one area had no change in the number of commuters, and one had no change in the share). Overall, these 33 areas had a 10 percent increase in the number of transit users, although their transit percentage share declined from 9 to 8 percent, and their share of the nation's transit commuters increased from 31 to 37 percent. Transit agencies rarely can do anything about metropolitan growth and development, so they cannot be expected to control the denominator of the percentage transit share. They can, however, work on the numerator (the size of their market) even if they cannot affect the share. The absolute increase in transit commuters in these 33 areas (as a group) shows that they are doing their jobs. In Table 13 the five areas that account for the greatest losses in transit commuters also had a growth rate for total workers during the decade only one-sixth as great as the national average (5 versus 29 percent) and accounted for only 2.5 percent of the actual growth in total workers.

Finally, we must look at Fulton's prescriptions. First, he recommends that the geographic coverage of transit service must be increased. One of the financial problems of transit was its past overextension into low-density areas where it cannot be operated efficiently. The transit industry has learned that lesson. The many forms of ride-sharing, such as vanpooling, club buses, and normal carpooling, are more effective in such areas than conventional, fixed-route transit.

Second, Fulton tells us that we must make policies in coordination with and in anticipation of demographic trends instead of depending on the alteration of such trends, and that the data underscore the deterministic effect that demographic changes have on the use of public transportation. To treat demographics as immutable natural forces ignores all other potent influences that shape urban areas and the provision and use of transit: local economic conditions, local fiscal capacity, transit support from other levels of government, politics, the price of automobile use (social and out-of-

*Metropolitan Transportation Commission, Berkeley, California.

Table 13. Comparison of workers, transit users, and transit share, 1970-1980.

	1970		1980		Change from 1970 to 1980	
Area	Number (000s)	Percent	Number (000s)	Percent	Absolute (000s)	Percentage
Population Greater than 1 Million						
All workers	33,412	44.9	41,427	43.1	8,015	+24.0
Transit users	5,445	81.7	4,994	80.9	-451	-8.3
Transit share		16.3		12.1	-[a]	-25.8
New York, Philadelphia, Newark, Chicago, and Detroit						
All workers	10,981	14.8	11,522	12.0	541	+4.9
Transit users	3,368	50.6	2,712	43.9	-656	-19.5
Transit share		30.7		23.5	-[b]	-23.5
Population Greater than 1 Million, excluding New York, Philadelphia, Newark, Chicago, and Detroit						
All workers	22,431	30.2	29,905	31.1	7,474	+33.3
Transit users	2,077	31.2	2,282	36.9	205	+9.9
Transit share		9.3		7.6	-[c]	-18.3
Population Less than 1 Million						
All workers	40,970	55.1	54,742	56.9	13,772	+33.6
Transit users	1,217	18.3	1,182	19.1	-35	-2.9
Transit share		3.0		2.2	-[d]	-26.7
United States						
All workers	74,382	100.0	96,169	100.0	21,787	+29.3
Transit users	6,662	100.0	6,176	100.0	-486	-7.3
Transit share		9.0		6.4	-[e]	-28.9

Note: All statistics are derived from Table 4. Differences from other tables are due to rounding.
[a]-4.2 percent. [b]-7.2 percent. [c]-1.7 percent. [d]-0.8 percent. [e]-2.6 percent.

pocket costs), urban infrastructure investment policies, service standards, nontransportation objectives (social mobility, air quality, and energy conservation), and all other demand and supply variables that affect human choices. Surely, these must bear some weight against the juggernaut of demographic trends.

Third, Fulton argues that the census data do not support the planners' assumption that mass transit would provide the ultimate remedy to the urban transportation problem by reshaping urban form and modifying consumer behavior. The only instances of trying to affect urban form are the big rail projects, like the Bay Area Rapid Transit (BART) in San Francisco and Metro in Washington. These are not routinely prescribed for all urban areas, even by the most starry-eyed planners. As for modifying consumer behavior, I think most in the transit community would agree that their task is to deliver a service that gives people more choices for travel, not to restrict them.

REFERENCES

1. Census of Population and Housing, 1980: Summary Tape File 3. U.S. Bureau of the Census, U.S. Department of Commerce, 1980.
2. Census of Population and Housing, 1970: Vol. 1 Reports. U.S. Bureau of the Census, U.S. Department of Commerce, 1970.
3. Census of Population and Housing, 1970: Fourth Count Summary Tape. U.S. Bureau of the Census, U.S. Department of Commerce, 1970.
4. Census of Population and Housing, 1970: Special Tabulations. U.S. Bureau of the Census, U.S. Department of Commerce, 1970.
5. Census of Population and Housing, 1980: Standard Metropolitan Statistical Areas and Standard Consolidated Statistical Areas: 1980. U.S. Bureau of the Census, U.S. Department of Commerce, 1980.
6. The Journey-to-Work in the United States: 1975. U.S. Bureau of the Census, U.S. Department of Commerce, Current Population Reports, Series P-23, No. 99, July 1979.
7. The Journey-to-Work in the United States: 1980. U.S. Bureau of the Census, U.S. Department of Commerce, Supplement to the 1980 Annual Housing Survey, unpublished.

Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.

Streamlining Collection and Processing of Traffic Count Statistics

DAVID T. HARTGEN AND JOHN H. LEMMERMAN

Traffic volume counts provide basic information for transportation analysis and forecasting, as well as for facility design, monitoring, and operation. The traditional methods of organizing a traffic count program have changed little since 1965. Basically, they include a system of continuous counter stations for developing seasonal adjustment factors, seasonal control stations to aid in factor development as well as in determining the seasonal assignment of coverage count stations, and the coverage count stations themselves. Such a program works well but is quite costly. Thus, the New York State Department of Transportation examined its program and procedures and looked at new technology to streamline the process and reduce costs. Seasonal adjustment factors for coverage counts were revised and the number of factor groups was reduced. The method of determining which sections should be counted was also revised, based on changes in traffic volumes, and telemetry systems and methods of collecting additional traffic data (truck weight, vehicle classification) were examined. Implementation of these improvements yielded a 35 percent reduction in counting with little or no loss of information.

Today, as in the past, traffic counts continue to provide basic information for transportation analysis and forecasting as well as for highway design, monitoring, and operation. The importance of current and reliable traffic counts for sound project planning has always been recognized, and their use in supporting regional or corridor studies through vehicle miles of travel (VMT) estimation, impact assessment, and traffic flow has also been identified and expanded. FHWA has traditionally supported the collection, processing, and dissemination of traffic volume data through its regular Highway Planning and Research Program (HPR) planning funds and its various monitoring requests.

Traffic count programs serve many client needs and, therefore, have historically been expanded. Many states and communities have attempted to satisfy a wide range of client needs through the use of a centralized master file of counts large enough to handle the most detailed of questions. Further, counts have been conducted periodically at numerous locations instead of relying on sampling principles or sample statistics for overall reliability. In New York State as recently as 1978, for instance, traffic counts taken on the state highway system for all purposes totaled 10,600, cost more than $1 million in HPR funds (out of a total HPR budget of $6.2 million), and required more than 60,000 hr of staff effort.

Such procedures were certainly sufficient, and perhaps even appropriate, to support highway planning activities in the 1960s and 1970s, but they are not properly scaled to current programs. Increasing fiscal austerity at all levels of government in all program areas means that agencies must ensure that adequate information is obtained at reasonable cost. This had led governments to institute goal-oriented count programs in which counts are taken for specific purposes or to achieve an agency goal. For example, FHWA's recent guide (1) suggests that count program goals (e.g., programming, operational improvements, safety, planning, policy assessment, environment or energy, and revenue allocation) be identified against problem scale (e.g., single location, cordon or cutline, corridor or subarea, or region). Then, the sample size necessary for a required level of accuracy is determined. Program layout (site identification, data collection, and processing) then follows readily from the sampling plan. The specific problems of a given case would need to be assessed individually; however, the examples in this guide clearly suggest that count programs of many agencies have, historically, been too extensive.

In recognition of such concerns, the New York State Department of Transportation (NYSDOT) recently set as a goal a 30 percent reduction in its 1982 count program relative to 1981. This program would continue a recent downward trend from 10,600 counts in 1978. Recent work undertaken by NYSDOT to achieve that goal, commensurate with modern count needs, equipment, and fiscal constraints, is described in this paper. First the count program is placed in perspective, and then recent work in improving factoring procedures, count scheduling, and timing is reviewed. Results of the program are described and some advantages of implementing a telemetry-based system are discussed.

BACKGROUND

A good summary of the early development of the current methodology employed in estimating annual average daily traffic (AADT) on highway sections has been published by the Department of Highways, Ontario, Canada (2). Based on work by Petroff (3-5) in cooperation with several highway departments, the Bureau of Public Roads (BPR) published its Guide for Traffic Counting Manual (6) in 1965. For high-volume roads (AADT > 1,500), the method produces AADT estimates that have a standard deviation of estimate of ±10 percent. The type of program described is based on three levels or types of traffic counting operations: continuous permanent counting stations (PCSs), seasonal control stations where counts are taken 4, 6, or 12 times a year for periods of from 48 hr to 2 weeks, and coverage stations where counts are taken usually once a year for a 24- or 48-hr period for factoring to AADT.

The BPR procedure was based on grouping together the PCSs that have similar patterns of monthly traffic variation, which then provide the basis for the adjustment of coverage counts. The seasonal control stations help define areas and similar individual facilities for purposes of assignment of coverage-count sites to a particular factor group. PCSs are grouped in such a way that, for the stations within a given group, the difference between the smallest and largest monthly factor for any month considered does not exceed 0.20. The monthly adjustment factor for month i is

$$F_i = AADT/MAWDT_i \quad (1)$$

where MAWDT is the average weekday traffic for month i. Then, for each of the groups so defined, the average of the factors for each month is computed to arrive at the month group mean factors.

Bodle (7) indicates that, in most states that use this procedure, three to five PCS groups are defined. He points out, however, that several sources of error are possible in that method--monthly factors at the coverage-count station generally will not be exactly equal to the group mean, coverage counts may differ from the average weekly count, and the road section on which a coverage count is taken may have been assigned to the wrong group.

Darrell, Dale, and Hayne (8), using the BPR method on low-volume (<1,000 vehicles per day) local roads, and farm-to-market and resort roads, found that 48-hr counts between May and October produced AADT estimates as good as those based on 7-day counts in the same period.

Drusch (9), reporting on tests of the BPR method in Missouri, found that the grouping of PCSs by that criterion resulted in too many groups. But, by averaging factors over 4 years, five pattern groups were obtained. The New York State experience began in a similar fashion, with 12 factor groups in the 1960s. That number was reduced to eight in 1975, and further reduced to four in 1982 based on a study of the 1980 PCS counts and analysis of the proximity and overlap of the error-of-estimate bands associated with the various seasonal curves (10).

Great Britian established a countrywide system of 50 PCSs (50-point traffic census) during the late 1950s. Bellamy (11) used data from 1968 to 1973 to perform an analysis of variance for the ratio of average daily traffic flow for each month to AADT by day of week. Based on that analysis she calculated five sets of ratios for each site, which represented the seasonal variation for Mondays, Tuesdays-Thursdays, Fridays, Saturdays, and Sundays. Nevertheless, a clear-cut set of groupings of sites that were valid over all days of the week and that allowed a simple definition of site classification was not found.

The sites were then classified somewhat subjectively by their known characteristics as follows:

1. Urban or commuter,
2. Low flow (<1,000 vehicles per day) nonrecreational rural,
3. Rural long distance, and
4. Recreational.

This grouping masks the intersite differences between Monday and Friday and the rest of the weekdays, and hence, one pattern was obtained for all weekdays. The variability of the estimates for the grouping ranged from about 2 percent in the case of the group 1 sites up to a maximum of around 20 percent for the group 4 sites. In general, the coefficient of variation was lowest for counts taken in April, May, September, and October and highest for those taken in August. These results were similar to those obtained in the latest analysis of New York State PCS data and described in this paper.

Phillips (12) takes this process one step further. He uses data obtained from the 50-point census for 1974 and 1975 to analyze the ability to estimate traffic at specific sites based on short (<24 hr) counts. He concludes that reasonable accuracy can be achieved through counts of 2 hr or more, with a diminishing return in accuracy for periods over 6 hr. The counts, however, should be taken in late spring or early autumn or, in the case of recreational routes, June is also suitable.

In 1977 Hillier, Mathews, and Moore (13) wrote a review of trends in traffic data collection and analysis in which they described equipment and techniques both available currently and desirable. They stressed the increasing need for vehicle classification, speed, axle weight, and headway information and the use of microprocessors in handling these data in an efficient and timely fashion. In 1981 Moore (14) described characteristics and applicability of various types of road sensors for traffic data collection.

A year later Wyman of the Maine Facility Laboratory and Lyles of Richard Lyles Associates conducted a study sponsored by FHWA to evaluate vehicle classification equipment (15). Recording systems and appropriate sensing devices were tested for basic accuracy, internal logic, and continuous use at various Interstate sites. In general, except for a prototype unit from England, all currently marketed systems had serious problems in performing the classifications.

In 1981 John Hamburg and Associates, Inc., (16) prepared a report on Improved Methods for Vehicle Counting and Determining Vehicle Miles of Travel. Initially, the research had attempted to develop a count program based on a sampling of traffic count sections. The final plan concluded, however, that counts should indeed be taken on all links over a cycle period. The report recommends that highway links be stratified by functional class, urban-rural categories, and growth class and then be reverse-ordered within a growth stratum by year of most recent count to ensure that the links that have the oldest counts are counted first. If the growth on a particular link is estimated to be more than 10 percent, then that link should be counted in the current year rather than wait for the cycle. This is not unlike the system currently being implemented in New York State.

IMPROVEMENTS IN PROCEDURES

Overview of NYSDOT Traffic Count Process

Figure 1 describes the steps involved in collecting and factoring traffic counts in New York State. Basically, raw data from road-tube counters are collected for a short (48-hr) period but are generally not adjusted to account for multiaxle vehicles. The raw count is adjusted for seasonal variations to estimate AADT and design hourly volume (DHV). Traffic forecasting (for individual sections) is then undertaken by a variety of procedures.

The operation of a traffic count program requires a number of inputs. First, of course, are the raw traffic counts, taken periodically at the various locations throughout the state. The New York State program (until recently) counted each highway section once approximately every 3 years to estimate current traffic volumes, and two or three times (once during the spring or fall, and once during the summer) every 9 years to check and possibly change the seasonal factor group assignment of each section. The count factoring procedure recommended by FHWA and used by NYSDOT (17) uses factors for each month for each group of highways. These factors consist of the ratio of monthly average weekday traffic (MAWT) to AADT. (Because 5-day weekday traffic is usually greater than weekend traffic, the factors average to greater than 1.0.) Data to estimate these ratios are determined from the department's 59 permanent continuous counters located throughout the state (Figure 2). To determine the correct seasonal factor for a given short count, the short count's summer ratio was compared with similar ratios for the continuous counters, and the short count was then assigned to the group that had the most similar ratio.

Counts are taken on state highways by using portable road-tube counters. They are normally set out early in the week and picked up late in the week, thus providing a good midweek traffic count. Factoring is then done by dividing the short count by the ratio for the month or group in which the short count falls.

Output (paper tapes) from the portable counters at coverage count locations are collected by NYSDOT regional personnel and forwarded to the main office for processing. The tapes are checked and translated

Figure 1. Typical traffic counting, factoring, and forecasting process.

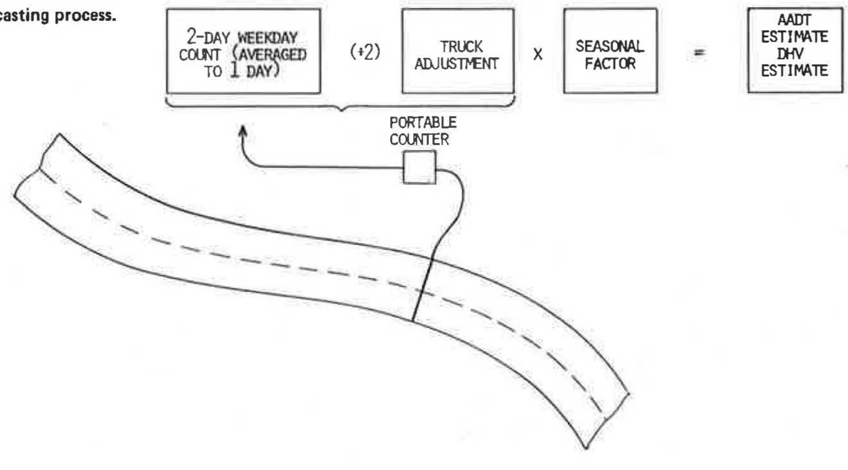
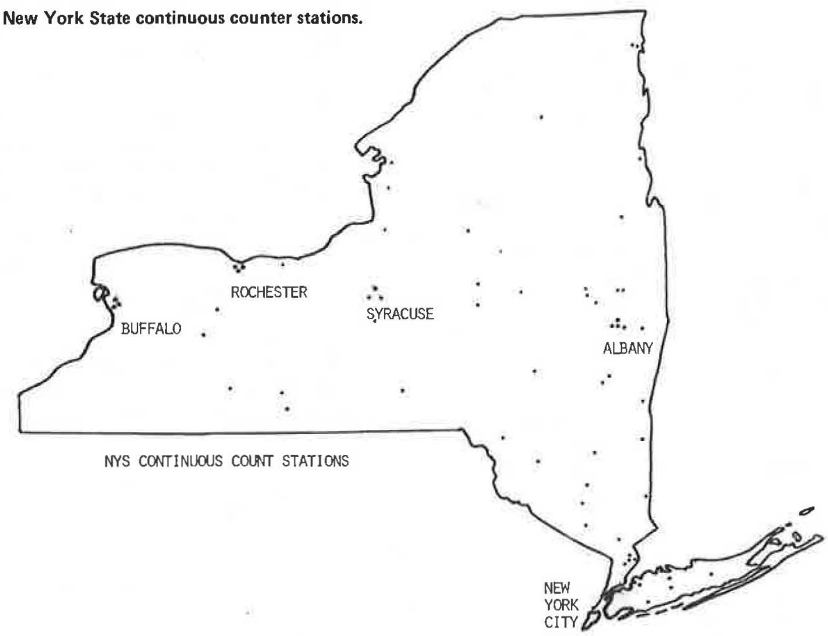


Figure 2. Map showing New York State continuous counter stations.



to a computer-compatible medium, and the traffic count volumes are recorded and factored. Once the year's counting and processing have been completed the annual traffic volume report (18) is published, the seasonal variations of count locations that have multiple counts are examined, and the factor group assignments are adjusted if necessary. Those adjustments are then reflected in subsequent editions of the traffic volume report.

Truck Adjustments

Some vehicles are multiaxled; therefore, traffic will be overestimated from raw road-tube counts. Axle correction factors should, therefore, typically be applied to raw counts. Although FHWA (1) provides standard formulas for the computation of such factors, NYSDOT has not, traditionally, undertaken such corrections unless the percentage is large (e.g., above 20 percent multiaxle vehicles). Recent concern for trucks as a factor in pavement deterioration and highway capacity has increased attention on this subject. Multiaxle vehicle proportions tend to vary widely by location and facility type and reflect local conditions such as truck terminal centers, so considerable research is needed to establish the nature of such variations. NYSDOT is currently undertaking this work. In the interim adjustments are made when needed but not generally for most counts.

Accuracy of Raw Counts

Various mechanical devices and manual counts will not always yield identical results. Differences in clock time, sensitivity of recording devices to vehicle pass-over, multiaxle vehicle proportions, and mechanical equipment failure or malfunction all combine to introduce error.

In August 1981 a comparison of four separate counting methods (manual counting, portable road-tube counters, a fixed mechanical continuous counter using induction loops, and a telephone-based counter system using the same induction loops) was conducted at a single location for a 24-hr period (19). Without elaborating on some of the details of the experiment and recognizing that hazards are associated with generalizations based on the observations and situations encountered at one station of such short duration, some findings are presented in the list that follows.

1. There is no such thing as an accurate traffic count: clock error, machine error, percentage of trucks, and other factors are likely to cloud the reliability of any count.
2. Manual counts are likely to contain considerable errors, particularly if conducted by inexperienced or unsupervised personnel.
3. Counts taken with a road-tube counter will overestimate traffic volume depending on the percentage of trucks. Such counts should be adjusted for the multiaxle truck percentage and checked closely for clock accuracy.
4. Counts should be taken for at least 24 hr to minimize overall clock error. Even during longer count periods (1 day to 3 weeks), counters are not likely to give similar results, but the differences will be smaller.
5. Both the continuous counter and the telemetry system show very similar, but not identical results.

This test also showed that portable road-tube counters are reasonably accurate and can be relied on to accomplish most of the department's traffic counting needs.

Counts Based on Need

The department's system for scheduling traffic counts was designed around the concept of periodic (cyclical) counting, supplemented by a double count every 9 years. This approach results in some unnecessary counting because changes in volume may be slow at many sites and even slower for many factor groups. This approach also delays needed counts on fast-growing sections.

To deal with these problems NYSDOT instituted a count program based on need rather than on just cyclical frequency. Need to count was defined in terms of two different growth criteria.

1. Traffic growth--For each section scheduled for counting in 1982, trend data on traffic growth were obtained by reviewing previous counts and calculating growth rates for each section. These were then sorted according to growth rate, and the fastest-growing sections were identified. To further aid in this procedure the format of the traffic volume report has been revised to show the previous three counts on a section, the year in which they were taken, the percentage difference between the two most recent counts, and the latest count and year taken.
2. 1970-1980 growth rates of town and county population, households, and automobile registrations (20)--Traffic count growth data for rural towns and counties were regressed against growth data on population, households, and vehicle registrations. The best regression line (one for each functional class) was then used to estimate town-level traffic growth rates based on the 1980 census. The fastest growing towns were then identified for focus.

NYSDOT used these two procedures to determine those sections in which historical growth rates would be greater than 5 percent/yr, 2 to 5 percent/yr, and less than 2 percent/yr. (By comparison, total traffic in New York State grew at about 2 percent/yr from 1980 to 1981.) A bonus on the fastest growing sections (5+ percent/yr) would reduce the required count program by approximately 30 percent. Lists of sections to be counted were then prepared and transmitted to regional offices along with instructions to focus the count program on the fastest-growing sections.

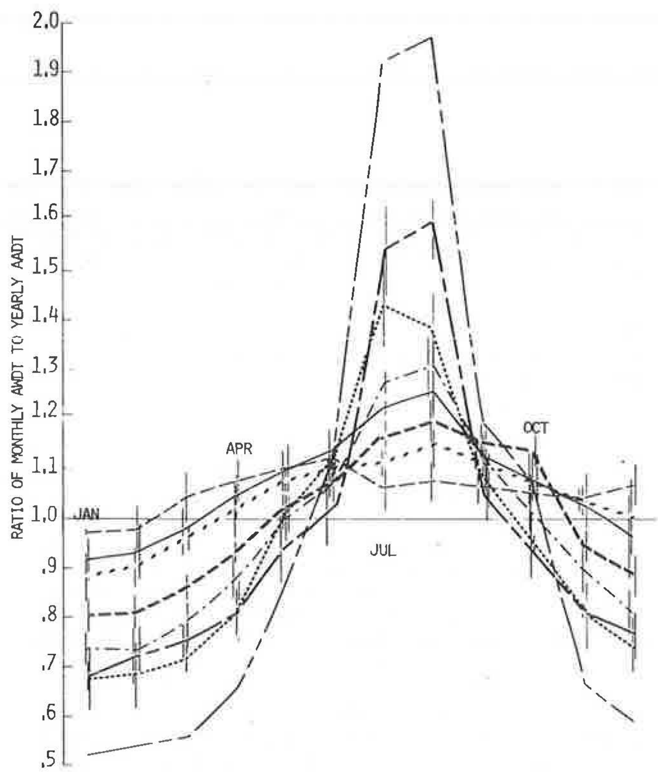
Seasonal Adjustment of Traffic Counts

Over the years a number of changes and improvements have been made in the procedures for seasonally adjusting traffic volume counts. In 1975 NYSDOT examined and updated the traffic count factor groups and factors within those groups. That analysis used 1972 volume data because the 1973 and 1974 data were not typical because of the energy crisis during those years. The 1975 study resulted in a reduction in the number of factor groups from 12 to 8 and an updating of design hour factors to correspond with the new factor groups. The count program and procedures were generally left intact.

Figure 3 shows the 1972 base factor groups used by NYSDOT. The eight factor groups represent groupings of roads based on their seasonal patterns. Note also the small vertical lines that represent the standard deviation about each point. They show that the factors for adjacent groups for some months cannot often be determined exactly, and so the distinctions among factor groups are blurred. Therefore, the factors have unnecessary false precision in group assignment. Actually, a smaller number of groups might be more efficient if grouped carefully.

Knowledge of the factor group for a particular count is often unnecessary if counting is scheduled for certain months, for example, June or September. Similarly, the ability to discriminate among groups 1, 2, and 4 is unnecessary if the count can be scheduled in the January to June period. Other such patterns and overlaps are readily apparent. The meaning of this is as follows: Precise knowledge of factor grouping is important only if the factors will be different; if the count can be scheduled such that group overlap is maximized (e.g., January

Figure 3. Traffic count factor groups, 1972.



to June for groups 1, 2, or 4), then grouping requirements can be reduced without loss of accuracy. These requirements must, of course, be traded off against the practical scheduling of counts in the field.

Aware of these problems and concerned with rising costs and limited staff resources, NYSDOT undertook an extensive review of its count program in 1981. This time the goal was to assign count sections to a seasonal factor group on some basis other than similarity to a particular pattern as determined by the multiple count taken once every 9 years. Revisions focused on a number of problems described earlier, and generally paralleled recent work by Sharma and Werner in Alberta, Canada (21).

Data for 1980 were first obtained for 55 of NYSDOT's continuous counters. They showed substantially less peaking than the 1972 data because of the dampening effect of the energy crisis on traffic in rural and recreational areas (10). These data were for the most recent year for which complete data were available.

The analysis approach used was to cluster the stations according to similar patterns of MAWT/AADT ratios. Several programmatic approaches to clustering data were employed, but results from the BMDP/2M statistical software (22) were found to be the most satisfactory. This procedure initially considers each observation as a separate cluster. Cases or clusters of cases are then joined in a stepwise process until all cases are in one cluster. The procedure uses the Euclidean distance between nearest neighbors as the criterion for joining clusters and cases. In the application the distance measure was based on the Euclidean distance in n-dimensional space defined by the MAWT/AADT ratios for the 12 months. The result is a most efficient (statistical) grouping. The method is conceptually similar to that used by Sharma and Werner.

Different numbers of clusters, from two to eight, were tested and compared. The standard deviations and coefficients of variation of the monthly factors showed that four clusters provide a good representation of the seasonal patterns, and that little accuracy is gained by further increasing their number. Figure 4 shows this four-group cluster pattern. The standard error of each number (vertical lines) is generally wider than that of the eight-cluster (1972) groups but overlaps in the error bands are considerably less. This means that the four-group design discriminates better among clusters than did the old eight-group design.

The next step was to investigate the locational, design, or use characteristics each of the sections had in common with each of the other sections within its own group. The purpose of this investigation was to allow the assignment of counts to factor groups by these characteristics rather than on a strict 24-hr/AADT ratio. Variables examined included:

1. Location characteristics--urban or rural, area type, culture, and terrain;
2. Design characteristics--number of roadways, number of lanes, access control, lane width, pavement width, and shoulder width; and
3. Use characteristics--functional class, proximity of alternate routes, population growth, and volume.

No combinations of these variables would distinguish the groups perfectly, but the urban or rural nature (defined by FHWA procedures) and access control of each highway section appeared to have the greatest potential for making an initial distinction.

In general, once these two characteristics of a highway are determined, the choice of which factor group to assign is narrowed to two (as given in Table 1), with a greater probability of one of these over the other. Some knowledge of the area or the facility can then aid in the final selection. If this is not possible the more probable of the factor groups (in terms of numbers of continuous count stations assigned to that group) can be assigned, the count can be scheduled to minimize error (see below), the factors of the competing groups can be averaged to estimate AADT, or the section can be multicounted and analyzed.

This design shows how the various traffic count factor groups can be assigned to highway sections, based on the urban or rural nature and the access control prevalent on that highway. In the event that the factor group for a particular section cannot be determined, then the count for that highway section should be made during certain months to minimize the probability of error in estimating AADT. This is shown in Figure 5.

Figure 4. Traffic count factor groups, 1980.

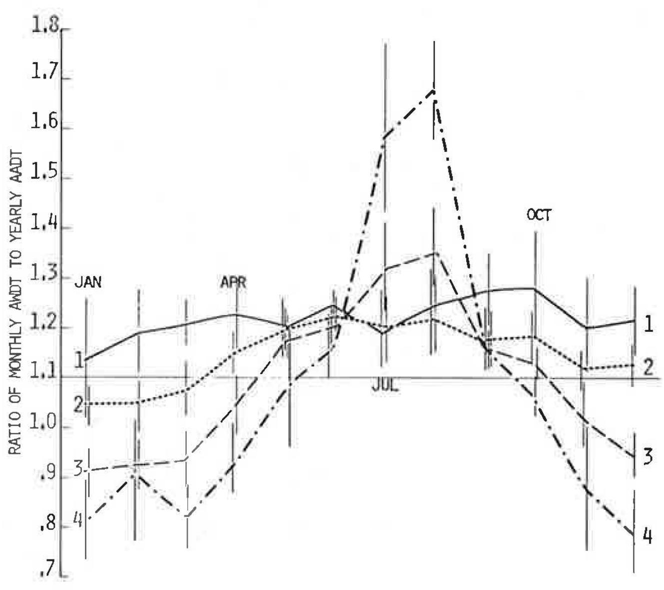


Table 1. Characteristics of factor groups.

	Access Control	
Group	Full or Partial	No Control
Rural		Factor groups 2, 3, and 4, greater probability of 3
Small urban, pop <50,000	Factor groups 3 and 4, greater probability of 3	Factor groups 1 and 2, greater probability of 2
Large urban, large metropolitan area	Factor groups 1 and 2, greater probability of 2	

Note: Most facility mileage falls in rural with no access control group.

Figure 5. Scheduling rules.

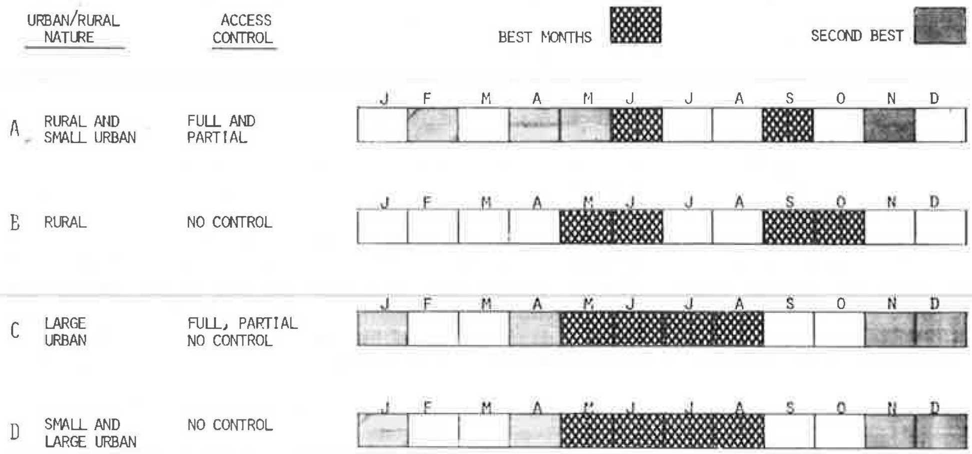
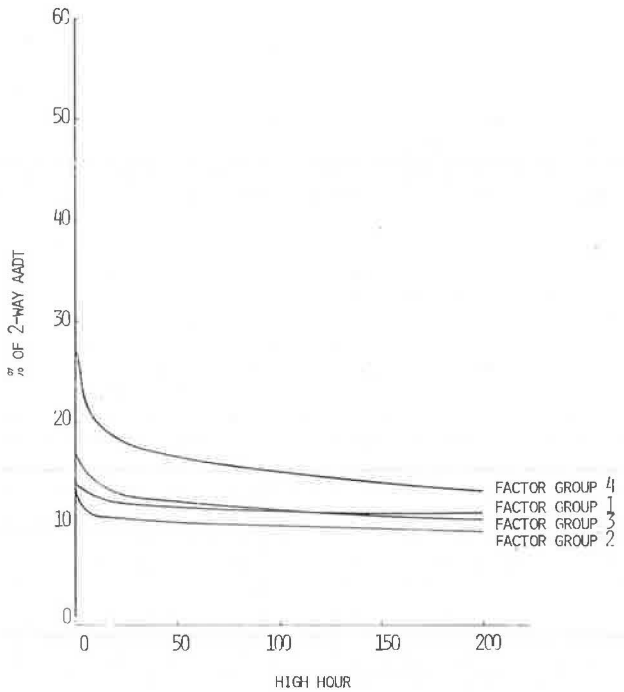


Figure 6. Peak-hour traffic from NYSDOT continuous counter stations, 1980.



Design Hourly Volume

Along with the analysis of seasonal traffic variation and the means of adjusting for it, a study of peak hour and design hourly traffic volumes (DHV) was undertaken. Peaking characteristics and DHV are more useful items than AADT in the design of a highway; however, little information exists to document their development. Early efforts in the area were described by Walker (23). By examining the year's 30th highest hourly volume (30 HV) factors he found that they were generally lower on high-volume roads and that, as traffic increased over time, they tended to decrease. Bellis and Jones (24) found, in addition, that 30 HV increases and K decreases as population increases, and the capacity of a facility seems to have little effect on these factors. The 1965 Highway Capacity Manual (25) reports these findings; and, even though it does not recommend the use of 30 HV for rigid adoption, it goes a long way toward institutionalizing it.

For most (nonlocal) two-lane highways DHV is now usually the 30th highest hourly volume that is expected to occur in some future design year. For multilane highways the directional design hourly volume (DDHV) is used. DHV is established by multiplying the projected future AADT by a K factor (K = DHV/AADT), and the DDHV (if required) is determined by multiplying DHV by a directional facor (D) (26).

Most of the analyses of peak and design hour traffic use data available from permanent counting stations, but in practice the ability to apply the correct factor or select the X-highest hour requires data that are not generally available. In determining design volumes, NYSDOT's Regional Planning Offices use a variety of methods. NYSDOT's annual traffic volume report contains an estimated DHV for each highway section. This volume is an estimate of the year's 30th highest hour based on the seasonal traffic variations exhibited and the factor group to which the section has been assigned. Figure 6 shows New York's 1980 curves for the 200 highest hours.

In the absence of more detailed information, and especially if the traffic volumes are not particularly high, NYSDOT planners use the DHV figures directly from the current traffic volume report. For a projected volume they will apply the appropriate K factor to the future year's AADT. Where greater reliability is needed, special traffic counts are taken at the site and in the vicinity of the planned highway project. (Depending on the type of traffic served, the count may or may not include the weekend.) In addition, manual classification, turning, and directional counts are taken where applicable during the weekday morning and evening peak hours. The planning engineer will use this count information as well as the necessary economic and growth forecasts to develop a design year traffic volume for the facility. Then DHV is calculated for the facility either by developing K as the two-way high hour for the week counted divided by AADT or by using the average value developed from the permanent count stations for the factor group to which the section is assigned.

In reviewing the development of DHV for one recent highway project, concern was expressed as to what HV was actually represented by the volume developed. This led to a comparison of the New York State high hour curves with those referred to in the development of that project. The latter (Figure 7) were taken from the Transportation and Traffic Engineering Handbook of 1976 (26), which, in turn, took

Figure 7. Peak-hour traffic.

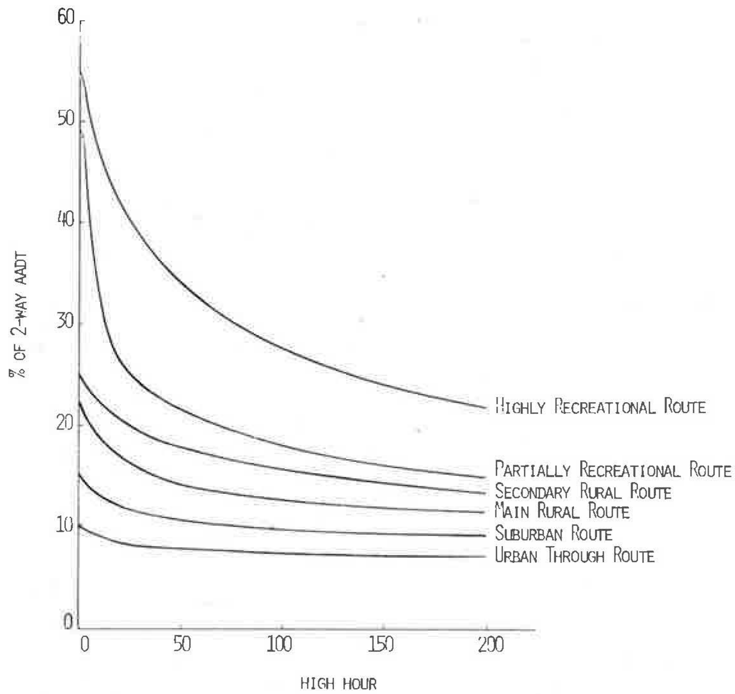


Table 2. Results of traffic count reduction program.

	Machine Counts		Person-Hours Used		
Year	Number	Percentage Change	Number	Percentage Change	Ratio (hr/count)
1978	10,572		60,185		5.7
1979	10,754	1.7	55,698	-7.5	5.2
1980	7,806	-27	50,059	-10	6.4
1981	7,762	-0.6	40,957	-18	5.3
1982	5,040	-35	37,250[a]	-9	7.4

[a]Estimate based on 22 of 26 accounting periods.

them from Traffic Engineering of 1955 (27), which adapted them from a 1947 Connecticut State Highway Department publication. Thus, the current standard high hour curves are really about 35 years old. Some additional work appears to be required in this area.

RESULTS

Resources Saved or Diverted

Implementation of the primary elements of the reduced traffic count program began in 1980 and is continuing. Preliminary analysis of town and county population growth rates and section-specific traffic growth rates were provided to NYSDOT regional offices with instructions to begin the count scheduling process. Factoring and scheduling procedures were transmitted shortly thereafter.

Table 2 compares New York State's programs for the past 5 years. Overall, the goal of a 30 percent reduction for 1982 was achieved; several of the regions are implementing reductions of more than 50 percent. Compared with 1978's counts, the reduction is much more striking.

Reductions of this magnitude may cause greater than anticipated flexibility on the part of traffic count personnel and may lead to staff capacity that is underused. Table 2 suggests that this may be the case with NYSDOT's count program. Large count reductions in 1980 (-27 percent) did not result in the same reduction in hours spent, which implies a decrease in productivity. In the next year (1981), however, an additional 18 percent reduction in hours with little actual reduction in counts did occur, which implies some lag time before the newly available person-hours can be reprogrammed efficiently to other duties. Part of the difficulty is that the amounts and timing of hours saved may be such that little usable time is actually generated during the count season. Indeed, the required layout and scheduling of counts may convert some of the savings to travel time. In any case, the new-found time is available for reassignment to other duties.

Options Available

Several options are available for further reductions in the count program should this become necessary or desirable. The first, of course, is to reduce the magnitude of the coverage count program further and either skip certain count sections (perhaps on some rotational basis), lengthen some of the sections, or change the threshold criteria that determine when highway sections should be counted. Any of these could be implemented by first determining the size of the reduction and then fitting the program into it.

A second means of effecting cost reductions is through sampling of representative sections and expanding the sample to the universe. FHWA recommends this type of procedure for the estimation of

VMT (1). A similar approach (16) has been developed by NCHRP. A third approach is interagency and intergovernmental cooperation, some of which is already taking place through the metropolitan planning organizations (MPOs) in large urban areas. By coordinating programs each group or agency could work in areas where it is most adept.

New Technology

The future of traffic data collection appears to lie in automation of the recording, transmittal, and processing functions. Telemetry systems that count vehicles have been used in a number of states for more than a decade. These systems must generally be queried on a daily basis and can be polled automatically from a central system or called as desired. (During the 1979 fuel crisis FHWA queried a number of such systems directly in several states to produce up-to-the-minute estimates of travel and fuel consumption for crisis- and policy-evaluation purposes.)

In the test of systems mentioned earlier the telemetry system was found to be superior to the mechanical system in speed of delivery because pickup time is eliminated. In addition to the collection of timely data, the telemetry system can be queried by telephone to check its operation. This reliability feature can ensure a minimum loss of data in a faulty operation. On the other hand, the mechanical system requires a site visit to check its operation--in New York State, the counter is checked at about the midpoint of each monthly recording period.

The cost advantages of the telemetry system are also substantial, primarily the result of a much lower labor cost to collect the data. A recent comparison of the costs to convert New York State's network of continuous count stations versus the costs to maintain the present type of system, including labor, over a 20-year period, showed that the telemetry system would cost about half as much as the electromechanical system (19).

Technological advances are currently being made in the capability and diversity of new recording equipment. Depending on the detector arrangement, the new equipment is capable of classifying vehicles (based on vehicle length and number of axles), measuring speeds, and counting vehicles and axles. Some of the new recorders are also programmable (at least one even by telephone) with respect to which data to collect and how to record them. These multiple capabilities of the new recording equipment make it even more valuable and would considerably lessen the need to launch separate operations to collect other than traffic count data throughout the state.

The need for traffic count information will continue to exist. The process of collecting it, however, will have to consume fewer resources than it currently does. Cooperation, technology, process efficiencies, and an analysis of true need should all contribute to bringing about this result.

ACKNOWLEDGMENT

We wish to express appreciation to Alfred J. Neveu for his help and patience in working with and interpreting the output from various computer software packages. Thanks also to Tammy Zitzmann, Sue Carbone, Dianne DeVoe, Tedi Toca, and Donna Brown for their work in the preparation of the manuscript.

Discussion

David L. Greene*

The authors have written a timely paper that addresses an important topic (how to produce good traffic statistics in a time of shrinking budgets). It is full of good ideas; perhaps, it even has too many ideas for one paper. Two papers in one were presented here. One is about the accuracy of traffic counting methods, and the other is about more efficient procedures for grouping road segments for traffic counting. The paper makes contributions in both areas.

The data on the comparative accuracy and reliability of manual counts versus portable, versus continuous, versus telemetry counters should be of considerable interest to all who work in this area. It contributes to one of the most important needs of vehicle travel data--the need to know the accuracy of the data. One source of error is the measurement of the count, and we are given some important information on that score.

The paper also describes some clever analysis of AADT factor groups, by using cluster analysis to decrease the number of groups and still maintain homogeneity within groups. Interestingly, a single-linkage method was used to form clusters rather than a p-means method. The first joins nearest neighbor clusters until the desired number of clusters is left. The latter explicitly maximizes variance between group clusters, which is equivalent to minimizing the sum of variance within groups about the cluster centroids. I point this out because the latter is one optimal way of forming strata for stratified random sampling. I will come back to this point later.

As pointed out earlier, the paper is divided roughly into two parts. The first part deals with data accuracy, or data validity. (Let me state my bias here by saying that I approach this problem from a statistician's viewpoint.) This first part quantifies the error of measurement of counts, deals inconclusively with the problem of multiaxle vehicles and their effect on the estimated number of vehicles, and then briefly touches on a large problem (the heart of the accuracy problem, in my opinion) by discussing VMT growth and frequency of sampling (called need to count). The idea is that segments where traffic is increasing rapidly need to be counted more often that stagnant sections. This was said to decrease the required counts by 30 to 40 percent, but exactly how and with what, if any, loss in accuracy were not stated.

This brings me to the second (closely related) issue addressed--efficiency. The real issue here is, if you can only sample part of the system part of the time, how should it be done to give the most accurate estimate of travel in such a way that the accuracy of the estimate can be quantified for a given level of resources (or one could fix the accuracy of estimate and minimize the resource). Let me reemphasize this because it is the crux of the matter. Not only do we only look at part of the system in each year, but, with the exception of PCS, we only look at it for typically 1 to 7 days.

*Transportation Energy Group, Energy Division, Oak Ridge National Laboratory

To a statistician this is a classic problem of designing a statistical sample, but one that must be designed over time and space. From this viewpoint the problem of factor groups for coverage counts (and seasonal counts as well) is a somewhat altered problem of defining optimal sample strata. Whereas the definition of an optimal factor group is not entirely clear (this paper comes pretty close to a precise definition and a method for finding the groups as well), optimal sampling strata are those that maximize between-group variance and minimize within-group variance for the variable to be measured (traffic volume). The theory of stratified random sampling is well worked out (e.g., Cochran's text on sampling) and, if it is possible to define the cost of sampling for each stratum, then it is possible (for a given level of accuracy of estimate) to determine the optimal distribution of samples among the strata. This could even be used to determine which strata should have PCSs.

My last point about stratified random sampling is that it defines the problem well enough that, given a budget reduction, one can, in theory, define the optimal sampling strategy and tell how much accuracy has been lost because of the reduced level of effort.

In closing, let me say as a tribute to the richness of this paper that there are still many points I have not addressed; e.g., design hourly volume estimation, the overview of new, automated counting and classification equipment, and literature review of traffic count studies. In fact, the authors try to accomplish a bit too much, but they can be excused for this because they have taken a fresh and analytic look at a neglected but important information system, one that has increased in importance with concern for energy use and even more so with increasing concern over highway maintenance and user taxes. As better automated count and classification equipment become available, issues will arise with respect to jointly collecting data on vehicle volume and type, ensuring their validity, and measuring their accuracy. As usual, NYSDOT seems to have anticipated this problem (or perhaps opportunity is a better word) and taken a pioneering role in research in this area.

Authors' Closure

Dr. Greene's views help to clarify and sharpen the value of the paper to both transportation analysts and administrators. Single-linkage clustering was used in this case because of the small data base available. The entire cluster structure contains only 59 observations; with samples this small, more powerful clustering methods, based on analysis of variance, seem inappropriate. Ideally, this could be done in those states that have more stations. Clarification is also necessary concerning the reduction in counts achieved by the need-to-count approach. These reductions were achieved largely because only a small portion of traffic counts on individual highway sections appeared to be growing rapidly enough to warrant frequent counting. Thus, the need-to-count rule is a useful device for identifying sections that may require attention.

The particular goals of traffic count programs should determine their designs. These include the estimation of areawide VMT. Perhaps more important, however, is the need to have a reasonably current traffic count on most sections of highways so as to respond to questions concerning traffic volume. These data are also essential for project assessment. The state transportation departments often undertake extensive counting programs for these purposes.

Sampling plans for estimating VMT or other areawide parameters can be designed efficiently by using the sampling concepts suggested by Greene. Statistical sampling frames for traffic, such as the HPMS sample (which consists of 2,700 observations in New York State) are available and are monitored periodically. Our small sample of 59 continuous counters may also be considered to be a sample of traffic on state roads; it is used as such to establish traffic trend by season and year. The other goals (providing current data for project analysis) cannot be achieved, however, unless additional counts are undertaken. The department's program is a compromise among the need for current information, the importance of areawide information, and administrative and staffing realities.

To expect one system of traffic data collection to satisfy all goals and respond to all purposes may be unreasonable. A flexible arrangement of systems, in which background information is tracked on a steady basis and periodic information is collected less frequently, seems to be a reasonable compromise.

REFERENCES

1. Guide to Urban Traffic Volume Counting. Urban Planning Division, FHWA, 1981.
2. M.D. Harmelink. Estimation of Annual Average Daily Traffic (AADT) and Thirtieth Highest Hourly Volume (30 HV). Department of Highways, Downsview, Ontario, Canada, Rept. 140, Aug. 1968.
3. B.B. Petroff. Some Criteria for Scheduling Mechanical Traffic Counts. Proc., 26th Annual Meeting, HRB, Vol. 26, 1946, pp. 389-396.
4. B.B. Petroff and R.C. Blensly. Improving Traffic-Count Procedures by Application of Statistical Method. Proc., 33rd Annual Meeting, HRB, Vol. 33, 1954, pp. 362-375.
5. B.B. Petroff. Experience in Application of Statistical Method to Traffic Counting. Public Roads, Dec. 1956.
6. Guide for Traffic Volume Counting Manual, 2nd ed. Bureau of Public Roads, U.S. Department of Commerce, Feb. 1965.
7. R.R. Bodle. Evaluation of Rural Coverage Count Duration for Estimating Annual Average Daily Traffic. HRB, Highway Research Record 199, 1967, pp. 67-77.
8. J.E.P. Darrell, R. Dale, and W.J. Hayne. Minnesota Experience in Counting Traffic on Low-Volume Roads. Proc., 37th Annual Meeting, HRB, Vol. 37, 1958, pp. 396-417.
9. R.L. Drusch. Estimating Annual Average Daily Traffic from Short-Term Traffic Counts. HRB, Highway Research Record 118, 1966, pp. 85-95.
10. J.H. Lemmerman and D.T. Hartgen. Revised Procedures for Factoring Short Traffic Counts to Average Annual Daily Traffic. Planning Division, New York State Department of Transportation, Albany, Rept. 10, July 1982.
11. P.H. Bellamy. Seasonal Variation in Traffic Flows. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, Supplementary Rept. 437, 1978.
12. G. Phillips. Accuracy of Annual Traffic Flow Estimates from Short Period Counts. Transportation and Research Laboratory, Crowthorne, Berkshire, England, Supplementary Rept. 514, 1979.
13. J.A. Hillier, D.H. Mathews, and R.C. Moore. Trends in Traffic Data Collection and Analysis.

Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1977.
14. R.C. Moore. Road Sensors for Traffic Data Collection. Sensors in Highway and Civil Engineering, Institution of Civil Engineers, London, 1981.
15. Evaluation of Vehicle Classification Equipment. FHWA, Sept. 1982.
16. C.C. Francis, Jr.; John Hamburg and Associates. Improved Methods for Vehicle Counting and Determining Vehicle Miles of Travel. NCHRP, Prelim. Tech. Rept. (in press).
17. Guide for Traffic Volume Counting Manual, 3rd ed. FHWA, March 1970.
18. 1982 Traffic Volume Report. Planning Division, New York State Department of Transportation, Albany, 1983.
19. E.J. Snow and G. Jillson. Accuracy Cost and Responsiveness of Continuous Traffic Counting Systems: A Report on the Burnt Hills (Glenville) Route 50 Telemetry Demonstration Project. Planning Division, New York State Department of Transportation, Albany, Rept. 7, March 1982.
20. A.J. Neveu. Quick-Response Procedures to Forecast Rural Traffic. Planning Division, New York State Department of Transportation, Albany, Transportation Analysis Rept. 3, June 1982.
21. S.C. Sharma and A. Werner. Improved Methods of Grouping Provincewide Permanent Traffic Counters. TRB, Transportation Research Record 815, 1981, pp. 12-18.
22. P2M Cluster Analysis of Cases. BMDP Statistical Software, 1981 ed. (W.J. Dixon and L. Englemen, eds). Univ. of California Press, Berkeley, Calif. 1981.
23. W.P. Walker. Trends in the 30th-Hour Factor. HRB, Bull. 167, 1957, pp. 75-83.
24. W.R. Bellis and J.E. Jones. 30th Peak Hour Factor Trend. HRB, Highway Research Record 27, 1963, pp. 1-13.
25. Highway Capacity Manual, 1965. HRB, Special Rept. 87, 1965, 411 pp.
26. J.C. Oppenlander. General Traffic Characteristics. Transportation and Traffic Engineering Handbook, Institute of Transportation Engineers, Prentice-Hall, Englewood Cliffs, N.J., 1976.
27. T.M. Matson, W.S. Smith, and F.W. Hurd. Traffic Engineering. McGraw-Hill, New York, 1955.

Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.

Transportation Energy and Related Data Collection at State and Substate Level

BETTY J. YELICH, NATHAN S. ERLBAUM, AND K.W. PETER KOEPPEL

The collection and timely monitoring of transportation energy consumption and related data for use at the state and substate level have many problems. Much of the relevant data are available at the national level and must be disaggregated to the state or substate level to be meaningful. A comprehensive data directory has been prepared with funding from UMTA. The directory lists sources, level of detail, frequency of reporting, lag time, and quality of data. It describes data series and information systems in 10 broad categories. The frequency of reporting, as well as time lags, pose significant difficulties for the timely monitoring of transportation energy use and travel behavior. Deregulation of certain industries by the federal government (e.g., airlines) and the subsequent loss of reporting requirements are also a factor. In addition, proprietary and private data sources may limit data availability for monitoring activities. At the substate level sources of data for two urban areas in New York State are identified. Rochester is a large metropolitan area, the core city of a four-county standard metropolitan statistical area (SMSA) that has a population of more than 100,000. Oneonta is in a rural county and has a population of 50,000. One of the more important measures of transportation energy use (automotive fuel) at the substate level has no direct data source. Synthesis methods are used to obtain an indication of fuel use. Six methods are used to estimate the fuel use for Rochester, New York. Some of the reasons for the different results obtained by using each method are discussed.

Shrinking energy resources, world events, environmental pressures, federal and state regulations, and socioeconomic factors all have an effect on travel and transportation energy use. Therefore, timely, accurate collection and monitoring of the parameters associated with energy use and travel are of great importance. Planning for the repair and maintenance of deteriorating highways and bridges requires data collection and monitoring. Declining fuel tax revenue because of decreased fuel consumption increases the need for more timely collection and monitoring of transportation energy use.

Many railroads and transit operations have declining patronage. Concomitantly, political pressure is being exerted to expand services for the disadvantaged and maintain reasonable fare structures. This too requires data collection for the development of monitoring and forecasting methodologies. The New York State Department of Transportation (NYSDOT), Planning Division, Data Services Bureau is responsible for the collection of data used to support ongoing studies of transportation and energy-related issues.

The problems in New York State may be even more severe than those of the nation in general. New York has an aged industrial base, decreasing employment opportunities for a large unskilled portion of the population, and a shrinking state tax base.

NYSDOT'S DATA COLLECTION PROCEDURES

The reasons for collecting data are many and varied. NYSDOT collects data in response to legislative mandate, agency need (i.e., revenue forecasting), and client requests (i.e., interagency use and public relations). Before beginning a data collection or monitoring activity general basic questions should be addressed:

1. What is the purpose of the effort? Is the collection effort consistent with the purpose and need?
2. What types of data need to be collected?
3. What are the sources of the data and are they readily available at the level of detail desired? and
4. If the data are not readily available, can they be synthesized with secondary sources or estimated through some method?

Perhaps the most difficult of the preceding tasks outlined is task 3. To this end, the NYSDOT Data Services Bureau, Planning Division, under funding from UMTA, has developed a detailed Data Source and Reference Directory of Transportation and Energy Data (1). This directory describes major data series and information systems in 10 broad categories, as given in Table 1.

Each data category, series, or system in Table 1 is described in detail with respect to contents, contact person, system status, reference, and other documentation. Most of the categories, series, or systems described in the Data Source Reference Directory (1) contain historical data from 1976, although some are reported from earlier periods. The sources and collection of data, synthesis and estimation methods, and problems and resolutions associated with collection and monitoring are discussed in this paper.

Many of the categories noted in Table 1 may be used to monitor changes in transportation and energy, but the selection of actual parameters is often limited by

1. Data availability,
2. Time lag in reporting, and
3. Inconsistencies with levels of aggregation and disaggregation and subsequent noncomparability of data.

Cross-reference charts for the 10 categories in Table 1 give a comprehensive overview of the level of detail, political divisions, frequency of report, lag time, quality of the data (i.e., primary, secondary source, or synthesized), and other pertinent information. In Table 2 many of the available data are collected at the national level. These data, in general, are not ideally suitable for timely monitoring at the state or substate level. The category of fuel use is discussed in this paper as an example.

FUEL USE

In each of the 10 categories noted in Table 1, and in particular the example shown here, problems exist that are typical of any major data collection effort. In some instances a straightforward resolution is possible. In others, however, a variety of approaches are illustrated, each of which presents its own unique problems.

Case Study: Data Collection at Substate Level

The case study describes the data collection effort in two urban areas in New York State, Rochester and Oneonta. Rochester and Oneonta are quite different in terms of population and available transportation services. Rochester is a large metropolitan area located on Lake Ontario and is served by the New York Thruway. Oneonta is a small urban area served by Interstate 88. Both are in upstate New York. Rochester is the core city of a four-county standard metropolitan statistical area (SMSA); Oneonta is located in a rural county. Monroe County, in which Rochester is located, has a population of more than 700,000; Otsego County (Oneonta) has a population of 59,000. Local transit, intercity transit, and air passenger service are available for both localities; Rochester also has rail passenger service. A substantial array of transportation data is available for both areas, as given in Table 3.

In the table the obvious blanks indicate the lack of data. Possibly most important are those data relating to highway fuel consumption. NYSDOT has developed nine general methods for estimating substate gasoline use as follows (2):

1. Network,
2. Households,
3. Number of stations,
4. Retail sales at stations,
5. Partition of statewide sales,
6. Vehicle registration,
7. Trip,
8. Population, and
9. Continuous count per vehicle miles traveled (VMT).

These methods produce a fairly large range of estimates that have a standard deviation of up to 25 percent of the mean for a given area. Many of the methods rely on secondary data at the substate or statewide level (i.e., VMT, number of service stations, and percentage of gasoline sales to total retail sales).

The New York State Department of Taxation and Finance is the primary source of data for statewide

Table 1. Transportation data, data series, and data sources.

No.	Category	Description
1	Fuel supply	Jet fuel, aviation gasoline, diesel, gasoline, electricity
	Fuel use	Highway use, nonhighway use, aviation, rail and transit, bus, marine, automobile, truck, other
2	Fuel price	Gasoline, diesel, aviation, other
3	Vehicle registration	Automobile, truck, motorcycle, taxi, recreational vehicle, bus, fleet, government vehicle, driver license, other
4	Vehicle efficiency	Automobile, truck, bus, taxi, motorcycle, fleet, recreation, other
5	Transit service	Bus, rail-rapid, other
6	Intercity service	Passenger, freight, air, bus, rail, automobile, truck, water
7	Highway use	Automobile, motorcycle, truck, bus including school bus, other, general use statistics, operating characteristics, facility-related
8	Recreation	Marine, motorcycle, snowmobile, recreation activity, recreation vehicle, other
9	Socioeconomic	Population, household, consumer price index, income and wages, employment, unemployment, business activity
10	Travel data	Residence, work place, mode, carpooling, commuter movement, travel distance, travel time, sociodemographic characteristics, public transportation users, changes in use, trip purpose, trip rates per household, automobile ownership per household, vehicle occupancy rates, truck use, automobile use, vehicle use, rates, seasonal variation, automobile drivers, automobile owners, household characteristics, traveler characteristics, vehicle owner characteristics, tract- and zone-level data, employment characteristics, special interest questions, other

Table 2. Fuel use—data sources.

Source	Electricity	Jet Fuel	Avg Gasoline	Diesel
FAA, census of civil aircraft			Survey results by state from actual secondary sources	
FHWA, table MF-24 private and commercial nonhighway use of gasoline			Annual reports by state from actual and synthesized sources	
Ethyl Corporation		Monthly reports by state from actual secondary sources, lag time of 1 quarter	Monthly reports by state from actual secondary sources, lag time of 1 quarter	
FHWA, table MF-21 motor fuel use, highway and nonhighway				
FHWA, table MF-22 motor fuel				Monthly and annual reports state from actual and synt sized sources
FHWA, table MF-23 highway use, motor fuel				Monthly and annual reports state from actual and synt sized sources
FHWA, table MF-25 special fuels, highway use				Monthly and annual reports state from actual and synt sized sources
FHWA, table MF-26 highway use, gasoline				
FHWA, table MF-33G gross gallons of gasoline				
FHWA, table MF-2 motor fuel use				Annual reports by state fro actual and synthesized sou
FHWA, table MF-21A total gasoline use				
Port Authority of New York and New Jersey		Monthly regional reports from actual primary sources, lag time of 3 months		
FHWA form PRR 551M				
New York State Department of Taxation and Finance MT104.19		Monthly statewide reports from actual and synthesized sources, lag time of 2 months		Monthly statewide reports f actual and synthesized sou lag time of 2 months
New York State Energy Office	Annual statewide reports from actual secondary sources	Monthly statewide reports from actual secondary sources		Monthly statewide reports f actual secondary sources
State energy data report DOE/EIA 0214(78)	Annual national and statewide reports based on actual primary and secondary sources, lag time of 2 years	X	X	X
Monthly energy review DOE/EIA 0035(8)01		Monthly national reports based on actual primary and secondary sources, lag time of 1 quarter		X
Weekly petroleum status report DOE/EIA 028(81/4)		Weekly national and regional reports based on actual primary and secondary sources, lag time of 2 months		
American Public Transit Association (APTA) Transit Fact Book				X
APTA Transit Operating Report				X
Association of American Railroads statistics of railroads of class 1-U.S.				X
U.S. Department of Energy, U.S. Bureau of the Census foreign trade bunker fuel				Monthly substate reports ba on actual and synthesized sources, lag time of 3 mon
U.S. Department of Energy environmental impact assessment, energy data reports, sales of fuel oil and kerosene				
American Bus Association				X
Oak Ridge National Laboratory, region analysis of highway energy use				X
Gasohol, National Gasohol Commission				
Census of retail sales geographic area RC 77A-33				
Tax and finance survey reporting				
Preliminary Research Report 187				
Air Transport Association		Annual national report based on actual secondary sources and general statistics		
Census of retail sales miscellaneous subjects RC77-S-Z				

Note: X is a use designation.

Gasoline	Highway Use	Nonhighway Use	Aviation	Rail or Transit	Bus	Marine	Truck	Other
				X				
Annual reports by state from actual and synthesized sources		X						
Monthly reports by state from actual secondary sources, lag time of 1 quarter				X				
Annual reports by state and substate from actual and synthesized sources		X						
Monthly and annual reports by state from actual and synthesized sources	X	X						
Monthly and annual reports by state from actual and synthesized sources	X							
	X				X	X	X	X
Monthly and annual reports by state from actual and synthesized sources	X							
Monthly reports by state from actual and synthesized sources, lag time of 3 months	X	X						
Annual reports by state from actual and synthesized sources	X	X						X
Annual reports by state from actual and synthesized sources	X	X						X
			X					
Monthly reports by state from actual and synthesized sources, lag time of 2 months	X	X						
	X	X	X					
Annual substate reports based on estimates and actual secondary sources, lag time of 1 year				X				
X								
X								Factors, general statistics
X								
X					Annual national report based on actual secondary sources			
X					Annual national report based on actual secondary sources			
						X		X
						X		
					X			
					X		X	
X								
National, statewide, and substate survey results based on actual primary sources and general statistics								
Statewide and substate reports based on actual primary sources, lag time of 2 years								
Substate-level reports based on estimates								
National, statewide, and substate survey results based on actual primary sources and general statistics								

Table 3. Information sources.

Item	Rochester, Monroe County	Oneonta, Otsego County
Local transit	NYSDOT Transit operators	NYSDOT Transit operators
Intercity transit	Interstate Commerce Commission Transit operators Industry associations	Interstate Commerce Commission Transit operators Industry associations
Aviation		
Passenger certificated	I.P. Sharp	
Commuter	I.P. Sharp	I.P. Sharp
Freight	I.P. Sharp	I.P. Sharp
Rail		
Passenger	National Rail Passenger Corporation Station master	
Freight traffic	Consolidated Rail Corporation Chessie System U.S. Department of Transportation	 Delaware and Hudson U.S. Department of Transportation
Marine, freight tonnage[a]	U.S. Army Corps of Engineers	
Automobile registration	R.L. Polk New York State Department of Motor Vehicles	R.L. Polk New York State Department of Motor Vehicles
Truck registration	R.L. Polk New York State Department of Motor Vehicles	R.L. Polk New York State Department of Motor Vehicles
Drivers' licenses	New York State Department of Motor Vehicles	New York State Department of Motor Vehicles
Traffic counts	NYSDOT	NYSDOT
Thruway counts	New York Thruway Authority	
Automotive fuel use		
Automotive fuel prices	Local survey Platt and Lundberg Surveys	Local survey

[a]Rochester port was disbanded in 1974.

fuel sales (3,4). These data are based on tax receipts for motor fuel sales at the wholesale distributor level. The lag is approximately 6 weeks between the sale as reported by New York State Department of Taxation and Finance and the actual transfer of the fuel to the consumer.

The New York State Energy Office has prepared a proportional county use profile based on wholesale fuel sales data (5). The source document is not available to the public; however, estimates of county-level gasoline data may be developed by using various other sources of information such as the following:

1. Census of retail sales (6) reports by state, county, and SMSA the number of establishments and the amount of retail sales by business category (published every 5 years); and
2. Miscellaneous reports from the census of retail sales (7) give detailed data on the gallon sales of gasoline and other automotive fuels by state and SMSA.

Remember that census data are based on samples, and some information may be withheld for reasons of confidentiality.

The first six methods for estimating substate highway fuel consumption are described in detail in the following sections.

Network Approach

Gasoline is estimated by extracting travel data from computer network assignments, traditionally VMT, and multiplying these data by an estimate of vehicle efficiency. For example,

$$\text{Gal of gasoline} = [1/\text{MPG}]\ [\text{VMT}] \quad (1)$$

where MPG is miles per gallon.

The Results File for NYSDOT traffic assignments in urban areas shows the gallons of gasoline consumed per average 24-hr period by ring, district, and zone. By aggregation and multiplying by 330 (the number of equivalent days in the assignment model year), the number of gallons for each appropriate study area can be obtained.

$$\text{Gal of gasoline} = (\Sigma\ \text{Districts or zones})\ (330) \quad (2)$$

Household Approach

A statewide estimate of travel per household (in miles) is multiplied by the average vehicle efficiency and the number of households in the study area. Travel per household is defined as the average annual mileage per vehicle times the average number of vehicles per household. Number of vehicles and miles per gallon are weighted to reflect differences between passenger cars and trucks.

$$\text{Gal of gasoline} = [1/\text{MPG}]\ [\text{Annual avg miles/Vehicle}] \times [\text{Avg number of vehicles/Household}] \times [\text{Number of households}] \quad (3)$$

Stations Approach

Gasoline consumption is estimated by multiplying the average gasoline sales volume per station by the number of stations in the study area. The number of stations varies depending on the source. The sources in this study included actual service station counts (6) and estimated number of stations (8) based on population, New York State Energy Office gasoline consumption, vehicles registered per square mile, and an average of these three. The general equation is

$$\text{Gal of gasoline} = [\text{Avg gasoline sales volume/Year}] \times [\text{Number of stations}] \quad (4)$$

Retail Sales Approach

Gasoline consumption is estimated by multiplying the retail sales at service stations by a factor of 0.75 to reflect the portion of sales that corresponds to gasoline and then dividing by the average monthly or annual gasoline price. Data on retail sales at service stations are available in the 1972 and 1977 Census of Retail Trade or from the state tax department on a more current basis. The equation is as follows:

$$\text{Gal of gasoline} = [\text{Retail service station sales (\$)}]\ (0.75) \times [1/(\text{Avg \$/Gal})] \quad (5)$$

This method has several critical assumptions that affect the estimation of gallons sold. First, the percentage of gasoline sales is estimated as a portion of total service station sales. Second, retail sales taken from the census are limited to establishments where gasoline sales are 50+ percent of retail sales.

Gasoline is sold not only through gasoline and service stations (other merchandisers sell gasoline; e.g., Sears, and major users may obtain supplies directly from wholesalers); therefore, this method will underestimate sales and will not be useful for ongoing monitoring because of the long intervals between up-to-date source data.

Another variation of the retail sales approach is to use localized sales tax reports, where available. In New York State some counties and cities have local sales taxes. In these instances the New York State Department of Taxation and Finance is able to calculate the dollar amount of local sales of goods and services at gasoline retail stations. Proceeding as shown in Equation 5, the current gasoline sales at the substate level can be estimated. The New York State Department of Taxation and Finance will not publicize the sales tax collection data, however, because of departmental confidentiality rules. On the local government level, such data can be obtained from local treasurers' offices after receipt of the tax collected from the New York State Department of Taxation and Finance.

This method holds the greatest promise of accurate and timely reports of substate fuel sales for local government entities that have a sales tax of their own. Note that, in New York State, this procedure addresses the problem of a major distributor that serves a large area made up of several local government entities. The sales tax in New York State is levied at the point of delivery, not at the location of the office of the seller. Thus, if fuel is delivered to the location of a major user, it would be reported at that location.

Partition Approach

The New York State Energy Office statewide total of gallons of gasoline consumed is partitioned by the appropriate study area share (5). Percentage shares are based on county sales tax revenue (which can be adjusted for counties that do not have a sales tax), number of service stations in the county, and sales revenue for businesses in the county categorized as gasoline service stations.

$$\text{Gasoline} = [\text{Total statewide gasoline consumption}] \times [\text{Partitioning share}] \quad (6)$$

Registration Approach

The registration approach obtains a measure of regional fuel consumption from vehicle registrations, annual miles driven per vehicle, and average fuel efficiency. Clearly, on the substate level, all three of these measures may be difficult to obtain. Vehicle registration data are (at considerable expense) available from R.L. Polk and Company if unavailable from state motor vehicle departments. Fleet efficiencies can be calculated by using the historical new car Polk registration data and new vehicle efficiencies reported in annual U.S. Department of Energy and U.S. Environmental Protection Agency Gas Mileage Guides. National fuel efficiency numbers may also be used.

$$\text{Gal of gasoline} = [1/\text{MPG}]\ [\text{Avg annual miles/Vehicle}] \times [\text{Number of vehicles}] \quad (7)$$

Annual miles driven per vehicle is probably the most difficult variable to obtain for this approach and the one to which this method is most sensitive.

This method appears to have considerable merit for ongoing monitoring efforts where the focus is an assessment of the impact of changing fuel efficiencies on fuel consumption in a certain area. In this case annual miles driven per vehicle could be held constant to conduct a ceteris paribus analysis.

Results of Case Study

Table 4 gives a summary of results for Monroe County by using six of the nine methods developed by NYSDOT. A comparison of the results of the methods shows that the partition approach gives an estimate at the high end of the range, and the retail sales approaches give the lowest estimates. The range of estimates is large compared with the mean of the methods. A presentation of the complete nine methods is shown elsewhere (2). This pattern would point to systematic errors in the procedures. To resolve which procedure would give more accurate results it might be useful to estimate fuel consumption for each area in the state by each method, aggregate each method's total, and compare the difference between individual totals and a statewide total reported from other sources. This test is beyond the scope of this paper.

Many of these methods deal with fuel sales; however, our focus is usually on fuel use. The available data reflect sales within a geographic area; however, sales in an area do not necessarily reflect consumption in that specific area. For example, price differential due to the effects of local sales taxes may affect border crossings; corridors of interstate travel and tourist areas may also be affected. Sale versus use distortion for a particular political subdivision can possibly be explained by service station location, proximity to highway systems, and density of development.

The estimation of diesel fuel use at the substate level is a more complex problem even though the registration, retail sales, or partitioning approach

Table 4. Summary of estimates of gasoline consumption by six NYSDOT methods for Rochester area.

Method	Year of Estimate	Gallons (000s)
Network	1977	235,759
Household	1979	265,735
Number of stations		
Census of retail trade	1977	161,874
New York State Department of Environmental Conservation	1979	248,270
Retail sales		
New York State Department of Taxation and Finance	1978	144,532
Census of retail trade	1979	174,270
Partition		
New York State Energy Office, tax based	1978	331,731
Fraction of retail sales at service stations	1977	321,190
Fraction of total service stations	1977	262,284
Registration	1978	306,152
Range		187,199[a]
Mean		254,234

[a]Calculated by 331,731 - 144,532 = 187,199.

would appear to be applicable. For now diesel consumption by the passenger automobile sector is not a significant factor. Diesel use is largely dependent on trucks (found in R.L. Polk and Company reports), buses (found in annual transit operating reports of the American Public Transit Association), and nonhighway vehicles; therefore, diesel fuel use does not have the same relationship as the one between automobiles and gasoline sales. Diesel city bus operations can be identified easily from annual operating reports. The location of other users of diesel fuel is difficult to determine, however, because the service provided, and consequently the fuel used, may not be in the location in which the vehicle is registered or the fuel is purchased. As a result, estimates of diesel fuel use may vary significantly with respect to actual use. In recognition of the obvious limitations in using truck registration, retail sales, or a partitioning approach, the estimates obtained should be used judiciously.

Discussion of Results

This case study covers only one of the 10 categories noted in Table 1; however, it is characteristic of most of the problems encountered in attempting to identify data for monitoring purposes.

The activities of NYSDOT have not been limited to the preparation of the Data Source and Reference Directory (1) and the subsequent examination of data reporting problems. NYSDOT maintains an extensive computer data base that encompasses most of the 10 categories noted in Table 1. In addition, NYSDOT prepares a transportation statistics report that monitors and tracks trends in key measures of transportation activity. Most of the data readily available reflect direct or indirect measures of transportation system background activities. For example, a direct measure is gasoline sales, which reflect, to a large extent, travel actions or purchase actions of the motoring public. Employment affects transportation use as well as trip generation. This type of sociodemographic measure represents a somewhat more indirect measure of the transportation background. By monitoring the activities of interrelated parameters one can keep track of the transportation system's performance at varying points. In some instances likely near-future performance based on comparison with earlier data can be anticipated; however, bumps in the data or trend lines are not always easily explained.

Of the 10 categories listed in Table 1, six basic parameters emerge that can be monitored to best detect changes in system travel, travel behavior, and the subsequent change in transportation energy use. These are most directly related to transportation system performance:

1. Fuels use--gasoline and diesel,
2. Fuel prices,
3. Vehicle registrations,
4. Vehicle efficiency,
5. Use of transit and transportation service data, and
6. Highway use.

The other categories noted in Table 1 affect travel background and transportation energy use, but not to the extent of these six variables.

GENERAL FINDINGS

Data collection and monitoring are extremely resource-intensive, both in terms of direct dollar expenditure and time involved in finding appropriate sources. A good starting point would be a compendium of available statistical data sources, such as the Data Source Directory (1). This directory should be comprehensive in sources and level of detail available. Some sources may no longer be readily available because of the current trend toward deregulation of certain transportation industries (e.g., demise of the Civil Aeronautics Board and proposed trucking deregulation). The available commercial data sources, such as R.L. Polk and Company and I.P. Sharp and Associates, should be examined. These sources may have the required information, but the cost factor could be significant. The cost element would be a serious consideration for many state and county departments involved with transportation and transportation energy.

Another concern is that of the data being disaggregated to the proper level. The collected data series should match the data requirement; i.e., gasoline use at the county or urban level. Problems with estimation have already been discussed. A third difficulty for data users is consistency of the data over time. Any inconsistency (because of changes in collection or preparation procedures) can lead to serious difficulties in interpretation of trends.

ACKNOWLEDGMENT

We are deeply indebted to all those in NYSDOT who were associated with the compilation of the Data Source and Reference Directory of Transportation and Energy Data. In addition, we wish to acknowledge the patient typing of Dianne DeVoe, Charron Many, Tedi Toca, and Judy Khachadourian.

REFERENCES

1. N.S. Erlbaum. Data Source and Reference Directory of Transportation and Energy. New York State Department of Transportation, Albany, Draft Rept., Task A of UMTA Study, Oct. 1981.
2. N.S. Erlbaum, C.E. Meyers, and D.T. Hartgen. Procedures and Data Needs for Estimating Urban Area Gasoline Consumption. New York State Department of Transportation, Albany, Prelim. Res. Rept. 187, Aug. 1980.
3. Motor Fuel Monthly Gallonage and Revenue Rept. New York State Department of Taxation and Finance, Albany, Form MT 104.9.
4. Monthly Motor Fuel Consumption. FHWA, Form PR 551M.
5. M. Schultz. Estimated Annual Gasoline Consumption by County, Methodology Used in Preparing County Gasoline Estimates. New York State Energy Office, Albany, memoranda, Nov. 14, 1980 and Dec. 19, 1980.
6. 1977 Census of Retail Sales. Bureau of the Census, U.S. Department of Commerce, Geographic Area Series for New York, RC77A-33, Table 7, Aug. 1979.
7. 1977 Census of Retail Sales. Bureau of the Census, U.S. Department of Commerce, Table 3, Miscellaneous Subjects, RC77S-2, Aug. 1979.
8. D.C. Perley. Number and Distribution of Bulk Storage Tanks in New York State. New York State Department of Environmental Conservation, Albany, Department Rept., Aug. 1980.

Publication of this paper sponsored by Committee on Statewide Multimodal Transportation Planning.

Computerized Method for Updating Planning Data Bases Used in Travel Demand Forecasting

LARRY W. McPHERSON, CLINTON L. HEIMBACH, AND LARRY R. GOODE

Data bases that have been created for use in urban transportation studies describe the urban environment for only one point in time. With the exception of simple manual techniques, no provision has been made for updating these planning data bases on a continuing basis. The objective of this study was the design and implementation of a computerized information system capable of supporting the continuing socioeconomic data requirements for the travel demand forecasting phase of the urban transportation planning process. Agency operating and administrative records from state and local governments serve as system input. The system performs a geographical analysis that determines the home base locations for these records and, hence, the data that they contain. Subsequent to this geographical analysis, the system aggregates the data on a small area basis required by the travel demand models in the planning process. Requirements for data confidentiality established by law for certain socioeconomic data led to the development of a mathematical model that predicts the income variable. The model is based on harmonic analysis and is specified as a Fourier series. The transportation planning information system developed during the course of this research is capable of synthesizing, on a small area basis, the demographic and employment data used in the transportation planning process for urban study areas that have populations less than 500,000. This conclusion is supported by the findings from system implementation and testing in Greensboro, North Carolina.

Data bases created for urban transportation studies describe the urban environment for only one point in time. With the exception of simple manual techniques, no provision has been made for updating these planning data bases on a continuing basis. The difficult, and sometimes impossible, task of continually updating the transportation planning data base is reflected in both direct and indirect costs. Direct costs are straightforward expenditures of time or money. An example of direct cost is the 8 person-months required to manually update the employment and dwelling unit inventory for an urban area that has a population of approximately 200,000 persons. Indirect costs accrue whenever the transportation planner cannot respond to the transportation needs of the urban area because the planning data base is not kept in an accurate, complete, and current condition. Indirect costs tend to be a function of how well the initial planning data inventories are updated.

Currently, no means exist, except on an ad hoc basis of staff estimates, for keeping the planning data inventories of household size, household income, and car ownership updated. Even though quantification of these indirect costs is difficult, over time the indirect costs would exceed the direct costs involved in updating the planning data base. Therefore, this study focused on the development of a computerized system capable of monitoring and recording those changes in the urban environment that can then be used to update the transportation planning data base periodically. The updated transportation planning data base can then be used to produce estimates of travel demand derived at the household level, or any other higher level of aggregation desired.

LITERATURE REVIEW

Writers in the field of transportation engineering generally agree that the ultimate success of the urban transportation planning process (UTPP) depends in large part on an information system that can provide a source of current, accurate, and timely planning data. These same writers are in less agreement, however, about the design specifications for such an information sytem. Regardless of the various features that could be manifested in such a system design, the consensus was that any information system designed to support UTPP must be capable of monitoring and reflecting socioeconomic change that may occur in single-family households throughout the planning area under consideration. The single-family household is the basic unit of analysis used in the travel demand phase of UTPP and is the reason that any well-designed data monitoring and retrieval system must be capable of quantifying change in any of several household characteristics, including (a) family size, (b) family income, and (c) family car ownership. Street addresses were the suggested means for linking or referencing these data to specific household units in the planning study area under consideration.

The literature review disclosed no general theory of information system design that could serve to establish an operational framework for a particular data system capable of supporting the data requirements for transportation planning. Recent work, however, in the disciplines of information and computer science allows a logical and common-sense design approach to the urban transportation planning information system (UTPIS). One such conceptual design by Peat, Marwick, Mitchell and Company (PMMC) demonstrates such an a posteriori approach to UTPIS development (1).

The major findings of the PMMC study that are germane to the research reported herein are as follows:

1. The potential for using secondary data sources (i.e., government records) to update transportation planning data bases and
2. The use of computerized geocoding techniques for linking these data sources to geographical areas of interest.

METHODOLOGY

The design and development of the North Carolina Urban Transportation Planning Information System (NCUTPIS) required two analytical techniques. They can be listed in order of importance to this research effort as: (a) computerized geocoding and (b) harmonic analysis.

Computerized Geocoding

The single most important element in the design of NCUTPIS was the geocoding subsystem that can assign locational (areal) codes to government source-data records automatically. A geocoding system provides the mechanism for linking together a variety of diverse data sources that may have only one common link (e.g., street addresses). After the data have been linked to specific areal units (e.g., traffic zones) they can be aggregated to any other higher level of spatial detail. The geocoding process provides a convenient framework for continuous monitoring and updating of transportation planning data. The address matching and geocoding process is illustrated in Figure 1 (2).

Figure 1. Address matching and geocoding process.

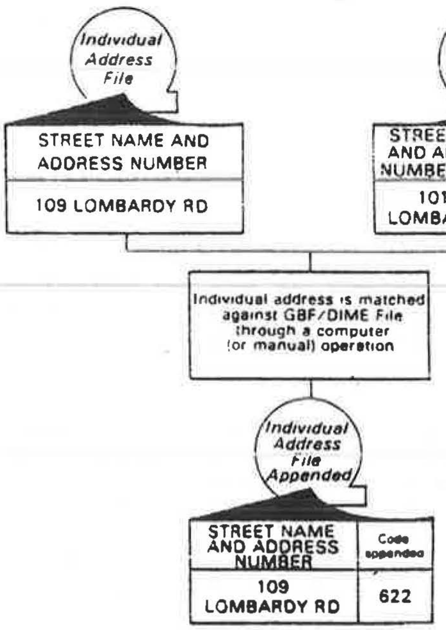


Figure 1 shows that address matching is simply the process of matching data records in two files (address and reference) on the basis of street name and number. In the example 109 Lombardy Road (address file) is first matched to a range of addresses 101-199 along Lombardy Road (reference file). After the match has been made the geographic identifier (geocode) 622 is appended to the matched record. This process of geographic identification is referred to as geocoding.

Once a file of individual addresses has been geocoded, any data related to those addresses can be summarized according to any geographical or areal units of interest. For example, the total number of vehicles can be tabulated according to the traffic zones in which they are registered.

A computerized geocoding system is inherently complex but potentially powerful. An example of such a system is the UNIMATCH geocoding procedures developed by the U.S. Census Bureau. UNIMATCH is complex but has a powerful potential for geocoding. The U.S. Department of Transportation (USDOT) sponsored a project in 1978 aimed at facilitating use of that system. The UNIMATCH system includes procedures for address matching, printer graphics, generalized record linkage, and programmable text generation (3). The technology transfer project sponsored by the USDOT consisted of installing these procedures in the Urban Transportation Planning System (UTPS) environment for convenient use by transportation systems analysts, engineers, and planners. This interagency project was successful and has resulted in the most sophisticated, generalized record linkage system currently in existence in the field of transportation (4). These address matching-geocoding procedures were first field tested successfully at the North Carolina Department of Transportation's (NCDOT's) Computing Center in June 1979.

Harmonic Analysis

Harmonic analysis was the second most important analytical technique used in developing NCUTPIS. Harmonic analysis was used to express the income-estimating model as a Fourier series. Functional expansions in Fourier series and their uses are part of a branch of mathematics known as orthogonal functions. These functions have many important engineering applications because of their curve-fitting characteristics. Often a Fourier series is written for some function whose values are given numerically. The process of finding the Fourier coefficients for such a function is called harmonic analysis (5).

Many applied mathematical texts treat the situation where a given function [f(X)] can be expanded over the interval (-L, +L) and into a series of the type:

$$f(X) = (a_o/2) + \sum_{n=1}^{\infty} a_n \text{Cos}\,[(n\pi X)/L] + \sum_{n=1}^{\infty} b_n \text{Sin}\,[(n\pi X)/L] \quad (1)$$

and a_n and b_n are given approximately by

$$a_n = 2/K \sum_{P=o}^{K-1} f(X_p) \text{Cos}\,[(n\pi X_p)/L] \quad (2)$$

and

$$b_n = 2/K \sum_{P=o}^{K-1} f(X_p) \text{Sin}\,[(n\pi X_p)/L] \quad (3)$$

when the expansion interval (-L, +L) is divided into K equal parts of length $\Delta X = 2L/K$ and X_p (P = 0, 1, 2, . . . K-1) are the finite division points.

Income Model Construction by Harmonic Analysis

Manual construction of income models can be a tedious and error-prone task. The computer program HARANAL was written to help with this developmental task. The program is designed to read tabular percentages, fit the data with a harmonic series, and prepare a printer plot of the results (6).

Data for the calibration of income models are usually available from comprehensive origin and destination surveys. Other sources of data are the 1970-1980 Census Urban Transportation Planning Packages. Car ownership-income models have been developed from census data for a number of urban centers across the United States, including the North Carolina cities of Greensboro, Winston-Salem, High Point, and Raleigh (7). None of these models, however, was based on mathematical analysis.

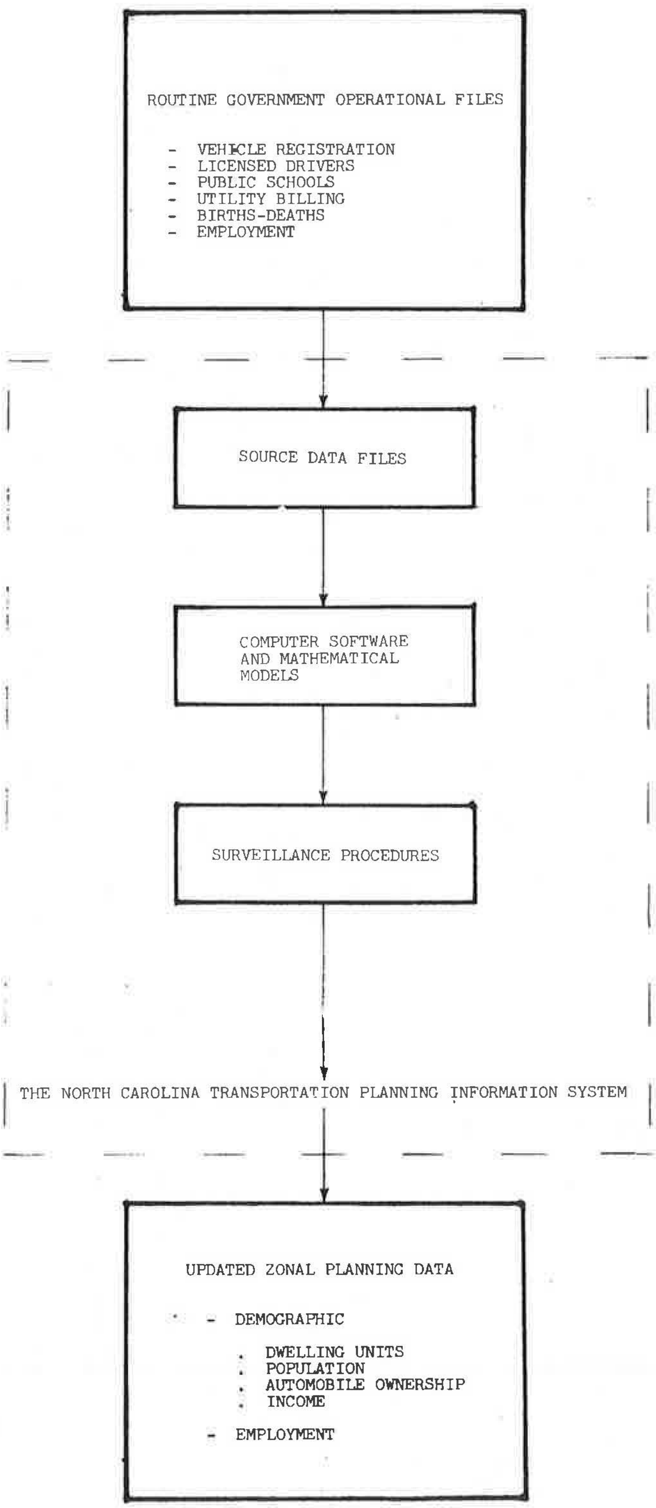
CONCEPTUAL STRUCTURE OF NCUTPIS

Although the conceptual structure of the transportation planning data system developed during this research was based on the PMMC study, it is distinguished from that work in three important aspects.

1. In the design of NCUTPIS major emphasis was placed on the travel demand data requirements of North Carolina's current urban transportation planning process. No attempt was made to develop a system that would meet all of the urban center's potential land use and transport planning data requirements. Design emphasis was placed on meeting the immediate needs of the continuing UTPP and also on providing the flexibility for future expansion of the system to include the collection of other high-priority data items.
2. An attempt was made to make NCUTPIS more user-oriented by use of the geocoding procedures found in the familiar UTPS environment.
3. NCUTPIS includes computerized techniques for development of an income model.

The conceptual structure of NCUTPIS that was designed to produce current planning data used for travel demand is shown in Figure 2. Figure 2 illustrates four major components. The first component consists of six source data files. All of these source files can be obtained from the routine operating or administrative records of state or local government agencies. These six files are used to feed original source data to NCUTPIS . Table 1 summarizes the data contents of each file and also indicates ownership and geographic coverage.

Figure 2. Conceptual structure of NCUTPIS.



The second major component of NCUTPIS consists of all the computer programs needed to extract, format, merge, sort, and summarize the planning data. This software is contained in the UTPS Macro Procedure Generator System (UGEN) or maintained in the utility procedure library of NCDOT's Computing Center.

Mathematical models make up the third major system component. As expected, limitations exist on the use of original source data. In many situations legislation imposes confidentiality requirements on the use of agency data that relate to employment, income, revenue, or sales. Mathematical models can be constructed and used in a manner that complies with established confidentiality requirements. The income-estimating model used in NCUTPIS is a good example of the synthesis of sensitive data.

The fourth system component consists of two surveillance subsystems designed to produce current demographic and employment data on a small zonal basis. The demographic subsystem is designed to produce the following data for each traffic analysis zone in the study area:

1. Total number of single family housing units,
2. Total population,
3. Population less than 6 years old,
4. Population 6 to 17 years old,
5. Population 18 to 64 years old,
6. Population more than 64 years old,
7. Total number of vehicles,
8. Number of vehicles by age group,
9. Number of vehicles by weight classification,
10. Number of trucks,
11. Number of pickup trucks,
12. Number of motorcycles, and
13. Average income.

The employment subsystem is designed to produce estimates of employment by place of work for each traffic analysis zone broken down by the standard industrial classification (SIC) categories.

Source Data Files

Six magnetic tape, source data files are used as input into NCUTPIS and are listed in Figure 2. Table 1 summarizes the data content of these source files.

Motor vehicle registration was obtained directly from NCDOT's Division of Motor Vehicles vehicle master file. Data contained in this file can be geocoded and summarized by traffic analysis zone. These summaries can serve as direct input for travel demand models used in the continuous planning process. They also serve as the primary input to the income estimating model.

Data on dwelling units were obtained directly from public utility water billing records. Other public utility billing records, including telephone and electricity, can provide additional housing data. These billing records are geocoded and summarized by traffic analysis zone. The housing counts derived from these billing records serve as direct input into travel demand models.

Data on the population age group 16 years and older were gleaned from the NCDOT's Division of Motor Vehicles' licensed driver master file. This file required significant modification before it was used in NCUTPIS. The driver file normally resides in disk storage. The driver records had to be extracted and output as a tape file before they could be geocoded by NCUTPIS. Population data on the age group 5 to 15 were obtained from public school registration records. Usually two separate school files must be geocoded, one from the county and one from the city. Population data on the age group birth through 5 were obtained from the Department of Human Resources.

Table 1. Data file source, ownership, geographic coverage, and use restrictions.

Source of Data	Ownership	Type of Data	Geographic Coverage	Restrictions
North Carolina DOT's Division of Motor Vehicles vehicle master file	NCDOT	Vehicle registration by type and owners' addresses	State and county	Confidential data—release controlled by Division of Motor Vehicles
Local utility billing master file	City	Dwelling unit addresses	City	None
North Carolina DOT's Division of Motor Vehicles licensed drivers master file	NCDOT	Licensed drivers' addresses	State	Confidential data—release controlled by Division of Motor Vehicles
North Carolina Public Schools	County and city	Student addresses	County and city	Confidential data—release controlled by county or city public schools
North Carolina Department of Human Resources birth/death file	Department of Human Resources	Addresses of births and deaths	State	Confidential data—release controlled by Department of Human Resources
North Carolina State Employment Security Commission summary file	N.C. State Employment Security Commission	Address of employment by SIC code	State and county	Confidential data—release controlled by state law through Employment Security Commission

Employment data are the single most difficult data that the transportation planner has to assemble on a zonal basis. The most promising sources of employment data are the files maintained by state departments responsible for administering unemployment insurance programs (i.e., the State Departments of Labor and Employment Security). NCUTPIS extracts employment data from the Employment Security Commission's master file.

Collectively, the six source data files shown in Figure 2 provide the current, raw data input to NCUTPIS. This system extracts and then, on a zonal basis, summarizes the socioeconomic data found in the source files. These data can then be used directly or indirectly as input to travel demand models.

Computer Software

The set of computer programs used to support the operations of NCUTPIS includes two major components:

1. The address matching and geocoding procedures and
2. The generalized utility procedures used for the sorting, merging, summarizing, manipulating, and analyzing tasks.

These procedures are currently maintained at the NCDOT Computing Center. This center will supply program documentation on user request.

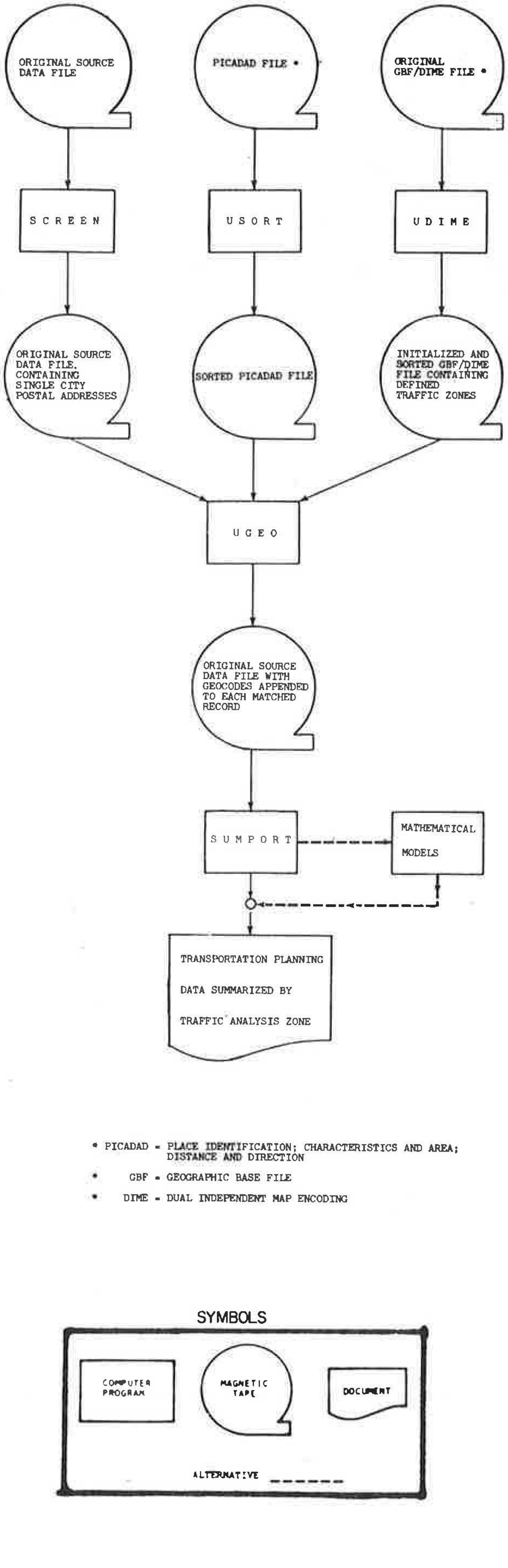
Address Matching and Geocoding Procedures

The process of record linkage is sometimes referred to as file matching. File matching involves the transfer of data from one file to another when certain matchkeys and predefined criteria are met. Address matching is an important special case of record linkage. Four UTPS computer programs accomplish the address matching and geocoding tasks (3):

1. UGEN--The UTPS procedure generator (UGEN) was written so that some severe limitations of IBM procedures (PROCS) could be overcome. These limitations include (a) no conditional generation of data description (DD) cards, (b) PROCS could not be used to generate or modify program data cards, and (c) PROCS cannot invoke other PROCS. UGEN overcame these limitations by providing a method for the UTPS user to enter familar UTPS Job Control Language (JCL) and parameters and have the required JCL and parameters for the record linkage programs generated. UGEN input may consist of several jobs where each job may contain several steps. The EXEC cards may contain references to programs, catalogued procedures, or MACRO PROCS (any procedure recognized by UGEN) that can be interpreted by UGEN. UGEN outputs a file that is read by the operating system and executed as a sequence of one or more batch jobs.
2. USORT--USORT is a MACRO PROC designed to prepare a single sort JOB step for execution. USORT calculates required work space automatically. This PROC requires no SYSIN file for the sort parameters because they can be given as EXEC keywords. The USORT PROC is used to sort the PICADAD file in preparation for input into the UGEO PROC.
3. UDIME--The UDIME MACRO PROC prepares GBF/DIME files for input to the UGEO MACRO PROC. The original GBF/DIME files are created by the CUE (correction, updating, and extension) program from the census. These files must be initialized by the ZIPSTAN and GBFSPLIT programs and then sorted. UDIME accomplishes these job steps with a minimum of user specifications. UDIME outputs a permanent file that becomes the GBF input file for the UGEO program. UDIME also provides the single most important design feature for NCUTPIS. It provides the user with the capability of defining traffic analysis zones in terms of the geography contained in the GBF/DIME file. As with the other MACRO PROCS, UDIME must be invoked with UGEN.
4. UGEO--The UGEO MACRO PROC, like other UTPS PROCS, has a compilation and an execution phase. The execution of UGEN constitutes the compilation phase. The execution phase consists of a series of nine complex steps of computer processing available through a single EXEC card and several keywords. UGEO requires two reference files and one data file as input. The reference files are PICADAD and GBF/DIME for the particular urban center for which data are being geocoded. PICADAD is used to obtain place codes for each city name. UGEO outputs a file that contains the original data records, each of which are appended with a matchkey composed of standardized address fields and any geocodes that the analyst has requested.

In summary, UGEN was developed to simplify the use of computerized address-matching systems. The computer processes in the system are inherently complex. System users required automation of these processes in order to prevent a labor-intensive and error-prone operational task. Hence, UGEN was designed as the interface between the user and the computer operating system (OS). UGEN prevents much of the drudgery, complexity, and redundancy involved in user specification of sophisticated computer pro-

Figure 3. Flowchart for address matching and geocoding process within UTPS environment.



cesses. The UTPS user prepares a card deck with the step that invokes UGEN via a catalogued JCL procedure. The data for this step include the user's whole job. UGEN reads the user's JCL and UTPS parameters, generates a JCL job stream, and later submits this job to the computer OS if no errors are detected (3).

Generalized Utility Procedures

Several utility programs were developed to support the UTPS computer programs discussed. The general purpose of these utility programs is to prepare the source data for input into NCUTPIS and to summarize and report on the geocoded output data. One special purpose program provides the user with a generalized curve-fitting procedure and is capable of preparing a computer plot of a user-specified income model.

Currently, four utility programs support the UTPS geocoding programs:

1. SCREEN--SCREEN reduces a source data file that contains addresses for several cities to a data file that contains addresses for a single city. This file can be geocoded much faster than can the unreduced file.
2. REFORMAT--REFORMAT reformats records from the data file that may not be suitable for input to the UTPS geocoding programs. For example, the street name field may precede the street number field. The UTPS geocoding programs require the switching of these fields by the REFORMAT program.
3. SUMPORT--SUMPORT provides for the summary, reporting, manipulation, and analysis of the data that are located in the geocoded records output from the geocoding programs.
4. HARANAL--HARANAL provides a generalized curve-fitting capability that is based on Fourier series and harmonic analysis. The output includes a harmonic equation of the fit and a printer plot of the calculated and observed tabular values. This procedure can prepare a printer plot of a user-specified income model.

The sequencing of both the utility and UTPS programs is shown in Figure 3. This sequencing is intended to manifest the essential aspects of the data acquisition tool (NCUTPIS) that was designed during this study to support the travel demand models used in UTPP.

Income Models

Development of a mathematical model to estimate income became necessary because of confidentiality requirements established by law. Federal and state revenue agencies are prohibited by law from releasing their income tax files to other agencies, even for planning purposes. If not for this constraint, an income file would have existed from which zonal incomes could have been extracted directly. Fortunately, the confidentiality constraint can be overcome by use of the income-estimating model.

Surveillance Subsystem

The fourth system component was designed to collect zonal summaries of urban transportation planning data on the two functional categories of demography and employment. The structure of each subsystem is shown in Figures 4 and 5.

Demographic Subsystem

The demograhic summary procedures will produce 13 items of demographic data annually for each traffic analysis zone in the study area.

The demographic subsystem structure requires five of the source data files (shown in Figure 4) as subsystem input.

Employment Subsystem

The employment subsystem is designed to produce zonal summaries of employment broken down by SIC codes. The input to this subsystem was the North Carolina State Employment Security Commission master file, which includes the number of employees for all establishments except small businesses and government agencies. The employment subsystem structure is shown in Figure 5.

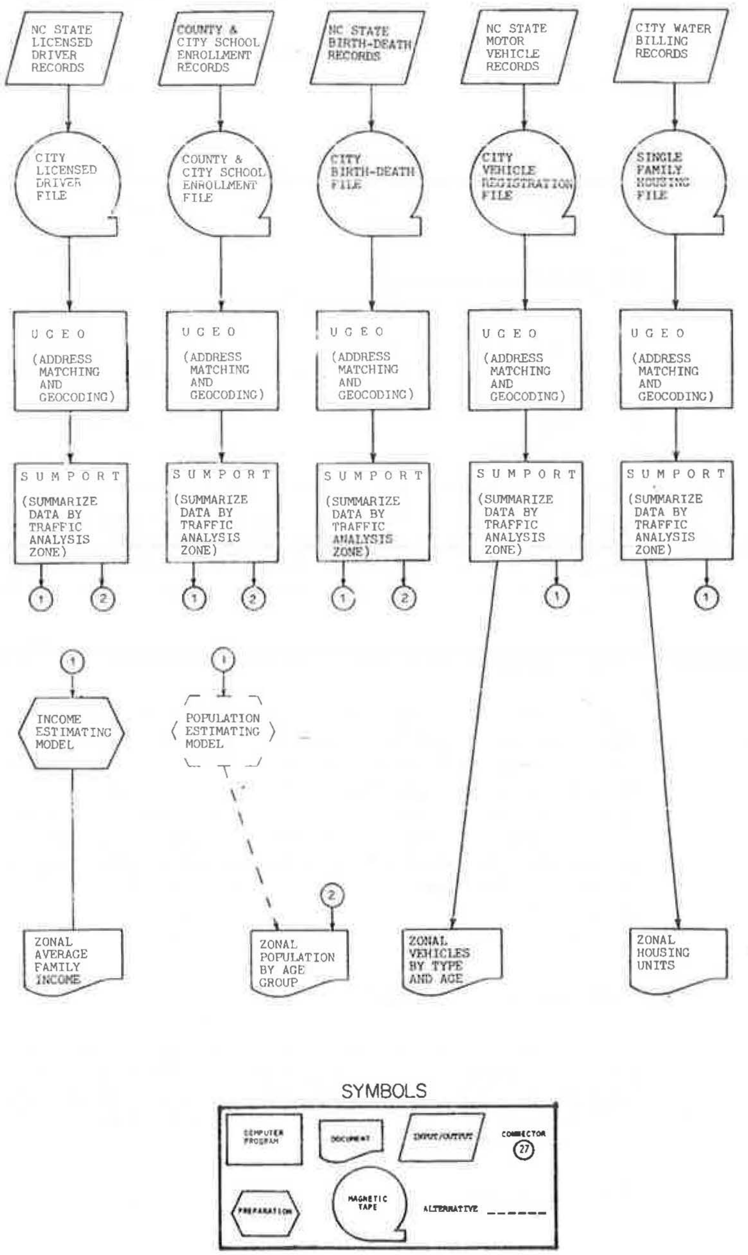
IMPLEMENTING AND TESTING NCUTPIS

The North Carolina urban center chosen for initial implementation of NCUTPIS was Greensboro. For the purpose of analysis the Greensboro transportation planning study area was divided into 302 traffic zones. Testing of NCUTPIS involved nothing more than determining if the system could provide summaries of population, single-family housing units, vehicle ownership, average family income, and employment for each of the 302 traffic zones in the Greensboro transportation planning area.

Demographic Subsystem

The demographic subsystem was initially tested by using the vehicle registration file for Guilford County as source data. The total number of vehicle records in that file was more than 200,000. The program SCREEN was used to reduce this to 122,727 records that contain only Greensboro addresses. This source data file along with the PICADAD and geographic base (GBF) files for Greensboro were

Figure 4. Demographic surveillance files.



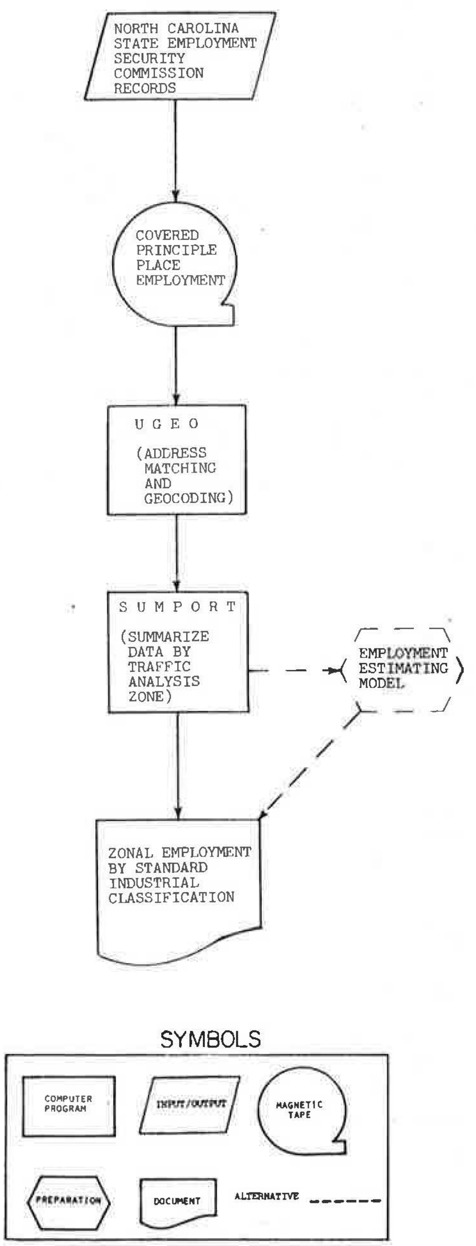
input to the UGEO procedure. The output results from this procedure are shown in Figures 6 and 7.

The entire demographic subsystem of NCUTPIS was tested in a similar manner. Each new source data file was input to the UGEO procedure. The output from this procedure was then summarized over each traffic zone. The first 25 zonal summaries for the demographic subsystem are given in Table 2. Average family income for these traffic zones (as estimated from the Greensboro income model) are also included in this table.

The following findings are suggested relative to implementation of the computerized data retrieval system (NCUTPIS) and are based on the data sets and analytical procedures used during this investigation.

1. The match rate for the vehicle registration file was 80.5 percent. The matched records total included 89,844 matched with certainty and 8,973 matched with uncertainty.
2. The match rate for Greensboro's water billing file for dwelling units was 93.2 percent. The matched records total included 51,421 matched with certainty and 1,821 matched with uncertainty.
3. The match rate for licensed drivers in the Greensboro area was 77.2 percent. The matched records total included 96,916 matched with certainty and 15,950 matched with uncertainty.
4. The match rate for county students living in Greensboro could not be determined because of the many non-Greensboro addresses contained in the file. The file contained a total of 28,031 student records of which 5,090 were matched to Greensboro addresses with certainty and 4,570 were matched with uncertainty.
5. The match rate for Greensboro city students was 93.8 percent. The file contained a total of 22,835 records of which 20,241 were matched with certainty and 1,170 were matched with uncertainty.
6. The match rate for births in the Greensboro area was 65.9 percent. This birth file contained 10,325 records from 1974 through 1979 of which 6,233 were matched with certainty.
7. Harmonic analysis provided a convenient mathematical framework for the construction of the income model.
8. Estimates of zonal income averages are easily obtained from the income model.
9. NCUTPIS is completely user-oriented relative to the UTPS environment.

Figure 5. Employment surveillance files.



Employment Surveillance Subsystem

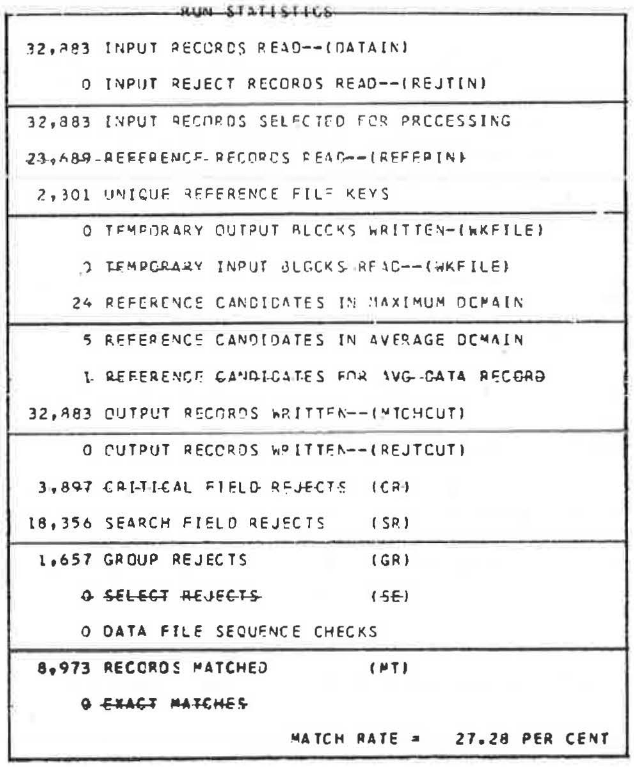
The employment subsystem was tested in the same way as the demographic subsystem. Greensboro employer records as maintained by the North Carolina State Employment Security Commission served as the input data file. A match rate of only 49.2 percent was

Figure 6. Vehicle registration address matching run statistics matched with certainty.

```
                    RUN STATISTICS
122,727 INPUT RECORDS READ--(DATAIN)
      0 INPUT REJECT RECORDS READ--(REJTIN)
122,727 INPUT RECORDS SELECTED FOR PROCESSING
 23,689 REFERENCE RECORDS READ  (REFERIN)
  5,559 UNIQUE REFERENCE FILE KEYS
      0 TEMPORARY OUTPUT BLOCKS WRITTEN-(WKFILE)
      0 TEMPORARY INPUT BLOCKS READ--(WKFILE)
     20 REFERENCF CANOIDATES IN MAXIMUM DOMAIN
      3 REFERENCE CANOIDATES IN AVERAGE DOMAIN
      2 REFERENCE CANDIDATES FOR AVG DATA RECORD
 89,844 OUTPUT RECORDS WRITTEN--(MTCHOUT)
 32,883 OUTPUT RECORDS WRITTEN--(REJTOUT)
  3,897 CRITICAL FIELD REJECTS  (CR)
 28,014 SEARCH FIELD REJECTS    (SR)
    972 GROUP REJECTS           (GR)
      0 SELECT REJECTS          (SE)
      0 DATA FILE SEQUENCE CHECKS
 89,844 RECORDS MATCHED         (MT)
      0 EXACT MATCHES
                   MATCH RATE =   73.20 PER CENT
```

obtained during the geocoding process. The primary reason for this low match rate was the large number of postal addresses found in the record address field. Postal addresses cannot be processed by the present version of NCUTPIS's matching procedure. Until this address-matching problem is solved, NCUTPIS cannot be used in extracting zonal employment data. The only other method that can be used to obtain employment data is the manual inventory or survey.

Figure 7. Vehicle registration address matching run statistics matched with uncertainty.


RUN STATISTICS

32,883 INPUT RECORDS READ--(DATAIN)
0 INPUT REJECT RECORDS READ--(REJTIN)

32,883 INPUT RECORDS SELECTED FOR PROCESSING
23,689 REFERENCE RECORDS READ--(REFERIN)
2,301 UNIQUE REFERENCE FILE KEYS

0 TEMPORARY OUTPUT BLOCKS WRITTEN-(WKFILE)
0 TEMPORARY INPUT BLOCKS READ--(WKFILE)
24 REFERENCE CANDIDATES IN MAXIMUM DOMAIN

5 REFERENCE CANDIDATES IN AVERAGE DOMAIN
1 REFERENCE CANDIDATES FOR AVG-DATA RECORD
32,883 OUTPUT RECORDS WRITTEN--(MTCHOUT)

0 OUTPUT RECORDS WRITTEN--(REJTOUT)
3,897 CRITICAL FIELD REJECTS (CR)
18,356 SEARCH FIELD REJECTS (SR)

1,657 GROUP REJECTS (GR)
0 SELECT REJECTS (SE)
0 DATA FILE SEQUENCE CHECKS

8,973 RECORDS MATCHED (MT)
0 EXACT MATCHES
MATCH RATE = 27.28 PER CENT

SUMMARY

The main function of transportation planning is to predict future traffic flows and patterns so that a system plan can be developed that will provide an acceptable level of service relative to the predictions. Tremendous advances have been made in the art of travel forecasting during the last 30 years. The cost of performing travel forecasting has also been tremendous. A major part of this cost involves the collection of socioeconomic data that are the determinates of existing traffic patterns in the urban area.

The objective of this study was the development of a computerized data retrieval system that would obviate the need for costly and time-consuming surveys. The development of such a data system required the construction and calibration of an income-estimating model.

The analysis of system implementation suggests the following conclusions:

1. The implementation of NCUTPIS shows it to be capable of collecting the demographic and some of the employment data used in the continuing UTPP. NCUTPIS, therefore, accomplishes the objectives that were set forth for this research. No attempt was made to provide a system design that could serve all the data requirements for all types of urban planning. The data system, however, can serve as a basis for any future expansion that may be desirable.

2. Conclusions as to how well NCUTPIS performed its function of data acquisition tool could not be drawn because no equivalent data base exists to serve as a bench mark for measuring its performance. What can be concluded is that NCUTPIS provides an operationally feasible method for maintaining, on a small area basis, an urban transportation planning data base by using data extracted from the routine operating and administrative records of state, county, and local government agencies.

Table 2. Demographic subsystem output.

Greensboro Traffic Zone No.	Total Vehicle Registration	Total Single Family Dwelling Units	Licensed Drivers	County Students	City Students	Births from 1974 to 1979	Avg Family Income ($)
1	340	421	130	63	1	10	4,800
2	201	80	17	4	0	0	22,500
3	35	60	20	1	0	0	3,600
4	80	150	101	83	9	6	3,500
5	188	104	85	39	8	10	16,000
6	173	154	68	24	2	2	6,400
7	107	148	68	5	7	2	4,500
8	204	155	268	2	169	22	7,500
9	502	352	706	62	29	32	8,100
10	62	15	73	4	3	6	22,500
11	661	447	793	171	94	50	8,500
12	231	81	105	9	11	9	22,500
13	494	366	521	2	54	24	7,500
14	502	332	595	11	40	27	8,900
15	898	723	1,050	42	142	103	6,800
16	681	416	482	13	209	32	10,500
17	60	109	94	0	24	10	3,550
18	42	94	31	0	0	0	3,500
19	14	78	16	0	8	0	3,500
20	1,377	998	1,227	63	411	73	7,800
21	277	311	471	4	484	35	6,100
22	330	241	443	13	275	16	8,900
23	1,122	867	1,273	16	317	128	7,400
24	112	77	66	0	18	12	8,300
25	168	227	222	37	14	4	4,600

3. Implementation of NCUTPIS demonstrates the importance of intergovernmental and interagency working agreements. The overall effectiveness of this data system depends to a great extent on these agreements.

This study should provide useful guidelines for implementation of similar information systems in other urban centers. Note that the necessity for developing an income model was revealed in this study. Income data were not previously available on a continuing basis from any other source. If income is to continue to be the primary predictive variable used in the simulation of urban travel demand, then additional research must be performed to substantiate the findings of this investigation. If subsequent substantiation cannot be furnished by research, then a suitable substitute for the income variable should be found. This substitute must be easily monitored and collectible on a small area basis. One possible substitute that has these characteristics is family life-cycle.

REFERENCES

1. Peat, Marwick, Mitchell and Co. Conceptual Design for an Urban Transportation Planning Information System. Office of Highways Planning Division, FHWA, Jan. 1973.
2. GBF/DIME System Description and Uses. Geography Division, Bureau of the Census, U.S. Department of Commerce, Feb. 1978.
3. M.A. Jaro, J.R. Baker, and R.M. Somers. Introduction to Census/UMTA Technology. Statistical Research Division, Bureau of the Census, U.S. Department of Commerce, Nov. 1978.
4. R.B. Dial. The Urban Transportation Planning System (UPTS): Philosophy and Function. Office of Transit Planning, UMTA, June 1975.
5. R.E. Gaskell. Engineering Mathematics. Dryden Press, New York, 1958.
6. L.W. McPherson. Toward the Development of an Urban Transportation Planning Information System. North Carolina State Univ., Raleigh, PhD Dissertation, 1981.
7. Urban Planning Division, Federal Highway Administration. Trip Generation Analysis. U.S. Government Printing Office, Aug. 1975.

Publication of this paper sponsored by Committee on Transportation Information Systems and Data Requirements.