BETTER USE OF EXISTING TRANSPORTATION FACILITIES

REFER TO:	ACT	INF.	SIG.
MATERIALS ENGR.		1	
OFFICE MANAGER			
QUALITY CONTROL		2	
PROJ. RECDS. & SUMM.			
ASSOC. MATLS. ENGR. II		2	
PROJ. DEVELOPMENT			
SOILS & FOUND. ENGR.		3	
SOILS			
SOIL MECHANICS			
CHIEF GEOLOGIST		4	
DRILL CREWS			
ASSOC. MATLS. ENGR.		6	
RESEARCH		8	
TESTING ENGR.		5	
CHEMISTRY			
ASPHALT			
STRUCTURES			
ASPHALT MIX			
AGGREGATES			
RECEIVING			
MAINTENANCE			
E. I. T.			

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
SPECIAL REPORT 153

1975 TRANSPORTATION RESEARCH BOARD

BETTER USE OF EXISTING TRANSPORTATION FACILITIES

Proceedings of the Seventh Summer Meeting of the Transportation Research Board in cooperation with the Florida Department of Transportation
August 5-7, 1974
Jacksonville, Florida

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
SPECIAL REPORT 153
Washington, D.C., 1975

Transportation Research Board Special Report 153
Price $9.80
Edited for TRB by Marjorie Moore

subject areas
14 transportation finance
15 transportation economics
40 general maintenance
51 highway safety
52 road user characteristics
53 traffic control and operations
55 traffic measurements
82 urban community values
83 urban land use
84 urban transportation systems

Transportation Research Board publications are available by ordering directly from the Board. They are also obtainable on a regular basis through organizational or individual supporting membership in the Board; members or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

National Research Council. Transportation Research Board.
Better use of existing transportation facilities.

(Special report—Transportation Research Board, National Research Council; 153)
1. Transportation—Congresses. I. Florida. Dept. of Transportation. II. Title. III. Series: National Research Council. Transportation Research Board. Special report—Transportation Research Board; 153.
HE11.N36 1975 380.5 75-17628
ISBN 0-309-02381-5

Contents

Foreword

This Special Report contains papers presented at the Seventh Summer Meeting of the Transportation Research Board held in Jacksonville, Florida, August 5-7, 1974. The objective was to present timely and useful information on the theme of the meeting, Better Use of Existing Transportation Facilities. This theme was selected because of the need for increased transportation capacity in the face of growing resistance to new construction.

The Task Force responsible for the meeting program decided that, before the meeting turned to solutions applicable to specific areas and elements, those in attendance should be reminded that a metropolitan area is a transportation unit.

Each of the papers in part 1 is concerned with an entire metropolitan area. In the first paper, Elmberg reports on the effects of a compartmentation system in Gothenburg, Sweden, that restricts private automobile movement in the city center. Estep and Moskowitz define the operations, operational analysis, and planning of operational improvements performed by the California Department of Transportation and discuss their application on the Los Angeles freeway network. Hall addresses the multidisciplinary approach to transportation management used in Phoenix.

The papers in part 2 reflect current interest in capacity and service level aspects of urban freeways and signalized intersections. McCasland summarizes Texas experience with improving and controlling the level of service on freeways and describes methods such as geometric roadway improvement, detection and management of incidents, and improvement in passenger-vehicle ratios. DeShazo describes the successful Dallas park-and-ride demonstration program and discusses future plans for expansion of the program to include dial-a-ride passenger collection systems in residential areas.

One of the most expedient means of increasing the people-carrying efficiency of a street is to provide bus-only lanes. In part 3, Muzyka describes a traffic simulation tool that has been developed to predict bus service and general traffic performance if special bus priorities are implemented. Various performance measures can be computed for each priority strategy to form a basis for the selection of the most promising system.

In part 4, May discusses the current status of approaches to determining intersection capacity and emphasizes research needs. Lee, Crowley, and Pignataro present an advanced signal control strategy, queue-actuated control, as a solution to problems of frequent oversaturation of signalized intersections. Lieberman and Goldblatt present a new approach to signal warrants based on delay and use UTCS-1 simulation to evaluate signal and stop sign control under differing volume conditions.

In part 5, Butler discusses the use of advance planning of traffic lane closures for maintenance so as to achieve minimum disruption to motorists. Rohde explores the idea of paying a premium to accomplish work at an accelerated rate, McLean describes traffic handling procedures to minimize confusion during the construction process, and Grimsley and Morris describe an increasingly widespread procedure for replacing broken areas of portland cement concrete pavement with precast slabs. Procedures for handling emergencies cannot be completely preplanned, but general guidelines can be prescribed and crews trained to handle emergencies before they occur. Meister describes the California process and reports on actual case studies where the plans were tested.

The need for better methods for detection and management of freeway traffic incidents led to the research reported in part 6. Dudek delineates the nature and scope of the problem and gives several challenges still to be met by research and implementation in arriving at effective incident management. Extensive operating experience is the basis for McDermott's report of surveillance and control for incident detection and management on Chicago's high-volume expressway system. California experience with development and implementation of alternate route schemes has led to a procedure of preplanning for required capability when incidents cause closure of freeways. Roper, Murphy, and Zimowsky make a strong case to support this systems approach. McCasland tells of Houston's incident detection and response capability and points out that, although the program is effective, it could be improved by better coordination and further commitments of resources by the agencies involved. Yagoda and Adler concentrate on the system challenges of incident management programs and point out that such programs can no longer be left to chance in the face of today's urgent need to derive maximum benefit from existing facilities.

The papers in part 7 suggest alternatives for better use of present transportation resources by reducing the demand for trips, particularly vehicle trips, during weekday peak hours. O'Malley examines the effect that staggering work hours has had on reducing transportation congestion in the Manhattan and New York-New Jersey central business districts. Mohring investigates the role of price in the use and development of transportation facilities and discusses the nature of the price that governs consumer decisions and what costs society incurs. Hoch is concerned with patterns of change in population and automobile ownership, trends in total population growth, and trends in the distribution of population by locale. Lessieu and Karvasarsky discuss the extent to which driving has been reduced and how much of the reduction might be attributed to the increase in gasoline prices. Special attention is given to the observations of differential change in traffic volume in peak periods versus off-peak periods.

The papers in part 8 are all concerned with highway safety improvements. Goodwin and Phillips trace the history of legislation that has influenced programs at the federal, state, and local levels of government. The impact of the Highway Safety and National Traffic and Motor Vehicle Safety Acts, both passed in 1966, is discussed, and particular attention is given to safety standards promulgated under the acts. Jorgensen describes the evaluation of highway safety improvement programs that was developed under NCHRP Project 17-2A. The paper stresses the need for more accurate traffic accident data and discusses a system to evaluate improvement projects. Bennett discusses four categories of highway-related safety programs: identification and surveillance of accident locations; highway design, construction, and maintenance; traffic engineering services; and accommodation of pedestrians. Goodson discusses how traffic law enforcement agencies today are confronted with increasing demands for service without comparably related increases in manpower and budget allocations. The use of special emphasis patrols by the Arizona Department of Public Safety is described along with other innovative programs. The paper by Waters describes the Jacksonville, Florida, emergency medical system. The success of the coordinated emergency ambulance service and the work of emergency medical technicians are detailed.

Compartmentation as a Tool to Reduce Traffic Congestion and Improve the Environment

Curt M. Elmberg
Gothenburg Transit Authority, Sweden

A system of compartmentation or zoning was used in Gothenburg, Sweden, to reduce traffic congestion, to improve public transit operation, and to improve environmental quality in the central core of the city. This system is described, and the effects it has had on vehicle flow, parking, public transit operation, accidents, and noise and air quality are discussed. Polls of citizen reaction to the compartmentation revealed that two out of three citizens had favorable reactions to the system. Although Gothenburg has had difficulty securing funds for follow-up studies, several other Swedish cities have been impressed with the Gothenburg system and have adopted a similar compartmentation system.

The word compartmentation is far from easy to understand, for it is not found in most dictionaries. Literally, it may be defined as the act of making compartments or the act of dividing something. In medieval defense strategy, compartmentation was a very useful tool for keeping the enemy at a suitable distance. In modern times the same tool can be used as a strategic weapon to keep another enemy at a suitable distance: traffic congestion in the core of the urban area.

Practical experiences in using this strategy over a lengthy period of time are rare, but fortunately some large-scale experiments have been conducted. The two most known are probably the traffic restraint schemes in Bremen, West Germany, and Gothenburg, Sweden. As a basic principle both schemes have adopted the technique of compartmentation. The Gothenburg scheme is described to demonstrate what can be accomplished with this tool.

On the west coast, Gothenburg is the second largest city in Sweden and is more than 350 years old. Statistical data related to population and area are given in Table 1, and data on automobiles and public transportation are given in Table 2.

As with many other cities built centuries ago, the central area was not designed to handle the increasing number of motor vehicles. Different from many other old cities, Gothenburg has a few very wide streets in the core because they were once part of a canal system that was later filled.

In the early 1960s public transportation officials who had studied the successful implementation of the traffic restraint scheme in Bremen, West Germany, approached city and traffic planners with measures for rearranging vehicle traffic in the central area, primarily with the objective of improving the regularity and speed of public transit. Although interest was shown, the task of planning for the change from left-hand traffic to right-hand traffic had to be given priority up to 1967.

Table 1. Population and area of Gothenburg.

Year	Population	Area (km²)
1650	4,000	—[a]
1700	6,000	—[a]
1750	8,500	—[a]
1800	12,300	—[a]
1900	130,600	25
1950	353,700	179
1960	412,100	183
1970	451,800	451
1971	454,400	451
1972	441,500	451
1973		
Within city limits	431,500	451
Central business area	4,000	1.2
Central area	35,000	7.0
Region of Gothenburg	680,000	2,640

[a]No records available.

Table 2. Automobile ownership and public transit use.

Year	Population per Automobile	Transit Trips per Population	Year	Population per Automobile	Transit Trips per Population
1935	59	252	1960	7.0	239
1940	43	291	1965	4.6	206
1945	55	304	1970	3.6	175
1950	48	334	1971	3.5	175
1955	18	296	1972	3.4	170
			1973	3.3	183

Figure 1. Central business area with compartmentation.

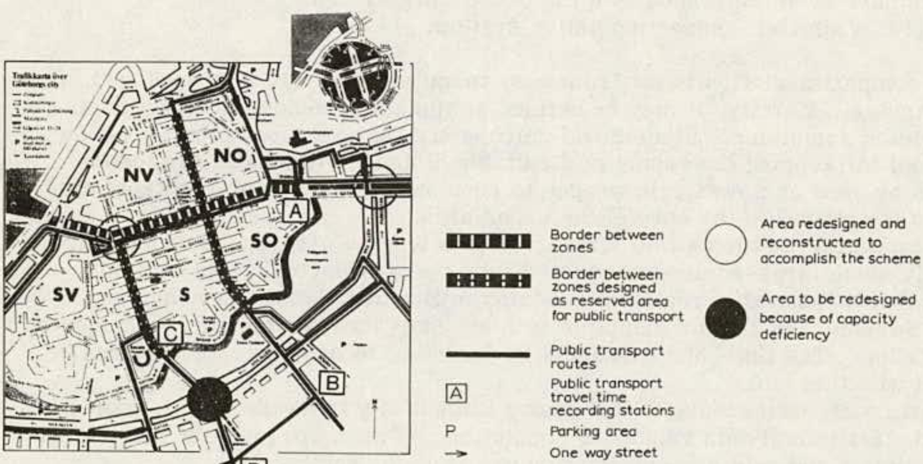


Figure 2. Physical barrier used to direct traffic and outline zone border.

City officials and politicians, however, became more and more alarmed by the ineffectiveness of the traffic system, especially in the central area, although large sums were invested in private and public transport. The environmental impact of the situation also became an issue.

In late 1969 the City Council decided that a small commission of city officials representing various authorities should seriously look into the problem and propose measures to relieve it. The commission was given extraordinary power to use the offices of the city administration for detailed preparation of necessary plans and surveys and to implement any measure.

The aims of the work carried out by the commission were to

1. Improve the environment for people living, working, and walking in the central business area,
2. Increase the practicability, regularity, and reliability of the public transport system, and
3. Propose measures possible to accomplish rapidly.

PROPOSAL

To reduce the volume of traffic in the central area, two schemes were considered: permanent abolishment of private cars and permanent restriction of route selection to the area. The commission rejected permanent abolishment of automobiles because of the great legal aspects.

Interest was turned to the scheme for a permanent restriction of movement, and the principal features of such a scheme were outlined as follows:

1. The CBD should be divided into compartments or zones.
2. Direct contact between zones by car should be eliminated.
3. Contacts between zones should only be permitted along ring routes outside the CBD.
4. The number of entrances and exits for each zone should be reduced to a minimum.
5. Within each zone, vehicular movement should not be restricted.
6. The current parking policy should prevail.

Consequently, the central business area was divided into five zones, named after the cardinal points (Fig. 1). Only public transit vehicles but not taxis could cross the borderline between zones. Wherever a borderline between zones divided the direction of a transit route, reserved lanes for transit were introduced.

The proposal was submitted to various authorities and corporations for their consideration, and in spring of 1970 final approval was given to implement the scheme and necessary funds were made available. The total cost of implementation of the scheme was Kr 1,130,000 or $220,000 at the 1970 exchange level. An additional sum of $40,000 was later added for impact studies.

Before the scheme was implemented, all households in the Gothenburg region, all offices, department stores, shops, movie theaters, filling stations, and the like received multicolor pamphlets explaining the reasons for and the outline of the scheme. The scheme was sold to the public not as a restraint but as an environmental improvement.

IMPLEMENTATION

The scheme was implemented on Tuesday, August 18, 1970. On the night of August 17, the central business area was completely reshaped. Borderlines between zones were painted with solid white lines or were reserved for transit. At certain locations the painted lines were supplemented by physical barriers (Fig. 2). The painted lines were

Figure 3. Methods of outlining borders between zones: (a) painted lines, (b) concrete string in middle of border area, and (c) concrete barrier glued to pavement.

Figure 4. Larger area with no traffic used for transit stops.

Table 3. Relative change in vehicle volume on an average weekday.

Recording Station	After 2 Weeks	After 8 Weeks	Recording Station	After 2 Weeks	After 8 Weeks
Scheme			Ring		
11	-35	-35	21	+20	+20
12	-50	-70	22	+24	+25
13	-34	-35	23	+23	+20
14	-1	-1	24	+9	+10
15	-7	-5	25	+1	+1
16	-17	-15	26	+6	+10
17	-48	-45			
18	+45	+45			

Table 4. Change in vehicle volume after compartmentation scheme.

Recording Station	Vehicles per Hour		Percentage of Change	Recording Station	Vehicles per Hour		Percentage of Change
	1970	1973			1970	1973	
Scheme				Ring			
11	1,200	1,230	+3	21	1,625	1,745	+7
13	1,360	1,240	-9	22	1,230	1,075	-13
14	1,975	1,760	-11	23	1,935	2,190	+13
15	705	535	-25	24	1,895	2,010	+6
17	300	245	-18	25	1,285	1,215	-5
18	790	725	-8	26	995	990	0

not sufficient to prevent motorists from illicit crossing. After some experiments the solid lines outlining the edge of a reserved area for transit were transformed into neat concrete strings (Fig. 3).

The rather bulky concrete barriers directing traffic were efficient though not aesthetic. Therefore, later those spots were redesigned by reshaping sidewalks and curbs to give an appearance of a permanent structure. Areas closed to motor vehicles except transit were redesigned as convenient and spacious stops for transit (Fig. 4).

The commission responsible for the implementation of the scheme asked the various authorities involved in the project to conduct surveys to determine whether the scheme had resulted in a positive or a negative impact on the environment.

Impact on Vehicle Flow

About 2 months after the introduction of the scheme the redistribution of vehicle flow stabilized, and no significant changes have taken place since (Table 3). The change in vehicle volume between 1970 and 1973 is given in Table 4. (Figure 5 shows the recording stations referred to in Tables 3 and 4.)

Along the main shopping street vehicle volume was reduced by 70 percent and along other important shopping streets by 50 percent. The ring routes surrounding the same area registered an increase of as much as 25 percent at certain spots.

Impact on Parking

No changes in the parking policy were introduced when the scheme was inaugurated, nor were the number and availability of parking places in the central business area altered. There were 4,900 places, 600 of which were in a multistory garage and the remaining were metered street parking.

Three months after the introduction of the scheme, the City Council adopted a new pricing policy on parking in the central business area. The hourly parking fee was raised 100 percent, which had an immediate effect on the parking turnover (Table 5). The new parking fee was Kr 2.00 ($0.40) per hour, and the maximum parking time was 30, 60, or 120 min depending on the location of the parking space. In 1973 this area of the new pricing policy was extended beyond the central business district.

Impact on Public Transport

One of the objectives of the scheme introduced was to improve the reliability of public transport with the possible anticipation of a change of the modal split. Track-bound public transport dominates in the core area. Of the nine tram routes, six pass through the core and two touch the area. Two city bus routes pass through and a few regional bus routes, too.

Surveys revealed that a noticeable reduction in travel time was not forthcoming, which, in fact, was not too surprising because only 5 percent of the entire route length was affected by the scheme. On the other hand, the alterations of the standard deviation and the range (Table 6), which might be regarded as measures of the state of regularity, have been favorably noticed by the customers and the transit authority.

Public polls among transit users revealed that because of the scheme 6 percent more weekday trips to the central business district were made by public transport and 2 percent more trips on Saturdays. Most of the respondents gave the difficulty in finding a parking place as the reason for this change; if parking is filled in one zone, considerable driving is required to reach the next one.

Rationing of gasoline in the beginning of 1974 and a drastic price raise on gasoline caused a shift in travel habits: There are now almost 10 percent more riders on public transit than before although there is now no shortage of fuel.

Figure 5. Layout of central area with compartmentation.

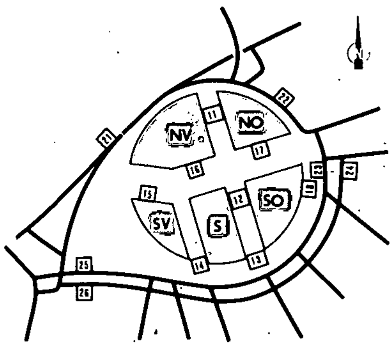


Table 5. Percentage of duration of occupancy of parking places.

	30-Min		60-Min		120-Min	
Zone	1969	1970	1969	1970[a]	1969	1970[b]
NO	49	47	51	51	—	—
NV	62	63	65	59	65	57
S	65	65	70	60	66	56
SO	61	63	68	65	71	65
SV	57	59	66	62	—	—

[a]Compartmentation. [b]Compartmentation and parking fee increase.

Table 6. Public transport travel times (in minutes).

		Morning Peak			Afternoon Peak		
Route[a]	October	Mean	SD	Range	Mean	SD	Range
A-B	1967	4.6	0.70	2.7	6.5	0.86	3.9
	1970	4.7	0.71	2.9	5.4	0.78	3.6
	1971	3.9	0.63	2.5	5.7	0.62	3.2
B-A	1967	5.2	0.98	4.5	6.7	1.03	5.0
	1970	4.8	0.84	3.0	5.6	0.71	2.8
	1971	4.6	0.55	2.4	5.4	0.82	2.6
A-C	1967	5.1	0.72	3.2	5.6	0.94	5.5
	1970	5.0	0.68	2.9	5.2	0.86	3.0
	1971	5.4	0.66	2.7	5.8	0.98	3.8
C-A	1967	5.3	0.85	4.0	6.7	0.88	5.0
	1970	4.7	0.71	3.3	5.6	0.69	3.6
	1971	4.4	0.72	2.6	4.7	0.66	3.2

[a]Routes between recording stations shown in Figure 1.

Table 7. Number of accidents after initiation of compartmentation.

	Ring Route Area			Scheme Area			Entire City		
Year	Injury	Noninjury	Total	Injury	Noninjury	Total	Injury	Noninjury	Total
First year of operation									
First 3 months	3	19	22	-3	-4	-7			
First 6 months	8	19	27	-6	-43	-49			
First 9 months	11	-14	-3	1	-65	-64			
First 12 months	16	-19	-3	6	-91	-85			
1969	84	485	569	70	267	337	1,201	4,350	5,551
1970							1,229	4,964	6,193
1971	110	558	668	69	205	274	1,325	4,679	6,004
1972	89	492	581	37	189	226	1,130	4,299	5,429
1973	75	351	426	33	146	179	1,131	3,981	5,112

Impact on Accidents

During the first year of operation there was a significant reduction in the number of accidents within the scheme area (Table 7). After that, however, accidents of all types increased, probably because improved practicability resulted in higher speeds. This very unfortunate trend seems to have changed during 1973.

Impact on Air Pollution

Along the main shopping street, the mean amount of carbon monoxide for a period of 30 min has changed from 60 to 5 ppm after the scheme was introduced. Maximum values recorded during 30 sec have decreased from 200 to 10 ppm.

Shopkeepers report that they need to clean display windows only once a week now instead of once a day as before the scheme was introduced.

Impact on Noise

The noise level at the main shopping street, which previously served as a thoroughfare through the central area, was 75 dBA according to surveys taken by the Public Health Authority. After the introduction of the scheme, the noise level stabilized at 71 dBA.

Impact on Retail Trade

A survey was taken of 33 shopkeepers, and the results are given in Table 8. Public polls conducted by the Gothenburg Retail Association established that 93 percent of the customers living within the city limits purchased the same amount or more after the introduction of the scheme. The frequency of visits to the central business area increased for 9 percent and decreased for 23 percent. Of the 23 percent, only 5 percent mentioned the compartmentation as the main reason. It should be noted that a huge shopping center with a discount store and several thousand free parking places opened a few months before the scheme was introduced, followed by another still larger shopping center in early 1972.

The complete change of the retail structure during a very short period of time, with the market almost unaltered, has created problems within the CBD. In addition, new department stores have grown in zone NO as a consequence of a comprehensive redevelopment scheme. The structural change, in many cases resulting in reduced turnover for smaller shopkeepers, has created a feeling among them that free-flowing private vehicles and more parking facilities would change the situation.

For years an underground parking garage in zone S has been debated, and the project has been accepted by the City Council by a very small majority.

Public Reaction

Compartmentation has been received favorably, even by the motorists. Polls among citizens revealed that two out of three were definitely in favor of the scheme, whereas one out of 10 was definitely negative. About 85 percent of the respondents found no difficulties in finding their way to the central business area; 15 percent were of the opposite opinion. Table 9 gives the effects of the scheme that respondents ranked most favorably.

Initially, taxi drivers complained bitterly about the scheme. A journey from a taxi stand to an address within the area could not be made by the shortest route and likewise for trips between different zones. Privilege for taxis to use lanes exclusively reserved for public transport was rejected on the grounds of capacity and safety. Relocation of taxi stands and permission to cross the borderline between zones NO and S through an opening for buses have reduced the complaints.

Table 8. Change in retail sales after compartmentation.

Trade	Positive	Negative	No Change
Books	0	1	0
Clothing	4	2	6
Food	3	2	0
Footwear	2	0	1
Hardware	1	0	1
Jewelry	0	2	1
Photo	0	1	1
Radio	0	2	0
Sport	0	0	1
Newspapers	1	1	0
Total	11	11	11

Table 9. Percentage of respondents giving favorable response.

	Respondent's Residence	
Impact	Within City	Outside City
Easier for pedestrians	42.3	32.2
Reduced vehicle traffic	23.0	21.7
Improved environment	17.4	18.6
Improved practicability	5.0	5.8
Miscellaneous	12.3	21.7

Table 10. Swedish cities having compartmentation.

		Scheme Area		
City	1973 Population	Type	Population	Employment
Göteborg	431,500	Central	4,000	35,000
Stockholm	681,300	Central/residential	19,300	23,000
Lund	74,200	Central	8,300	7,000
Halmstad	75,600	Central	1,100	5,500
Umeå	73,000	Central	5,000	9,000
Uppsala	136,000	Central	17,700	14,600
Västerås	117,900	Residential	8,000	700

FUTURE EXPANSION

The scheme introduced in August 1970 has been regarded as the first stage in a comprehensive plan to encompass an area about six times as large as the one in operation.

According to the original plans, the provisional opening between zones S/SV and NO/NV due to capacity deficiency on the northern ring route should be closed when the reconstruction of this route is completed. Because of the structural change of the CBD, decision makers are now hesitant to approve such an action until the entire plan for the central area is proposed.

A four-lane, partly six-lane, city expressway will encompass the central area. Part of this system is already completed, but the circumferential route system will not be completed during this decade. Existing streets will be used to divide the entire central area into zones—the technique of compartmentation—and the streets encompassing such a zone will consequently serve as ring routes and distributors. The streets within each zone will be closed to through traffic.

Different from the present scheme, areas with residential development might encounter a less favorable environmental impact along the streets designated as ring routes. Therefore, in the study of the ultimate scheme serious consideration is given to determine how ring route alternatives would affect environment and residents.

AFTEREFFECTS

The success of the Gothenburg compartmentation scheme has had an unfortunate negative aftereffect. Because of the rapid stabilization of impact patterns, very limited interest has been shown to appropriate funds for more comprehensive follow-up studies.

On the other hand, the success of the scheme has also had a positive aftereffect. Several other Swedish cities have adopted the technique of compartmentation (Table 10).

CONCLUSIONS

Compartmentation is not the proper remedy for every urban area and its CBD. Experience gained in Germany and Sweden has, however, proved that compartmentation can be a useful tool, among many others, if proper conditions exist.

Aided by a comprehensive origin-destination survey and a computer a skilled traffic engineer should be able to play the game of compartmentation at his desk by testing whether his particular CBD or subdivision hits a jackpot. If so, he must then have the courage to pursue the task, a challenging but far from easy job.

Getting the Most Out of a Freeway System

A. C. Estep and Karl Moskowitz*
California Department of Transportation

Responsibility for the operations, operational analysis, and planning of operational improvements of the California state highway system lies with the Department of Transportation. These three activities are defined, and their application in a program to upgrade and control the Los Angeles freeway network is discussed. Methods of controlling freeway traffic (ramp metering, freeway widening, new construction) are described.

California is fortunate enough to have constructed a major portion of a freeway system that provides a higher level of transportation service for more people than has ever been provided anywhere. Nearly 40 percent of all ground transportation in California now occurs on 3,800 miles of completed freeways. We owe this to an exceptionally farseeing legislature that not only established the freeway and expressway system in 1959 but followed through with a resolution calling for a study of how this system would be operated. This study (1) was based on the premise that "the state has the continuing responsibility to assure that the use of highway plant shall be efficient, safe, convenient."

Among the conclusions reached in this study were the following:

1. Traffic operations covers those things that are done after a highway is built that enable motorists and pedestrians to move safely and with a minimum of delay. All departments of the transportation agency contribute to the fulfillment of this function.
2. In the years ahead, the highway operations function will become increasingly important, and it will consume a larger share of the total funds available for highway purposes.
3. The so-called freeway problem is actually an aggregate of several problems. Some occur only on certain sections of the system; others are more or less general. Among them are recurrent congestion, congestion due to special events or accidents, accident prevention, disabled vehicles, directional signing, and law enforcement.
4. The transportation agency should consider the use of helicopter patrol of urban freeways, particularly during peak periods of travel.
5. The Division of Highways should consider the installation of a communications cable as a part of the freeway system.
6. Plans should be developed for emergency closing of freeway entrances and exits by remote-controlled electronic devices.

*Deceased.

7. The California Highway Patrol and the Division of Highways should undertake studies and experimentation to determine methods for faster clear-up of congestion on the freeways.

8. A freeway operations task force should be established under the general supervision of the highway transportation agency with representation from the Division of Highways, Highway Patrol, and Department of Motor Vehicles.

9. Freeway traffic control is concerned with traffic control decision making. It covers the entire traffic control process from data collection to communication of control decisions to motorists.

From that legislative mandate, the present policy of the California Department of Transportation has evolved.

OPERATIONAL RESPONSIBILITY

Providing safer, more expeditious travel is the ultimate goal of both highway construction and operation. Although new highway construction (including widening projects) is the major means available to alleviate traffic congestion and accidents, operations and operational improvements can, where appropriate, accomplish the same goal at relatively low cost.

It is the policy of the California Department of Transportation to make maximum effective use of the state highway system through a program of traffic operations and operational improvements. The California Highway Patrol has responsibility for controlling traffic, communicating with motorists, and coping with emergencies; generally these activities are handled by mobile forces. To the California DOT, which is fundamentally responsible for the physical aspects of the highway, operations, operational analysis, and planning operational improvements are separate but interrelated activities. All are necessary to operate the highway system effectively.

For the purposes of this paper, which is concerned with facilities as distinguished from enforcement, these three activities are defined below.

Operations

Operations does not mean the construction of operational devices nor modification of the geometry of a highway. Rather, operations includes surveillance and control of traffic and response to changing traffic conditions in real time.

Instructions to drivers (such as the color of a traffic signal) and information for drivers (such as a radio message) are changed by the moment, based on what is happening at a particular time. (Conventional highway signs also instruct or communicate with drivers, but the real-time factor is absent.)

The decision to install a traffic signal and the installation itself are planning, design, and construction in the traditional sense. The timing of the traffic signal, i.e., adjusting the dials, is operations. For another example, preliminary engineering and constructing a freeway ramp control project are not operations, but maintaining continuous surveillance of traffic on the affected freeway and adjusting ramp-metering rates are.

Operational Analysis

Operational analysis is the study of highway traffic. It can include things such as an inventory of freeway congestion, analysis of what happens to one bottleneck when another one is unplugged, and evaluation (ahead of time) of alternative methods of alleviating bottlenecks. Mathematical relationships that take into consideration items such as rate of flow, cumulative storage of vehicles by location in the system, highway capacity, density, and travel time come under this heading. This kind of analysis often becomes preliminary engineering for an operational improvement. It is also necessary input for real-time traffic control such as traffic signal settings. It is a kind of engineering somewhat different from conventional civil engineering, which has comprised the bulk of our work in the past.

Operational Improvements

The program for operational improvements involves specific projects designed to add to or better the existing system. It includes

1. Installing ramp-metering devices,
2. Placing changeable message signs,
3. Adding auxiliary lanes,
4. Providing left-turn storage lanes,
5. Restriping highways to provide added lanes,
6. Placing channelization devices,
7. Widening bottleneck locations, and
8. Installing traffic signals.

Although operational improvements were routinely accomplished when deficiencies became obvious, the process of identifying the possibilities and proposing improvements can be more organized and systematic. Responsibility for conducting operational activities and developing operational improvements must be clearly defined. Although the specific organizational structure for accomplishing this work may vary among transportation departments, or even districts, it is essential that the function be identified and adequately staffed. Only in this way will worthwhile improvements not be overlooked and will proper priority be established for each project in order to maximize the payoff in this area.

PROGRAM TO UPGRADE AND CONTROL THE LOS ANGELES FREEWAY NETWORK

In accordance with the concepts defined, a freeway operation unit was organized in Los Angeles in 1965. During the first year, the two-man staff mainly defined the scope of work and made recommendations for an action plan and proposed staffing, scheduling, and costs. The action plan called for operational projects that included a freeway congestion inventory to determine

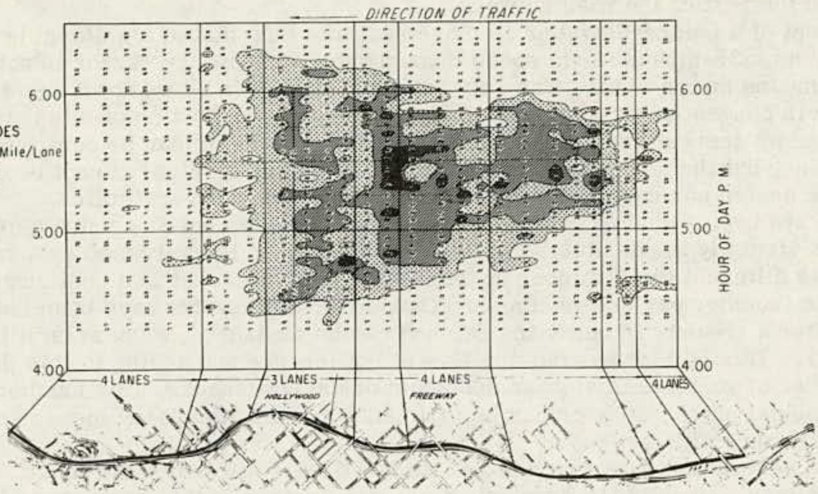
1. Location of problem spots,
2. Duration and extent of problems,
3. Quantitative estimates of travel time and speeds on various sections of the system,
4. Estimates of traffic demand at bottlenecks (where existing counts do not represent demand), and
5. The relationship of one bottleneck to another;

an analysis of specific bottlenecks (using inventory as base data); and a correlation of design and planning functions. The action plan also called for the following planning and research projects:

1. Evaluation of methods of ramp control,
2. Survey of nonrecurrent congestion, and
3. Communication with the motorist (signs, radio, etc.).

The first analysis of specific bottlenecks was done in 1966, concurrently with the development of scope of the study, and resulted in a manually operated metering signal at one ramp and peak-period closure (also manned) of another ramp on the Hollywood Freeway. This project was followed, in 1968, by a system of five metered ramps and two ramp closures on the Harbor Freeway. These initial projects were very successful

Figure 1. Typical density chart.



in reducing overall delay on the freeway and in the involved corridors as a whole.

In 1969 the weekday congestion inventory was completed, and it was determined that 170 directional miles of freeway (out of approximately 700 directional miles at that stage of development) were subject to recurrent congestion during peak periods. This inventory was primarily based on peak-period aerial photography, using 35-mm hand-held cameras. The results are summarized on density contour maps showing time of day on the Y-axis, location along the freeway on the X-axis, and isodensity contours on the Z-axis. Forty-seven separate areas of congestion containing 106 individual bottlenecks and resulting in 1,000,000 vehicle-minutes of daily delay were identified and quantified in this initial study (1967-69 data). (A sample density chart is shown in Figure 1.)

The individual inventory reports on these bottlenecks were the basic data for a report issued in July 1970 (2), which concluded that "current manifest demand" could be accommodated by eliminating those bottlenecks at a cost of \$115 million. The report was updated in 1972, and, like everything else, the cost went up. In 1972 the estimated cost was \$160 million. We are now estimating \$172 million, but it will probably exceed \$200 million by the time we are through. This must be viewed in comparison with the \$2.5 billion that was originally invested in the freeway system. In other words, we found that we could make the \$2.5 billion system work for an additional investment of about 8 percent. In this context, work means essentially free-flowing traffic for 24 hours a day instead of 21 or 22 hours a day.

The ramp control and interim widening stages of this program are scheduled to be complete in 1977 or 1978. Preferential treatment for buses is an integral part of the program. In cooperation with the Regional Transit District, bypass lanes are provided at appropriate controlled entrance ramps, so that the advantages of exclusive bus lanes are obtained without reducing the capacity of the freeway. In fact, bus travel on the controlled sections is actually faster than it would be in a designated reserved lane on an uncontrolled freeway, and the thorny problem of access to the reserved lane is eliminated.

BALANCED FREEWAY SYSTEMS

It is promising that we may be able to construct and operate a truly balanced system of freeways. In this sense, balance means the relation between demand and capacity so that operating conditions, or quality of flow, on the freeways will be more uniform geographically and by time of day. Balance also includes balancing the load between freeways and the rest of the road network.

The concept of a balanced system is more realistic than that of a uniform level of service such as a 35-mph uniform speed throughout a system. The latter is not achieveable by any means known to us. The only reason traffic ever slows down below 45 mph is that there is congestion or stop-and-go traffic ahead. Congestion of this nature is always caused by demand exceeding capacity. The arrival rate can be controlled by ramp metering, but the demand-capacity ratio, or level of service, cannot be controlled by geometric design nor can it be uniform along a stretch of several miles.

It has always been a goal to balance demand and capacity, but an accurate forecast of demand is virtually impossible. [In this paper, demand is expressed as a rate of flow. This is different from number of trips per day or hour. It is a continuously changing rate (number per unit of time).] Therefore, bottlenecks have been inevitable. However, after a freeway is built, the demand can be measured, even when it is greater than capacity. This is done by counting flow at bottlenecks and adding to this the time-related number of vehicles in storage upstream of the bottleneck. This number can be obtained by aerial photography or can be estimated mathematically by measuring speeds and/or density along the freeway.

There are several reasons why demand rate of flow is difficult to forecast. First, of course, is man's inability to forecast anything. Another is that land use and travel patterns almost always are radically altered after the forecast is made by transportation planners. Furthermore, even if perfect forecasts could be made, rate of flow is by definition a quotient of number (of vehicles) divided by the time used to service this number. It is this latter concept that this discussion addresses.

At this point it may be desirable to define some terms. Manifest is a term that can be applied to demand, congestion, and delay. Counting traffic on the road or in the network, photographing congestion and counting stored cars, and measuring speeds for the purpose of describing quality of flow are all measurements of manifest phenomena—things we can see or get our hands on.

Latent can be defined most simply by example. If the service rate in the morning peak period is improved, people simply leave home later. The demand was always there, but it was invisible. There is a latent demand for trips on many urban freeways to be compressed within a shorter time period although the number of trips remains the same. This increases the required rate of flow and creates manifest congestion. If the latent demand is dampened or held constant, the congestion will not be manifest or will be manifest at a different location. When the time spread of trip demand approaches zero, the capacity necessary to absorb the flow rate will approach infinity.

Similarly, if one route becomes more attractive (has less delay) than another, the demand for the attractive route will rise until the capacity of the less attractive route is reached and delay occurs on the attractive route. Conceptually it is possible to keep adding capacity to one route or system of routes, thus making it more and more attractive, until there is no traffic (or almost no traffic) on the less attractive route or system. It is not economically possible to provide enough capacity on one route to dry up all competitive routes. Freeway entrance ramp control offers a solution to this dilemma.

FREEWAY CONTROL

Ramp control or metering consists of traffic signals at entrance ramps that control freeway input at a rate the freeway is capable of handling. It does not increase the capacity of the freeway, but, by diverting or storing traffic, it allows more traffic to use the freeway upstream of that point. Thus, the throughput of the freeway or the rate of accommodating vehicle-miles of travel is increased. Use of electronic sensors and computers makes the signals more responsive to fluctuations in traffic flow. Because the rate of flow on the freeway is less than capacity at all points along the controlled section, the level of service or travel speed is relatively high. Speed seldom drops below 40 mph and, if it does, only for short stretches.

Ramp Control on Existing Freeways

A congestion inventory of existing freeways should be made and kept current. From this inventory, an estimate can be made of the potential increase in throughput (vehicle-miles per unit of time) that can be achieved by controlling the flow rates on entrance ramps. An estimate is also made of the effect on surface streets and other routes that may be used as alternates by traffic diverted at the ramp entrances. In many instances, the majority of traffic diverted at controlled ramps has actually been accommodated on the main line of the freeway being controlled. In other words, many of the same trips use the freeway, but enter it at some upstream point, thus getting a longer and much faster freeway ride. This is accomplished by increased throughput. In all cases, the actual number of vehicles using a controlled ramp during a peak period decreased. (We say this to dispel the fear that ramp queues will be intolerably long.)

Estimates are also made of the savings in travel time that will be experienced by present users of the freeway and new users, and these are compared with the detriment in travel time suffered by the traffic stored at the ramp signals or diverted to other routes. If the net savings is large enough, the metering system should be installed (3).

The congestion inventory will also reveal certain imbalances in demand-capacity ratios that can be corrected by geometric improvements (generally by widening or adding auxiliary lanes in bottlenecks).

Metering systems should be installed on some freeway sections that are not currently suffering peak-period congestion but that prove through continuous surveillance to be approaching this state. The reason for this is to obviate the shock in travel patterns that occurs when ramp metering is installed on a congested freeway.

The magnitude and suddenness of changes in trip patterns that can happen were illustrated when on-ramp volumes on 5 miles of the Harbor Freeway were reduced by 1,400 vph the day the ramp signals were turned on. Although the travelers readily adapted to the changes and were pleased with the results, a gradual imposition of control would have caused less disruption.

New freeways should be planned so that surveillance and ramp control can be implemented with minimal revision. In certain cases, the hardware can be installed as part of the major construction contract. If metering is inaugurated before congestion sets in, the advantages that are always observed when a freeway is first opened to traffic will be preserved for an indefinitely long period, with little noticeable deterioration to the metered traffic.

As noted earlier, geometric bottlenecks are inevitable because of unforeseen changes in demand, among other reasons. Input can be controlled so that congestion does not occur even when geometric imbalance exists. However, this is undesirable because it results in underutilization of a large portion of the freeway. The desirable thing to do is to measure the actual demand at all locations along the freeway and to keep this measurement up to date. When geometric bottlenecks are identified they should be corrected.

Freeway Widening

When a freeway is widened to accommodate existing congestion, ramp control should be installed. This is the only way to ensure that latent demand will not congest the newly widened freeway. The ramp-metering plan can be very liberal to begin with, resulting in short or nonexistent ramp queues and delay. Geometric modifications of the ramps and electrical conduits should be an integral part of the widening contract. The only problem with this procedure is that the user benefits of the ramp control cannot be evaluated independently of the benefits of widening.

FREEWAY SURVEILLANCE BY COMPUTERS

There is nothing in the California experience to indicate that computer surveillance of traffic flow will drastically affect highway safety. Computer surveillance can assist in early detection and management of traffic accidents. The earlier an accident or incident is removed from the roadway, the less chance there is for a second accident to occur as a result of congestion created by the first one.

But nothing will ever take the place of efficient incident management, and this can only be accomplished by a management team, the key member of which is the law enforcement officer or highway patrolman on the scene.

The entire California freeway system of some 3,800 miles is under routine surveillance by the California Highway Patrol. On 42 miles of this system, the patrol is assisted by electronic surveillance. The computer is only one element of electronic surveillance, albeit a key element.

The Los Angeles Area Freeway Surveillance and Control Project—the 42-mile loop—is a large-scale experiment to determine what can be done to reduce delay, reduce accidents, and relieve motorist frustration caused by nonrecurrent congestion. It is both an operating system and an experimental system. Planning for this experimental system started in 1968, and it became operational in late 1971. We have now had 3 years of operating experience.

To a large extent, the project was a response to mounting public concern that everything possible be done to maximize the operational effectiveness of the existing highway plant through the use of state-of-the-art electronic technology. Forty-two miles of the heaviest traveled urban freeways in California were selected as the site of the project in order to make the test valid and to make it large enough in scope to be realistic. It should be noted, however, that 42 miles is only 10 percent of the Los Angeles urban freeway network, and in this respect the experiment was conservative in scope.

All the 42 miles are eight-lane or 10-lane freeways, and about 700,000 trips per day use one part or another. The highest traffic volume in this system reaches 240,000 vpd, although the average for the whole loop is about 120,000 to 150,000 vpd.

Although operations engineers have always kept track of what is happening on freeways (i.e., traffic volumes, accidents, and congestion problems), this project differs in that we get information by the moment, instead of by the day or the year, and all over instead of at random locations. The project consisted of four major elements.

Continuous data are collected for operations research. We hoped to obtain much more accurate estimates of congestion and of the daily variation in flow and congestion than had previously been available. Actually, data processing became so complex that we have been unable to produce meaningful summary data that could be used by researchers or management.

Incidents that have caused a difference in flow or congestion are detected early. It was our goal to detect 90 percent of the incidents within 5 minutes, with an average of 1.5 minutes. We came close to that goal, at the expense of some false alarms. Incidents include accidents and events such as stalled vehicles and gravel spilling from a truck, as well as major problems such as a truck tipping over or a landslide.

Early detection is not so important as management of incidents. Electronic surveillance per se does nothing to improve incident management. Communication between the data surveillance center and the field command must be greatly improved if the surveillance is going to be meaningful.

Information for the motorist is transmitted via commercial radio. However, the changeable message signs are used so infrequently and it is so difficult to make them timely, meaningful, and accurate, that we are not planning to expand their use. Even when they are timely and accurate there is usually nothing the driver can do that he could not do in the absence of such signs.

We are able to make ramp control more responsive to random changes in flow. This feature of the experimental surveillance system is being continued on an operational basis.

Surveillance, especially electronic surveillance, means keeping track of what is happening in real time, and in the Los Angeles project it was addressed primarily to unusual events or nonrecurrent congestion, as opposed to recurrent congestion.

Electronic surveillance of a freeway system does not affect traffic flow except very indirectly. The only way to control traffic on a freeway is by controlling it before it gets on with ramp signals. As opposed to surveillance, ramp control has a positive and dramatic effect on the flow of traffic on the freeway, and everybody knows that something has been done. Real-time surveillance is not necessary in order to run a ramp control system. It is necessary, however, to keep track of operating conditions and to adjust ramp-metering rates from time to time. For this reason, it might be desirable to install electronic surveillance even if it is used only to furnish data for off-line adjustments of the metering plan.

CURRENT AND FUTURE RESEARCH

Some of the problems that we are working on can be solved by research and some are engineering design problems that can only be solved by experience (or experiment). Some may never be solved.

We are now developing our third generation of ramp controllers. The first was a pretimed controller run by clockwork and had a maximum of three metering rates. This controller could not read demand or passage of a vehicle and therefore needed a yellow interval, for it could turn green when a car was still a long way upstream and then turn red before the car arrived.

For the second generation, the controller has three adjustable metering rates and can read whether a detector upstream or downstream of the ramp signal is occupied. Thus we can use red-green sequence with no yellow and get more definite one-by-one metering. The adjustable metering rates can be preset by time of day or can be called up by remote supervisory computer control.

The signal goes from steady green (in the off-peak mode) to a 3- or 4-second yellow when the metering mode starts. From then until the metering period is over, the signal rests on red so that approaching cars see a red light if there is no queue, and the yellow is unnecessary. The signal turns green when (a) there is a car waiting (on the calling detector) and (b) sufficient time has elapsed since the last green so that the allowable metering rate is not exceeded. The light stays green until the car to be served crosses the canceling (passage) detector. This sequence of decisions allows for variable reaction times among drivers and at the same time the second car in line cannot start while the light is green (because the first car has not moved). The second generation controller costs about $1,200.

It should be noted that the second generation controller requires external equipment and telemetry to change metering rates in real time (although it can operate independently if metering rates are set by clock time). The third generation, which is being specified for projects now under design, is a microprocessor that costs $1,500 to $2,000 and can do everything the second generation does plus the following:

1. It can adjust its own metering rate in response to local traffic parameters;
2. It can batch data for transmission to a central computer (thus making telemetry requirements less troublesome); and
3. It can check out malfunctions in local hardware.

One of our earliest projects (Chula Vista) used a homemade analog processor that had continuously variable metering rates responsive to occupancy on the freeway upstream of the ramp.

Of the hundred or so ramps under control in California, about 80 percent have preset metering rates. The only ramps that have metering rates responsive to main-line fluctuations in flow, on a system basis, are the unique main-line meter on the Bay Bridge and two systems on the 42-mile loop that are supervised by a large computer (Sigma 5) connected by leased telephone lines. We feel that interconnection and traffic responsiveness should be beneficial and that there are more efficient methods than using large computers with large data transmission requirements. But when we start designing alternative interconnect plans, they turn out to be so expensive that there is considerable doubt about their cost effectiveness.

One of the reasons for systemwide control with surveillance (feedback of traffic parameters) is to save labor. That is to say, the manpower requirements for manual off-line adjustments of 500 to 1,000 metering plans are formidable. But now we are beginning to wonder whether electronic surveillance will actually enable us to reduce total manpower requirements. First, it takes considerable manpower to operate and maintain the surveillance hardware. And this particular type of manpower is very difficult to train and keep. Second, we do not know exactly how to read the traffic parameters that electronics can measure, nor what algorithms to use to change or update control strategies. Third, a surveillance system that would tell us what is going on on the surface street portion of the corridor would be so complex and costly that we would never get it built.

We need theoretical or analytical models to tell us what to do to optimize flow in the corridor.

Main-line metering, or metering freeway-to-freeway connectors, has barely been scratched.

In conclusion, I would say that we are making progress but there are enough problems remaining to keep research teams as well as operations engineers busy for many years.

REFERENCES

1. Services of the State Government for the Motoring Public. 1963.
2. A Program to Upgrade and Control the Los Angeles Freeway Network. California DOT, July 1970.
3. Freeway Ramp Control—What It Can and Cannot Do. Traffic Engineering, July 1969.

Making the Most of What We Have

Edward M. Hall
Deputy City Manager, Phoenix

Because of the demands being placed on available funds, construction of new transportation facilities is rare. It is suggested that a separate Transit Trust Fund be set aside for development of public transit facilities. In the meantime, the solution to the problem is to make the most of what is already in existence. This paper discusses, first, the capacity and demand of highway and transit facilities in Phoenix and, second, how these are being improved with available funds. Programs to balance the expedient with the long term are addressed.

Meaningful long-range transportation planning is becoming increasingly complex in urban America. This is particularly true if the long-range plans include the construction of facilities to serve the public. Further complicating transportation planning and programming are the ever-increasing demands placed on available funds. These demands are not only for additional uses but for an ever-increasing number of refinements and furtherance of this or that well-intended national program.

A separate Transit Trust Fund could provide needed additional funds for urban transportation. Fortunately, efforts are continuing in the Congress by enlightened legislators to provide a separate funding source and program for transit.

From the total city perspective, there are great demands for all kinds of social, general government, and capital programs that compete for limited available funds at the local level. Then there is the taxpayer who pays for all this.

The compounding of all the above, along with the nation's serious inflationary problems, points to the need to develop short-term plans of 8 to 10 years and to implement them by capital programs of 3 to 6 years. Hopefully these short-term plans will be based on long-range plans of about 20 years. The simple conclusion is that we must make the most of what we have.

WHAT HAVE WE?

A brief overview of one very rapidly growing urban area will provide some insight into the opportunities and challenges facing the local transportation administrator.

The 1,200-square-mile Phoenix urban area has a population of 1.2 million people. The city of Phoenix is nearly 270 square miles in area and has a population of more than 750,000. Density in the city is about 2,800 people per square mile. The city's growth rate is demonstrated by the fact that in 1960 there were 439,000 people. This is a growth of more than 300,000 people in 14 years. Another graphic measure is that, just in the city of Phoenix, the building permits valuation last year was nearly $400 million.

The Council of Governments is made up of 18 cities and towns, Maricopa County, and the Arizona Department of Transportation. One of the key elements of this organization is a continuing, cooperative, and coordinated transportation planning program. This program is based on the comprehensive plans of the several agencies in the Maricopa Association of Governments.

We have had the advantage of high-level urban transportation planning by both top-notch consultants and local staff for more than 2 decades. Additional studies are under way concerning the location of a key segment of Interstate 10 to the west of Black Canyon Freeway (Interstate 17).

Although they have been thoroughly studied, our in-service ground transportation facilities are quite modest. The basic major street system is slowly evolving from a 312-mile system of concrete roads constructed by farsighted farmers in 1919-1920. Most of these were 16 feet wide, a few 18, and most have been widened by force account methods over the years. Some, of course, have been reconstructed to modern standards by the various jurisdictions.

Of the 420 miles of the major street system, 155 miles are critically deficient. Further, there is a demonstrated need for eight railroad grade-separation structures and numerous bridges over rivers and washes that carry sizeable amounts of water when it rains.

Since 1960, the city has built 65.4 miles of major streets and two railroad crossings at a cost of $36.3 million. Under way are an additional 9 miles. This was largely made possible by a bonding program of the city and state legislation in 1963 that provided funds specifically for construction and reconstruction of major streets. These modern major streets provide an opportunity for the city traffic engineer to implement a number of effective techniques for maximizing capacity and safety. The need for traffic engineers to use every possible technique is emphasized by the fact that 41 miles of major streets carry over 30,000 vehicles per day—1 out of each 10 miles—and 1 out of every 4 miles of major streets carries over 20,000 vehicles per day. Traffic volumes are increasing by about 10 percent per year.

There are only 28 miles of freeway open to traffic in the entire urban area. This is the lowest ratio of freeway mileage per capita in an urban area of over 50,000 people. Our transportation plans to date have been based on the assumption that approximately 40 percent of all vehicle-miles would be carried on modern freeways. Just as a modern freeway construction program was about to move forward, it was stopped in May 1973 by an advisory vote on a key segment. In Phoenix, one of the fastest growing cities in America, not 1 inch of new freeway has been constructed since October 1968. Those freeways in operation are already carrying near-capacity volumes.

A bottleneck removal program was initiated by the city traffic engineer in 1958. This program included 262 city-financed projects completed at a total cost of about $2.2 million. In addition to this, the program included seven federal-aid TOPICS projects costing about $1.6 million plus a computerized central corridor signal project funded at $1.5 million.

In March 1971, the Phoenix City Council accepted the responsibility for ensuring the continuation of transit service for the Phoenix urban area. The city has had a management contract with the American Transit Corporation to provide service since that time. This has been a highly successful arrangement. There are 586 route-miles of service providing transit service within one-quarter mile of about 470,000 people. However, only about 15,000 to 16,000 passengers per day use the entire system; 18,000 rode during the gasoline shortage. On April 29, 1974, express service on the freeway was inaugurated from a major regional shopping center. The Roadrunner Express now carries nearly 500 passengers per day.

In short, we are hurting. We will continue to experience rapidly increasing congestion. Only a token number of people use public transit: approximately 0.5 percent of about 3 million person-trips each day. Long-range plans for a modern high-capacity highway system are basically stalled in study and controversy. It is hoped that we will begin to make progress in this area when an ongoing study is completed within a year or so.

The level of traffic service measured by the average regional travel speed during the peak period is decreasing. About 10 years ago, the average peak-period speed was about 29 mph. It has now decreased to 21 mph and is forecast to decline to about 15 mph by 1980.

An important step toward progress was made during the 1974 state legislative session. The legislature increased, placed all in one pot, and redistributed the State Highway User Revenues. This included a 1-cent increase in the gasoline and diesel tax, increased weight fees, and so on. The redistribution provides much needed assistance to the cities and towns. This forward looking action will enable Phoenix to approximately double the major street construction program to a rate of about 12 miles per year. This construction program will allow the city to correct existing deficiencies in a 20-year period, assuming constant dollars. We are hopeful that this program can be accelerated so that existing deficiencies can be corrected within about 15 years. This goal is important because of the long lead time that now exists before substantial additional miles of modern high-capacity transportation facilities can be anticipated to serve the urban area.

WHAT ARE WE DOING?

The preceding picture is not encouraging. It is presented as an overview of the problems and to emphasize the need for a balanced program to make the most from what we have. The following are highlights of a dozen major transportation programs of the city of Phoenix.

Major Street Construction

The newly accelerated Six-Year Major Street Program will build approximately 12 miles per year. It is anticipated that over the next 6 years approximately $78 million, including federal aid, will be invested in this program to construct 74.5 miles of major streets, one costly railroad grade-separation structure, and several bridges over washes and canals. This program is based on adopted and published street policies that are geared to the adopted Functional Classification Map and adopted Minimum Right-of-Way Standards Map for all streets in the city. These maps were adopted in 1960 and 1961 respectively and are periodically updated. A significant innovation that has expedited the processing of our federal-aid program is the use of the adopted Minimum Right-of-Way Standards Map and corresponding city standards as the basis for standard section requirements.

Public Transit

A fleet of 110 buses provides service on 586 route-miles. During 1974 the operating support was about $830,000. With the assistance of a 1972 $1.9 million UMTA Capital Grant, matched by $0.9 million in city funds, we have purchased 40 new buses, ordered 15 more, equipped the entire fleet with exact-change fare boxes, and built 10 passenger sun shelters. We are now in the final stages of applying for another UMTA Capital Grant for 48 new buses, a downtown off-street transfer station, 70 passenger sun shelters, radio equipping the fleet, and other improvements. This will be a $5.3 million grant if it is approved.

Bottleneck Program

The bottleneck program has completed 262 projects at 160 intersections. Further, seven major costly projects have been included in the federal-aid TOPICS program.

Central Corridor Computerized Traffic Control System

The Central Corridor Computerized Traffic Control System is a part of the TOPICS program. It is a $1.5 million project that will interconnect 258 signalized intersections and be expandable to 410 signalized intersections. This is a traffic-responsive system, and we are hopeful that it will be on line in 1975.

Signal Modernization

This is an effective program for modernizing mast arms, controllers, left-turn arrows, and other traffic engineering equipment.

Resurfacing and Sealing

We must maintain those facilities that have been constructed. Most engineers assign first priority to this function. The city initiated a major resurfacing and sealing maintenance program in the early 1960s. The city has pioneered in the use of asphalt, rubber seal, precoated hot-chip seal to reduce dust, nighttime sealing to minimize traffic interference, and a number of other innovative programs. This program is funded at a range of $550,000 to $750,000 a year and is a primary means by which existing facilities are preserved in order to maximize their use and to keep hand labor maintenance to a minimum.

Street Improvement Districts

The Neighborhood Street Improvement District Program and the Petition Street Improvement District Program have been actively pursued. More than 256 miles of modern local and collector streets have been built by these programs since 1960, and 21 miles are under construction. New subdivisions are now constructed to proper standards.

Fall 1973 Accelerated Program for Moving Traffic

In fall of 1973 the City Council, recognizing the need and importance of accelerating programs for moving traffic, requested that the city traffic engineer develop a 60- to 90-day accelerated program. The program that was recommended to City Council, adopted, and carried to completion consisted of the following key elements:

1. Parking restrictions—Along with the already 124.4 miles of restrictions, additional peak-period and all-day parking restrictions were recommended.
2. Turn restrictions—Along with the 25 intersections with prohibitions on left turns, four additional intersections were recommended for this treatment.
3. Bottleneck removal program—Five additional projects were recommended costing $217,000.
4. Traffic signal installations—This program included installation of 26 additional traffic signals and the Central Corridor Computerized Traffic Control System currently under construction.
5. Traffic safety improvements—The new Traffic Safety Division in the Traffic Engineering Department has completed 68 projects that show a 45 percent decrease in total accidents, a 49 percent decrease in injuries, and a 12 percent increase in traffic volumes at the studied intersections. This program included 19 traffic safety projects.
6. Reversible lanes—The city has 1.3 miles of reversible lanes in operation. A study to determine whether reversible lanes could be installed on additional streets showed that one of the most promising major streets for reversible lane treatment could be widened within existing rights-of-way for less money. This was the action recommended by the traffic engineer.

7. Medians—Along with the existing 19 miles of concrete and 65 miles of painted medians, 3 additional miles of painted medians on major streets were completed. Without having the full improved modern street to work with, this improvement would not have been possible.

8. A third lane in the outbound peak traffic direction—A third lane was provided on 12.5 miles of major streets with parking restrictions. The recommended streets required resurfacing to satisfactorily block out the existing striping and marking before the new pattern was painted. This program has proved effective in moving the outbound peaks on congested major streets.

9. Street lighting—Phoenix now has more than 25,300 lights in the city. The accelerated street lighting program included installation of more than 600 high-intensity lights to provide an additional 33 miles of continuous street lighting and 700 residential lights. Also, some incandescent lighting was upgraded to mercury vapor lighting.

10. Staggered work hours, car pooling, and construction traffic control—Publicity to further the use of staggered work hours was an important part of the program. During last winter's fuel shortage, major emphasis was placed on car pooling. This effort produced good results that appear to be continuing. The city traffic engineer for many years has made major efforts to minimize delay in construction zones. A construction traffic control manual has been kept up to date and is widely distributed.

Citizen Committees

The Citizens Streets Advisory Committee has provided farsighted recommendations on street matters to the City Council since 1960. The sustained interest and enthusiasm of the committee are quite remarkable.

Recently the Mayor and City Council appointed a Citizens Advisory Transportation Committee to reappraise the broad transportation program. It is hoped that this committee will assist in the development of public support of key principles and programs so that progress can again be made in total transportation for the city and the urban area.

Restudy of Interstate Location

As a result of the advisory vote held in May 1973 and subsequent actions by various policy agencies, a restudy is under way of the Interstate 10 location to the west of existing Interstate 17 in Phoenix. It is hoped that this restudy will lead to the completion of I-10.

Advance Transportation Planning Team and Public Transit Administrator

Phoenix has had a multidisciplinary, full-time team of urban planning, traffic engineering, and public transit professionals in operation since January 1, 1961. In addition to these disciplines, other supportive disciplines such as economics and architecture have input. In 1972 the City Council authorized the new position of Public Transit Administrator. This team provides a management arm for obtaining answers ranging from immediate-action problems related to major new developments to long-range plans in which key emphasis is shifting to short-range planning. One of the most recent products of the team was the 1980 Transit Capital Program for the Urban Area, which was submitted to the Maricopa Association of Governments. The team previously did the 1972 National Transportation Needs Study transit element and of course has conducted many freeway location and area transportation studies as well as airport master planning work over the years.

Cooperation

A most important ingredient in the program is the cooperation that exists among the city of Phoenix, the Arizona DOT, and the local division of the Federal Highway Administration. We have also had excellent cooperation from UMTA in the Capital Grant and Technical Studies Programs.

The spirit of cooperation in the highway program has enabled us to move forward in an effective manner. We were able to use all of the new urban system funds allocated to the city of Phoenix last fiscal year ($3.1 million) and in fact were able to move forward on relatively short notice with one additional project costing approximately $1 million. This additional project was made possible by an outstanding effort of the city engineering staff plus the highest possible level of cooperation by the Arizona DOT and the local office of the Federal Highway Administration. The FHWA division engineer notified us on April 1 that the additional funds would be available if we could complete the necessary processing of an additional project by May 24. We did it!

YARDSTICKS AND COMPARISONS

There is a need for yardsticks and comparative measures with which to evaluate the level of transportation service. It is essential that these be simple.

One of the simplest and most effective measures of level of service is peak-period average speed. About 10 years ago, the average peak-hour speed in Phoenix was about 29 mph. It has now decreased to about 21 mph. This is a regional measure that is an effective tool.

Certainly traffic volumes, both daily and peak period, and accidents are important comparative measures. The city traffic engineer uses the number of intersections of extreme congestion as a yardstick. Over the past 10 years this has increased from 32 to 82. The traffic backup during the peak period is over 400 feet long, and the delay requires a motorist to wait more than three signal cycles to pass through the intersection.

Other factors suggested by the traffic engineer are population, vehicle registration, citizen traffic complaints, total traffic accidents, miles of major streets up to modern standards, and deficient miles.

Some feel that these level-of-service yardsticks are too simplistic. Local administrators, however, need simple, understandable, and easily obtainable measures and yardsticks. For example, capacity considerations have become highly sophisticated. In fact, I wonder whether these calculations do not lead to enlarged projects by theoretical calculations, which in turn tend to delay the project because additional funds are needed to meet higher and higher standards. Sometimes these calculations and others almost seem to prove that "it just won't work" or "it can't be done."

Although these comments may be misunderstood, I want to convey the urgent need for research and development of yardsticks, parameters, or comparative measures that can be easily, economically, and rapidly applied by the urban transportation administrator. He has not got the staff nor is he allowed the luxury of the time to develop some of the more sophisticated techniques. Surely he does not want to be provided with information that constantly adds to cost and delays on needed projects, while technicians argue the merits of the various calculations and formulas.

CONCLUSIONS

Major streets in urban areas need to be improved. These improved streets in turn provide the traffic engineer with the basics for other improvements to maximize use of street systems. This means there must be continued efforts to ensure that policy makers are aware of the need for adequate funding at the federal, state, and local levels for construction of modern street facilities.

We must balance the approach of interim betterments with a long-term effort to build a great city and thus to make possible some of the more effective traffic engineering techniques through the construction of modern street facilities. In short, we need to balance the expedient and the long term. One technique we have developed is to require that no interim bottleneck improvements be made within 3 years of a programmed major street improvement.

It is important to look at the total perspective in the allocation of funds for major street construction and transit improvements. We need to look for a maximum return on the investment; i.e., we must be cost effective. The same is true in the allocation of funds among street maintenance, traffic operations, and transit operations. These are difficult balances to achieve from a common source.

A separate Transit Trust Fund would do much to relieve the problem of trying to spread limited funds to cover ever wider needs and programs. We need to add more funds for urban transportation capital improvements. A Street and Highway Trust Fund plus a separate Transit Trust Fund could go a long way toward achieving that end.

Federal guidelines are tightening. The lag time to final construction of freeways, parkways, and major transit facilities is growing longer. Although we have achieved a high level of cooperation in our local area, it is discouraging and frustrating to see the ever-increasing requirements. Often these start out as simple guidelines but are requirements that then get rigidly interpreted and cause significant delays. Delay obviously costs money because of inflation and, perhaps worst of all, wastes time of staff as months go by while technical questions are argued between various levels of professional engineers at the various levels of government. We have been able to solve some of these problems by the top administration of the city, state, and federal highways communicating quickly with one another when problems are discovered.

The need to achieve the maximum cost effectiveness of the total transportation program was never greater. Part of the difficulty is that we really do not have a good definition of maximum cost effectiveness in terms of the total mobility of persons and goods. Further, it probably varies from urban area to urban area and with the stage of development in any given urban area. The development of relatively simple yardsticks to help achieve maximum cost effectiveness would be beneficial to the local administrator.

On top of all of this is the fact that the final decisions for the allocation of funds and thus the program balance rest with elected officials at the city, county, and state levels.

Perhaps it all boils down to the need for the transportation administrator to make the most effective use of funds appropriated by the policy makers and thus to make the most from the available transportation systems, techniques, and construction programs. It is an important challenge and assignment for all of us.

Review of Research on Controlling the Level of Service on Urban Freeways

William R. McCasland
Texas Transportation Institute

This paper presents a review of the research and implementation of traffic control systems in Texas. The magnitude of the problem is illustrated through data from freeways with the heaviest traffic flow in Texas. Systems that both control the level of service and only improve the level of service are discussed. Several approaches to entrance and exit ramp control are reviewed, and the entrance and exit advantages and disadvantages are discussed. Control of freeway demand is outlined in terms of the problems of operation and implementation and the advantages of the control. The methods and procedures for improving the level of service of operation on urban freeways broaden the scope of the paper to include geometric expansion of the roadway, detection and removal of capacity-reducing incidents, and improvement of vehicle occupancy rates. All of these methods are directly affected by the availability and use of traffic surveillance, control, and motorist communications systems.

A goal of the research and implementation program in Texas is to control the level at which a freeway will operate, but the scope of this paper is broadened to include those systems, procedures, and physical changes in the freeways that are implemented simply to improve the level of service, for they are a part of the procedure for gaining control of a freeway so that desired levels of service can be achieved.

This paper is directed principally toward the work that the cities and the Texas Highway Department are doing in conjunction with the research program of the Texas Transportation Institute.

Texas is fortunate to have a well-developed urban freeway system. The parallel service roads and the arterial street system in most instances provide adequate capacity for urban traffic demands. Traffic congestion, for the most part, is of short duration. This is not to say that there are no problems nor that the urban transportation problem is solved for the next 10 years. Rather, the situation is that there is room to grow with proper planning and control. There is time to develop other forms of public transportation that will be needed. Texas has a very fine foundation on which to build an urban transportation system.

MAGNITUDE OF THE TRAFFIC PROBLEM

The magnitude of the freeway problem in Texas today is illustrated by statistics

Table 1. Traffic problem on urban freeways in Houston.

Site	ADT	Traffic Flow	Time: Minutes	Time: Percent	Percentage of Traffic
1	170,000	Free	16¼	68	51
		Heavy[a]	5	20	25
		Congested[b]	2¾	12	24
2	130,000	Free	15	64	34
		Heavy	8	33	58
		Congested	1	4	8

[a]1,500 vehicles/lane/hour. [b]Speed reduced to below 40 mph.

given in Table 1 from the Houston-Galveston Regional Transportation Study for the urban freeways of Houston. The table shows the percentage of time that the most critical freeway sections in Texas experience congestion.

During the peak period, 24 percent of the traffic experiences congestion and an additional 25 percent is in heavy traffic at site 1. If only 2 percent of the peak-period congested traffic is diverted and if strict control of traffic flow can be maintained, then no traffic will experience congestion and 47 percent will be in heavy traffic flow conditions.

If the control is much stricter so that 25 percent of the peak-period heavy plus congested traffic is diverted and if strict control of traffic flow can be maintained, no traffic will experience congestion and no traffic will experience heavy traffic flow.

How could 25 percent of the peak-period traffic be diverted?

1. Divert truck traffic, which is 1 to 2 percent of the traffic.
2. Increase the car pooling rate from the present 1.3 to the desired 1.45. This would produce a 10 percent reduction of vehicles and no change in person volume.
3. Use control to divert 750 vehicles per hour, which would produce a 10 percent reduction.
4. Increase ridership on buses from 150 to 375 passengers per hour—a reduction of 2 to 5 percent (no change in person volume).
5. Try other methods.

These methods would yield a total diversion of 25 percent of the vehicles but only 10 to 12 percent diversion of persons.

MEASURES TO CONTROL LEVEL OF SERVICE

Some of the measures taken in Texas to reduce or eliminate traffic congestion on urban freeways and to control the level of service provided the motorists are discussed below.

Ramp Metering

Now into its second decade, ramp metering continues to offer benefits to urban freeway operations by increasing efficiency and reducing accidents. It has been the experience in Texas that ramp metering alone cannot control the level of service. It delays the development of congestion, shortens the length of congestion, reduces overall travel time, and reduces accident experience and the resultant costs of property damage and delays. It establishes a proper travel pattern for existing conditions. It may, in some instances, actually eliminate traffic congestion, but often congestion continues because

1. Control strategies often have optimum throughput as a goal, which depends on exit volume demands that are not controlled;

2. There are ramp control overrides, such as queue length and maximum waiting time, that must be accommodated; or

3. There are ramp demands that must be accommodated, regardless of freeway conditions.

Even so, ramp metering continues to be a very cost-effective control system.

Closure of Entrance Ramps

Experiments and studies on the effects of ramp closures have been conducted many times during the last 15 years. They have proved that entrance-ramp closures improve the level of service more than metering does, but they can only be used when good alternate routes are available and when the political consequences are favorable.

Ramp closure may not be so efficient as metering in improving overall traffic operations if it is not responsive to changing traffic conditions on freeways and arterials. Ramp closures do have the advantage of providing priority control to buses and perhaps car pools.

Ramp closures should be considered as a part of a ramp control program, in combination with ramp metering where possible.

Closure of Exit Ramps

On occasions, exit-ramp closures can be used to improve freeway operations. Permanent closure is the most acceptable from an operational view, but, because of access rights established over the years, it may not be possible. Relocation of ramps as part of a betterment program is being used successfully in Houston.

Real-time control of exit ramps has been tested in Dallas and shown to be effective. In the North Central Expressway corridor, there are adequate alternate routes adjacent to the freeway to provide a bypass to critical bottlenecks. Closing exits just downstream of the bottleneck caused a new traffic pattern to develop in which large volumes exited the freeway upstream of the bottleneck.

Methods of operating the ramp closure are still a problem, and so manual closure was used in the tests.

On-Freeway Control

The last element of urban freeway traffic to be controlled is the on-freeway traffic, i.e., main-lane traffic. This concept has been tested on special controlled-access facilities such as tunnels, bridges, and, to a certain extent, toll roads.

Texas is currently trying to control traffic demand with traffic signals to reduce the density of traffic downstream. At this time, there are many more questions than answers, which means considerable study and testing remain to be done. The basic question is, can this type of control be applied to urban freeways in a safe and efficient manner? This same question was being asked several years ago when ramp metering was being developed.

Questions

1. Can the control be operated safely?
2. Will the control be recognized by motorists as a valid control measure?
3. Will on-freeway control be effective in achieving the desired objective?
4. Can it be applied to all urban freeways?

In answer to question 1, there is no other way that it will be done. If advanced signals, signs, markings, and so on cannot provide reasonable assurance of safe operation, then the control should not be used.

With regard to motorists, adequate design with strict enforcement will be necessary. The same is true of other unique control measures such as ramp metering and lane assignment to car pools.

Question 3 is the easiest to answer. Texas has conducted both theoretical and field simulations of this type of operation, and data have been collected that support this type of control.

As with ramp metering, each freeway has different operational characteristics that must be considered in the design and operation of a control system. Careful consideration must be given to the first test site for this control. Only after experience in acceptance by the public and the operational achievements of the control has been obtained can question 4 be answered.

Objectives

The purpose of on-freeway controls is to control traffic demand so that, on the freeway lanes downstream of this location, a desired level of service can be maintained. This desired or designed level of service may be set to maximize throughput of a downstream bottleneck; it may be set to provide a high level of service to all vehicles downstream of the control site, particularly buses and car pools that enter from nearby park-and-ride terminals; it may be set to distribute total delay in a more equitable manner, with some freeway capacity being saved for downstream entrance ramps; or it may be set to encourage diversion from the freeway because of normal traffic congestion or because of capacity-reducing incidents.

Problems of implementation

As with any new type of control system, there will be problems in implementing the traffic signals. First, all operating agencies at the local level must approve the control. The policing agency must be convinced that the signal is enforceable. The state and federal agencies that fund and install the system must be convinced that it is a proper application of traffic control principles. Finally, a suitable test site that has the highest probability of success must be located.

Alternatives

Following are some of the alternatives to total freeway control that allow control of traffic demand to achieve desired levels of service.

1. Ramp metering—Generally, ramp metering cannot achieve the necessary level of control to reduce freeway densities that have increased above a critical level.
2. Ramp closure—Closing a ramp is more effective than ramp metering, but may be impractical because of the lack of alternate routes or political considerations.
3. Lane closure—Lane closure has been used to provide a transition to reversible flow, to balance merge and capacity with demand, to improve merge area operations, and to warn motorists in advance of lane blockage. The use of lane closure to control demand on the main lanes of the freeway has not been tested. Merging one lane of traffic into adjacent lanes that will be congested is difficult.
4. Lane assignment—Reserving a lane for buses or car pools or both could have the effect of reducing the input to a downstream roadway section. It is less restrictive than lane closure, but has the same drawback in operation and is an enforcement problem.
5. Speed control—It has been suggested that strict approach-speed control could reduce demand. This has not been very effective on urban freeways when motorists have been requested to do so voluntarily. Speed control into tunnels or bridges is more effective.

Advantages

Control of all freeway lanes by a standard traffic signal should be much easier to enforce than lane closure, lane assignment, and speed control. Also, motorists should understand what is expected of them.

On-freeway control should require minimum time inasmuch as the control directly affects flows on the order of 100 to 130 vehicles per minute. This type of control would also probably be less expensive than any of the alternatives.

On-freeway control could be further enhanced by the application of lane assignment upstream of the control site. That is, a lane reserved for buses or car pools could move through the freeway signal with little or no delay.

MEASURES TO IMPROVE LEVEL OF SERVICE

This paper has concentrated on the subject of on-freeway control because it provides the opportunity to close the loop on the freeway control system and to apply the principle of control of level of service.

There are other activities under way in Texas that apply more to improvement of level of service.

Expansion of Roadway

Temporary or permanent expansion of the roadway is always an excellent approach to solving a capacity-demand problem. However, the addition of a new lane to a freeway may be financially or physically impractical.

The temporary use of a roadway shoulder as another travel lane has been practiced for several years, but often this approach has been rejected because shoulders are not continuous over bridges. After freeways are constructed, traffic patterns change. The use of the shoulder over short sections of freeway may prove very beneficial in relieving a bottleneck section. Further expansion of the roadway may permit the return of the shoulder to the use of an emergency lane.

The same consideration is being made in the merging areas of major interchanges. The use of the shoulder for travel can move the point of merge some distance from the interchange and reduce the conflicts caused by sight restrictions. The improvement in accident experience is expected to be significant.

Detection and Removal of Disabled Vehicles

Maintaining existing capacity is most important. The management of traffic incidents is one area that requires more emphasis in terms of controlling level of service, for one minor incident can render all control systems, priority systems, or expansion systems useless.

Texas is promoting the use of service patrols in an attempt to define what is cost-effective and how to best use the few vehicles that are made available for this function.

The research program is investigating electronic systems for detecting incidents at low and high volumes. The highway department has taken the positive research results in the Accident Investigation Site (AIS) system and is promoting AIS procedures that do not require expensive capital investments, but do require cooperation and involvement of the policing agencies. Time spent on accident reports at or near the scene of the accident is too long and affects both directions of traffic flow. The provision of a parking area out of sight of freeway traffic can be accomplished by new construction on freeway rights-of-way or on adjacent streets and parking lots.

Assignment of Facilities to High-Occupancy Vehicles

One method of controlling on-freeway traffic demand is the use of lane assignment for high-occupancy vehicles. Its application to the control of the level of service should be examined.

If a freeway lane is taken away from general use and assigned to buses and car pools, the remaining lanes will in most instances operate very poorly at a lower level of service. If the reserved section of the lane is relatively short, compared to the average trip length, and if the reserved lane reduces demand to the downstream section and causes all lanes to operate at a high level of service, the overall level of service may be improved. If the served lane is added to the direction of flow (either new construction or counterflow operation), the level of service of the other lanes is improved by a reduction in demand and elimination of buses.

The approach in Texas has been to improve bus operations by improving operations of all vehicles on the facility where possible, but priority (lane bypass, entrance ramp, signal preemptions) has been provided, where necessary, at the expense of passenger vehicles. In Dallas, entrance ramps are closed to all vehicles except buses, some ramps have a bypass lane, and the traffic signal at intersections is preempted by approaching buses.

In addition, consideration is being given to the assignment of a freeway lane upstream of the control station.

Improvement of Vehicle Occupancy Rates

Reducing vehicle demand by increasing bus and car pool use relates directly to the control measures mentioned. The computer matching activity is being promoted by cities in Texas as elsewhere in the United States. Incentives other than improved level of service are also being offered.

CONCLUSIONS

The basic control systems and procedures for improving and controlling the level of service on urban freeways have been or are now being tested. Implementation of proven systems has taken longer than expected because of the need for an agency to operate the systems. Also, there are many agencies and jurisdictions involved in urban traffic that must be coordinated.

Finally, it is important to realize that in most areas urban freeways are considered the backbone of the public transit system and the surveillance and control systems that are being developed are of greater importance when the public transportation vehicles are considered.

ACKNOWLEDGMENT

The research discussed was conducted by the Texas Transportation Institute and the Texas Highway Department in cooperation with the Federal Highway Administration, the Urban Mass Transportation Administration, and the City of Dallas. The contents of this paper reflect the views of the author who is responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration or the Urban Mass Transportation Administration. This paper does not constitute a standard, specification, or regulation.

Park and Ride: The Dallas Program

John J. DeShazo, Jr.
Dallas Department of Traffic Control

As part of the Dallas Urban Corridor Demonstration Program, a park-and-ride facility was developed as an incentive for automobile users to switch to transit. The goals of the program were to improve the level of service on the North Central Expressway and to divert some of the 160,000 trips projected for 1990, but improvements in air quality and energy consumption were also achieved. Because of the success of the initial park-and-ride facility, other facilities were developed. This paper presents the criteria used for selecting and evaluating potential park-and-ride sites. Other data such as demand analyses and on-bus surveys were used to determine the need and desirability of park-and-ride facilities. Another factor in site selection was the pending opening of the Dallas/Fort Worth Airport.

Dallas was one of the cities selected by the U.S. Department of Transportation to participate in the Urban Corridor Demonstration Program (UCDP). The primary thrust of the program was to relieve congestion in central business district travel corridors during peak periods by diverting trips from automobiles to high-occupancy vehicles. This effort was to be accomplished by creating incentives, i.e., providing preferential treatment to buses, to attract trips to transit.

The North Central Expressway corridor was selected as the location for the Dallas UCDP. The North Central Expressway, a controlled-access freeway that extends from downtown Dallas generally northward to Richardson, a distance of approximately 12 miles, is the principal travel facility in the corridor. Interchanges on the expressway are diamond types except for a full cloverleaf interchange at Loop 12 (Northwest Highway) and a directional interchange at I-635 (L.B.J. Freeway). The expressway has three lanes in each direction from the downtown area to Mockingbird Lane (approximately 5 miles) and two lanes in each direction from that point north. One-way frontage roads, generally three lanes wide, are continuous except at the Mockingbird, Loop 12, and L.B.J. interchanges. Traffic entering the freeway is controlled by a ramp-metering system that has been in operation for 3 years. Also additional control systems are interfaced with the ramp-metering system in an effort to optimize travel throughout the corridor.

The specific UCDP projects approved by the Department of Transportation for implementation in the North Central Expressway corridor (1) are

1. A freeway control system giving preferential treatment to transit vehicles and

2. An arterial signal control system allowing transit vehicles to preempt signals.

As a complement to these control strategies a park-and-ride facility was developed at the northern end of the corridor. Although the initial emphasis was to improve the level of traffic service on the freeway and to divert some of the 160,000 trips projected for the facility in 1990, recent air quality control and energy conservation measures have become important factors as well.

PLANNING EFFORT

Design and development of a park-and-ride facility were based on the following objectives:

1. Optimize the use of each mode of travel in order to minimize urban congestion and travel time, improve air quality, and conserve energy.
2. Develop a strategy for an integrated transportation system and examine its potential as a valid mode of transportation for work trips in the Dallas area.
3. Identify the necessary conditions for an effective policy on park and ride.
4. Develop a facility that serves trips destined to both the CBD and the new Dallas/ Fort Worth Airport.

The general location for the park-and-ride facility was based on studies that identified the centroid of trip origins for air travelers living in the North Dallas area and to serve residents in the North Central Expressway corridor bound for destinations in the Dallas CBD. These studies showed that the intersection of L.B.J. Freeway with Central Expressway was the focal point of these two considerations. Since there were several candidate sites available near the intersection of these freeways, the following criteria were developed as a basis for evaluating each site:

1. Cost,
2. Accessibility,
3. Visual exposure to users,
4. Compatibility with long-range transit plans,
5. Compatibility with desired SURTRAN routing,
6. Compatibility with Texas Highway Department freeway access plans,
7. Compatibility with desired CBD expressway routing,
8. Availability of property,
9. Time required for purchase and access revisions,
10. Compatibility with adjacent land uses,
11. Compatibility with land requirements, and
12. Effect on other traffic operations.

The sites shown in Figure 1 were rated according to the criteria. Table 1 gives a comparison of the sites and shows that site D is the most desirable location for a park-and-ride facility.

DEMAND ANALYSIS

The park-and-ride facility in the North Central corridor was planned to serve as a park and ride for person-trips to downtown Dallas and also as a terminal for the surface transportation system, SURTRAN, to the Dallas/Fort Worth Airport. This multifunctional capability made the project especially compatible with regional public transportation plans.

The demand for parking facilities was estimated in part by relating the diversion of potential CBD work trips to fringe or intermediate parking locations. During the Dallas Center City Transportation Project, special interviews were conducted at

Figure 1. Site priority.

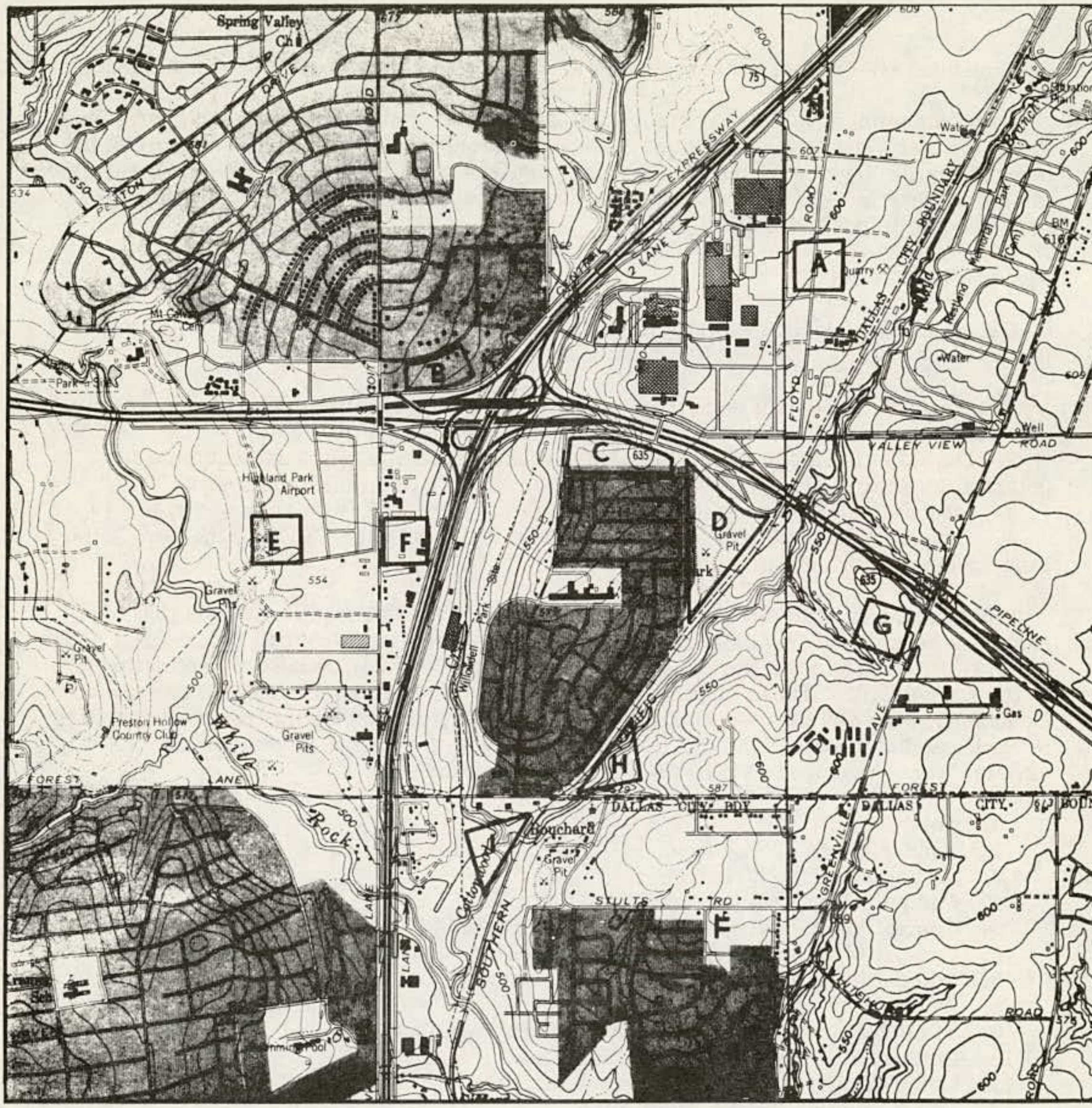


Table 1. Evaluation criteria for terminal site selection.

Criterion	A	B	C	D	E	F	G	H	I
	Terminal Site								
Accessibility	3	1	3	1	4	2	2	3	2
Visual exposure to users	4	1	2	2	4	2	2	4	3
Compatibility with long-range transit plans	1	4	3	1	5	4	3	1	1
Compatibility with THD freeway access plans	3	2	3	1	5	4	1	2	2
Compatibility with desired SURTRAN routing	3	1	2	1	4	3	2	3	2
Compatibility with desired CBD express routing	3	1	3	2	3	1	4	4	2
Availability of property	2	2	2	3	3	3	2	1	2
Time required for purchase and access revisions	3	2	2	3	5	4	2	2	2
Compatibility with adjacent land uses	1	1	3	2	2	1	2	1	2
Compatibility with land area requirements	2	2	3	1	2	3	1	2	1
Effect on other traffic operations	2	4	3	2	3	3	2	3	3
Total	27	21	29	19	40	30	23	26	22

Note: 1 = best; 5 = worst.

major office buildings in downtown Dallas (2). The results of this survey indicated the potential for a 20 percent diversion. In 1970, the occupancy level of CBD parking spaces by work trip parkers was about 30 percent of the 50,000. This figure of 15,000 spaces was established from an analysis of trip purpose, duration, and accumulation data. Thus about 3,000 spaces or 20 percent of 15,000 spaces should be established for design.

The demand for the park-and-ride facility was substantiated by the need for a high level of ground transportation to the Dallas/Fort Worth Airport to serve passengers, visitors, and employees. Based on the 1972 Terminal Movement Study forecast and the 1969 on-board bus survey, approximately 35 percent of the air passengers in the Dallas/Fort Worth region have origins or destinations in northern Dallas. About 81 percent of the region's air passengers using bus or limousine service will have origins or destinations in downtown and northern Dallas.

Employee needs for service to and from the airport are entirely different and constitute a large portion of public transportation service. About 50 percent of the daily employee population at the airport has origins or destinations in the Dallas area of which 15 percent originates or terminates in the North Dallas area.

JOINT PARK-AND-RIDE/SURTRAN TERMINAL

In October 1973 the city of Dallas decided to implement another park-and-ride facility on an interim basis pending approval by the Texas Highway Department and the U.S. Department of Transportation for a grant to cover a portion of the cost of the proposed terminal. This decision was made in an effort to meet the target date for the opening of the Dallas/Fort Worth Airport and a commitment to the Environmental Protection Agency to establish park-and-ride service as a means of reducing pollution caused by automobiles (3). The city entered into a 2-year lease on 4.5 acres of land on the west side of Central Expressway just south of I-635 at a cost of $50,000 per year. The lease included use of an existing building for a terminal and a paved parking area accommodating 450 automobiles. Repairs to the building and parking area cost an additional $20,000. It is anticipated that the permanent park-and-ride terminal will be completed at the end of the lease period. The city is recovering some of these costs from a rental charge to SURTRAN and revenues from parking, car rental operations, and food and beverage concessions.

The facility was opened 2 days after President Nixon appealed to the nation to conserve energy. Nonstop bus service was inaugurated from the terminal to the CBD. The operation was a success from the outset. The facility was overrun with automobiles on its first day of operation. The 692 parkers overflowed the facility onto an adjacent unpaved area. The Dallas Transit System carried 1,058 passengers inbound and 986 outbound on the first day. The high interest continued during the next 2 days of operation when 679 and 628 parkers respectively used the facility. Transit ridership was comparable to that observed on the first day. On the fourth day of operation, use dropped to 479 and the total riders (inbound and outbound) to 1,550. This decline was anticipated, for coupons had been distributed for free bus service on the first 3 days of operation. The use was up to 560 parkers and 1,686 riders on the following Monday. The city leased and surfaced an adjacent area to accommodate several hundred additional automobiles after it became apparent that the service was a success. Use of the facility continued at this level until the opening of the Dallas/Fort Worth Airport on January 13, 1974. After the initiation of SURTRAN service at the facility, the parking facility was unable to accommodate the vehicles left by both air travelers and commuters. A search began for another facility that could provide an instant solution to the need for additional parking space. There was no room for further expansion of the original facility. Consequently, on January 22, 1974, the North Dallas Park-and-Ride Terminal was moved across North Central Expressway to the Gemini Drive-In Theater, which could accommodate 1,000 cars. Currently, more than 1,600 passengers ride the express buses, which have headways of 5 minutes

during rush hours. More than 600 automobiles are parked in the theater lot each day.
The SURTRAN terminal continues to be a success at the original location. Use of the SURTRAN parking facility ranges from a low of about 225 vehicles on an average Sunday to more than 450 on an average Wednesday.

ADDITIONAL PARK-AND-RIDE TERMINALS

The North Central Expressway Park-and-Ride Service had such a favorable response from the public that the city next initiated a program to identify and assess potential park-and-ride sites in other sections of the city. Information on some 30 sites was evaluated to varying degrees of detail. Locating and evaluating potential park-and-ride terminal sites were based on the following criteria.

1. A site should require limited improvements; that is, the parking area should be already paved, lighted, and fenced if possible and have a bus loading area available.
2. A site should have ready access to a freeway to the CBD.
3. A site should be located at the centroid of an area that would have the potential for attracting park-and-ride patrons.
4. A site should be located to offer as many Dallas residents as possible the opportunity to use park-and-ride services.
5. A site should be available at minimum expense to the city, and lease, improvement, and bus operating costs should be fundable from operating revenues if possible.
6. A site should be located near an existing bus line in order to use that service during off-peak periods.
7. A site should be available for use as soon as possible.

Pleasant Grove Facility

Based on these criteria and results of site assessment, the Pleasant Grove Athletic Field parking area containing 710 parking spaces was selected as a site to serve the Pleasant Grove area of Dallas. This facility is owned by the Dallas Independent School District and is used only for athletic events during a limited time of the year. The city leased this parking facility at an annual rental of $15,000. Because only minimal improvements were required for lighting and shelter, the first year lease and improvement cost amounted to $23,300, and the second year would require only $15,000 for the rental fee.

The facility was opened and express bus service initiated on January 23, 1974. Although initial use and transit ridership were less than had been anticipated, both have climbed steadily from an initial daily average of 165 parkers and 516 inbound and outbound riders to 337 parkers and 745 riders in June 1974.

Oak Cliff Facility

A park-and-ride site adjacent to South R. L. Thornton Freeway, just north of Loop 12, was opened in April 1974. It serves residents of Oak Cliff and cities to the south and southwest of Dallas and can accommodate 625 cars. The 5-acre site was leased for 2 years at an annual rental of $48,164.80. The site is paved, lighted, and fenced.

Although this facility has been less successful than the two earlier ones, the average number of daily parkers increased from 115 in April to 140 in June. The total inbound and outbound ridership increased from 274 to 334 in the same period (Table 2).

PARKING AND TRANSIT COSTS

A fee of 25 cents is charged for all-day parking at each of the facilities. The bus fare is 50 cents each way or $1.00 for the round trip.

Figure 2. Proposed sites for park-and-ride facilities.

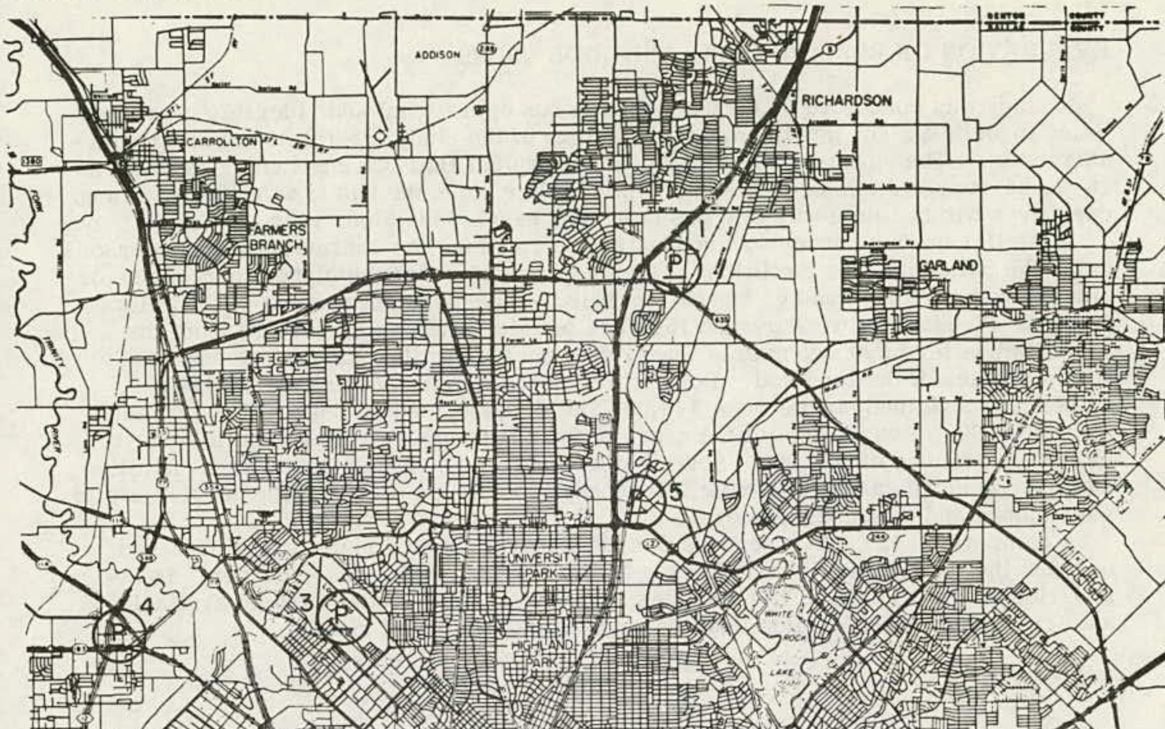


Table 2. Average daily riders and number of cars at park-and-ride facilities.

Month	North Dallas		Pleasant Grove		Oak Cliff	
	Riders	Cars	Riders	Cars	Riders	Cars
December	1,437	1,090				
January	1,794	1,220	558	400		
February	1,621	1,168	573	420		
March	1,619	1,160	688	580		
April	1,554	1,114	697	630	274	230
June	1,639	1,200	745	674	334	280

EVALUATION OF BUS AND PARK-AND-RIDE USERS

A study was conducted in 1972 to evaluate bus operational and ridership characteristics in Dallas. The most significant findings of the study are related to rider characteristics. The on-bus survey identified the typical Dallas transit rider as being a 16- to 24-year-old female. Also the typical rider rides the bus to work daily, has no car, lives within two blocks of a bus stop, and earns $4,000 per year.

A similar on-bus survey was conducted to determine the characteristics of persons using the park-and-ride facilities. The study revealed that the highest percentage of riders are between 35 and 65 and are married. Seventy-eight percent of the riders from the Pleasant Grove Park and Ride are female, whereas 61 percent from the North Dallas terminal are male. The typical user rides the bus to work and lives within 5 miles of the terminal. Forty-three percent have incomes under $10,000, 41 percent have incomes between $10,000 and $20,000, and 16 percent have incomes over $20,000. Two percent of the users of the Pleasant Grove Park and Ride have incomes in excess of $20,000. Seventy-three percent of the users are from families with two or more cars. The most frequent reasons given for using the service are to save money and to conserve energy.

A comparison of the two surveys indicated that there is a significant difference between the typical park-and-ride user and the typical Dallas transit rider. Park-and-ride users (a) have higher incomes, (b) have another mode of travel available, and (c) include fewer females than the regular Dallas transit riders.

FUTURE PLANS

Dallas recently completed a subregional public transportation plan that proposes additional park-and-ride facilities and express bus service between the terminals and the CBD. The locations of these facilities are shown in Figure 2. Two of these are scheduled for service in 1975 and the remainder in 1976.

Also the city in concert with the U.S. Department of Transportation will expand the Dallas UCDP to include operation and evaluation of dial-a-ride service between the North Central Park and Ride and the surrounding neighborhood. This service will use six small 18- to 22-passenger buses operating between the terminal and stops within a 1- to 2-mile radius.

Dallas has experienced the benefits that can be realized from park and ride and, therefore, plans to pursue an aggressive strategy to implement additional park-and-ride facilities in concert with other low-capital programs in an effort to meet its transportation requirements.

ACKNOWLEDGMENT

The assistance of R. W. Kelly and Sayed Omar, Traffic Control Department, City of Dallas, in the preparation of this report is acknowledged.

REFERENCES

1. Urban Corridor Demonstration Program for Dallas, Texas. City of Dallas, Texas Transportation Institute, and Wilbur Smith and Associates, 1972.
2. Center Cities Transportation Study. Wilbur Smith and Associates, December 1970.
3. Plans for Reduction of Hydrocarbon Emissions From Mobile Sources in Dallas County. Traffic Control Department, City of Dallas, October 1973.

Bus Priority Strategies and Traffic Simulation

Ann Muzyka
Transportation Systems Center
U.S. Department of Transportation

Simulation of corridor traffic (SCOT), a recently developed computer model that simulates traffic flow within a specified traffic system, has been used to predict the effect on bus service and general traffic performance of implementing candidate bus priority strategies. Numerical values of standard traffic performance measures were determined from computation of network vehicle trajectories. SCOT was calibrated to current peak-hour traffic conditions within an urban street grid representative of the central business district of Minneapolis. Data from city agencies were used in conjunction with field experiments to verify SCOT simulations. Various bus priority strategies designed to increase bus speeds by providing bus-only lanes were evaluated. The significant elements in bus travel time were shown to be frequency of station stops and red light signals. Further studies planned include identifying optimum bus station locations, developing bus progressive traffic signal timings, and evaluating various bus preemption of traffic signal strategies. Demonstration projects are under consideration to evaluate the simulation technique as a transit operations tool.

Until recently, traffic engineers were given the responsibility of moving vehicles safely and efficiently through city streets and highways. The bus operates at great disadvantage in the general traffic stream. If the objective of a transportation system is to move people and freight, and not simply vehicles, then the assignment of priority to buses and trucks in the traffic system must be evaluated. There is considerable activity in the transit industry to provide buses with priorities on the roadway and within the traffic control system commensurate with their people-moving capability. Some bus priority strategies already in effect or planned for early implementation are downtown bus-only streets, both direct-flow and counterflow bus-only lanes, bus bypass lanes for ramp-metered freeways, and bus preemption of traffic signals.

Before a proposed bus priority strategy is implemented, a preliminary analysis is conducted to estimate the expected benefits for the bus system and the disadvantages for the automobile and truck system. Experiments are conducted before and after the bus priority system is put into effect, and comparative bus and automobile-truck performance indexes are measured. Because traffic performance depends very much on weather conditions, time of day, day of the week, special events, and holidays, field observations must be made over a time span sufficient for statistical validity. In addition, driver adaptability makes properly controlled experiments difficult to design.

Computer simulation affords an ideal method for evaluation of candidate bus priority strategies, as well as general traffic control systems. The public is not inconvenienced, nor can it interfere with a computer experiment. Traffic conditions can be controlled to determine the systemwide effect of varying any individual component. Time can be compressed so that many strategies may be evaluated in a few weeks and only the most promising ones selected for eventual field test. Computer simulation can prove to be a valuable traffic engineering tool. However, it is, at best, an approximation to the real world and cannot replace field experiments. But it can provide the insight necessary to identify problem areas and lead to more successful demonstration projects.

STUDY PLAN

The objectives of this study were as follows:

1. To predict the changes in traffic performance and bus service caused by implementing various bus priority strategies,
2. To conduct a demonstration and thereby determine the validity of the preliminary analyses performed, and
3. To evaluate the usefulness of the computer model as a transit operations tool and to formulate an appropriate training program if the decision is made to recommend widespread use of the model.

The study area selected was the Minneapolis central business district. Most of the data needed to model the current traffic and bus system were made available by the traffic engineer of the City of Minneapolis and the operations authorities of the Twin-Cities Metropolitan Transit Commission. They cooperated in conducting experiments to acquire the remaining necessary data and selecting the candidate bus priority strategies. These strategies specified designation of bus-only lanes on two one-way streets. The first strategy specified that one lane be assigned to buses on the five middle blocks of two 10-block streets in the direction of traffic. This constraint required that passenger cars and trucks not be permitted to make right turns or to stop within this region. The second strategy specified that one counterflow lane be assigned to buses for the entire 10 blocks of two streets. No restrictions on turning movements were made, but one lane (of five) was used to separate the two directions of flow.

The study plan included using the computer model SCOT to simulate current traffic flow patterns for the morning and afternoon peaks. These would be the base or before cases and provide the quantitative traffic performance and bus service data. Once the model was calibrated to the study area, each bus priority strategy would be simulated under the same traffic conditions and the changes in traffic performance would be determined. Recommendations would then be made for a demonstration project for the more promising strategy.

COMPUTER MODEL

SCOT (1, 2, 3) is the amalgam of two computer programs: UTCS-1, urban traffic control system (4, 5), and DAFT, dynamic analysis of freeway traffic (6). In this study, the urban, microscopic part of the model was used because traffic flow within only the Minneapolis CBD was to be simulated. This traffic flow is characterized by different trajectories for passenger cars, trucks, and buses, governed by driver types, roadway geometry, presence of queues, downstream signal phase, and driver destination. Thus the trajectory of each vehicle must be individually computed. From these trajectories, traffic performance measures may be determined for each street. The traffic engineer can then identify the bottlenecks and the time and duration of congestion for each candidate bus priority and/or control system. Traffic demand can be held fixed, resulting in a controlled computer experiment—the ideal condition for evaluating a proposed change in bus service. Any component in the system may be systematically increased and the breaking point of the candidate strategy determined without public inconvenience.

The model logic specified vehicles entering the network at various points on the periphery, their flow rates, and automobile-truck ratios determined by field observations. As each vehicle enters the system, it is designated as a passenger car or a truck and randomly assigned one of 10 driver types. The driver set ranges from extremely conservative to reckless. These designations stay with the vehicle for the duration of its network trip and influence the selection of such quantities as acceleration (more sluggish for trucks) and desired cruising speed (highest for reckless). The vehicle speed profile includes acceleration, cruise, and deceleration segments. The desired cruising speed is selected on the basis of driver type from a statistical distribution about a mean free-flow speed observed for each network link. Crusing continues until deceleration is indicated. This mode is triggered by the presence of an obstacle in the vehicle's path (a slow-moving vehicle or a queue), a red or amber traffic signal downstream, or the need to make a right or left turn at the next intersection. The driver may change lanes when his lane is obstructed and an acceptable gap exists in an adjacent lane. The position and speed of each car are updated every second, and collisions do not occur. The deceleration regime is designed to ensure a safe stop. As a vehicle approaches an intersection, it may turn right or left, proceed through, or stop. Turning movements are in conformity with field site traffic statistics. Movement at an intersection depends on the traffic demand there and the signal currently displayed. The response of a driver to an amber signal depends on his speed and proximity to the intersection and which of the 10 driver types he is. Delays caused by pedestrian movement and determined from field data are added to the trajectories of right- and left-turning vehicles. Queues are discharged in a systematic way so that vehicles start up sequentially and arrive at the intersection at times and with speeds dependent on their position in the queue. Buses move through the network and stop at their assigned stations, which, in general, interferes with traffic. Their dwell times are computed from a statistical distribution about a specified mean value. Provision is made for intralink lane-blocking events. These simulate predictable temporary lane blockages, such as double parkers, taxi pickups, parking garage queues, and unloading trucks.

MODEL CALIBRATION: THE INPUT

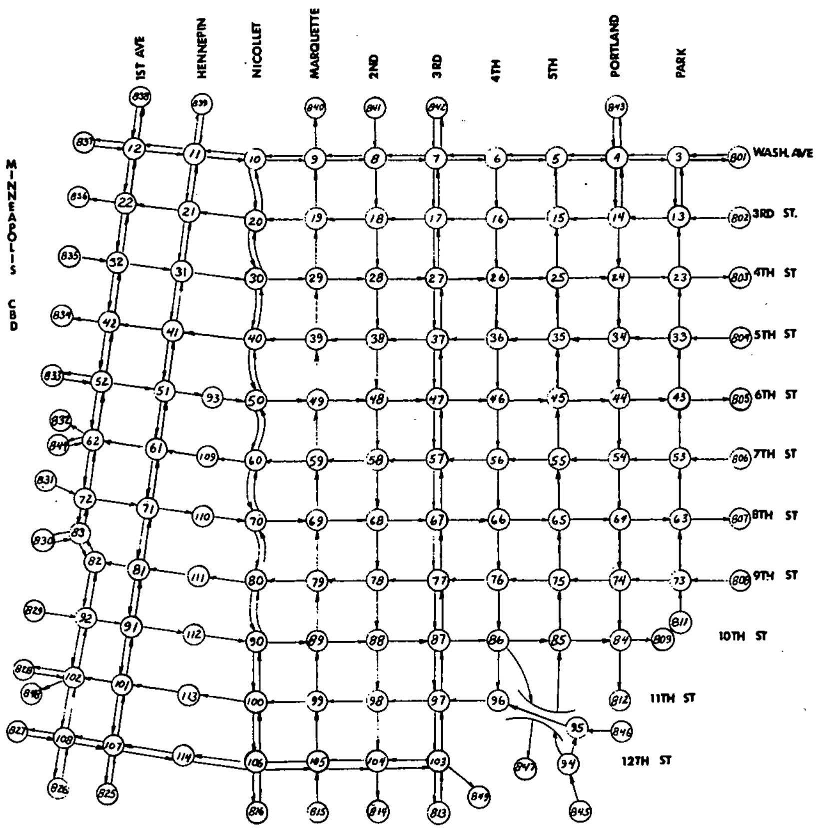
Before SCOT can generate traffic patterns, it must be calibrated to the study area. This means that the data describing the traffic system must be obtained and keypunched for input to the computer program. The study area selected was a portion of the Minneapolis CBD and is approximately a rectangle of 10 by 11 streets. It includes one- and two-way streets, a bus-only street, midblock and intersection traffic signals, parking lots, downtown and crosstown buses, pedestrians, and the downtown exit ramp of an urban freeway. This study area is modeled by a grid consisting of a set of unidirectional links that represent the street roadway and a set of nodes that represent the intersections. The information required to describe the study area traffic system falls into four sets: geometric, traffic demand, control system, and bus service data.

Geometric Data

The network model (Fig. 1) consists of 251 network links, 27 entry links, 27 exit links, and 114 network nodes. A two-way street requires two links. Two streets between Hennepin and Nicollet were not modeled as links, but their effects were represented by traffic sources and sinks.

For each street, the stop-line to stop-line span, the number of travel lanes, and the storage capacity of left- and right-turn pockets are needed. In this case the number of

Figure 1. Network model.



lanes per link varied from one to five; the spans varied from 254 to 1,114 ft, with the average span about 400 ft. The dimensions of the entire network are approximately 4,300 by 4,300 ft. Grades were not significant in this study area. The capacity of turning pockets ranged from two to 10 passenger cars.

Traffic Demand Data

The traffic conditions simulated were those occurring on a weekday morning between 7:30 and 8:30 a.m. and on a weekday afternoon between 4:30 and 5:30 p.m. From traffic counts at each intersection, turning movements and pedestrian counts were determined. Turning movements are assigned to each vehicle randomly at each intersection but in conformity with the field data; e.g., if the field data showed that 78 percent of the vehicles went straight through, 10 percent turned left, and 12 percent turned right, then a 100-sided die is cast for each vehicle as it enters a given link and, if the number lies in the set 1 to 10, the vehicle turns left at the downstream end of the link; if the number lies in the set 11 to 88, the movement is straight; and, if in the set 89 to 100, the vehicle turns right. Thus the turning movements are random, as in reality, and not uniform. Pedestrian volumes at each intersection are classified as negligible, 100 to 250, 250 to 500, and over 500 pedestrians per hour. Vehicles making right and left turns in conflict with pedestrians are given delays, and the amount of delay depends on the volume of pedestrian activity and the time within the green phase (heaviest at the beginning of the green phase).

The entry links bring vehicles into the network at observed uniform rates and with the observed truck-passenger car ratio. For the morning peak hour, the entry flow rates at the 27 entry points ranged from 21 to 3,165 vehicles per hour, the largest coming from the exit ramp of I-35W. Truck volumes ranged from 3 to 34 percent. Total entry volume was 19,999 vph. For the afternoon peak hour, the entry flow rates ranged from 24 to 1,509 vph, and total entry volume was 12,544 vph. Truck volumes ranged from 0 to 13 percent.

Minneapolis has many parking lots and structures that subtract vehicles from the traffic stream during the morning peak hour and add vehicles to the traffic stream during the afternoon peak hour. The total capacity of off-street parking within the study area is about 25,000 vehicles for 29 structures and 111 surface lots. This inflow-outflow activity was simulated by developing a sink model for the morning cases and a source model for the afternoon cases. This required identifying the location of each parking facility and the number of vehicles entering (or exiting) the lot from (or onto) each street during the morning (or afternoon) peak hour. The computer model subtracts (or adds) vehicles at the observed rate from each network link uniformly during the simulation period. The total morning parking lot inflow rate is 10,000 vph or 50 percent of the vehicles entering the network. In the afternoon, the parking lot outflow rate is 9,000 vph or 42 percent of the vehicles leaving the network.

Two minor streets that were not modeled geometrically as links were presented as sources and sinks in the streets they intersected. In particular, one dead-end street onto a major arterial provided a source of 486 vph to this artery. In this way, computer storage can be conserved.

Experiments were conducted at the field site to determine the free-flow speeds, driver reaction time (lost green time for the first driver in the queue), and queue discharge rates. These quantities can be specified individually for each link and are functions of local driving behavior, link geometry, and traffic demand. From the observations, free-flow speeds of 30 mph for the morning and 25 mph for the afternoon were designated for each street. The reaction time and queue discharge rates were both 2.3 sec for the morning period and were 2.2 and 2.5 sec for the afternoon period for each street. These data are averages for the desired cruising speed for the driver type and rates at which the queue at a red light dissipates during the green phase.

Field experiments were also conducted to determine the characteristics of lane-blocking events, such as passenger pickup or drop-off and truck loading. Short-term events are those lasting less than one traffic signal cycle length, here 90 sec. They can be specified for each network link. The mean duration of the short-term events was 15 sec, and the frequency of occurrence ranged from one event every 48 sec to one event every 900 sec, depending on location and time of day. The duration simulated is the mean duration multiplied by a factor ranging from 0.10 to 3.70 chosen randomly from a uniform decile distribution. Long-term events have specified durations, starting times, and locations, and ranged from 2 to 5 min. As vehicles traverse a blocked lane, delays are added to their trajectories as they seek and maneuver into gaps in an adjacent lane.

Control System

The traffic control system consists of fixed-time traffic signals at each intersection with a cycle length of 90 sec. Most of the signals have four phases with equal splits and 4-sec amber phases. During the peak hours, most offsets are zero. The timings of each phase and the offset for each signal were simulated as specified by the Minneapolis traffic engineer. There are two fixed-time midblock signals at pedestrian crossings. In addition, five intersections are modeled as nodes with no cross links because the traffic on the cross street could be modeled as sources and sinks and it was computer efficient to do so. There are 110 network traffic signals, and they have a maximum of seven phases, although the computer model can accommodate nine.

The control system also includes lane channelization, e.g., right-turn-only lane or bus-only lane. In addition, the turning movement may produce a de facto right- or left-turn lane, and these are specified as such even when the roadway is not so painted. Parking restrictions are taken into consideration by counting only travel lanes and adding right-turn pockets to model the downstream part of the curb lane at which parking is prohibited during the simulation period. In this way, the number of travel lanes is a function of the time of day and not simply of the road geometry.

Dual turning movements are common in the study area and were simulated. This refers to the practice of prescribing the curb lane to be right-turn-only (or left-turn-only) and vehicles in the adjacent lane may turn left (or right) or proceed straight. It is useful where heavy pedestrian volumes or turning movements occur.

Bus System

More than 50 bus routes traverse the study area during peak hours, and many routes have identical paths. The computer model now can accommodate at most 30 distinct bus routes. The Minneapolis bus system was therefore consolidated into 30 simulation bus routes, and headways were adjusted accordingly. This procedure affects only the statistics relating to bus routes and not those relating to bus activity on a street or at a bus station. The length of bus routes within the study area ranges from 3,972 to 9,610 ft, a total of 26 route-miles. The adjusted or simulation bus route headways ranged from 1.8 to 27.5 min. During peak hours, 350 buses enter the study area.

There are 99 bus stations in the study area, which is the maximum number the model can now accommodate. The term station does not imply an enclosure but merely differentiates from bus stops, which refer to the number of instances in which the bus came to rest, whether to serve passengers at stations, to join a queue, or to comply with a red traffic signal. Bus routes through the study area serve from 4 to 10 stations. A bus need not stop at every station on its route. The mean dwell time is specified for each station together with one of six decile distributions. Two of these distributions do not permit zero dwell times, but the other four do, thus simulating the skipping of a bus station with different probabilities. Experiments were conducted at the field site to measure dwell times. These data were reduced, mean dwell times were determined, and the appropriate distribution was selected. The afternoon mean dwell times ranged from 14 to 82 sec, more than three times longer than corresponding morning service durations.

In addition, the distance of each bus stop from the downstream intersection is needed to compute each bus trajectory. The capacity of each station, i.e., the number of buses that can service passengers simultaneously, is used to determine how long the station is overloaded and buses must wait in queue. This time period is part of the output as is the time during which each station is empty. Station capacities ranged from one to six buses. The computer model distinguishes between bus stations that permit the bus to pull out of traffic when serving passengers and those that do not and thereby produce lane blockages. This information is needed to ensure a realistic simulation and an accurate measure of the interference between buses and other vehicles.

RESULTS: THE OUTPUT

Performance Measures

The function of SCOT is to simulate traffic at a specific field site for a given traffic demand and control system. Trajectories for each vehicle (passenger automobile, truck, or bus) are computed, and, from these data, the numerical values of a set of performance measures are determined. These quantities are then displayed in the computer printout. Sixteen quantities are listed for each street in the network, averaged over the simulation period, e.g., 8:00-8:15 a.m. These quantities include average speed, average delay per vehicle, vehicle-miles, vehicle trips, average number of vehicles on the street, and average number of stops per vehicle. Similar quantities are listed for the network as a whole. Bus statistics are computed and listed separately. These statistics describe the bus activity on each street and at each bus station and include delay time, number of stops, and the time interval during which the capacity of each bus station was exceeded and during which it was empty. In addition, for each completed bus route, the total travel time, total dwell time, and average speed are listed.

Two important performance measures are worth special attention. The time, location, and duration of each instance of spillback are recorded. Spillback is the condition that occurs when a queue on a street is so long that it extends into the upstream intersection and blocks cross traffic. The number of cycle failures at each intersection is also listed. Cycle failure occurs when the green phase is too short to discharge the entire queue and some vehicles must sit out another red phase.

It is extremely useful to be able to plot each bus trajectory in order to identify the reason that each bus comes to a complete stop. These include scheduled station stops, midblock queues or lane blockages, and downstream red traffic signals. The information printed for each bus includes the time, location, duration, and reason for each stop.

With this extremely detailed printout, the city traffic engineer and transit operator have a bird's-eye view of what has happened on each street during the simulation period. They can thus identify bottlenecks and underutilized streets for each candidate control strategy. The candidate control strategy can then be improved, another simulation run made, and the performance observed. Through repetition, the control strategy can be adjusted until it merits field implementation.

Tabulation of Results

The bus priority strategies selected were confined to two adjacent streets, Marquette and Second Avenues. It was then found adequate and computer efficient to use smaller study networks.

The morning peak traffic demand was simulated for the 15-min period 8:00 to 8:15 a.m. The results, averaged over this period, are given in Table 1. The afternoon peak traffic demand was simulated for 4:30 to 4:45 p.m. The results, averaged over this period, are given in Table 2.

A typical afternoon bus trajectory was selected (Fig. 2).

Bus Route and Network Performance

Tables 1 and 2 give the changes in bus performance over completed routes. Each entry is an average of all buses completing a route during the 15-min period, and these range from 17 to 20 buses. Dwell times are not identical because dwell time distributions are stipulated at each bus station. The average dwell time per bus route includes all scheduled station stops. Average moving times can be compared by subtracting average dwell times from average travel times. From Tables 1 and 2 it is seen that providing a reserved bus lane in the direction of flow on the five middle blocks of these two 10-block streets for either the morning or afternoon peak traffic conditions results in no significant decrease in route travel time. Data given in the tables show that, when the two streets have 10 blocks of exclusive counterflow lanes, average moving time decreases 9.5 percent in the morning and 4.6 percent in the afternoon; this amounts to only 39 sec in the morning and 16 sec in the afternoon.

Tables 1 and 2 also show the changes in automobile and truck traffic performance for the entire network. Morning results are averaged for 2,840 vehicle trips and afternoon results for 3,330 vehicle trips completed during the 15-min simulation periods. In the morning, removing a lane from use by general traffic resulted in a 20 percent improvement in average delay per vehicle. The reason for the suprising development was that the direct-flow bus lane strategy restricts right-turning movements. The base case had two large right-turning movements that conflicted with heavy pedestrian movements. Eliminating them could only improve the traffic flow. This conclusion is reinforced by the elimination of 13 cycle failures also caused by these right turners.

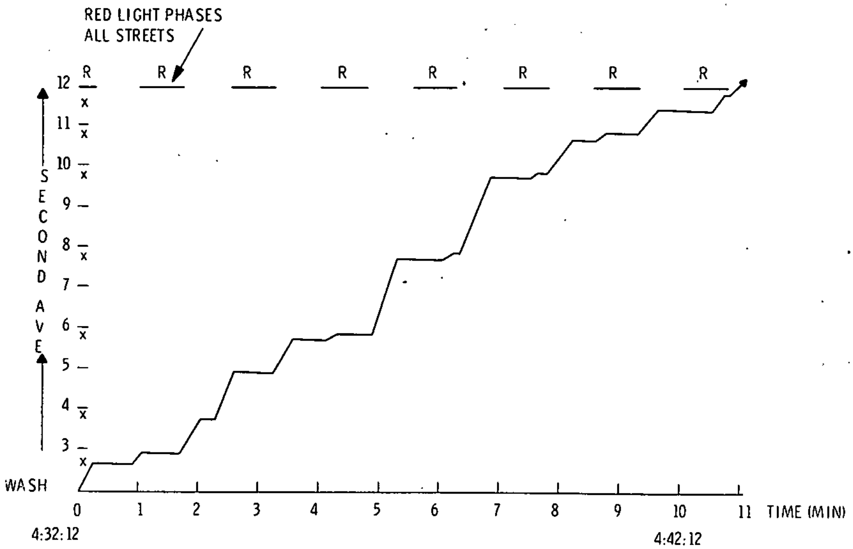
Table 1. Bus route and general traffic performance for 8:00 to 8:15 a.m. peak.

Traffic	Statistic	Base Case	Direct-Flow Bus Lane	Counterflow Bus Lane
Bus	Average travel time per bus route, minutes	8.14	8.00	7.29
	Average dwell time per bus route, minutes	1.26	1.16	1.06
	Average speed, mph	5.8	6.1	6.6
Automobile-truck	Average speed, mph	11.1	12.0	10.9
	Average delay per vehicle, seconds	44	35	46
	Average stops per vehicle	1.12	0.96	1.20
	Stopped delay per total delay	0.71	0.67	0.71
	Cycle failures at link location	7 at 48, 58 6 at 79, 69 1 at 58, 59	1 at 58, 59	3 at 79, 69 4 at 58, 59
	Spillback	None	None	None

Table 2. Bus route and general traffic performance for 4:30 to 4:45 p.m. peak.

Traffic	Statistic	Base Case	Direct-Flow Bus Lane	Counterflow Bus Lane
Bus	Average travel time per bus route, minutes	9.75	9.91	9.65
	Average dwell time per bus route, minutes	4.02	3.89	4.19
	Average speed, mph	4.9	4.8	5.2
Automobile-truck	Average speed, mph	11.1	9.8	9.7
	Average delay per vehicle, seconds	36	41	46
	Average stops per vehicle	1.16	1.13	1.29
	Stopped delay per total delay	0.73	0.77	0.76
	Cycle failures at link location	1 at 38, 39	2 at 48, 58 4 at 58, 68 2 at 89, 79 10 at 49, 39	3 at 48, 58 2 at 79, 69 5 at 69, 59 1 at 59, 49
	Spillback at link location, seconds	None	24 at 48, 58 10 at 49, 39	70 at 48, 58

Figure 2. Bus 142 trajectory simulation along Second Avenue from Washington Avenue to 12th Street at 4:32:12 to 4:42:53 p.m.



In the afternoon, the situation is different because of the increased general traffic volumes and pedestrian conflicts. Table 2 shows a 12 percent decrease in average speed and a 14 percent increase in average delay per vehicle for the direct-flow bus lanes and a 13 percent decrease and 28 percent increase for these quantities for the counterflow bus lanes. For both strategies, there is a marked increase in cycle failures and spillback. The afternoon volumes cannot afford to lose a lane to either strategy. The direct-flow bus lane strategy forbidding right turns included a provision for a left loop (one left turn followed by two right turns) to achieve equivalent vehicle destinations. These heavy left turns conflicted in the afternoon with heavy pedestrian movements, a condition not occurring in the morning. It seems clear that drivers will find better maneuvers than the left loop assumed here and the flow will not be so poor in the field.

Link Bus Performance

For direct-flow bus lanes within the priority region, right-turning vehicles have not proved to impede bus movements strongly enough to degrade trip times significantly. In the afternoon peak period, bus performance is again not consistently better for the direct-flow bus lane strategy either in the priority region or on the entire route.

For the counterflow bus lane strategy, improvements in bus performance are consistent: Bus speeds increased by as much as 26 percent (from 5.4 to 6.8 mph) for morning northbound buses. The inability of buses to pass each other in this mode does not appear to be significant at these headways.

Bus Station Activity

The time that a bus station is empty seems to decrease when station capacity is increased. For southbound buses, the average empty time is 6.5 min out of the 15 for the counterflow bus lanes where the entire block is a bus stop. This contrasts with 9.9 min for the base case and 9.2 min for the direct-flow bus lanes where two out of the seven bus stations are a full block long. The northbound bus stations are empty about 11 minutes for all three cases.

Bus Trajectories

We plotted bus trajectories to determine why the provision of reserved bus lanes did not decrease the average bus travel time significantly and consistently. Bus trajectories are computed in the model, so all information regarding the time, location, duration, and reason that each bus stopped was available and printed.

A typical bus trajectory is shown in Figure 2. The X's on the vertical axis denote the locations of the seven bus stations on Second Avenue. The horizontal bars at the top of the figure, marked with R's, represent the duration of the red signal phase. The cycle length is 90 sec and the red phase is 45 sec long. The signals turn red simultaneously at all streets from 3rd St. to 12th St.; there is no progression during the peak periods.

Bus 142 stopped at each of the seven bus stations for 13 to 52 sec of dwell time. It made seven stops for red lights for durations of from 5 to 38 sec. At no time was it impeded midblock by queues. Travel time for the 10-block route was 10 min 41 sec for an average speed of 4.4 mph. The total dwell time was 4 min, and total red light stopped time 2 min 40 sec. This leaves a moving time of 10:41 - 4:00 - 2:40 = 4 min 1 sec. In the model bus acceleration profile, the minimum moving time to touch base at each bus station when all lights are green is 3 min. This is for the acceleration-deceleration travel path.

Providing Bus 142 with its own lane would do nothing to decrease the time used in accelerating and decelerating at the seven stations and seven red lights. In this case slow bus speeds are caused by frequent service and red light stops and long dwell times, not by general traffic interference.

CONCLUSIONS

For the traffic conditions described in this study, including street geometry, traffic inflow rates, turning volumes, parking lot volumes, pedestrian movements, predictable lane blockages, traffic signal timings, bus routes, headways, and bus station dwell times, the provision of either direct-flow or counterflow bus-only lanes cannot significantly decrease the route travel time. The dominant factors in these low bus speeds are the frequent number of bus stations (seven in 10 blocks), the probability of encountering red traffic signals, and long dwell times. General traffic tends to avoid the right curb lane where buses travel, except right-turning vehicles and passenger drop-offs. Also, bus drivers are efficient in maneuvering around lane blockages. Traffic volumes would have to be considerably higher than those examined here before their effect on bus speeds could be determined and the provision of bus lanes would be expected to improve bus speeds significantly.

CURRENT PROJECTS

Plans are being made to validate this simulation study in Minneapolis. Experiments will be designed to compare before and after general traffic and bus performance at the field site with simulation predictions.

SCOT is being expanded to accept traffic demand described by origin-destination data as well as intersection turning movements. Vehicle movements are determined by computing current minimum time paths. Such a simulation will yield the time-optimal traffic pattern for a given control strategy. It will identify the ceiling or the best obtainable from a candidate control system. Once the necessary origin-destination data are collected, another simulation study will be made.

SCOT also has the ability to simulate bus preemption of traffic signals; i.e., under specified safety constraints, buses may extend the green phase of the downstream traffic signal long enough to permit them to pass through the intersection, or they may shorten the red phase and decrease their delay time. A study is planned to show the benefits to be expected from this bus priority strategy.

A study is under way to design fixed-time bus progression traffic signal schedules. These timings will be based on estimated time of arrival of buses at the intersection determined by mean station dwell times and distributions and the kinematics of bus trajectories. The design will be evaluated by simulation and, if the results are promising, tested at a field site.

ACKNOWLEDGMENTS

This study was supported by the Urban Mass Transportation Administration, U.S. Department of Transportation. The author is grateful for the assistance and encouragement given by R. J. Fisher, the sponsor's project manager.

The data necessary to model the study area were provided by D. Koski and D. Schroeder, Traffic Engineering Department, City of Minneapolis, and R. Neetzel, A. Issacs, and W. Marshall, Twin Cities Metropolitan Transit Commission. Their efforts and enthusiasm in working with the author are very much appreciated.

Major modifications to the computer model to accommodate the study area were made by P. Connolly, whose computer expertise was invaluable throughout the project. David Schmieder has made many extremely useful contributions to this study.

TSC is pleased to participate in the Work Study and Co-operative Student Plans to provide meaningful employment on a temporary and part-time basis for college students. The data reduction, network designs, computer operation and output analyses were performed by E. Drottar, T. Woo, J. Sinnot, M. Davis of Tufts University, M. Pembroke of Harvard University, B. Woo of Massachusetts Institute of Technology, P. Cleary of

Northeastern University, L. Hiller of Simmons College, N. Harris and A. Jones of Alabama Agriculture and Mechanical University, and N. Gilbert of Hampshire College. Their meticulous work is the foundation of this study.

REFERENCES

1. Muzyka, A. Urban/Freeway Traffic Control. IEEE Conference on Decision and Control, New Orleans, December 1972.
2. Lieberman, E. B. Simulation of Corridor Traffic: The SCOT Model. Highway Research Record 409, 1972.
3. Traffic Flow Simulation Study, SCOT Model. General Applied Science Laboratories, Inc., Westbury, N.Y., Technical Rept. 769, February 1972.
4. Lieberman, E. B., Worrall, R. D., and Bruggeman, J. M. Logical Design and Demonstration of the UTCS-1 Network Simulation Model. Highway Research Record 409, 1972.
5. Bruggeman J., Lieberman, E., and Worrall, R. Network Flow Simulation for Urban Traffic Control System. Federal Highway Administration, Office of Research, Rept. FH-11-7462-2, June 1971.
6. Lieberman, E. B. Dynamic Analysis of Freeway Corridor Traffic. Joint Transportation Engineering Conference, Chicago, October 11-14, 1970.

Intersection Capacity 1974: An Overview

Adolf D. May
Institute of Transportation and Traffic Engineering

This paper presents an overview of intersection capacity as of 1974. A brief background is given of the 1950 and 1965 U.S. Highway Capacity Manuals, and comparisons are made of the English and Australian manuals. The activities of the TRB Intersection Capacity Subcommittee are traced. These activities include the evaluation of the current HCM approach and investigation of alternative approaches. The Intersection Capacity Subcommittee identified and selected top-priority research needs. The results of the evaluation process and the 10 top-priority research topics are given. Each of the 10 topics is discussed. Comments are given and some initial analytical work is described.

Articles on the subject of intersection capacity date back at least to the early 1920s. In the late 1940s, several hundred approaches to signalized intersections were studied. The 1950 Highway Capacity Manual (1) included a chapter on signalized intersection capacity and was based primarily on the field studies of the 1940s. To attest to the international recognition of the 1950 Highway Capacity Manual, over 26,000 copies were distributed and it was translated into at least 10 languages.

In 1954, the highway capacity committee began revising and updating the 1950 manual. Some 1,600 intersection approaches were studied nationwide during 1955-1956. These data were analyzed graphically and by multiple regression techniques (2, 3). The 1965 Highway Capacity Manual (HCM) included a chapter on at-grade intersections and was based on this nationwide study as well as the results of individual research studies and the original manual. The HCM (4) is now in its sixth printing, and more than 27,000 copies have been distributed. It has been translated into French, German, Italian, and Spanish.

Capacity seminars have been held in every state and major city of the United States and in numerous foreign countries to present and discuss the philosophies and procedures described in the HCM. A bibliography of research published before the distribution of the HCM was prepared in connection with the capacity seminars in California (5). In addition to published reports, a number of research projects with primary emphasis on the subject of intersection capacity are currently under way (50-52).

Methods for determining the capacity of signalized intersections have also been developed in England and Australia. An excellent summary of the English method, developed by the Road Research Laboratory, was published in 1966 (6). The Australian method was developed by the Australia Road Research Board and was published in 1968 (7).

INTERSECTION CAPACITY SUBCOMMITTEE

In 1972, an Intersection Capacity Subcommittee was appointed by the TRB Committee on Highway Capacity and Quality of Flow. The subcommittee's main tasks were to review the HCM approach and to evaluate alternatives. In reviewing the current HCM approach, committee members and HCM users were requested to submit comments, criticisms, and suggestions about the chapter on at-grade intersection capacity. The responses were summarized in a 16-page document (8). In January 1974, the Intersection Capacity Workshop was held. The current HCM approach was reviewed, and alternative approaches, identified through literature searches, were evaluated (9).

IDENTIFICATION OF RESEARCH NEEDS

In April 1974, a questionnaire was distributed to all Highway Capacity Committee members and selected HCM users as an aid in identifying needed research on intersection capacity. Twenty-two areas of research were listed, and each individual was requested to rank the 10 that he felt had highest priority. A summary of questionnaire results is given in Table 1. The 10 topics receiving the highest priority are

1. Width of approach versus number of lanes,
2. Left-turning movements,
3. Load factor versus delay evaluation,
4. Overall urban arterial capacity,
5. Signal timing,
6. Special turn lanes and/or phases,
7. Total intersection evaluation,
8. Parking,
9. Pedestrians, and
10. Saturation flow studies.

Research problem statements are now being prepared and are to be included in a forthcoming special publication.

HIGH-PRIORITY RESEARCH TOPICS

Width of Approach Versus Number of Lanes

Both the 1950 and 1965 U.S. manuals use approach width rather than number of approach lanes. Although initially the 1950 manual was to have used number of lanes rather than approach width, trial analysis of the field measurements revealed that intersection capacity varied in almost a direct ratio with the width of the approach, and thus approach width was used. The 1965 manual reported that the width of the approach, rather than the number of traffic lanes, had proved to have the more significant bearing on the capacity of a typical approach.

The English method (6) developed by the Road Research Laboratory also proposed the use of approach width rather than number of lanes. The saturation flow rate was expressed in terms of the approach width by the following equation:

$$S = 160\ W$$

where

S = saturation flow in passenger cars per hour and
W = approach road width in feet.

The Australian method (7), however, found that saturation flows are related to the

Table 1. Priority ranking of capacity research topics.

Research Topic	Total Score	Priority Ranking
Approach width versus number of lanes	99	1
Parking	50	8
Load factor versus delay	84	3
Peak-hour factor	15	
Metropolitan area population	14	
Location within metropolitan area	16	
Right-turning movements	18	
Left-turning movements	98	2
Special turn lanes and/or phases	60	6
Trucks and through buses	2	
Local transit buses	13	
Signal timing	78	5
Marking of approach lanes	22	
Unsignalized intersections	25	
Pedestrians	49	9
Total intersection evaluation	51	7
Urban arterial capacity	79	4
Ambient conditions	8	
Geographical regional factors	6	
Grades	12	
Saturation flow studies	41	10
Nationwide data collection	9	

Figure 1. Comparison of English and Australian methods (approach width versus number of lanes).

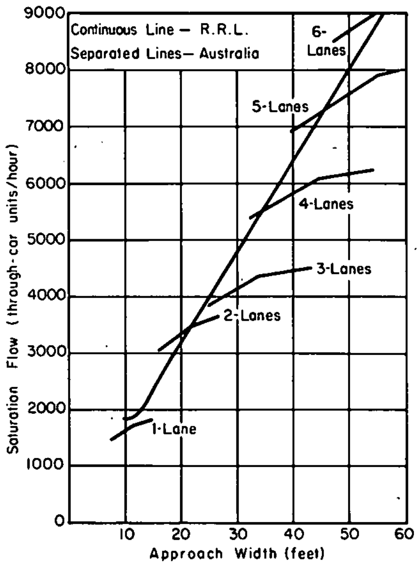


Figure 2. Effect of left turns on capacity at two-lane signalized approach.

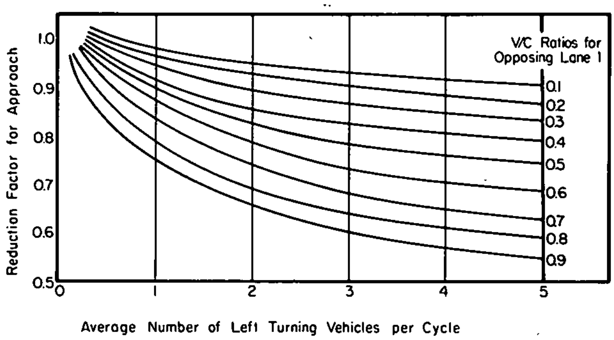


Figure 3. Relationship between average individual delay and load factor.

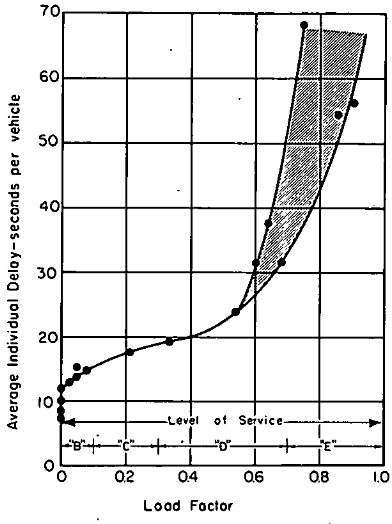
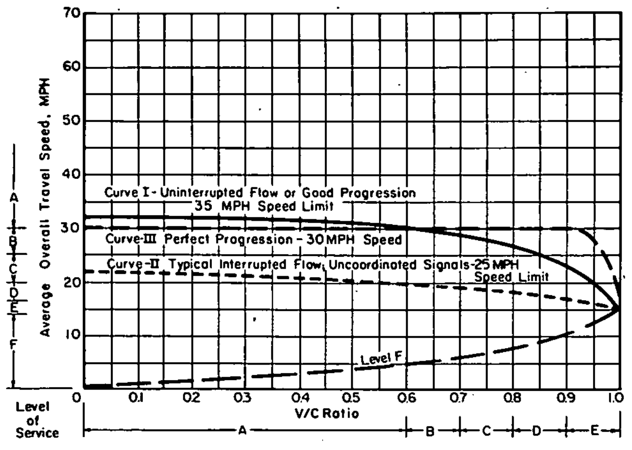


Figure 4. Typical relationship between volume-capacity ratio and average overall travel speed.



number of lanes, not to the approach width. It was determined that lane width had very little effect on saturation flow rate or capacity over a range of 10 to 13 ft. Figure 1 shows a comparison of the English and Australian methods and illustrates the difference between the approach width and number of lanes procedures. If the number of lanes rather than approach width could be shown through use of U.S. data to be a better method (or as good) for determining capacities, the U.S. method could be simplified.

Left-Turning Movements

A study at Northwestern University (48) demonstrated with field measurements that the HCM does not reflect the effects of different levels of opposing traffic volume for the basic case of no separate left turn lanes or signals. This study also identified the need for further study of the effect of number of moving lanes, presence of a left turn storage lane, prohibition of parking and standing in the approach, cycle length, actuated versus pretimed control, width of exit roadways, and turning radius. Figure 2 shows the overall effects of left turn vehicles on capacity as determined in the Northwestern University study.

Load Factor Versus Delay Evaluation

The introduction of the level-of-service concept is one of the important contributions of the HCM. Load factor is used as the measure of level of service at signalized intersections. Although it is relatively easy to measure in the field, some users have proposed that a measurement of delay be used to represent level of service. The results of a study at the University of California (25), however, raised questions about the validity of an assumed direct relationship between load factor and delay. Figure 3 shows the results obtained in a simulation study of the relationship between average individual delay and load factor.

Urban Arterial Capacity

Chapter 10 of the HCM provides only approximate means for determining overall urban arterial capacity and levels of service. In fact, the 1965 manual suggests intersection by intersection analysis supplemented by judgment. There is obviously a need for further research because of the significant amounts of travel on major arterials and because most individual trips pass through more than one signalized intersection.

The recognition of the importance of research in this area is undoubtedly related to the recognition of the need for further research into the influence of signal timing (priority 5) and total intersection approach to capacity (priority 7). Figure 4 shows the relationship of average overall travel speed to volume-capacity ratio in one direction of travel on arterial streets as given in the 1965 U.S. manual.

Signal Timing

In the study of an existing signalized intersection, one variable that the traffic engineer can control directly and that has a significant effect on capacity and level of service is signal timing (i.e., phasing, cycle length, and offsets). The relationships of offsets to overall urban arterial capacity (priority 4) and of signal phasing to total intersection approach capacity (priority 7) have been established.

The English method uses an effective green time in computing the G/C ratio rather than the green phase length. Figure 5 shows the relationship between effective green time and the green phase length and its importance when saturation flow (priority 10) is used.

Special Turn Lanes and/or Phases

At multiapproach intersections and at intersections with a relatively high percentage of turning movements, special turn lanes or phases can improve the level of service. Although the HCM describes methods for analyzing special turn lanes and/or phases, these methods are only approximate and are based on limited field data. More attention needs to be given to vehicle arrival distributions, number of turning lanes, loss time between phases, turn curvature, and associated user delay and capacity. Procedures are needed to determine when multiphase signalization is warranted. This research is also related to the effect of left-turning movements (priority 2). Figure 6 shows a proposed structure of special turn lanes and/or phases.

Total Intersection Evaluation

All existing methods for calculating intersection capacity are accomplished on a single approach basis. The G/C ratio used for this single approach affects the available G/C ratio for the opposite and crossing approaches. The HCM suggests that each approach be analyzed separately; the optimum solution, as the G/C ratio, is obtained manually through trial and error procedures. A procedure should be developed for determining the effects of intersection capacity and levels of service on a total intersection basis, which would result in the optimum signal settings. This research topic is closely related to influence of signal timing (priority 5).

The method developed by the Road Research Laboratory (6) proposes a procedure for determining optimum cycle length and green times. The cycle length that results in the minimum delay is obtained by differentiating the overall delay equation:

$$C_o = \frac{1.5L + 5}{1 - Y}$$

where

- L = total lost time per cycle in sec,
- Y = summation for the whole intersection of the y values,
- y = maximum ratio of flow to saturation flow for a given phase, and
- C_o = optimum cycle length in sec.

The green times that minimize delay were derived from the overall delay equation:

$$\frac{g_1}{g_2} = \frac{y_1}{y_2}$$

where

- g_1, g_2 = effective green times of phases 1 and 2 in sec, and
- y_1, y_2 = maximum ratio of flow to saturation flow of phases 1 and 2.

There are two limitations to this approach. First, only the most critical approach is considered in each phase. Second, if the selected cycle length is significantly different from the optimum cycle length (i.e., in the case of pedestrians crossing wide streets), the calculation of green times may not result in minimum total intersection delay. Figure 7 shows the effect of cycle length and green times on total intersection delay.

Parking

The HCM analyzes the influence of parking for only two conditions: parking and no parking (i.e., no parking within 250 ft of the intersection). The English method handles the influence of parking as a function of the distance between the stop line and the nearest parked car (Fig. 8). The effective loss of approach width is given by

Figure 5. Relationship between effective green time and green phase length.

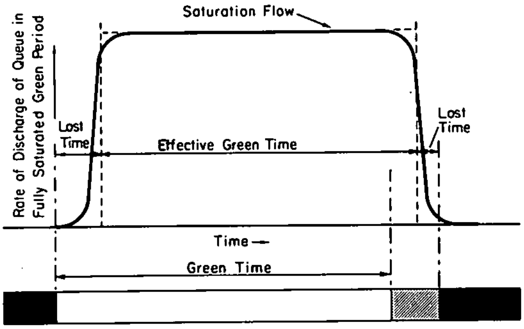


Figure 6. Structure of special turn lanes and/or phases.

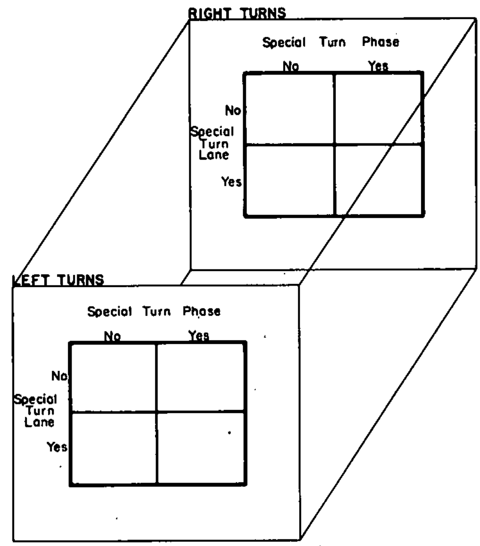


Figure 7. Delay for total intersection approach.

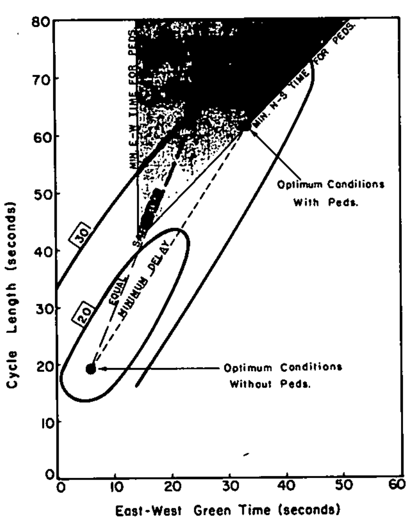
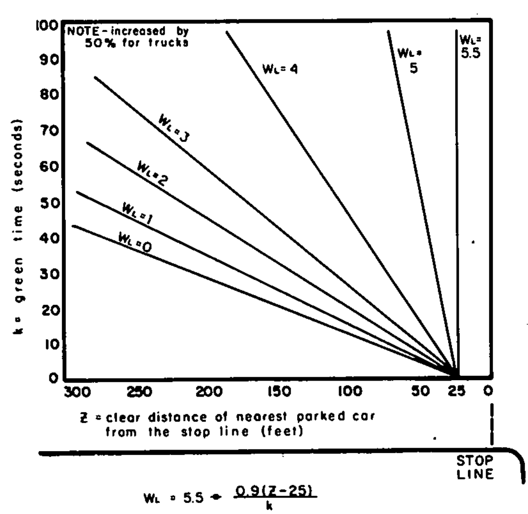


Figure 8. Effective loss in approach width due to parking.



$$W_L = 5.5 - \frac{0.4(z - 25)}{k}$$

where

Z = clear distance of the nearest parked car from the stop line in feet (if Z is less than 25 ft, let Z = 25),
k = green time in sec, and
$W_L \geq 0$ = effective loss in approach width in feet.

The English method also proposes that the effective loss be increased by 50 percent if the parked vehicle is a truck.

Other concerns about the 1965 manual include the inconsistencies between the parking and no-parking approach volume-approach width relationships, the lack of handling angle parking, and the lack of attention to the effect of parking maneuvers on traffic delay.

Pedestrians

The 1965 manual only makes mention of pedestrian influence, and the effect of pedestrians is only indirectly handled by the use of the location in metropolitan area factor. This area of research is related to signal timing (priority 5) in determining green times, cycle lengths, and special phasing required for exclusive pedestrian and/or left turn movements. In addition, no attention has been given to the level of service provided to pedestrians and to the combining of pedestrian and vehicular level of service.

Saturation Flow Studies

The English and Australian methods use saturation flow in determining intersection capacity. Saturation flow is defined as the flow that would be obtained if there were a continuous queue of vehicles and they were given 100 percent green time. It is generally expressed in vehicles per hour of green time. The capacity is obtained by multiplying the saturation flow by the G/C ratio (note G is effective green time, not necessarily green time) and by other factors (i.e., turning, grades). The English equation is

$$CAP = f[160(W - PVF)], (RT), (LT), (VCF), (G/C), (PF), (GF), (SF)$$

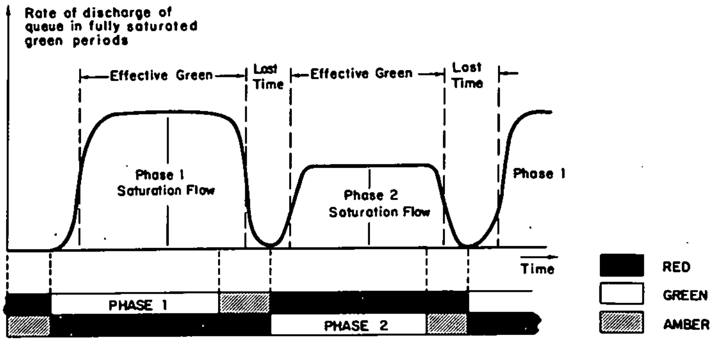
where

CAP = capacity in vph,
W = approach width in feet,
PVF = adjustment factor based on parked vehicle clearance distance,
RT = adjustment factor based on percentage of right turns,
LT = adjustment factor based on percentage of left turns,
VCF = adjustment factor based on vehicle composition,
G/C = effective green time to cycle time ratio,
PF = adjustment for off-peak period,
GF = adjustment factor based on percentage of grade, and
SF = adjustment factor for site characteristics.

The HCM uses a similar procedure except the term is referred to as approach volume and is hidden empirically by combining it with other influencing factors. The similarity between the two methods can be seen by comparing the previous English equation with the equation below.

$$SV = f(AV_{W,LF}), (POP, PHF), (LIM), (RT), (LT), (TF), (BF), (G/C)$$

Figure 9. Saturation flow concept.



where

SV = service volume in vph,
$AV_{W,LF}$ = approach volume based on approach width and load factor in vehicles per hour of green,
POP, PHF = adjustment factor based on metropolitan area size and peak-hour factor,
LIM = adjustment factor based on location in metropolitan area,
RT = adjustment factor based on percentage of right turns,
LT = adjustment factor based on percentage of left turns,
TF = adjustment factor based on percentage of trucks and through buses,
BF = adjustment factor based on local buses and bus stop type, and
G/C = green phase time to cycle time ratio.

It is argued that the saturation flow approach is more fundamental and represents the base condition. Capacity influencing factors are then empirically determined, and modify the saturation flow value. This approach to signalized intersections is somewhat similar to the HCM approach to freeways (i.e., the capacity of a freeway under ideal roadway and traffic conditions is 2,000 passenger vehicles per hour per lane and then influencing factors are applied).

Research into the saturation flow approach may not only improve the accuracy of capacity calculations but also provide the mechanism for better understanding of the phenomena of intersection capacity (Fig. 9).

SUMMARY

This paper has provided an overview of intersection capacity and proposes the direction for future research. The capacity committee welcomes your comments, suggestions, and criticisms. Furthering knowledge on capacity depends on the research that we undertake now.

REFERENCES

1. Highway Capacity Manual. Bureau of Public Roads, 1950.
2. Final Report on Intersection Traffic Flow. CEIR, Inc., 1960.
3. Normann, O. K. Variations in Flow at Intersections as Related to Size of City, Type of Facility, and Capacity Utilization. HRB Bulletin 352, 1962, pp. 55-59.
4. Highway Capacity Manual—1965. HRB Special Report 87, 1965.
5. Selected References on Highway Capacity. Institute of Transportation and Traffic

Engineering Library, University of California, Berkeley, Library Reference 31, July 1966.

6. Webster, F. V., and Cobbe, B. M. Traffic Signals. Road Research Laboratory, Technical Paper 56, 1966.
7. Australian Road Capacity Guide: Provisional Introduction and Signalized Intersections. Australian Road Research Board, Bulletin 4, 1968.
8. Summary of Comments on Chapter 6: At-Grade Intersections. HRB Intersection Capacity Subcommittee, July 1973.
9. Notes From the Intersection Capacity Workshop. Transportation Research Circular 161, Dec. 1974.
10. Argue, C. T. Computer Program for Intersection Capacity. Cornell Univ., 1966.
11. Savage, R. H. Capacities of Some Signalized Intersections in the City of Chicago. Northwestern Univ., 1966.
12. Intersection Capacity. Dallas-Fort Worth Regional Transportation Study, Texas Highway Dept., 1967.
13. Leisch, J. E. Capacity Analysis Techniques for Design of Signalized Intersections. Public Roads, Vol. 34, No. 9, Aug. 1967; No. 10, Oct. 1967.
14. May, A. D., Jr., et al. A Computer Program for Intersection Capacity. Traffic Engineering, Vol. 38, No. 4, Jan. 1968, pp. 42-48.
15. New Method of Calculating Intersection Capacity Is Developed in West Germany. Highway Research News, No. 26, Winter 1967, pp. 42-43.
16. Taylor, S. S. Application of a Computer Program for Intersection Capacity. Los Angeles Traffic Engineering Dept., 1967.
17. Bendtsen, P. H. Capacity Analysis of Signal-Controlled Intersections in Denmark. Presented at HRB 47th Annual Meeting, Jan. 1968.
18. Blunden, W. On the Capacity of Signalized Intersections. HRB 47th Annual Meeting, 1968.
19. Leisch, J. E. Capacity Analysis Procedures for Signalized Intersections. Traffic Institute, Northwestern Univ., 1968.
20. May, A. D., Jr. Capacity at Signalized Intersections—International Technology and Concepts. Presented at HRB 47th Annual Meeting, 1968.
21. May, A. D., Jr., and Hom, R. E. Intersection Capacity of Exclusive Turning Lanes: A Computer Program and Nomographs. Traffic Engineering, Vol. 38, No. 6, March 1968, pp. 48-53.
22. Schlums, J. R. A Contribution Towards Determining Capacities of Signalized Urban Intersections. Presented at HRB 47th Annual Meeting, 1968.
23. Webster, F. V. Traffic Signal Capacity. Presented at HRB 47th Annual Meeting, 1968.
24. Jones, R. A. Computer Simulation in Conflicting Vehicle Storage Problems. Traffic Engineering, Vol. 39, No. 5, Feb. 1969, pp. 42-44.
25. May, A. D., Jr., and Gyamfi, P. Extension and Preliminary Validation of a Simulation of Load Factor at Signalized Intersections. Traffic Engineering, Vol. 40, No. 1, Oct. 1969, pp. 46-52.
26. Intersection Capacity Study. Missouri State Highway Commission, Study 69-2, 1969.
27. Sagi, G. S., and Campbell, L. R. Vehicle Delay at Signalized Intersections. Traffic Engineering, Vol. 39, No. 5, Feb. 1969, pp. 32-40.
28. Woods, D. L. Limitations of Phase Overlap Signalization for Two-Level Diamond Interchanges. Traffic Engineering, Vol. 39, No. 12, Sept. 1969, pp. 38-41.
29. Kell, J. H., and Johnson, B. C. Optimizing Street Operations Through Traffic Regulations and Control. NCHRP Report 110, 1970.
30. Lehoczky, J. P. Stochastic Models in Traffic Flow Theory: Intersection Control. Highway Research Record 334, 1970, pp. 1-7.
31. McShane, W. R., Yagoda, H. N., Pignataro, L. J., and Crowley, K. W. Control Considerations and Smooth Flow in Vehicular Traffic Nets. Highway Research Record 334, 1970, pp. 8-22.

32. Munjal, P. K., and Hsu, Yuan-Shih. Development and Evaluation of Real-Time Control Algorithms for Diamond Interchange Signalization. System Development Corp., No. TM-4601/003/00, 1970.
33. Nemeth, Z. A., and Vecellio, R. L. Investigation of the Dynamics of Platoon Dispersion. Highway Research Record 334, 1970, pp. 23-33.
34. Reilly, E. F., and Seifert, J. Capacity of Signalized Intersections. Highway Research Record 321, 1970, pp. 1-15.
35. Surti, V. H. Operational Efficiency Evaluation of Selected At-Grade Intersections. Highway Research Record 321, 1970, pp. 60-73.
36. Tidwell, J. E., Jr., and Humphreys, J. B. Relation of Signalized Intersection Level of Service to Failure Rate and Average Individual Delay. Highway Research Record 321, 1970, pp. 16-32.
37. McInerney, H. B., and Petersen, S. G. Intersection Capacity Measurement Through Critical Movement Summations: A Planning Tool. Traffic Engineering, Vol. 41, No. 4, Jan. 1971, pp. 45-51.
38. Munjal, P. K., and Fitzgerald, J. W. A Simulation Model for the Evaluation of Real-Time Computer Control of Diamond Interchanges. System Development Corp., No. TM-4601/004/00, 1971.
39. Pontier, W. E., et al. Optimizing Flow on Existing Street Networks. NCHRP Report 113, 1971.
40. Alsop, R. E. Estimating the Traffic Capacity of a Signalized Road Junction. Transportation Research, Vol. 6, No. 3, Sept. 1972, pp. 245-255.
41. Behnam, J. Gap Acceptance as a Criterion for Left Turn Phasing. Traffic Engineering, Vol. 42, No. 9, June 1972, pp. 40-42.
42. Brand, D. Urban Traffic Control: How Far Can We Take It? Traffic Engineering, Vol. 42, Aug. 1972, pp. 50-55.
43. Sagi, G. S. Signalized Intersection Capacity Parameters. Traffic Engineering, Vol. 42, June 1972, pp. 50-55.
44. Berry, D. S., and Gandhi, P. K. Headway Approach to Intersection Capacity. Highway Research Record 453, 1973, pp. 56-60.
45. Johnson, R. R., and Matthias, J. S. Investigation to Determine Capacity of Protected Left-Turn Movements. Highway Research Record 453, 1973, pp. 49-55.
46. Sofokidis, H., Tilles, D. L., and Geiger, D. R. Evaluation of Intersection-Delay Measurement Techniques. Highway Research Record 453, 1973, pp. 28-48.
47. Crommelin, R. W. Employing Intersection Capacity Utilization Values to Estimate Overall Level of Service. Traffic Engineering, Vol. 44, No. 10, July 1974, pp. 11-14.
48. Fellinghauer, E., and Berry, D. S. Effects of Opposing Flow on Left-Turn Reduction Factors at Two-Phase Signalized Intersections. Transportation Research Record 489, 1974, pp. 13-18.
49. Yagar, S. Capacity of a Signalized Road Junction: Critique and Extensions. Transportation Research, Vol. 8, No. 2, April 1974, pp. 137-148.
50. Lee, B. Control of Oversaturated Intersections. Polytechnic Institute of Brooklyn, New York.
51. Lee, C. L. Simulation of Traffic by a Step-Through Technique. Center for Highway Research, University of Texas at Austin.
52. Sword, E. C. Computer Program for the Computation of Signal Settings From the Computed Capacity and Estimated Flows at an Intersection. Joint Highway Research Project, Purdue University.

Better Use of Signals Under Oversaturated Flows

Bumjung Lee, Kenneth W. Crowley, and Louis J. Pignataro
Polytechnic Institute of Brooklyn

A measure of oversaturation is developed and validated with field data. A control policy for oversaturated intersections is developed. The results of this work have applicability wherever intersection oversaturation is frequently experienced. The primary objective of this control policy is to delay or eliminate intersection blockage, which is the outgrowth of oversaturation. The control policy developed in this paper (queue-actuated control) is of a tactical nature. It responds to queues (maximum queues allowed at both approaches of an intersection) similar to the way in which the current traffic-actuated control responds to demand. The new control policy can be applied to isolated intersections as well as to coordinated ones. The specific aspects addressed are field validation of the new measure of oversaturation, development of a delay model for queue-actuated control, and effects of queue-actuated control on adjacent intersections. Validation of the delay model and evaluation of queue-actuated control in alleviating the problems of oversaturation are performed with the UTCS-1 simulator.

Congestion on urban street networks has become a familiar occurrence in central business districts across the country and abroad. Urban street grids are commonly used for a variety of often conflicting purposes. They serve through traffic and circulate traffic with destinations on or adjacent to the street grid. Land access is provided to abutting properties. Pedestrians as well as vehicles are serviced. Conflicting vehicular and pedestrian movements are interfaced at grade, with little or no separation.

This system of conflicting movements, modes, and purposes is controlled by a variety of common traffic control devices and techniques, principally signals. The emphasis of current traffic control schemes is directed toward light traffic flows in which flexible coordination of signals will enable traffic platoons to progress along green waves and toward moderate congestion in which the capacity of bottlenecks will be maximized.

When volumes are excessively high, most of the present concepts of optimization become ineffective or invalid (1). As traffic flows increase, the system or points within the system are subject to breakdown or congestion. Saturation begins principally at intersections, the points of maximum conflict and, if unchecked, spreads to other parts of the system.

In saturated conditions, congestion is unavoidable, and therefore the control policy should be aimed at postponing the onset or minimizing the severity of secondary congestion, which is caused by the blockage of intersections that are not the originators of congestion.

This paper covers two major areas of discussion (2, 3) related to signal control of oversaturated flows:

1. Definition of oversaturation and development of measures of oversaturation and
2. Development of a signal control concept better suited to oversaturated flows.

OVERSATURATION

Necessary Definitions

In addition to standard traffic engineering terminologies, the following definitions are used in this paper.

Critical intersection

A critical intersection (CI) is one whose capacity is the limiting value of the capacity of a segment of roadway or an entire system. Thus, there may be more than one critical intersection in a network, each having a different value of capacity. As traffic flow increases through a network, congestion first develops at critical intersections and then spreads to other intersections, depending on the duration and magnitude of the increased flow and the physical characteristics of the roadway.

Tactical control measure

A tactical control measure is one that is responsive to traffic and is usually applied at one or a few critical intersections.

Strategic control measure

A strategic control measure is a control measure that is applied systemwide. It may or may not be responsive to traffic.

Characterization of Oversaturation

Oversaturation occurs when queues fill a significant portion of a block and interfere with the performance of an adjacent upstream intersection. Oversaturation may be experienced at single points, along major arterials, or throughout entire subsections of a network. On a system operating in an oversaturated mode, the problem appears to be areawide. However, the onset of the oversaturated state occurs at one or perhaps a few critical locations at which congestion first begins to develop. Thus, the problem of oversaturation begins as a localized problem.

Definitions of traffic performance states and the characterization of oversaturation were developed as part of the research under an NCHRP project (2).

Definitions of Traffic Performance States

It is not unusual to find the terms congested, saturated, and oversaturated all applied to the same situation, i.e., where one or more vehicles remained to be served at the end of the green phase. A more exact and nonredundant set of terms needed to be developed to characterize the various conditions of traffic performance that could occur. No new terms were envisioned, only a more specific definition of the existing ones. Thus the terms uncongested, congested, saturated, and oversaturated are used throughout this paper, but they have been defined somewhat differently as shown in Figure 1.

Specifically, the definitions have been constructed around a queue formation mech-

Figure 1. Traffic performance states.

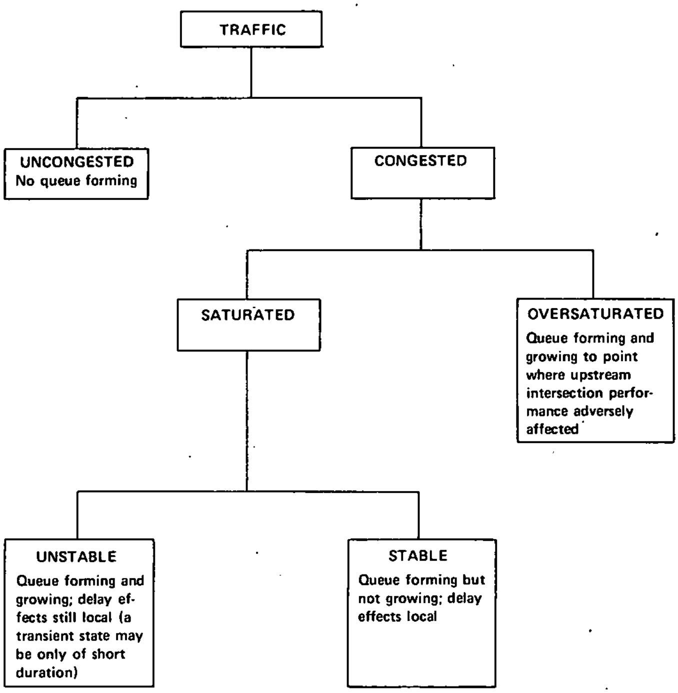
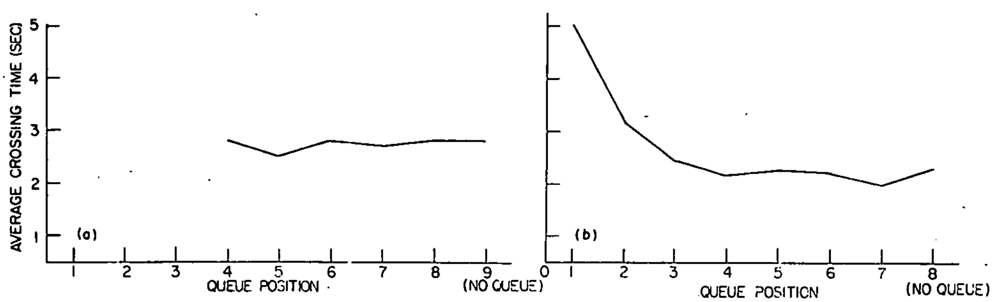


Table 1. Measures of traffic performance.

Measures	Descriptors	Descriptors and/or Control Parameters
Intersection	Load factor, saturation factor, maximum individual delay per vehicle, number of cycles to clear intersection, number of stops and starts to clear intersection	Queue length, total delay, average delay, link length minus queue length, ratio of queue length to link length, input-output, trapped vehicles, v/c ratio
System	Total delay, average number of stops per vehicle, density, mean velocity gradient, occupancy, total travel time	Density, occupancy, input-output

Figure 2. Average crossing time for entire approach: (a) site 1 and (b) site 2.



anism. The traffic performance definitions are all described in terms of one approach to a signal. As such they refer to the capacity of one signal's green time for the approach under consideration.

Measures of Oversaturation

Numerous measures of effectiveness (MOEs) are currently used to characterize traffic performance. Some measures describe traffic conditions at points in the system, whereas others describe overall conditions of the system or subsystem. Some of the measures only describe the state of traffic performance, and some have utility as control parameters.

Thus, before any attempt is made to develop a control concept to alleviate the problem of oversaturation, adequate MOEs must be determined. Proper MOEs must describe the conditions of oversaturation and have utility as control parameters if they are to be of any value. They must also be able to predict the onset of oversaturation: the transition from the unstable state to oversaturation. This is where the opportunities for delaying or preventing oversaturation lie.

A number of possible MOEs have been reviewed and evaluated (2) and are given in Table 1.

Point measures

The following measures were selected as the most promising: queue length, ratio of queue length to link length, and link length minus queue length. Each of these point measures of saturation can describe the defined traffic performance states and predict the onset of each state. The latter two measures, those explicitly incorporating link length, are considered to be desirable because they are more general forms.

A limited field study was conducted to determine whether queue-related measures adequately meet evaluation criteria in reality and which of these three point measures best describes the various levels of saturation. The field study and its results are discussed later.

System measures

Evaluation of several candidate system measures revealed that none of them adequately describes the conditions of saturated and oversaturated systems. An alternate measure has been proposed (2).

Data Check of Selected Measures

The primary purpose of a data check of the selected measures is to determine their adequacy in describing or modeling the properties of oversaturation, as well as the onset of oversaturation in reality. Based on an analysis of data collected at two pairs of intersections in Brooklyn, New York, the following conclusions were drawn.

1. The average crossing time (speed) of vehicles leaving the upstream intersection is significantly affected by the queue at the downstream critical intersection. When the clear distance to the end of the queues at the downstream intersection—one of the selected measurements of oversaturation—is less than approximately 230 ft, the quality of operation (expressed in speed) at the upstream intersection starts deteriorating. Data given in Figures 2 and 3 show that the effects are similar for vehicles that did and did not stop at the beginning of green. Queue positions 4 and 3 for sites 1 and 2 respectively both represent a clear distance of approximately 230 ft between the upstream intersections and the end of the queues at the downstream intersections. The clear distance of 230 ft is defined as the critical distance, and a downstream intersection is oversaturated if the distance between the end of the queue and

the upstream intersection becomes less than the critical distance. (The data used here were collected from two pairs of intersections with block lengths of 530 and 810 ft. Further field validation is desirable for blocks shorter than 500 ft.)

2. Although the quality of operation is reduced, the productivity of the upstream intersection is not affected by the queue at the downstream intersection until spillback occurs. The output will, of course, be reduced when spillback occurs. As shown in Figure 4, there are no discernible trends between the average headway and the queue length, for either stopped vehicles or not stopped vehicles. There is no statistically significant difference between the average headways. Because time headway equals space headway divided by speed and because time headways remain the same, space headway decreases as speed decreases. This implies that, as the queue reaches the critical length, the speed and space headways decrease so that the output remains the same. The decrease of space headways results in a more compressed traffic stream and a higher density.

3. Regardless of the queue length at the downstream intersection, the status of the downstream signal indication affects the quality of operation (speed) at the upstream intersection. However, it will not affect the selected MOE (Fig. 5).

4. Several other analyses were made to determine effects of the queue length at the downstream intersection, but are not presented here. These include start-up times, left turn movements, and others.

EVALUATION OF EXISTING CONTROL MEASURES

Developing a control technique better suited to the problem of oversaturation required that the effectiveness of existing control measures be thoroughly evaluated. This evaluation, therefore, included existing control measures that are well-known to practicing traffic engineers as well as advanced concepts that have not been fully tested in the field.

Control Measures Evaluated

In essence, in any form of traffic signal control only three parameters may be manipulated by the traffic engineer: cycle length, allocation of green to competing movements (split), and timing interrelationships between adjacent signals in a system (offset). Therefore, the control parameters evaluated can be classified as follows:

1. Cycle length (strategic) includes short cycle length versus long cycle length and a common cycle length for a system.
2. Split includes strategic control measures, such as uniform bandwidth and gradual change of bandwidth, and tactical control measures, such as Gazis' saturated intersection policy (4, 5) and Longley's control strategy for congested computer-controlled networks (6).
3. Offset (strategic) includes smooth flow theory (7-9), red progression (10), progression versus simultaneous, reverse progression, and combination of progression and simultaneous.

Evaluation Techniques

Inasmuch as the measure of oversaturation is a function of queue length, the primary evaluation criterion of the performance of a given control measure under oversaturated flows should be its ability to limit the queue length to some desired level. Average delay per vehicle is used as a secondary evaluation criterion.

Various existing control measures were tested on fictitious networks as well as at the sites used for MOE evaluation. Extensive simulation was done under oversaturated flow conditions with the UTCS-1 simulator.

Figure 3. Average crossing time for entire approach for vehicles not stopped at beginning of green: (a) site 1 and (b) site 2.

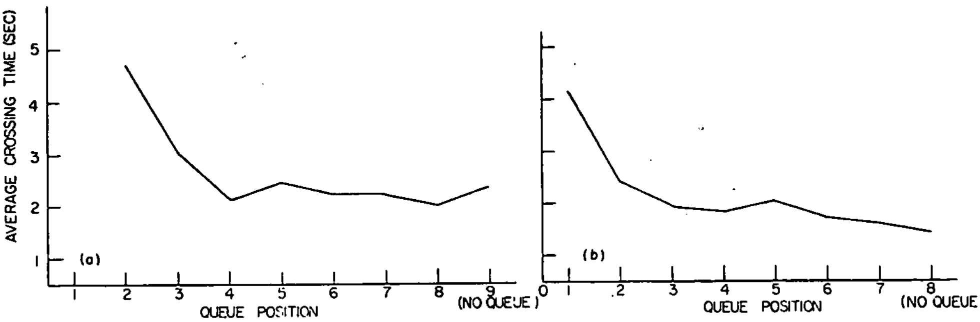


Figure 4. Headway versus queue position at (a) site 1 and (b) site 2.

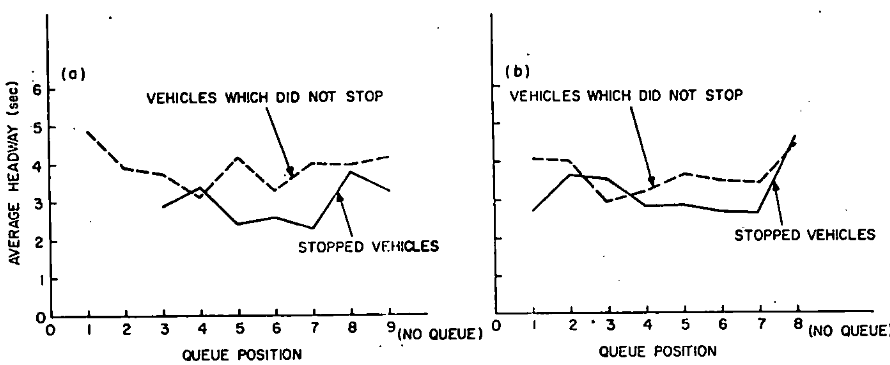
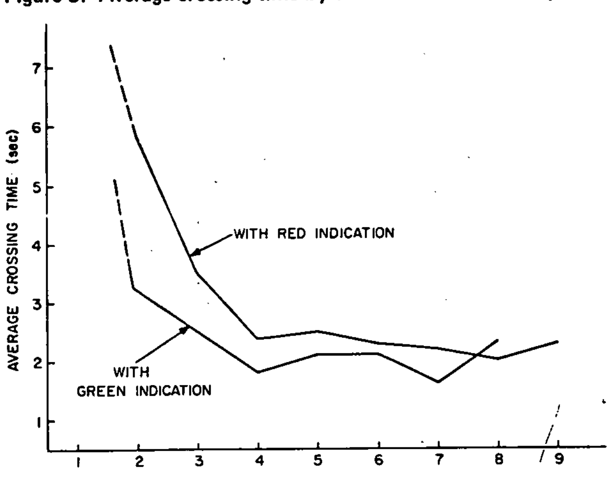


Figure 5. Average crossing time by downstream indication, site 1.



Results

Based on the results of simulation analysis of the current control measures, the following conclusions are made.

1. Some of the strategic control measures evaluated are very effective in delaying or eliminating oversaturation at critical intersections. They, however, tend to degrade the performance of adjacent intersections in the system. For example, a simultaneous offset plan significantly reduces the queue length at an oversaturated CI, but at the same time increases queue length and delay at other intersections within the system.
2. Strategic control measures generally reduce queues at the CI more than necessary to prevent oversaturation and reduce the productivity of the CI. Control measures such as smooth flow theory and reverse progression usually reduce queue length more than enough to keep a CI from being oversaturated or the upstream intersection from being blocked by extended queues. The productivity of a CI is significantly reduced by a strategic control measure.
3. Tactical control measures do not act positively to prevent the blockage of intersections, which is the prime concern during the period of oversaturation. Under oversaturated flows, a control measure that is more sensitive to intersection blockage is required because of the undesirable effects of such a blockage.
4. A new control concept is desired that not only ensures the prevention of intersection blockage as much as possible but also minimizes the degrading of other intersections in the system. This new control concept should be capable of maintaining high productivity at a critical intersection. Productivity is more important than the quality of operation during a period of oversaturation.

A NEW CONTROL CONCEPT

In oversaturated conditions, congestion at a critical intersection is unavoidable; therefore, the control policy should be aimed at minimizing the effects of queues at the CI on the upstream intersection. As volumes increase further, oversaturation of a CI may also become unavoidable. Then the control policy should be aimed at postponing the onset of oversaturation or preventing the blockage of the upstream intersection. Prevention of intersection blockage becomes more important with arterials or networks where traffic flow is heavy in all directions. In such cases, intersection blockage not only increases the delay through the system but also affects more segments of that system.

Queue-actuated signal control was developed to ensure, as much as possible, the prevention of intersection blockage. It also makes use of all available green time by creating continuous demand (queues) at all approaches. It allows more positive control of queue length at a given intersection.

It is a control policy in which an approach receives green automatically when the queue on that approach becomes equal to or greater than some predetermined length.

When the queue at one approach becomes equal to or greater than a given length (Q max), that approach receives green regardless of the conditions on the conflicting approaches, assuming that no other approach reaches Q max simultaneously. Thus, how much green one approach receives during a given time interval depends on the link length and the number of lanes, as well as the flow rates on both approaches.

It is obvious that, when one approach has a relatively low flow rate, drivers on that approach suffer long delays because the queue on the approach takes a longer time to reach the predetermined length. When volumes become extremely heavy on both approaches, the effective cycle length becomes very short, because less time is required for the queue to reach a given length. This problem can be reduced somewhat by selecting the proper Q max. As will be shown later, the effective cycle length is only a function of Q max, given the flow rates.

Cycle lengths that are too long or too short can be avoided by imposing maximum or minimum values of green time, which will determine the upper and lower bounds of cycle length and green time.

One of the advantages of this queue-actuated control is that it is a more positive way to control queue length and prevent intersection blockage than any other method reviewed.

Delay and Cycle Length at an Intersection Under Queue-Actuated Control

Newell (11) and Sagi and Campbell (12) expressed the delay at a signalized intersection under saturated flows in the manner shown in Figure 6.

When queue-actuated signal control is applied to a signalized intersection, the delay at both approaches during the i th cycle can be expressed as shown in Figure 7. Notation used for this and later figures is as follows:

- Q_{a1}, Q_{a2} = maximum queue lengths allowed at approaches 1 and 2, in vehicles or feet;
- Q_{1i}, Q_{2i} = queue lengths at approaches 1 and 2 at the end of green, in vehicles or feet;
- X_1, X_2 = flow rates at approaches 1 and 2, in vehicles per sec [X_1 and X_2 are assumed to have a Poisson distribution with expected values $E(X_1) = \lambda_1$ and $E(X_2) = \lambda_2$];
- t_{1i} = effective red phase at approach 1 during i th cycle, in sec;
- t_{2i} = effective green phase at approach 1 during the i th cycle, in sec;
- t_i = effective cycle length during i th cycle ($t_i = t_{1i} + t_{2i}$), in sec; and
- S = saturation flow (service rate), in vehicles per sec.

S is assumed to have a uniform distribution and to be equal to or greater than X_1 and X_2. When $S < X_i$, queues become infinite and the system breaks down, because $(S - X_i)$ is the net discharge rate from an intersection.

In addition, it is assumed that lost time due to start-up delay and the amber phase is uniform. Lost time is not included in the delay diagram, but is dealt with separately.

Because of statistical fluctuations of flows, four situations could arise even under oversaturated flow conditions.

Case 1

In case 1 (Fig. 8), switching from one phase to another occurs when queues at one approach reach a predetermined length (Q_{a1} or Q_{a2}). Queues at the end of green (Q_{1i} or Q_{2i}) are greater than or equal to 0 at the moment of the termination of green.

Case 2

In case 2 (Fig. 9), queues at approach 1 are completely cleared before queues at approach 2 reach Q_{a2} and the green phase at approach 1 is terminated.

Case 3

In case 3 (Fig. 10), queues at approach 2 are completely cleared before queues at approach 1 reach Q_{a1} and the green phase at approach 2 is terminated.

Case 4

Because case 4 (Fig. 11) would be likely under light flows, Q_{1i} is assumed to equal 0.

Probability of each case

Let P_1 be the probability of queues being cleared at approach 1 before green is terminated and P_2 be the probability of queues being cleared at approach 2 before green is

Figure 6. Delay at one approach of intersection.

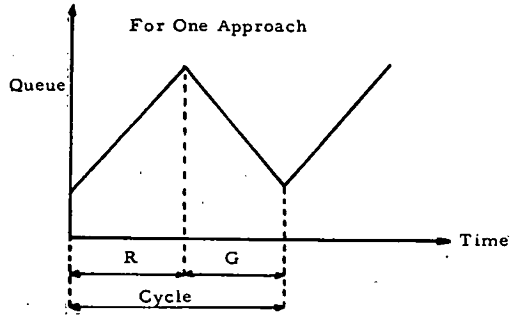


Figure 7. Delay at both approaches of intersection with queue-actuated signal control.

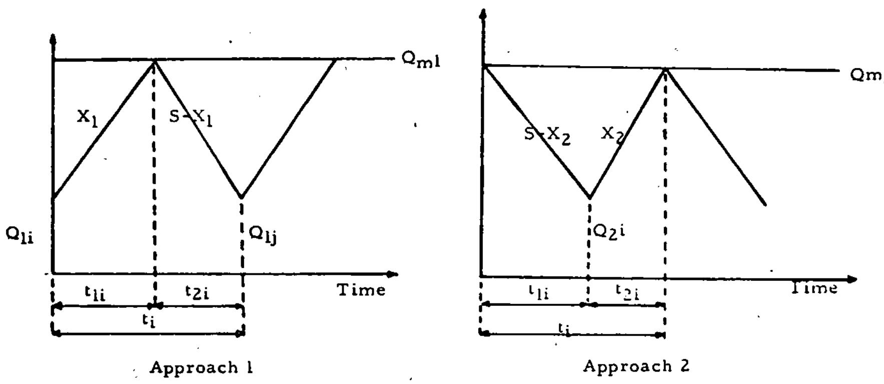


Figure 8. Oversaturated flow: case 1.

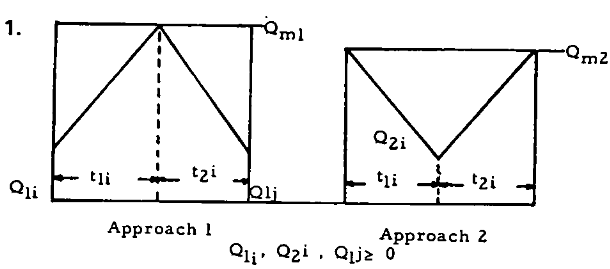


Figure 9. Oversaturated flow: case 2.

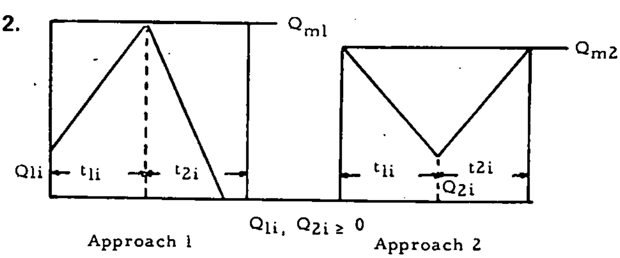
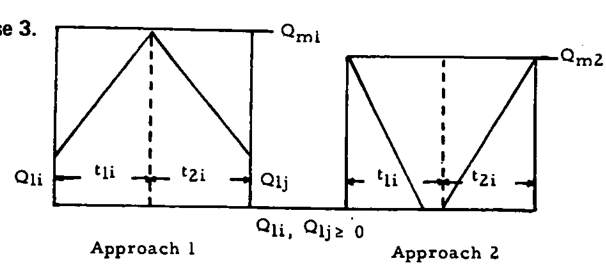


Figure 10. Oversaturated flow: case 3.



terminated. Then the probability of case 1 is $(1 - P_1)(1 - P_2)$, the probability of case 2 is $P_1(1 - P_2)$, the probability of case 3 is $(1 - P_1)(P_2)$, and the probability of case 4 is $(P_1)(P_2)$.

Under heavy flow conditions, cases 2, 3, and 4 do not likely occur because (a) every vehicle is expected to stop at least once, and there will not be a period of green without queue; and (b) their occurrence is indication that this control measure is not effective any longer.

The total delay during a given cycle caused by the existence of a queue is the summation of the shaded areas of the delay diagram shown in Figure 12.

The total traffic delay due to lost time (4, 5), such as start-up delay and the amber phase not used, for the i th cycle can be computed as

$$D_L = \frac{L}{t_i} \tfrac{1}{2}[t_i(X_1 + X_2)] = \frac{L}{2}(X_1 + X_2)$$

where L = lost time per cycle.

Thus, the total delay experienced by traffic on both approaches of an intersection is D_T = delay due to queue + D_L.

The total delay for each case of four different delay diagrams is summarized in Table 2. The table also includes the effective cycle length, as well as the effective lengths of green and red phases. For simplicity, the maximum queues allowed on both approaches were made the same ($Q_{m1} = Q_{m2}$). The expected delay and effective cycle length for all cases are

$$\text{Average delay} = \sum_{i=1}^{4} P(i)E[D_T(i)]$$

$$\text{Average effective cycle length} = \sum_{i=1}^{4} P(i)E[C(i)]$$

where

$P(i)$ = probability of case i,
$E[D_T(i)]$ = expected total delay for case i, and
$E[C(i)]$ = expected effective cycle length for case i.

The details of the derivation of the equations given in Table 2, evaluation of $P(i)$, $E[D_T(i)]$, and $E[C(i)]$, and a sample application of the delay model are discussed by Lee (3).

Comparison With Other Control Measures

The delay characteristics of an isolated intersection were also compared for three different control measures: fixed time, Longley's control strategy, and queue-actuated control (Fig. 13). The comparison demonstrated that, as volume increases, queue-actuated control results in the least average delay per vehicle, although at low volumes queue-actuated control produces the greatest delay, as expected.

Figure 11. Oversaturated flow: case 4.

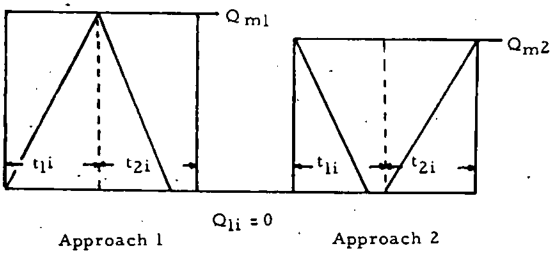


Table 2. Total delay and effective cycle length.

Case	Total Delay	Effective Cycle Length, Green and Red Phases
1	$D_T = (1/2)\left(\frac{S}{X_2}\right)\left(\frac{Q_a - Q_{1i}}{X_1}\right)\left[2(Q_a + Q_{1i}) - \frac{(S^2 - SX_1 - SX_2)}{X_1X_2}(Q_a - Q_{1i})\right] + \frac{L}{2}(X_1 + X_2)$	$t_i = t_{1i}\left(\frac{S}{X_2}\right)$ $t_{1i} = \frac{(Q_a - Q_{1i})}{X_1}$ $t_{2i} = t_{1i}\left(\frac{S - X_2}{X_2}\right)$
2	$D_T = (1/2)\left(\frac{S}{X_2}\right)\left(\frac{Q_a - Q_{1i}}{X_1}\right)\left[2Q_a - \frac{(S - X_2)(Q_a - Q_{1i})}{X_1} + \frac{(Q_a + Q_{1i})}{S}X_2 + \frac{Q_a^2X_1X_2}{S(S - X_1)(Q_a - Q_{1i})}\right] + \frac{L(X_1 + X_2)}{2}$	Same as case 1
3	$D_T = \frac{(Q_a^2 - Q_{1i}^2)}{2X_1} + \frac{2Q_a^2X_2 - Q_a^2(S - X_1)}{2X_2^2} + \frac{Q_a^2}{2(S - X_2)} + \frac{Q_a^2}{2X_2} + \frac{L(X_1 + X_2)}{2}$	$t_i = \frac{(Q_a - Q_{1i})}{X_1} + \frac{Q_a}{X_2}$ $t_{1i} = \frac{(Q_a - Q_{1i})}{X_1}$ $t_{2i} = \frac{Q_a}{X_2}$
4	$D_T = \frac{Q_a^2}{2X_1} + \frac{Q_a^2}{2(S - X_1)} + \frac{Q_a^2}{2(S - X_2)} + \frac{Q_a^2}{2X_2} + \frac{L(X_1 + X_2)}{2}$	$t_i = \frac{Q_a}{X_1} + \frac{Q_a}{X_2}$ $t_{1i} = \frac{Q_a}{X_1}$ $t_{2i} = \frac{Q_a}{X_2}$

Figure 12. Total delay caused by queue.

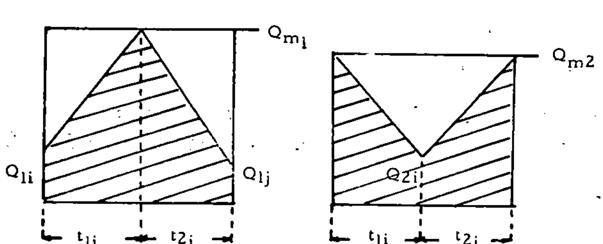
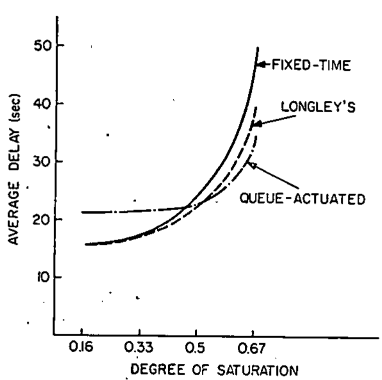


Figure 13. Delay comparison of different control measures.



Queue-Actuated Control in Coordinated Intersection

The underlying philosophy of coordinated systems is generally one of establishing signal timing to facilitate the uninterrupted movement of through vehicles along a roadway. In saturated and oversaturated operations, however, development of signal timing for such a purpose has little justification in that all vehicles must join queues and stop at least once somewhere upstream of the critical intersections. In practical terms there is no through traffic.

Queue-actuated control in a coordinated system of signals is functionally similar to the more conventional traffic-responsive control equipment, in that it places an essentially cycle-free traffic controller in the midst of a set of fixed-time signals.

It is worthwhile, however, to consider how the queue-actuated control policy performs when implemented at a critical intersection, which itself is in the midst of a coordinated system of signals. Inasmuch as this policy is queue-responsive, i.e., allocation of green time depends on some maximum queue length, the signal at the CI is driven by vehicle inputs from the immediate upstream intersections. Two extreme conditions can exist.

One extreme is when vehicle inputs are all provided by the through link upstream of the intersections upstream of the critical intersection. In this case, green at the CI is initiated as soon as through vehicles reach maximum allowable queue. Thus the relative offset between these two intersections looks like a forward progression. The other extreme is when all vehicles feeding the queue at the CI are turn-ins from the upstream intersection. In such a situation, the maximum allowable queue length will be reached before the initiation of green at the upstream intersection, and the relative offset looks like a reverse progression. For some combination of through and turn-in traffic, the relative offset appears to be simultaneous.

Thus, the actual initiation time and duration of green times at the CI under the queue-actuated control policy depend not only on the predetermined value for maximum allowable queue on the competing approaches and their respective volumes but also on the way in which these volumes input into the competing links.

The situation downstream of an initial intersection operating under this policy is different, in that a fixed-time signal with predetermined green time allocations receives traffic from an essentially cycle-free signal. In such a situation, in theory, a green phase can occur at the CI that will provide more traffic to the next downstream signal than it can handle, and thus create a saturated condition there. In such a situation the problem of saturation will not have been eliminated, only relocated. This can only occur, however, when traffic on the cross street at the critical intersection is very light; thus the main street is allowed an unusually long green time. In practice, however, this is quite unlikely, for one of the inherent features of the critical intersection is the presence of heavy traffic on both of the competing approaches.

The queue-actuated control in the midst of coordinated intersections was extensively tested by simulation on several networks, and its performance was compared with that of the existing control measures discussed previously. From this simulation, it was determined that (a) the adverse effects of the control are limited to the immediate downstream intersection; (b) although delay is increased at the immediate downstream intersection, queue-actuated control produces the least system delay; and (c) queue-actuated control results in the highest output at a critical intersection.

FURTHER RESEARCH

Further work is desirable on the following aspects:

1. The delay model needs further refinement, and more consideration should be given to the simplified assumptions used here.
2. The effects of errors in detection must be studied because of the limited accuracy of present detector systems available and because queue-actuated control must

rely on accurate detection of queue length.

3. The more desirable form of any control measure is one that could be used in oversaturated and undersaturated periods. Therefore the possibility of extending queue-actuated control to the period of undersaturation should be investigated.

4. An analytical model for the effects of queue-actuated control on adjacent intersections would help engineers better understand the effects of this control in the midst of coordinated signals.

REFERENCES

1. Improved Criteria for Traffic Signal Systems in Urban Networks. NCHRP Report 124, 1971.
2. Traffic Control in Oversaturated Street Networks. NCHRP Project 3-18(2), final report, 1974.
3. Lee, B. Signal Control in Oversaturated Street Systems. Polytechnic Institute of Brooklyn, PhD dissertation, May 1973.
4. Gazis, D. C., and Potts, R. B. The Oversaturated Intersection. Proceedings, Second International Symposium on the Theory of Traffic Flow, London, 1963.
5. Gazis, D. C. Optimum Control of a System of Oversaturated Intersection. Operations Research, Vol. 12, 1964.
6. Longley, D. A Control Strategy for a Congested Computer-Controlled Traffic Network. Transportation Research, Vol. 2, 1968.
7. Yagoda, H. The Control of Arterial Street Traffic. Presented at 1967 TFAC Control Conference, Haifa, Israel.
8. Chang, A. Synchronization of Traffic Signals in Grid Networks. IBM Journal of Research and Development, July 1967.
9. McShane, W. The Control of Vehicular Traffic on Urban Arterials and Network. Polytechnic Institute of Brooklyn, PhD dissertation, June 1968.
10. Yagoda, H., Stevens, E., and Sabala, E. A New Technique for the Optimal Timing of Two-Way Arterials. ASME Publication 70-Tran-39, 1970.
11. Newell, G. Approximation Methods for Queues With Application to Fixed-Cycle Traffic Light. SIAM Review, Vol. 8, 1965.
12. Sagi, G., and Campbell, L. Vehicle Delay at Signalized Intersections—Theory and Practice. Traffic Engineering, Feb. 1969.

A New Approach for Specifying Delay-Based Traffic Signal Warrants

Edward B. Lieberman and Reuben B. Goldblatt
KLD Associates, Inc.

A methodology is presented that uses both field studies and traffic simulation for developing delay-based traffic signal warrants. The field studies were performed to confirm the validity of the UTCS-1 simulation program. These warrants are based on a competition between two types of intersection control: traffic signals and two-way stop signs. Several of the worst common right-angle intersection configurations were analyzed over the entire spectrum of traffic volumes on their approaches. Analysis of the results led to the definition of criteria, based on volume and delay considerations, for selecting the appropriate type of control. The resulting warrants are expressed in graphical format, and accompanying specifications are given.

The Manual on Uniform Traffic Control Devices (MUTCD) states that a "comprehensive investigation of traffic conditions and physical characteristics of the location is required to determine the necessity for a signal indication..." (6). The assessment of traffic conditions for at-grade intersections is a continuing problem for traffic engineers. Solutions to this problem can take many forms ranging from purely empirical data collection techniques to development of complex theoretical models. Both approaches exhibit strengths and weaknesses. This project took a middle-ground approach in the form of a traffic simulation program validated by comparing the results with those obtained in the field.

The development of a simulation model represents a synthesis of theory and empiricism. To describe the highly variable, stochastic behavior of urban vehicular traffic requires that a microscopic simulation that properly replicates the process be developed. The computer model must be properly calibrated and then validated before it can be applied.

The UTCS-1 simulation model (1, 2) was developed for the Federal Highway Administration primarily to evaluate alternative traffic control policies on urban networks. After it was validated, it was modified (and reduced in size) for the purpose of addressing the single intersection problem. The objective was to use this program to provide the basic data necessary for the specification of delay-based traffic signal warrants. This paper describes the methodology and presents some representative results obtained in the initial phase of this study. Complete documentation appears elsewhere (3).

The basic approach was to specify traffic signal warrants that were firmly based on one or more operational measures of effectiveness (MOEs), which are of primary importance to the traffic engineer. Of necessity, the warrants were expressed in a format that can be applied in an unambiguous manner, avoids complex calculations, and requires a minimum of costly field data acquisition.

Survey results indicated that, although traffic volume data were available at all intersections considered, the direct measurement of MOEs such as vehicle delay and stops was rarely undertaken over an extensive period of the day on all approaches (3). From this information, it was clear that the relationship between the MOE on which the warrants would be based and the data that are directly available to the engineer (geometric descriptors, traffic volumes, etc.) had to be ascertained for the most common right-angle intersection configurations. The medium for providing this relationship was the modified version of UTCS-1.

USE OF SIMULATION TECHNIQUES

There have been an increasing awareness and acceptance of the UTCS-1 simulation model by the traffic engineering profession. In a study using the UTCS-1 as a medium for investigating traffic operations at intersections, Cohen stated (5):

> It can be concluded that the UTCS-1S Single Intersection simulation model has been successfully validated against field data from two intersections differing widely in geometry and location. This indicates that the model has, in addition to its flexibility, a sufficient degree of accuracy to enable it to be of considerable use in the geometric design and signal control of single intersections.

When the UTCS-1 stochastic simulation model is used, controlled experiments can be performed and sensitivity analyses can be conducted to identify the critical factors influencing traffic operations at intersections. Depending on field data alone is not appealing because

1. It is difficult to realize the full range of operating conditions in the field for all intersection configurations;
2. Manual data reduction is certain to introduce errors that may not be detected;
3. Field experimentation costs approximately twice as much as equivalent simulation analysis; and
4. The highly variable nature of traffic flow seriously compromises the validity of results that are aggregated over a time period of, say, 15 min.

This last factor is sometimes overlooked by researchers. It is well-known that the relationship between delay and volume is increasingly nonlinear as the load factor approaches unity. Hence, simple averaging of delay over 15 min yields inaccurate results in this range of volume, when cycle by cycle fluctuations in volume are pronounced.

Figure 1 shows a plot of volume versus time, where the field data are aggregated over four cycle periods (6 min). Even with this aggregation, fluctuations are so pronounced that it is impossible to hold volume constant over a sufficient time period to determine a statistically significant relationship between delay and volume. Applying simulation techniques within the framework of a controlled experiment avoids this problem.

The basic approach adopted was to conduct a competition between fixed-time signal control and two-way stop sign control. Hence, simulation was applied for both types of control at typical isolated, right-angle, four-legged intersections. The simulation results obtained for the two types of controls are presented separately. They are then synthesized in accordance with specified criteria for the development of the proposed traffic signal warrants.

The results of sensitivity studies led directly to the specification of base conditions for the warrants developed under the initial phase of the project. These conditions are

1. Random arrival of vehicles entering the approaches to an intersection,

Figure 1. Variation of traffic volume with time aggregated over 6-min time periods.

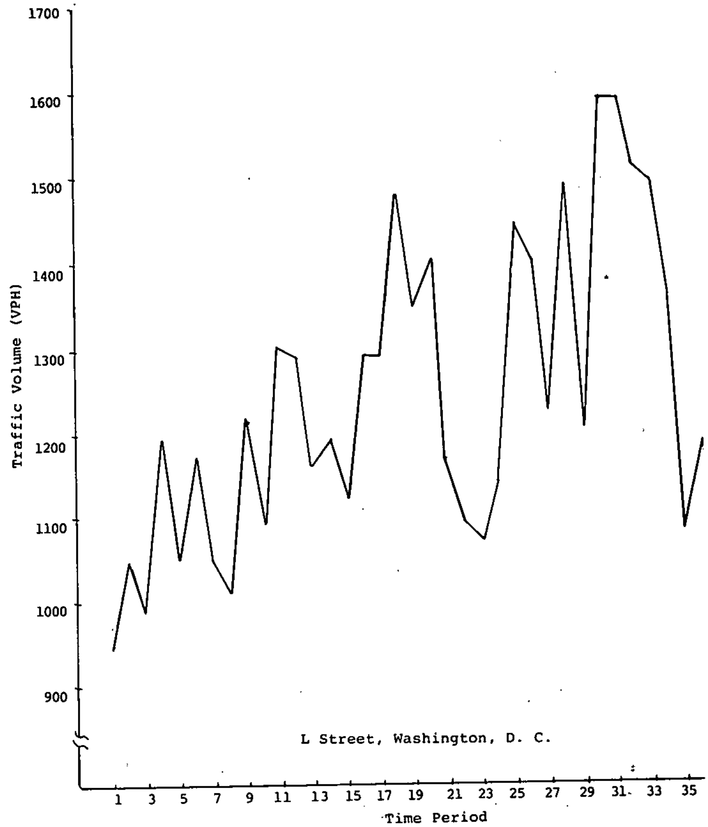


Figure 2. Delay per vehicle for various volume splits.

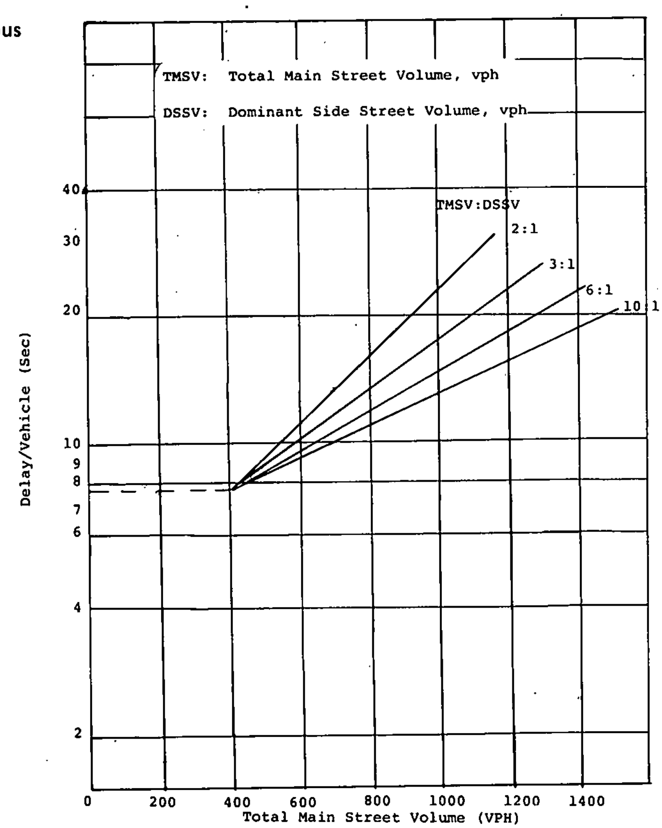


Figure 3. Total delay for various volume splits.

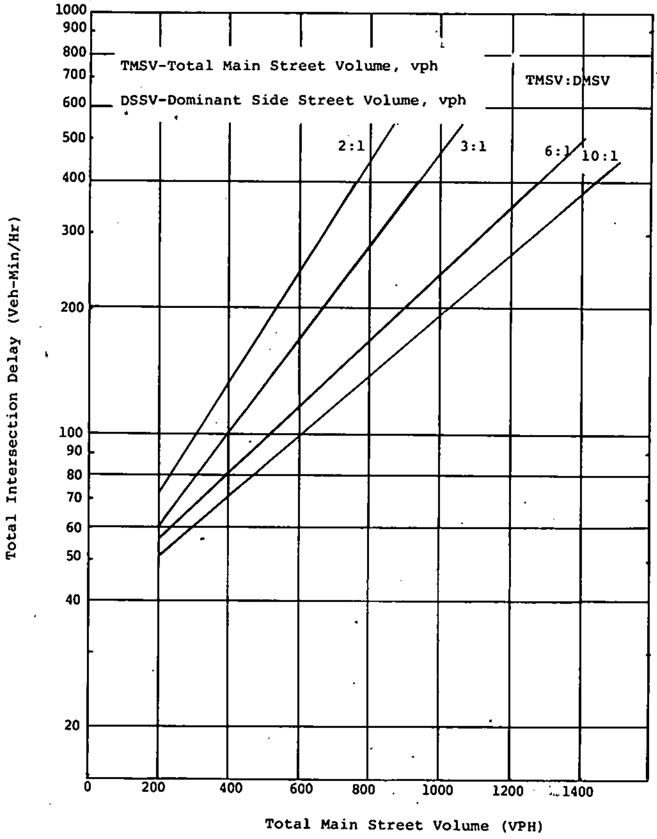
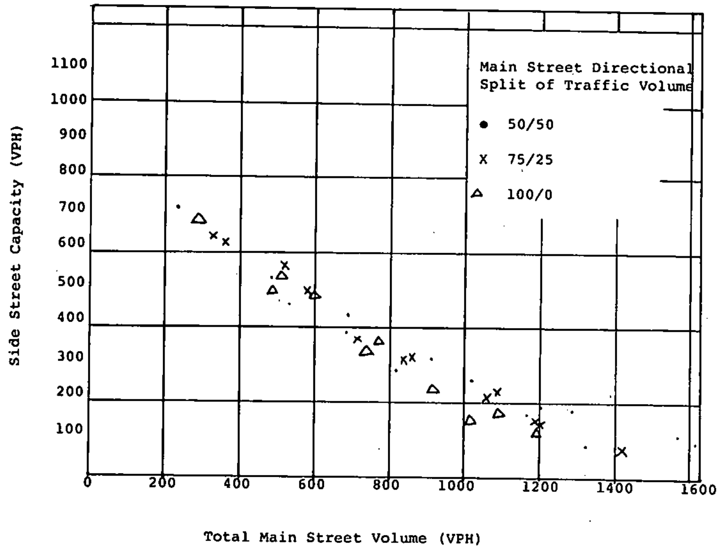


Figure 4. Capacity of a stop-sign-controlled approach.



2. Nominal turning movements and truck traffic,
3. Two values of queue discharge headway (2.1 and 2.4 sec per vehicle), and
4. Three values of mean gap acceptance for side street traffic (5, 6.5, and 8 sec) controlled by a stop sign.

Signalized Intersection Studies

These base conditions were used to perform a large-scale series of signalized intersection simulation runs. Delays specific to an intersection were obtained over the complete range of realizable (undersaturated) main street and cross street volumes.

From the simulation output, plots of per-vehicle and total intersection delay versus total main street volume were drawn for the traffic splits considered (Figs. 2 and 3). The intersection configuration used in these figures is an intersection of two two-way streets where each approach has one moving lane of traffic. For these studies, the values of optimal cycle length and of signal split (7) were implemented.

Two-Way Stop-Sign-Controlled Intersection Studies

Studies were performed to determine the capacity of approaches controlled by two-way stop signs. Once the capacity was determined, the range of volumes covering the region of undersaturated conditions was defined. It was then possible to develop stop sign delay curves through further simulation.

Figure 4 shows a typical capacity plot for side street approaches controlled by stop signs. This figure presents three sets of results, each representing a different directional split of traffic along the main street. As indicated, the effect of the directional allocation of main street traffic on capacity is minimal.

Various values of the mean acceptable gap $\bar{G}$ were also specified. (In Figure 4, $\bar{G}$ = 5.0 sec.) The results indicate that the capacity of an approach controlled by a stop sign is significantly reduced as $\bar{G}$ increases; capacity corresponding to $\bar{G}$ = 8 sec is 30 to 50 percent lower than for $\bar{G}$ = 5 sec over the range of main street volumes.

Based on data obtained from the stop sign capacity studies, the applicable range of side street and main street volumes for each configuration and mean acceptable gap was defined. Simulation studies were conducted to relate the delay measures (for side street vehicles only) to main street volume for each of three values of mean acceptable gap. Figures 5 and 6 show representative results for $\bar{G}$ = 6.5 sec.

ANALYSIS OF SIMULATION RESULTS

After the delay curves for both signalized and stop-sign-controlled intersections were developed, they were combined to form composite plots. Figure 7 shows a representative composite plot of delay per vehicle; Figure 8 shows a similar plot for total intersection delay. On all plots the same pattern is evident: When main street volumes are low, stop-sign-controlled intersections exhibit significantly less delay than do signalized intersections. As total main street volume increases, delay per side street vehicle (facing the stop sign) quickly exceeds the per-vehicle delay experienced by vehicles at signalized intersections. Total delay at intersections controlled by two-way stop signs, however, generally remains less than that at signalized intersections until main street volumes are quite high.

This relative disparity in behavior between mean delay per vehicle and total intersection delay is a key factor in the development of traffic signal warrants. It is clear that a trade-off is necessary to balance the dual objectives of minimizing total intersection delay and eliminating excessive per-vehicle delay.

Data shown in Figure 9 illustrate the approach taken. Figure 9(a) shows the variation in total delay experienced by vehicles on all approaches to an intersection controlled by a fixed-time, two-phase signal and by a two-way stop sign. Figure 9(b) shows the mean delay per vehicle for these two control types. These plots are con-

Figure 5. Delay per vehicle for various volume splits; side street controlled by stop sign.

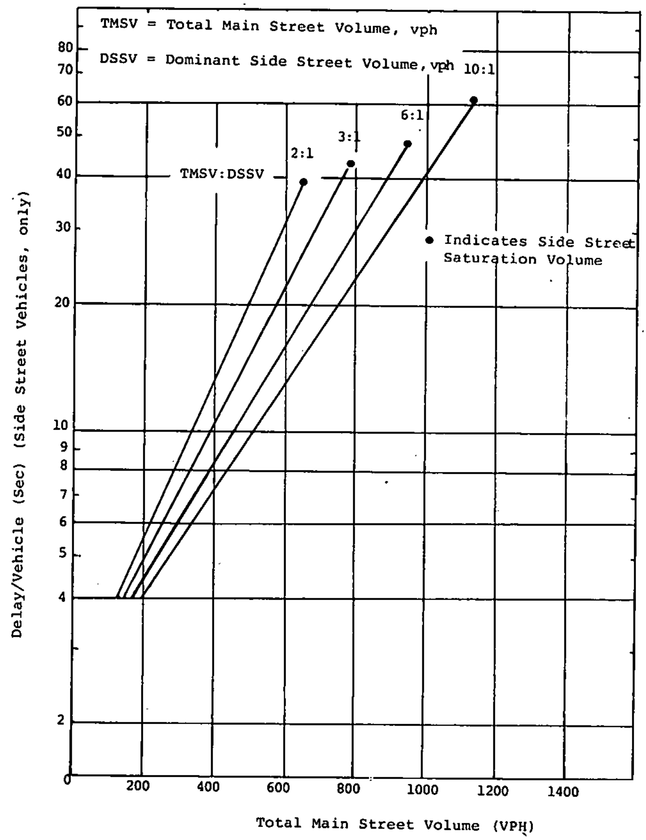


Figure 6. Total intersection delay for various volume splits; two-way stop sign control.

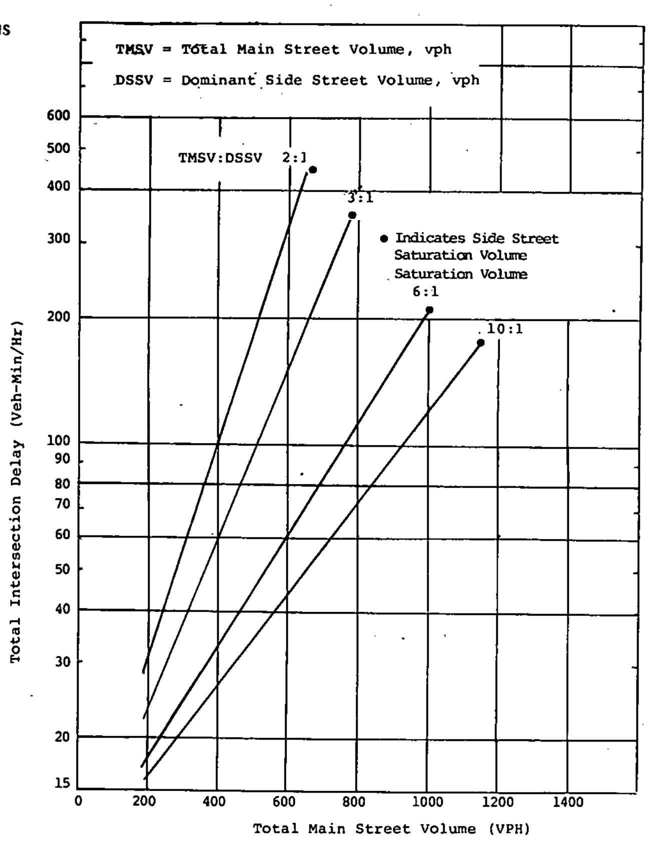


Figure 7. Composite delay per vehicle curves ($\overline{G}$ = 6.5 sec).

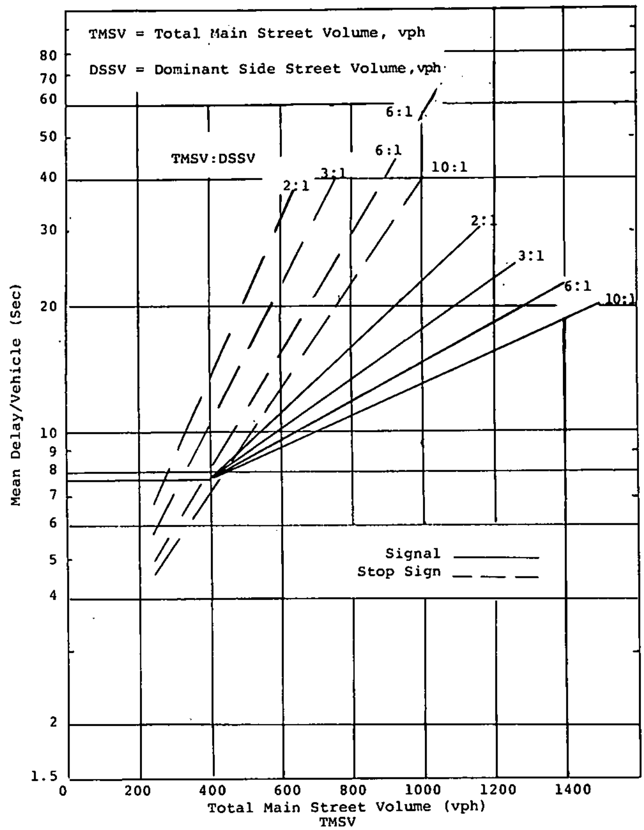


Figure 8. Composite total intersection delay curves ($\overline{G}$ = 6.5 sec).

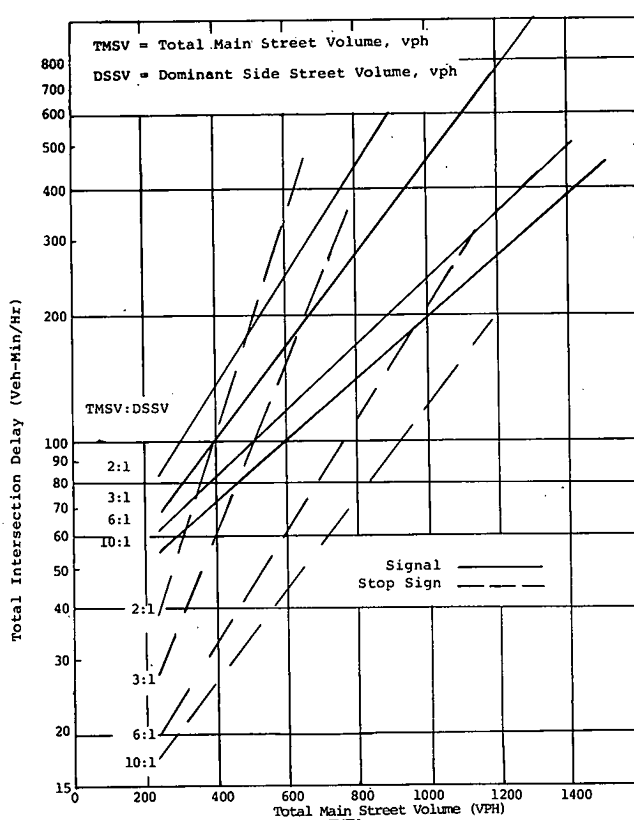


Figure 9. Conceptual representation of delay-volume relationships at isolated intersections.

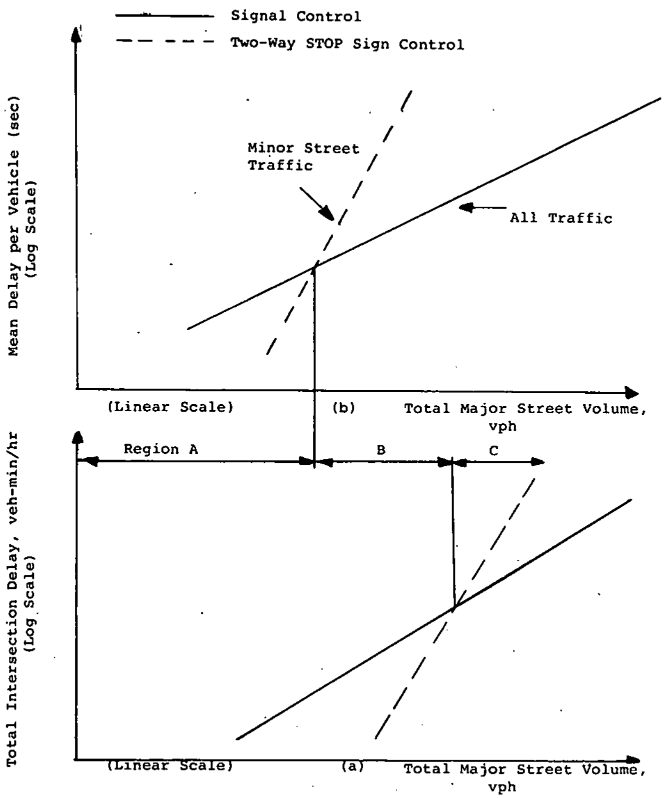
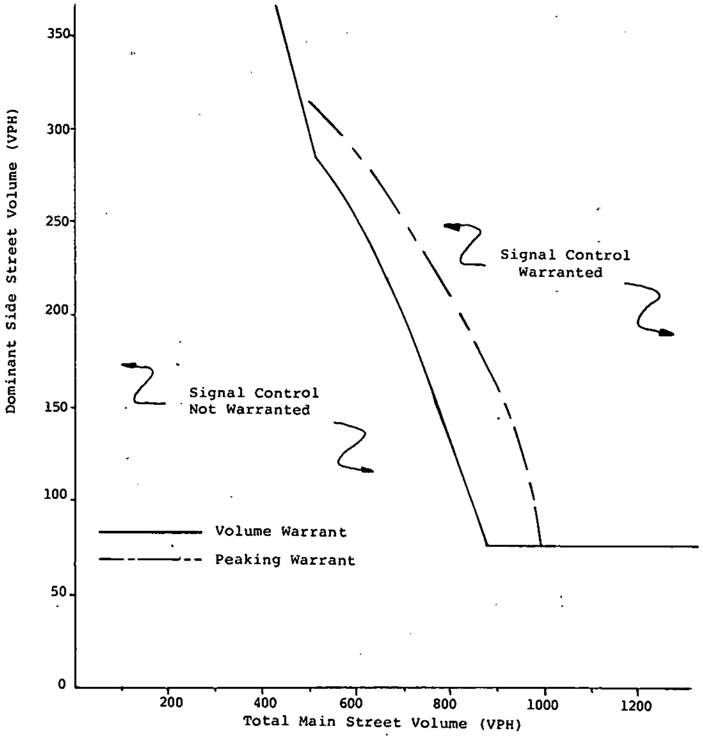


Figure 10. Typical volume and peaking warrants.



ceptual generalizations of those in Figures 7 and 8. Note that the value of main street volume that produces equal values of mean delay per vehicle for the two types of control is considerably lower than the value of main street volume that produces equal total intersection delay. This must always be the case when the total major street volume exceeds the total minor street volume. Three volume regions are identified in the figure.

Region A

In this region, the installation of a signal will degrade performance of all traffic. Both total intersection delay and mean delay per vehicle will be greater than that with two-way stop sign control. Clearly, the proper choice of control device in this region is the two-way stop sign.

Region B

In this region, installation of a signal will benefit the vehicles on the minor street approach but will produce disbenefits to the (larger number of) vehicles on the major street. It is in this region that the conflict between the two cited objectives prevails. Installation of a signal will act to equilibrate per-vehicle delay among those vehicles on the competitive approaches, thus satisfying the second objective, but will increase total intersection delay in the process, thus contradicting the first objective. Resolving this difficulty through an acceptable trade-off between the two objectives yields the basis for the proposed traffic signal volume warrants.

Region C

In region C, the installation of a signal satisfies both objectives; the proper choice of control device is clearly that of a traffic signal.

It is clear that the traffic signal warrants will be expressed in terms of those traffic conditions that lie within region B. (The current MUTCD volume warrants fall within this region.) It is necessary, then, to quantify the aforementioned trade-off, based on a concept of equity of service. Although it is necessary to assign priority to streets, it is unfair, impolitic, and conducive to unsafe driver behavior to impose excessive delay to a minor component of traffic in order to provide unimpeded service to a major component. A mean value must be determined that corresponds to tolerable delay for side street vehicles controlled by a stop sign. A survey of practicing traffic engineers across the country (3) yielded the mean value of 28 sec per vehicle.

Also, a number of traffic volume thresholds must be specified, including

1. A lower threshold for hourly side street traffic volume below which a traffic signal installation should not generally be considered and
2. A lower threshold for hourly total intersection volume below which a signal installation should not generally be considered.

It is generally not cost-effective to install a traffic signal when the peak-hour side street traffic volume is below some limiting value. Hence, the following thresholds have been adopted:

Lanes	vph
One	75
Two	100

Although these values are somewhat arbitrary, they represent a consensus of opinion and current practice.

PRESENTATION OF WARRANTS

Five criteria for the specification of delay-based traffic signal warrants have evolved from this study (3):

1. Minimum total intersection volume,
2. Minimum side street volume,
3. Equity of total intersection delay,
4. Tolerable side street delay, and
5. Equity of service.

Applying these criteria appropriately to the simulation delay data permits delay-based volume and peaking traffic signal warrants to be defined over the entire volume range of interest for intersection configurations that are most often candidates for signalization. The volume warrant specifies that a signal should be installed at an intersection if it is satisfied for each of any 4 hours. The peaking warrant is designed to assess short-term demands to determine whether a traffic signal is necessary; the need for a signal is indicated when this warrant is satisfied for any 2 hours.

Figure 10 shows the format for the volume and peaking warrants; a separate graph is provided for each intersection configuration considered. To apply the warrant, the engineer determines the total main street traffic volume and the associated dominant (higher) side street traffic volume for each hour. Hence, 1 hour of these data represents one point on the warrant figure. If any four of these points lie to the right of the solid line in the indicated region, then a traffic signal is warranted. Similarly, if two of these points lie to the right of the broken line, then a traffic signal is warranted, even if the first condition is not satisfied. If neither condition is satisfied, a traffic signal is not warranted on the basis of delay considerations.

The primary departures of these traffic signal warrants from those specified in sections 4C-3, 4 of the MUTCD (6) are

1. They synthesize MUTCD traffic signal warrants 1, 2, and 8,
2. They are defined over the entire range of admissible volumes in place of discrete tabulations,
3. A peaking warrant is presented, and
4. Warrants are based on the four highest volume or two highest peaking hours rather than on the eight highest volume hours.

Furthermore, it is believed that these traffic signal warrants exhibit the following advantages over those currently specified in the MUTCD:

1. They are based on that MOE most readily perceived by the motorist, i.e., delay;
2. The warrant curves are internally consistent in accordance with well-defined criteria;
3. The graphical presentation provides the engineer with a visual impression of the intersection status with respect to its need for a traffic signal (e.g., if data do not satisfy the warrant but there is a cluster of points near the curve, then this evidence could indicate a near-term requirement for upgrading the control there);
4. This graphical presentation could be an effective technique for responding to citizen groups that petition for a traffic signal where it is not warranted; and
5. The shorter time scale (4 versus 8 hours of data) reduces the engineer's data gathering requirements and associated cost and is more responsive to the characteristic peaking of traffic demand.

Ongoing research will extend the warrants to those intersections that experience heavy turning movements and will extend the number of configurations as well. In addition, a systems warrant will be developed for signals that are closely spaced.

ACKNOWLEDGMENT

This work was sponsored by the American Association of State Highway and Transportation Officials in cooperation with the Federal Highway Administration and was conducted under the National Cooperative Highway Research Program.

REFERENCES

1. Lieberman, E. B., Worral, R. D., and Bruggeman, J. M. Logical Design and Demonstration of UTCS-1 Network Simulation Model. Highway Research Record 409, 1972.
2. Bruggeman, J. M., Lieberman, E. B., and Worral, R. D. UTCS-1 Network Simulation Model, Final Report. Federal Highway Administration, Report DOT-FH-11-7885-4, 1973.
3. Lieberman, E. B., King, G. F., and Goldblatt, R. B. Traffic Signal Warrants. KLD Associates, Inc., Report 17, 1973.
4. Kennedy, W. J., Jr., Farr, R. M., and Labrum, W. D. Computer Simulation of Traffic Signal Systems. Traffic Engineering Division, Utah State Department of Highways, March 1974.
5. Cohen, S. L. Application of Network Simulation Models to the Analysis of Urban Intersection Performance. Federal Highway Administration, Report FHWA-RD-73-95, September 1973.
6. Manual of Uniform Traffic Control Devices for Streets and Highways. U.S. Government Printing Office, 1971.
7. Webster, F. V. Traffic Signal Settings. Road Research Laboratory, England, Technical Paper 39, 1958.

The Effect of Traffic Lane Closures on the Highway Motorist

B. C. Butler
Byrd, Tallamy, MacDonald and Lewis

The occupancy of roadways for maintenance and construction interferes with traffic. A computer program was developed to perform an economic analysis of the impact on the motorist created by lane closures associated with this roadway occupancy. The analysis accommodates hourly distributions of traffic by direction and generates motorist costs. Included are vehicle operational costs as a function of vehicle weight and speed and freeway alignment; time costs by trip purpose, income level, and time loss; and accident costs. Nomographs, derived from computer output, were developed to permit rapid determination of motorist costs for any given occupancy interval, traffic volume, and lane closure.

High traffic volumes on freeways frequently cause accelerated deterioration of the pavement surface. Consequently, early in its life, the pavement may need maintenance. Maintenance is difficult to perform in the presence of the high traffic volumes, and the time that the road can be occupied by such crews is limited. These constraints, together with hasty repairs, add to the costs of maintaining a freeway. Additionally, the occupancy of the roadway by maintenance crews creates conflicts with the motorist. These conflicts influence motorist operating costs and create motorist delays, increased opportunities for accidents, and increased levels of pollution. The Federal Highway Administration determined that one means of minimizing the cost associated with pavement maintenance is to produce a pavement requiring minimum maintenance or even no maintenance. The warrants for the design of this "premium pavement" would be the savings in maintenance and motorist costs realized by eliminating maintenance or severely limiting it.

The information presented in this paper is the result of a study made to develop traffic warrants for premium pavements. An economic analysis of roadway occupancy during maintenance was developed. This analysis encompassed, in addition to maintenance expenditures, the operating, time, and accident costs to the motorist caused by roadway occupancy. (The term roadway occupancy shall be used to mean occupancy of the roadway by maintenance crews.) The analysis is in the form of a computer program called EAROMAR, economic analysis of roadway occupancy for maintenance and rehabilitation. In general, the program (Fig. 1) performs three functions:

1. It establishes a data matrix of given and assumed information in an initialization routine;
2. It determines the specific hours that the roadway will be occupied by work crews

annually, together with the maintenance and rehabilitation costs associated with that occupancy; and

3. It determines the operation, time, and accident costs to the motorist caused by the roadway occupancy.

This paper is limited to those aspects of EAROMAR that relate to the impact of lane closures on the motorist.

FIELD DATA COLLECTION

Before the analysis and the relationships required to establish motorist speeds under different roadway closure conditions were developed, an extensive field study was made to develop quantitative data on speed profiles under different roadway occupancy conditions. Speed profiles were developed for both normal and roadway occupancy conditions so that the implications of the road closure could be quantified. The mechanics of the field data collection effort were automated as much as possible to improve the reliability of the data and to expedite data reduction efforts.

A simple frequency oscillator was built to be attached to the speedometer of the test vehicle. Oscillator output signals were monitored on small inexpensive portable cassette tape recorders. A midfrequency signal was used to indicate that the test was in progress. The oscillator was attached directly to the speedometer cable with a T-connector. The speedometer-excited frequency signal was a function of the speed of the vehicle (Fig. 2). A third signal was controlled by the vehicle driver and served as an event marker. In the development of speed profile data, the test vehicle was normally operated between interchanges. A test period was about 45 minutes, the length of the cassette tape.

In addition to the speed profile, two observers stationed at each end of the work zone collected volume, headway, and elapsed time data on vehicles passing through the work zone. Figure 3 shows a profile trace of a test vehicle passing through a work zone with observers located at each end.

DATA REDUCTION

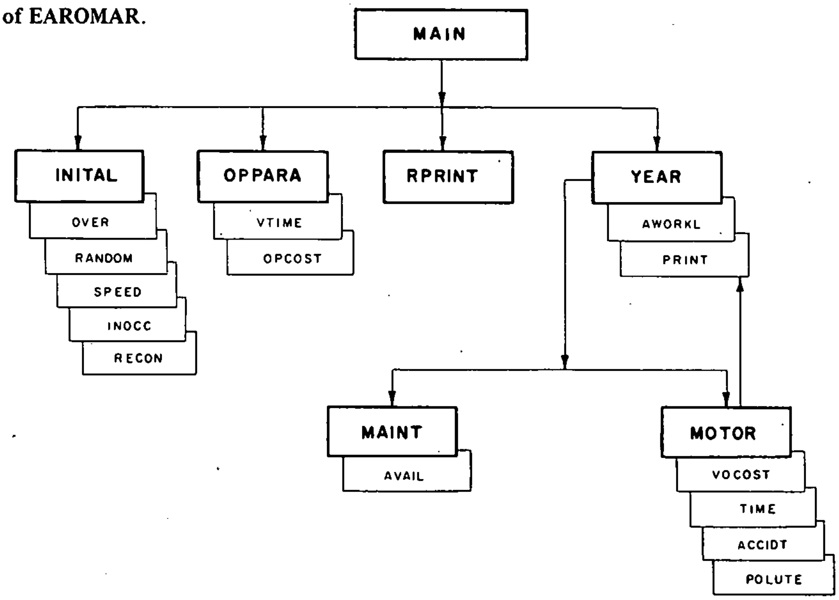
The data from the cassette tapes were consolidated on 7-in. reels of ¼-in. magnetic tape. These reels were processed through an analog computer to create a digital readout of times for each frequency event on the tapes. This was accomplished in the following manner.

1. An analog computer board was designed to identify three frequency signals on the tape.
2. A time pulser was set to operate $^1/_{100}$-second signals, which were counted by an accumulator and held.
3. As each signal was identified, it was assigned the value of the accumulated number residing in the pulse hold area.
4. The numbers for frequency signals above the test frequency were made negative to differentiate them from the frequency signals below the test frequency.
5. The resulting digital numbers were placed in a computer tape file.

Once the tapes were converted, programs were developed to convert the time information into speeds. These speeds were accurate to 1 percent. An example of the speed profiles developed is shown in Figure 4 for a two-lane closure on a six-lane divided facility. The period covered from 11 a.m. to 1 p.m. and the control zone was in the p.m. peak direction.

During the study no queues were observed in the presence of maintenance closures. This was not completely unexpected because as a general policy maintenance is scheduled during hours of low traffic volumes so as to minimize disruptions to traffic.

Figure 1. General flow of EAROMAR.



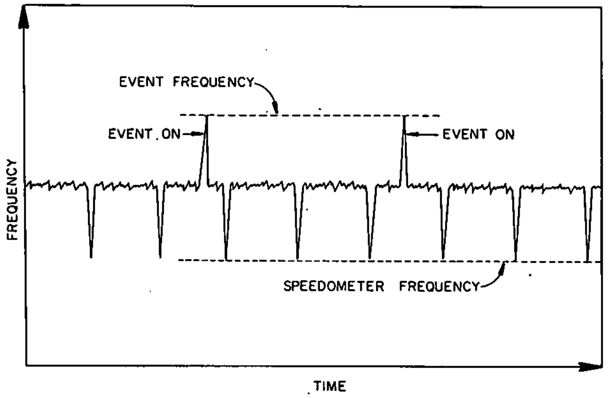


Figure 2. Schematic of speedometer (low-frequency) and event marker (high-frequency) signals.

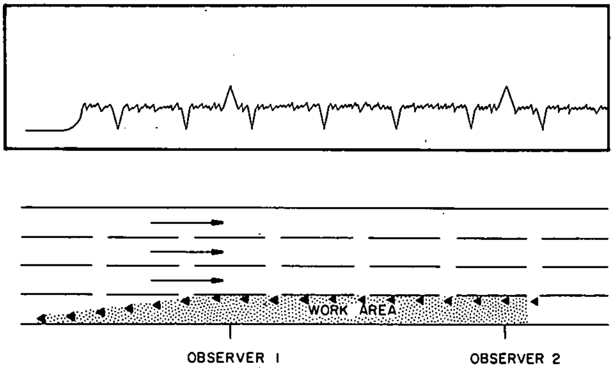


Figure 3. Schematic of signals generated during a typical test vehicle run through work zone.

Queue profiles were obtained from test runs made on Interstate 495 around Washington, D.C., during rush hour. A typical queue is created by traffic entering I-495 at the Va-123 interchange (Fig. 5).

SPEED MATRIX ALGORITHM

Based on the observations and data collected during the field studies and on established speed-volume relationships in the Highway Capacity Manual (8), an algorithm was developed to create a speed matrix as a function of freeway design speed, speed limit, and roadway capacity. Results of this algorithm are shown in Figure 6.

The user of the computer program EAROMAR provides a description of the freeway design speed and the speed limit and capacity for each feasible lane closure. The program generates a speed matrix for a range of volume-capacity ratios for each lane closure.

TRAFFIC

In the program a traffic matrix is initialized based on given hourly distributions of traffic by trip purpose. Each hourly distribution is for two directions, reflecting a.m. and p.m. distributions for the following trip purposes: work, social-recreation, personal business, vacation, school, and commercial. Figure 7 shows a typical hourly distribution for work trips in the peak a.m. direction.

VALUE OF TIME

The program generates a matrix in which the value of time is associated with each trip purpose, time loss in minutes, and income level. This is based on the work done at Stanford Research Institute by Thompson and Thomas (1).

OPERATION COST

The operation cost matrix created by the program is a function of roadway alignment, vehicle weight, speed and speed changes, and the consumption parameters—fuel, tire, oil, maintenance, and depreciation. The consumption parameter curves developed were based on the work of Winfrey (2), Claffey (3), and Sawhill (4).

Figure 8 shows typical fuel consumption curves. The consumption rate is a function of speed and grade. Figure 9 shows tire wear consumption in increments of 0.001 inch per speed change cycle as a function of speed change. Figure 10 shows the depreciation rate for a range of vehicle classes as a function of speed.

Unit costs are specified for fuel, tires, oil, maintenance, and depreciation as program input. The series of consumption models is used with inputs for vertical and horizontal alignment data, weight data on a range of vehicle weight classes, and the unit costs to generate two matrices of hourly operating costs, one for passenger cars and one for commercial vehicles. Operation costs are a function of speed in both matrices.

ECONOMIC ANALYSIS

EAROMAR performs an analysis annually by direction, work activity, and feasible lane closure. Once the actual hours of occupancy have been established, the impact of that occupancy on the motorist is computed in terms of changes in operation costs, value of time loss, increased accidents, and pollution effects.

Operation Costs

The operation costs are based on determining (a) the operating cost through the

Figure 4. Speed profiles when two of three directional lanes were closed to traffic.

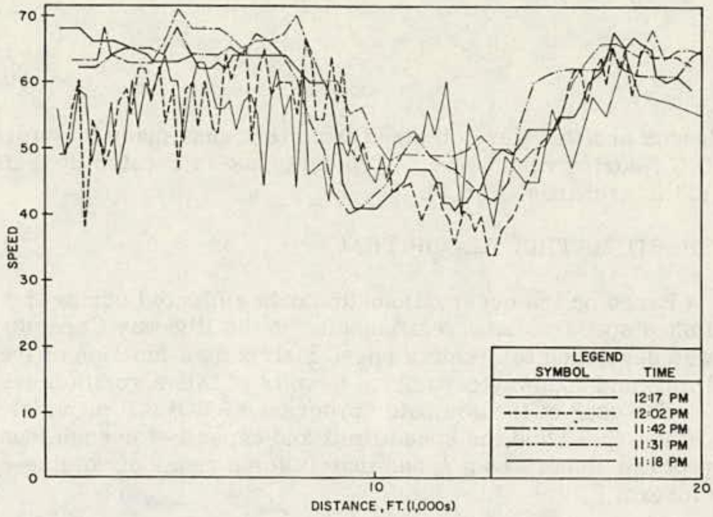


Figure 5. Example of queue on a four-lane freeway.

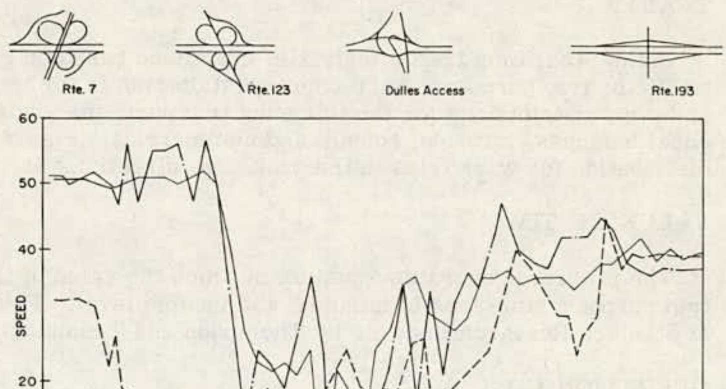


Figure 6. Speed curves for highway on which speed limit equals design speed.

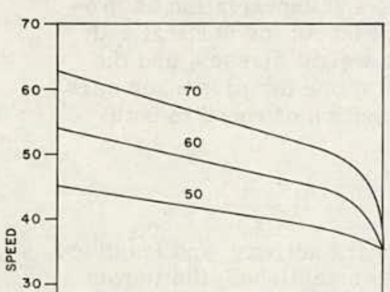


Figure 7. Distribution of all traffic in a.m. peak direction.

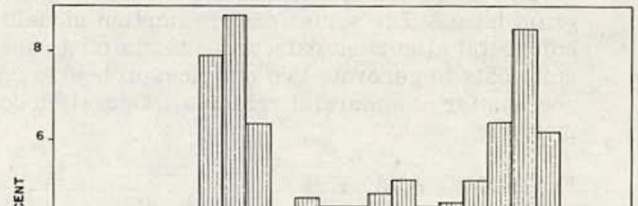


Figure 8. Gasoline consumption as a function of speed and grade.

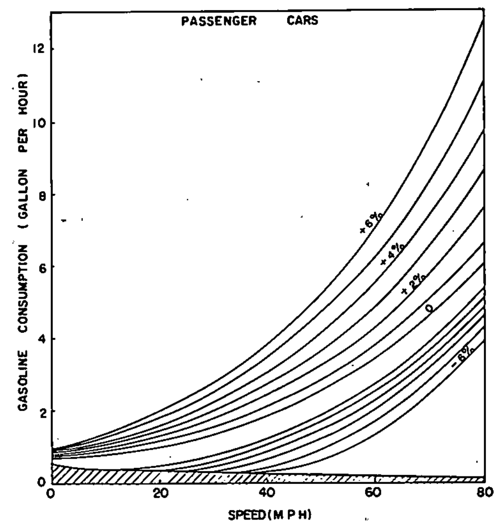


Figure 9. Excess tire wear as a function of speed change for passenger car.

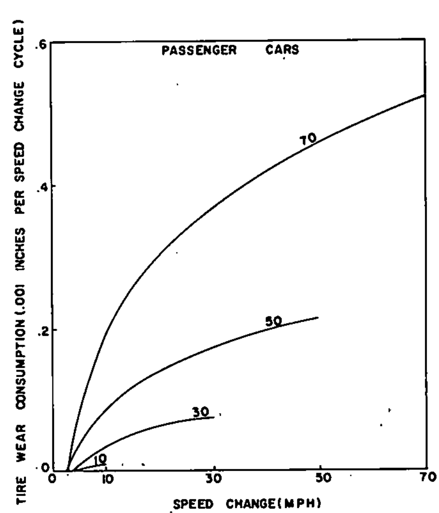
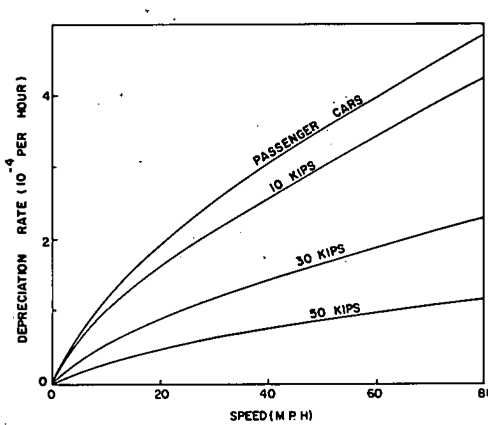


Figure 10. Depreciation rate for a range of vehicle weights as a function of speed.



influence zone of traffic under normal conditions, (b) the operating cost of traffic passing through the influence zone for the lane closure, and (c) the cost of the speed change cycle created when traffic slows from a normal freeway speed to a speed equal to 1.5 times the difference in freeway speed and influence zone speed. The costs for passenger cars and commercial traffic are separate, and adjustments are made to reflect a reduction in capacity as a function of the percentage of commercial traffic in the traffic stream. When a queue occurs, the operation costs for traffic operating normally over the distance queued and the operation costs in the queue are determined. The net operation costs associated with the occupancy are the occupancy costs less the normal costs.

Time

The value of time is a function of time lost and trip purpose (1). A distribution of traffic by trip purpose exists for each hour of the analysis. The lost time is the difference between the duration spent in the influence zone during normal operations and the duration spent during the closure. Additionally, any time spent in the queue is considered. The total time loss is used to compute a value of time for the traffic analyzed each hour.

Accidents

Changes in accident rates for normal and lane closure situations are determined two ways. An increased accident rate is associated with the deceleration process (5). Also, the accident rate is related to operating lanes (6).

Pollution Effects

The determination of pollution effects is based on emission factors for hydrocarbons and carbon monoxide. Slower speeds produce larger emission factors per unit of time. The pollution is presented as days of pollution, i.e., additional days of normal operation to create the additional emissions.

RESULTS

EAROMAR was used to generate motorist costs for a range of freeways and traffic volumes. The generated output was used to establish a series of nomographs. These nomographs are included in a manual developed to permit a noncomputer analysis (7).

Two sets of nomographs were developed. The first set, shown in Figures 11, 12, and 13, permits a determination of the motorist costs associated with road closures when there is no queuing. Illustrated is a roadway occupancy between 9 a.m. and 4 p.m. The motorist costs given in Table 1 were determined by using the three nomographs and an average directional annual traffic volume of 30,000.

Motorist costs increase substantially when a queue occurs. Figure 14 shows a nomograph used to determine motorist costs during hours of queuing when one lane is open to the motorist. The nomograph is used as follows:

1. A line is drawn from point A on the percentage scale to 30,000 on the directional traffic volume scale to create a point on the X-X line.
2. A line is drawn from point B through the point created on the X-X line to an intercepting point on the percentage scale.
3. A horizontal line is drawn through the interception point created on the percentage scale.

Every "hour of the day" bar that is crossed by the horizontal line is a queue hour for the 30,000 directional volume. The operation costs and value time losses must be

Figure 11. Nomograph for determination of increased vehicle operation costs.

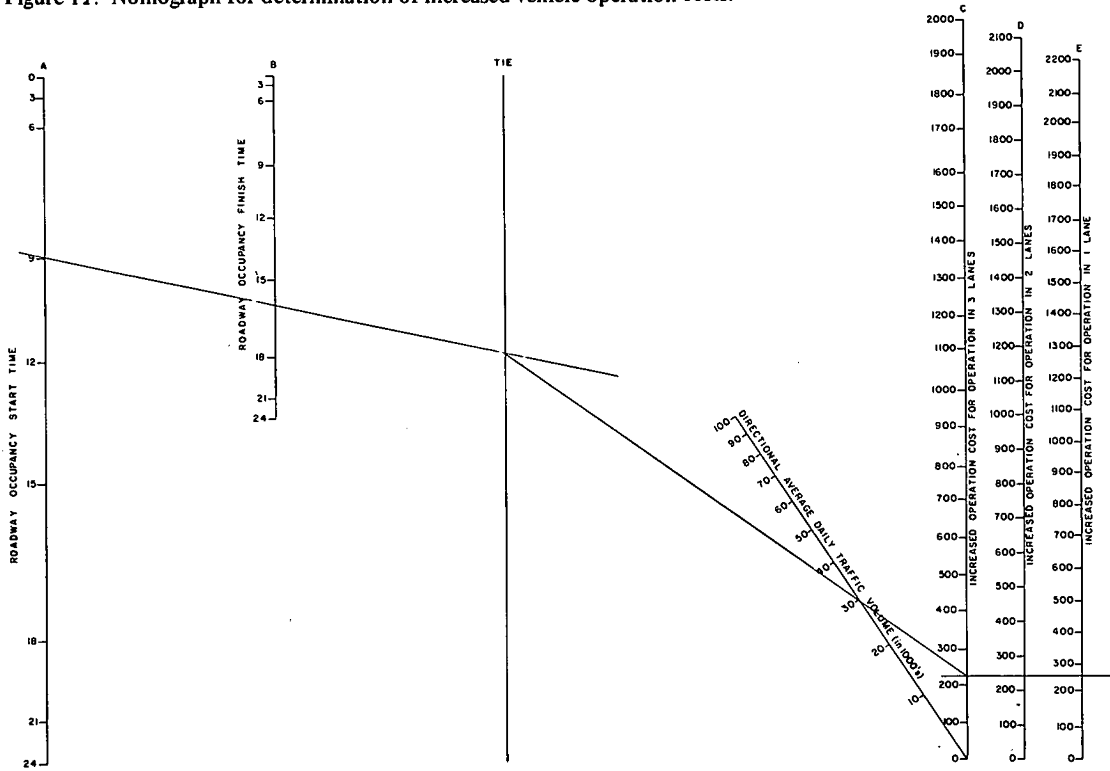


Figure 12. Nomograph for determination of time costs.

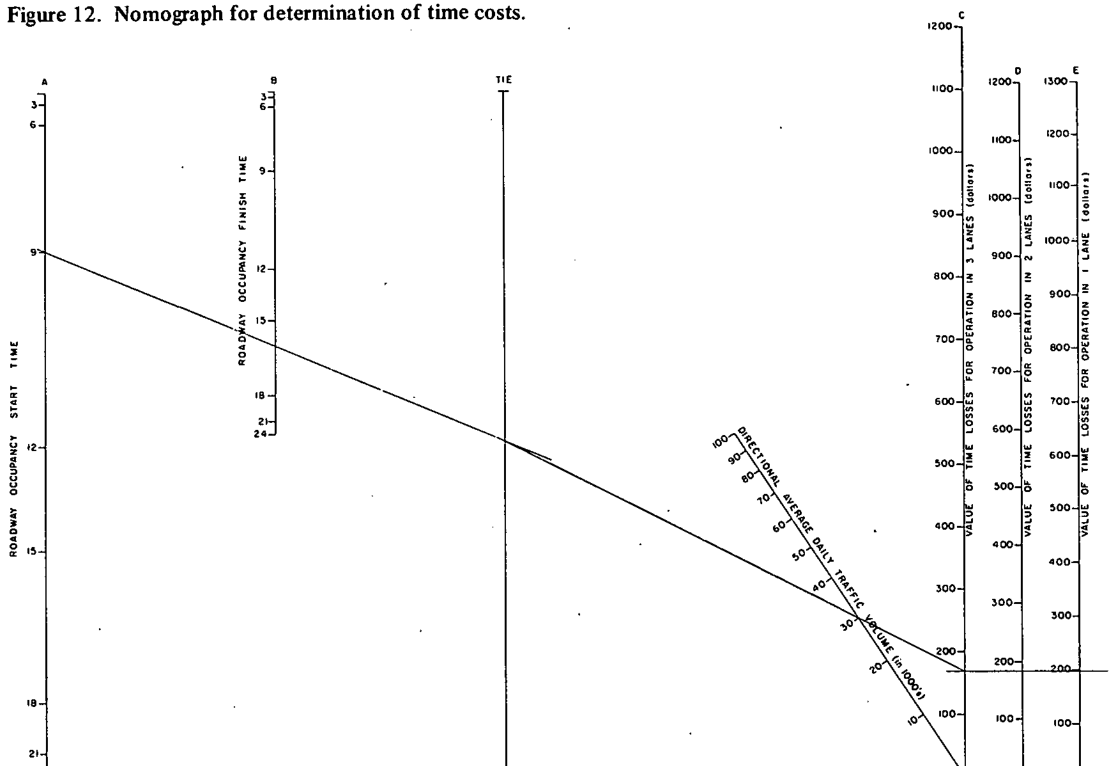


Figure 13. Nomograph for determination of accidents costs.

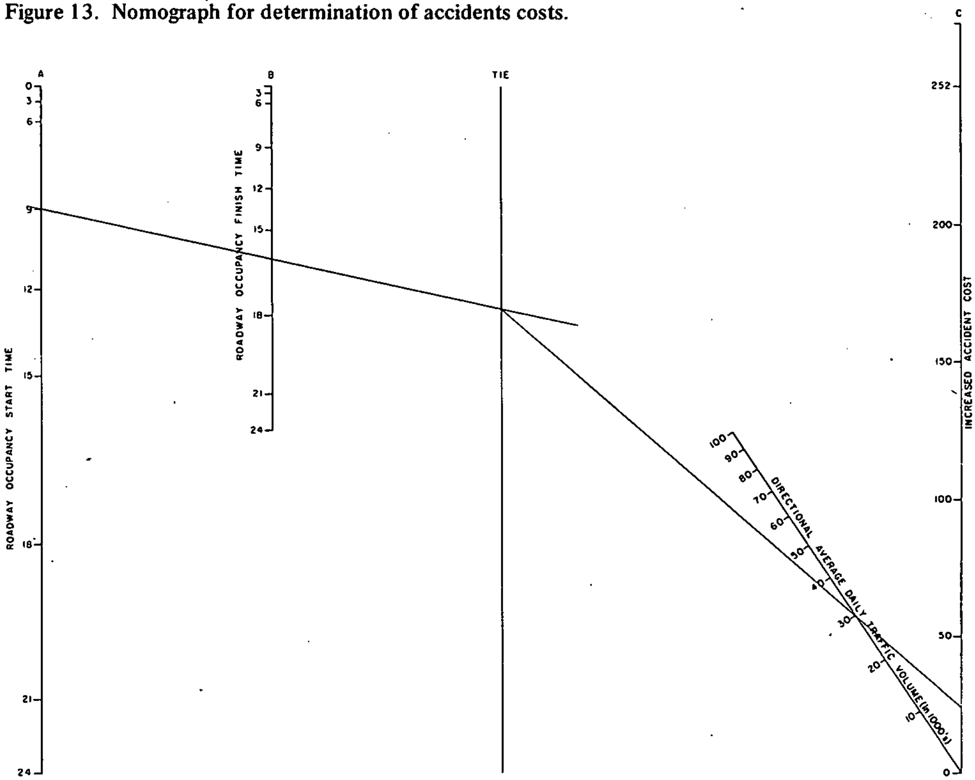


Figure 14. Nomograph for determination of increased operation and time costs during hours of queuing.

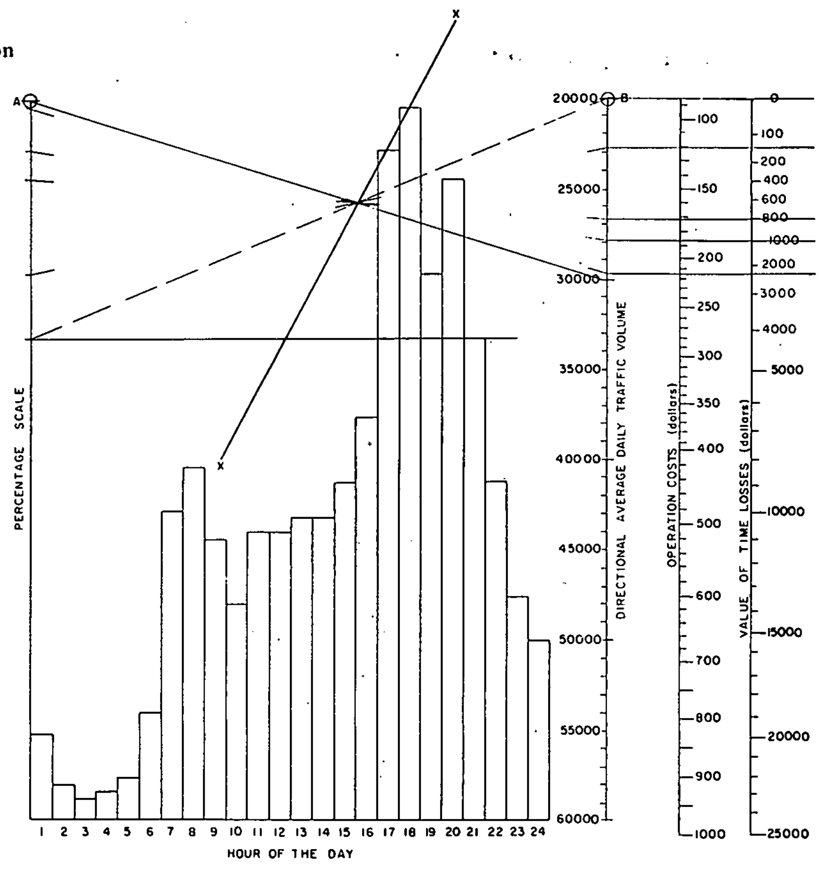


Table 1. Motorist operation costs during nonqueue conditions.

Lanes Open	Costs (dollars)		
	Operation	Tire	Accident
3	220	170	25
2	240	185	25
1	260	200	25

Table 2. Queuing costs for example conditions shown in Figure 14.

Hours	Operation Cost (dollars)	Value of Time Loss (dollars)
17	187	1,000
18	215	2,300
19	120	150
20	171	800
21	85	0
Total	778	4,250

determined separately for each queued hour. In the example shown in Figure 14, hours 17 through 21 are all queued. To determine the costs associated with the queued hours requires the following additional steps:

4. Draw a horizontal line from the top of each queued hourly bar to the percentage scale.
5. From each point created on the percentage scale for a queue bar, draw a line through the point created on the X-X line to an intercepting point on the directional traffic volume scale.
6. For each intercepting point on the directional traffic volume scale, draw a horizontal line through the operation costs and value of time losses scales.

The queue costs associated with queue hours 17 through 21 as taken from the nomograph are given in Table 2. Of course, these queued costs would only be determined if the actual roadway occupancy included the queue hours 17 through 21. The impact of the queuing in terms of increased costs to the motorist for operation costs is $778 and for loss time is $4,250; the total is $5,028. The cost to the motorist of having one lane open on a freeway between 9 a.m. and 4 p.m. is $485 because no queue hours are included in this interval.

CONCLUSION

An example of the motorist costs that are generated by the EAROMAR program when the roadway is occupied for maintenance and rehabilitation has been presented. The program permits an objective evaluation of the implications of roadway occupancy to the motorist. Although the program was designed to provide warrants for developing a premium pavement, it also has application in the development of general benefit-cost analyses, it can be used in development of closure strategies where roadways need to be occupied, and it can be used to forecast maintenance costs.

ACKNOWLEDGMENT

Contents of this report reflect the views of the author who is responsible for the facts and the accuracy of the data presented. Contents do not necessarily reflect the views or policies of the Federal Highway Administration.

REFERENCES

1. Thomas, T. C., and Thompson, G. I. The Value of Time Saved by Trip Purpose. Stanford Research Institute, Project MSU-7362, October 1970.

2. Winfrey, R. Economic Analysis for Highways. International Textbook Co., 1969.
3. Claffey, P. J. Time and Fuel Consumption for Highway User Benefit Studies. HRB Bull. 276, 1960, pp. 20-34.
4. Sawhill, R. B. Motor Transport Fuel Consumption Rates and Travel Time. HRB Bull. 276, 1960, pp. 35-91.
5. Heimbach, C. L., and Vick, H. D. Relating Change of Highway Speed per Unit of Time to Motor Vehicle Accident Rates. Highway Research Record 225, 1968, pp. 44-64.
6. Kihlberg, J. K., and Tharp, K. J. Accident Rates as Related to Design Elements of Rural Highways. NCHRP Report 47, 1968.
7. Butler, B. C. Traffic Warrants for Premium Pavements Requiring Reduced Maintenance. Report FHWA-RD-74, October 1974.
8. Highway Capacity Manual—1965. HRB Special Rept. 87, 1965.

Toll Agency Concepts of Project Acceleration

William Rohde
New Jersey Turnpike Authority

This paper highlights the approach taken by the New Jersey Turnpike Authority with regard to construction trade-offs and project acceleration. A recently completed $500 million widening of the New Jersey Turnpike is used as a basis to illustrate several examples. Traffic-handling procedures used are generally discussed, and comparisons are made between the approach of a toll agency and that of a government highway agency.

When a toll authority undertakes a major reconstruction project, it is faced with many considerations not immediately apparent to the public works engineer. Very often these considerations conflict with normal practice and are contradictory within themselves.

Financing, a prime consideration for any construction project, takes an added impetus. In the private sector, an Edsel or a C5A is allowed once in a while without total and complete disaster. In the public sector with tax money as financing, we can underestimate tax revenues and still complete the project. The Interstate System in New Jersey, for example, is a monument to missed schedules, lack of financing, and political manipulation; yet the end result will probably be a system substantially completed as originally planned. The fact that millions have been lost through cost escalations and inflation and the fact that the economy of the state has been seriously negatively affected by the nonexistence of an adequate highway system become lost or clouded through political campaign rationalizations and rhetoric.

The toll authority has to "put its money where its mouth is." Large sums of money have to be borrowed, and the only funds available to repay the indebtedness are the revenues the facility receives. The repayment, over a relatively long time period, must be correctly estimated initially. Everything, therefore, becomes dependent on time. The project must be completed on time because revenues must be received according to a fixed time schedule.

Additionally, the toll authority's ability to borrow money in the future is directly keyed to its performance on previous projects. Obviously, the agency that completes projects on time, without cost overruns, and that has completely met or bettered its obligations with respect to repayment will have an attractive position in the financial marketplace when funding is required for new projects. This is important, of course, with respect not only to the availability of money but also to the cost of borrowing the money.

In January 1970, the New Jersey Turnpike Authority (NJTA) completed a $500 million widening project. Construction took 3 years. This project can serve to illustrate several of the concepts of trade-offs that are available.

The widening project added six new lanes to an existing six-lane facility. The project was approximately 30 miles long. The end result is a dual/dual facility with four separated three-lane roadways. There are two three-lane roadways in each direction, each with its own separate exit and entry ramps to interchanges and service areas. Trucks and buses, incidentally, are limited to the outer roadways whereas passenger cars may use all roadways. In effect, therefore, it is a combined parkway and turnpike.

PROJECT ACCELERATION

Because all money is available at once, or at least available as required according to schedules developed by the engineer, it is possible and often advantageous to accelerate a construction project. In the case of the widening, design times were accelerated by dividing the project into 10 sections and simultaneously engaging the services of 10 engineering firms. Theoretically, design time was reduced by a factor of 10 and construction contracts could be scheduled so that work could begin simultaneously in all sections. Such a design approach is, of course, available to state highway departments as well. It is extremely doubtful, however, whether the approach would have been taken in this case if the project were financed through tax revenues. It does not seem feasible to spend more than $10 million to completely design a project when the financing of the project depends on future legislative actions based on fiscal considerations not even known when the project is begun. The design might be out of date when appropriations finally become available. A state highway department would have designed and constructed a section of the work, tailored to funds available at the time. The realities of politics dictate that a tangible result be shown as soon as possible. Several states have taken such an approach to the Interstate System; construction progresses on a haphazard basis with respect to the overall network, but relatively short stretches are opened, which alleviates local problems.

Another way to expedite a construction project is to acquire property in all areas at once. Real estate costs are spiraling everywhere and especially in the part of New Jersey immediately adjacent to New York City where this project was undertaken. The ability to immediately purchase all property at once, rather than to purchase on a section-by-section basis, is a great advantage.

Once the total project design is completed according to one integrated schedule, it is possible to begin construction in all sections simultaneously. This project took 3 years to construct. Had it been done in the sequence of completing a usable section, opening it to traffic, and then beginning construction of the next section, it would have taken about 12 years. Even this estimate assumes that design of the succeeding section would be completed during the construction of each section. The project would have to have been divided into six sections because of interchange locations, and the better part of two construction seasons would have been required for each section.

There were two not so obvious balancing factors at work in the project under discussion, however, that tended to cancel each other. By working on a crash basis on all sections simultaneously, we were, in effect, competing with ourselves for construction labor and equipment. This, of course, tended to increase construction costs. At the same time, we were in an inflationary period in which prices were escalating at about 12 percent per year. Conditions in the construction marketplace during the particular period in question must, of course, be analyzed for each project because these factors can fluctuate considerably.

Substantial savings can be made through better construction scheduling when the entire project is under way at one time and when the construction time is minimized. On the widening project, for example, it was possible to gain real savings in the area of hydraulic fill placement. Mobilization and demobilization for hydraulic fill contracts are substantial parts of the cost. Two dredging contractors were able to schedule their work under several contracts in more than one section so that mobilization and demobilization were minimized. Further, it was possible to more effectively balance upland cut and fill between all sections. Although this advantage was minimal on this particular project, on a road-building project in the Jersey meadowlands this concept was extremely important. Four of the sections required overload or surcharge in depths varying from 3 to 40 feet and for time periods of 9 months to 2 years. The 10 million yards of hydraulic fill used in the four sections was so scheduled that all overload was subsequently used as embankment. At the completion of the project there was no excess hydraulic material to be trucked away or stockpiled. This would not have been the case if contracts had not been scheduled in all sections simultaneously.

Another example of a cost-saving procedure through projectwide management is the ability to purchase scarce materials on a large scale. On this project we purchased all bearing piles required for the more than 100 bridges through a single purchase order. Piles were made available directly to the contractors, as required, on a cost basis agreed to by the authority and the vendor at the inception of the contract. This pile order, incidentally, was reported to be the second largest pile order in the history of the steel company.

SERVICE AND SAFETY

A toll authority must be very much aware of its patrons. If, through a maintenance or construction procedure, we diminish the road's ability to serve all the needs of patrons, we are obviously doing the NJTA a disservice. Contrary to the usual highway department approach, a toll authority cannot trade dollars for inconvenience. We cannot take the attitude that the motorist can decide for himself whether he wants to live with inconvenience or find an alternate route. Further, the contractor cannot be made responsible for convenience to the public or, more importantly, for safety. A toll agency, fully responsible under law and subject to legal action, cannot decide who can sue and under what conditions suit may be brought. NJTA is very proud of its safety record and is willing to assume the costs for the procedures and concepts outlined below.

Project acceleration provides the immediate advantage of minimizing construction time and, therefore, traffic exposure to construction conditions. Many of our decisions with respect to traffic handling, on this project and on all other Authority projects, are based on this concept.

The particular project under discussion did require three construction seasons, and, as is frequently the case, the construction season is unfortunately the season of highest traffic volumes. This project required changes in roadway direction, traffic detours from one roadway to another, bridge construction over live roadways, widening and reconstruction of bridges carrying traffic, and almost any other traffic-handling nightmare imaginable.

We determined at the outset that the following ground rules would apply universally in all sections.

1. Construction areas were always physically separated from traffic. This was accomplished by installing permanent steel guardrail, or, where that was impossible, temporary interlocked 12 by 12 timber barriers were installed adjacent to all construction areas.
2. Detours were designed for 60-mph speeds and had full 12-ft right shoulders and 5-ft left shoulders as did the existing turnpike. Except for some detour signing, the motorist did not necessarily know he was negotiating a detour from one roadway to another.
3. No lane closings were allowed between Memorial Day and Labor Day. Necessary asphalt overlays, ramp merges, and so on were scheduled for other time periods.
4. Traffic was stopped for a maximum of 5 minutes to allow for steel erection over live roadways. This requirement subsequently became unnecessary when we initiated

a traffic slowdown procedure. State police cars, traveling abreast, entered the roadway upstream from the work site and proceeded toward the area at approximately 30 mph. The gap between them and the 60-mph traffic in front of them could be regulated so that the contractor had the necessary time to erect steel. Obviously, a great deal of preplanning by the contractor and coordination by the engineer and the police was necessary. These procedures are now our standards and are used whenever necessary for construction or maintenance.

5. All standard turnpike authority lane-closing procedures and presigning for such lane closings were used. These procedures are too complex to describe here; however, an index of the importance we place on these procedures may be gained from the fact that 10 percent of our construction costs associated with work on live roadways are spent on traffic handling.

Besides traffic handling, there are several other areas worthy of mention in a discussion of acceleration and attendant pros and cons. We must, from time to time, make decisions where we knowingly accept less than optimum construction, time schedules, and patron convenience. Safety and revenue considerations combine to force us to make decisions where "first cost" or construction cost must take a back seat. This is acceptable as long as all facets of the problem have been evaluated and given their proper weight when the decision is made. We in the toll industry are exposed to these decisions as a consequence of the overriding necessity to meet completion dates.

During the course of this project, for example, we were faced with several serious soils problems in an area where the widening took the form of a new alignment through the Hackensack meadowlands. This is an area of virgin swampland where approximately 25 ft of peat or muck overlays clay that runs to a depth of over 100 ft. Although we used muck removal, sand drains, and overload, there were areas where acceptable consolidation simply could not be reached within our construction timetable. We accepted substantial settlements and the concept of asphalt overlays within 2 years of completion. This was done knowingly, and the cost of settlement correction and inconvenience were weighed against opening the facility to traffic and relieving congestion on the existing facility.

In the same vein we accepted a construction schedule that dictated finishing bridge and culvert construction at the same time paving operations were concluding. This led to relatively short paving areas and an almost checkerboard approach to paving. Obviously, such an approach is not conducive to the smoothest possible ride. We accepted the concept, however, as a trade-off against time and tolls. Although there are smoother roads, there are also similar or even rougher roads where neither tolls nor time was a factor.

Our standard, because of our need to attract patrons rather than to reluctantly accept traffic, is higher; therefore, our concerns with regard to smooth riding qualities are not shared by the public.

In the area of contract administration, too, we had to consider trade-offs. There were many instances in which differences in specification interpretation could not be fully adjudicated during the lifetime of the contract in question. When that occurred, the contractor proceeded as directed by the engineer and we accepted the exposure for claims inherent in such unilateral decisions. There were many claims, almost all of which have been satisfactorily settled. Again, inasmuch as the exposure was weighed against the overall benefit to the Authority and all other administrative procedures provided for in the contract had been exhausted, this procedure was correct.

In line with the concept just mentioned, several times we were faced with the necessity to accelerate construction. The contractor, through no fault of his own, was delayed. This could happen for a variety of reasons such as bad weather, unavailability of materials, or delay caused by an adjoining contractor. Under normal conditions, the contractor would have been granted an extension of time. Because time was the one commodity we could not expend, we were forced to conclude separate acceleration agreements wherein we accepted the costs associated with overtime, additional shifts, and the like.

These examples of contract administration procedures are important for another reason as well. We are a large construction user in New Jersey and our reputation with the contracting industry is very important. We attempt to present an image to the contracting industry of an agency with a no-nonsense approach to getting the job done on time, but at the same time an agency that will be entirely fair.

In summary, the differences in the approach to a construction project by a toll agency whose financing is based solely on revenue bonds and a government agency whose direction is set primarily by governmental and political considerations have been highlighted. The project discussed included a great deal of traffic involvement, but the comments would be pertinent to any construction project where the end result is a service facility.

Traffic-Handling Procedures During Maintenance and Reconstruction to Promote Safety and Minimize Delays

Charles McLean
Illinois Department of Transportation

In urban expressway maintenance projects, planning traffic control procedures requires as much time as making repairs. Therefore, these techniques must be constantly improved to minimize traffic delays and to accomplish the necessary improvements safely. This paper presents the Illinois Department of Transportation's urban Interstate maintenance project experience and uses the 1971 Kennedy-Dan Ryan Expressway rehabilitation project as a case study. Procedures for the planning of this project were based on experience gained from the 1966 Edens Expressway resurfacing project and involved a series of steps. Some essential components of this procedure include the formation of a task force to prepare the project contracts and the arrangement of meetings with various highway agencies and interested public and private transportation organizations. Proper timing of the elements of the planning effort ensures success of full news dissemination concerning the project. These efforts are designed to avoid congestion at the project's inception and to ensure that the motorist makes advance plans about alternate routes or modes of travel. The public information program is closely correlated with the success of the planned control measures.

The Illinois Department of Transportation has operational responsibility for the 135-mile (217-km) Chicago Metropolitan Expressway System, which carries more than 15 million vehicle-miles of travel daily and 25 percent of all of the vehicle-miles traveled within Chicago. Approximately 20 miles (32 km) of this system accommodates two-way daily traffic demands in excess of 200,000 vehicles (Fig. 1).

Drivers on urban Interstate facilities daily encounter peak-period delays caused by accidents, vehicle disablement, or lane closures for pavement repair. Figure 2 shows the peaking characteristics of an eight-lane urban expressway. Growth in urban Interstate travel demand has lengthened the peak periods, and public criticism of even off-peak lane closures for any type of maintenance is increasing. This public attitude plus the fact that the Interstate facilities are getting older and require more major preventive maintenance has emphasized the need for proper project planning. The public is more tolerant of congestion caused by accidents and disabled vehicles than it is of planned maintenance activities that require lane closures.

Figure 1. Maximum observed daily traffic volumes (in thousands of vehicles) in spring 1972.

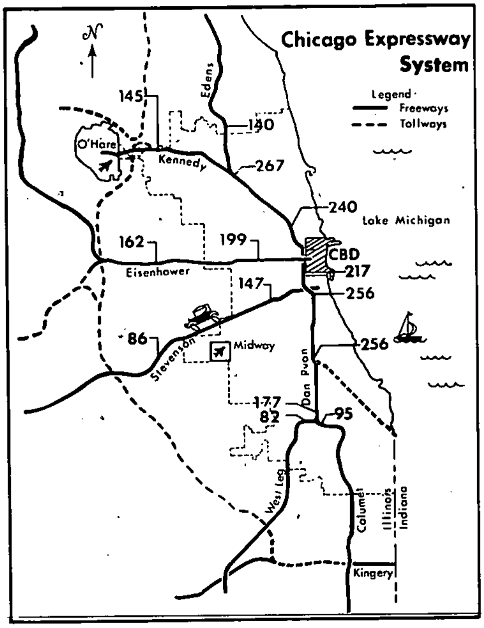
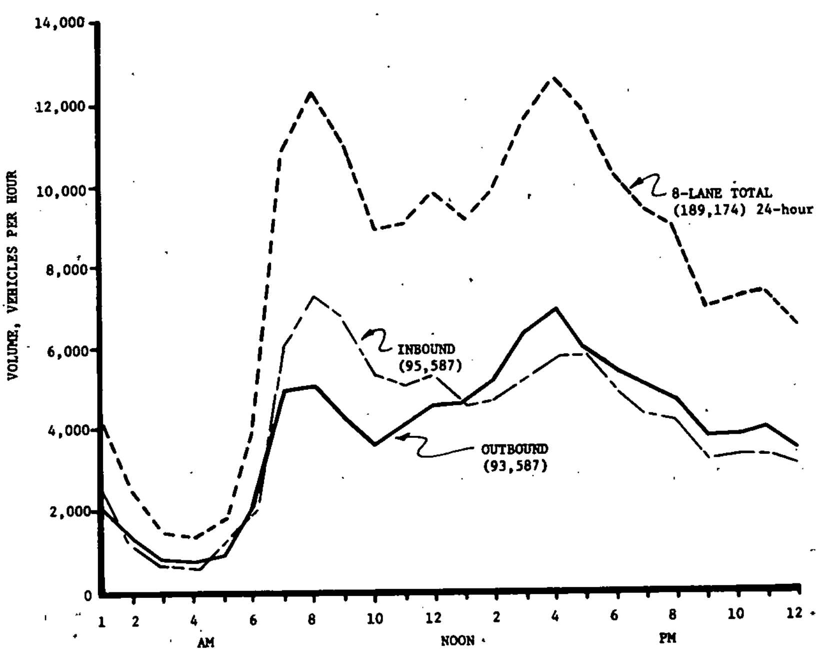


Figure 2. Hourly traffic volumes on Eisenhower Expressway at Sacramento Boulevard.



URBAN INTERSTATE EXPERIENCE

The Illinois DOT's urban Interstate maintenance experience began in 1966 when the Edens Expressway (I-94) was resurfaced. Because of the major traffic impact of this 90-day project for which two lanes of the six-lane facility were closed, we attempted to develop traffic procedures that would provide an acceptable level of traffic service and allow adequate working areas to expedite work completion. Use of critical path techniques for project evaluation allowed a review of the benefits of various means of work accomplishment from one lane to full roadway closures.

Since the Edens Expressway (I-94) project, we resurfaced the Eisenhower Expressway (I-90) in 1968, the Calumet Expressway (I-94) in 1969, and the Kennedy Expressway (I-94) in 1971. On these projects the work was phased within two lane closures. As funding allowed, we used the lane closures to accomplish all types of necessary maintenance and traffic work. Signing, guardrail modernization, bridge painting, sewer cleaning, bridge handrail construction, bridge waterproofing, pavement marking, and other improvements were performed during the resurfacing work.

Since the 1966 project, the motoring public has accepted the need for expressway maintenance. Project planning that emphasized traffic control techniques and public information efforts contributed greatly to this acceptance. Because we fully explained our intended actions before they were implemented and we provided informational assistance including control sketches and maps, we obtained cooperation from the news media. In fact, media representatives were disappointed when the expected chaotic congestion did not occur.

KENNEDY-DAN RYAN PROJECT

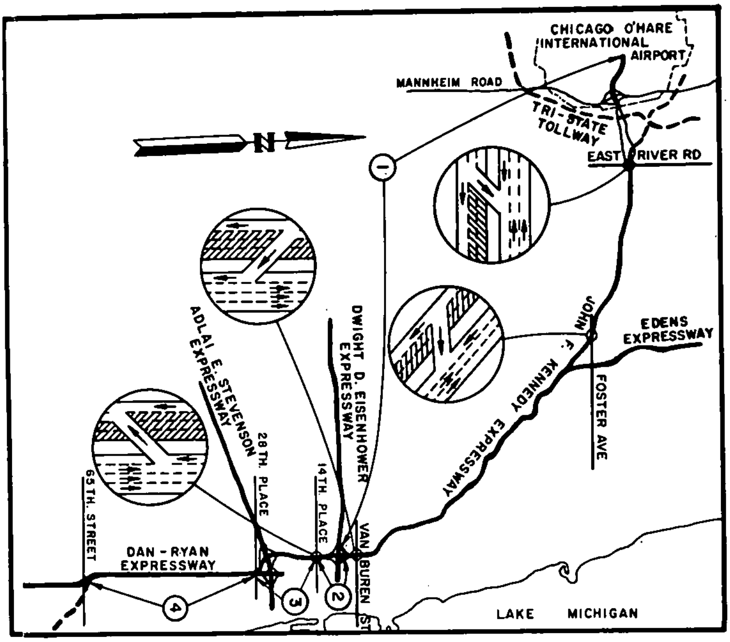
In 1971 the $16 million rehabilitation of the Kennedy Expressway (I-94) and Dan Ryan Expressway (I-90/I-94) was undertaken. To complete the project within the 15-week schedule, we planned two 10-hour shifts per day. The specific improvements (Fig. 3) are listed below.

1. Between Mannheim Road and the Eisenhower Expressway, the existing 10-in. (254-mm) portland cement concrete (PCC) pavement was structurally restored and resurfaced with 3 in. (76 mm) of bituminous concrete wearing surface. All main-line expressway bridges were repaired, and metal handrails were removed and replaced with PCC parapet walls. All cross-street bridges over the expressway were painted, and thermoplastic pavement marking was installed throughout the section.
2. Between the Eisenhower Expressway and 14th Place, the existing 8-in. (203-mm) continuously reinforced pavement was removed and replaced with 10 in. (254 mm) of continuously reinforced pavement on a 4-in. (102-mm) stabilized subbase; geometric modifications at the Chicago Circle interchange were included.
3. On the Dan Ryan bridge between 14th Place and 28th Place, the existing median guardrail was removed and replaced with a concrete barrier wall. Pavement surface and joint repairs were made.
4. From 28th Place to 65th Street, the existing 8-in. (203-mm) reinforced pavement in the local lanes was removed and replaced with a 10-in. (254-mm) continuously reinforced PCC pavement on a 4-in. (102-mm) stabilized subbase. No work was done on the express lanes in this area.

Project Procedures

Planning this 1971 project involved the following steps. A task force composed of design, construction, maintenance, and traffic personnel was formed to prepare the project contracts, to fully evaluate the necessary repairs, and to determine the traffic phasing and controls. A project traffic engineer was assigned for the duration of the project. A meeting was held with other highway agencies and interested public and

Figure 3. Crossover locations on Kennedy-Dan Ryan Expressway project.



private transportation organizations to discuss the impact of the proposed work and to identify possible conflicting work areas affecting alternate routes. Potential traffic control problems were reviewed with enforcement personnel.

Before contract bidding, a meeting was held with contractors to discuss the project work, phasing, and the importance of maintaining the required traffic controls. Briefings were held with regional communications center dispatchers, expressway emergency patrol supervisory staff, expressway surveillance personnel, and Kennedy-Dan Ryan Expressway maintenance yard supervisors to discuss the traffic control plans. In addition, individual radio traffic reporters were briefed on the details of traffic phasing, recommended alternate routes, and other modes of travel available.

The secretary of the Illinois DOT held a formal press conference to discuss justification for the Dan Ryan Expressway pavement replacement because of expected public interest. Film clips showing the specific problem areas of the roadway were distributed. Another press conference was held to announce the project approximately 1 week before work started. A department public information officer was designated to serve as liaison between news media and project personnel. Field trips were arranged after the conference to accommodate the individual media filming needs. On the day of the press conference, cross-street bridge-mounted signs read, KENNEDY-RYAN REPAIRS BEGIN JULY 5TH—ONLY 2 LANES OPEN. Advisory signs were also erected on approaching Interstate routes, including portions of the Illinois Tollway, to publicize alternate routes.

Private companies responsible for large changeable-message advertising signs adjacent to the Kennedy and Dan Ryan Expressways were requested to publicize the starting date of the project.

On the day the project began traffic engineers accompanied commercial radio helicopter traffic reporters to explain traffic control procedures. Up-to-date traffic reports were provided through the state regional communications center. The office staff was augmented to handle citizen inquiries about alternate routes, closed ramps, and access to high traffic generators.

Weekly work status reports were issued throughout the project, and any special project work features were highlighted. Field trips were arranged for the news media throughout the project to discuss project features with traffic engineers, construction engineers, and contractors.

Signs thanking the motorists for their cooperation were erected at the finish of the project. Press releases announced the project completion ahead of schedule, and the contractors earned bonuses for the early completion.

Project Communications Planning

The timing of some of the elements in the foregoing procedure was very important to ensure the success of full news dissemination regarding the project. We normally plan a press conference within a week of the start of a project and attempt to hold all project information so that maximum attention is ensured at the time of the formal announcement. If the press conference precedes the work start by too long a period, the press is not interested in republicizing the beginning of the work and highlighting the control details.

Because our efforts are oriented to avoid congestion at the project start and to ensure that the motorist makes advance plans about alternate routes or modes of travel, we give maximum publicity to the project immediately before the work start. Traffic control details such as which lanes and ramps will be closed, where the median crossovers are located, and what special traffic provisions are made for high traffic generators must be publicized extensively for motorist safety.

For a safe and congestion-free start on these large projects, we give great consideration to which day lane closures are initiated. There is a tendency to begin projects on a Monday. This detracts from the effectiveness of the informational program because the motorist has all weekend to forget the control information disseminated. He is more likely to remember radio messages regarding the project start if he hears them the previous day. Also, the heavy Monday and Friday traffic conflicts are avoided when the first barricades are placed.

There is a close correlation between the public information program and the success of the planned traffic control measures. A closed expressway lane is not a unique experience to the motorist. However, when an express lane is created by reversing traffic, ramps are closed, and median crossovers are operated over a 3-month period to facilitate traffic movement, good communications to the motorist are essential.

Lane Closure Staging and Control

The Kennedy-Ryan Expressway project required special staging of lane closures. The work was planned in two stages, each having two phases. We worked in one direction at a time, each direction representing one stage. Normally each work phase represented the number of major lane control shifts necessary to complete the work on one directional roadway. The staging shown in Figure 4 was complicated by the differing roadway geometrics, which varied from four lanes (area A near O'Hare field) to 10 lanes (area D where the median reversible lanes exist between two four-lane roadways).

It has been our practice to reverse traffic within a directional roadway (Fig. 5). A physical delineation created by tire-covered traffic cones establishes an express lane so that two traffic lanes are provided in each direction in a six-lane cross section having a two-lane work area. Signing legends for the typical median crossover are shown in Figure 5. Trucks were normally restricted from express lanes.

After the motorists gain experience in traversing the work area, traffic control problems do not develop until there are major project phasing changes. Traffic shifts to newly resurfaced lanes, changes in barricade placement, and redesignation of ex-

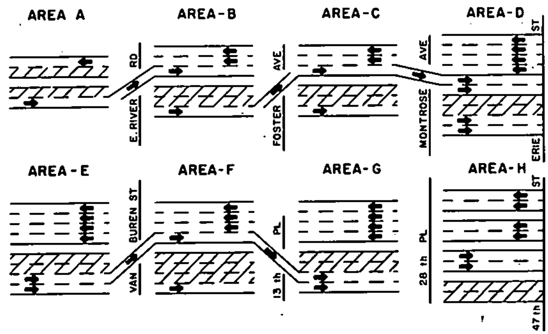


Figure 4. Traffic patterns for stage 1, phase 1 of rehabilitation project.

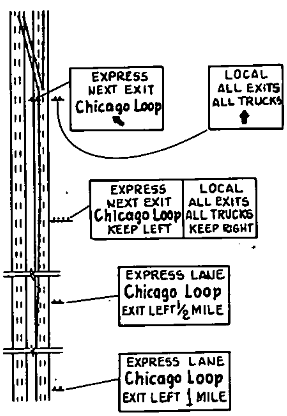


Figure 5. Crossover advance signing.

Table 1. Percentage change in weekday traffic counts.

Expressway Location	Normal ADT	Stage 1	Stage 2
Kennedy at Cumberland	130,000	-40	-25
Peak[a]		-50/-50	-40/-40
Kennedy at Cicero	125,000	-35	-25
Peak[a]		-35/-45	-25/-45
Kennedy at Ohio	210,000	-40	-45
Peak[a]		-40/-30	-35/-50
Dan Ryan at 55th	230,000	-35	-45
Edens at Wilson	115,000	-50	-35
Edens at Church	105,000	-5	-5
Eisenhower at Sacramento	190,000	0	+5
Eisenhower at East	125,000	0	+5

[a]Percentage change for a.m./p.m.

press reverse lanes create new control situations that require motorist understanding.

Speed limit reductions to 45 mph (72 km/h) within the construction area have been utilized on these resurfacing projects. Normally this limit is unnecessary during peak periods and too low in off-peak times. The enforcement agencies cooperate fully in providing the necessary control surveillance.

Expressway Volumes

The percentage changes in weekday traffic counts (both directions) were recorded during the 1971 Kennedy-Ryan resurfacing project (Table 1). Normally from midnight to 6 a.m. the volume change was slight. As indicated in Table 1, the biggest impact was during the peak periods. Traffic observations on the Edens Expressway reflected the effects of the adjacent Kennedy Expressway lane closure. Traffic increases on the Eisenhower Expressway resulted from traffic diversions to avoid the project congestion.

Special Traffic Activities

The Department of Transportation maintains a 41-vehicle fleet of emergency traffic patrol vehicles. Patrolling the expressway system on a 24-hour basis serves a valuable function during major reconstruction projects. Prompt removal of accident or disabled vehicles ensures minimum congestion on the already heavily used lanes. Patrolmen report work area traffic problems and any barricade or signing conditions that require remedial action by the contractor.

The Expressway Surveillance Project with its traffic detection and ramp-metering system provides continuous traffic flow information. This is utilized to evaluate work-site control effects and to report traffic conditions to the media.

The traffic control procedures developed during these urban expressway reconditioning projects require as much planning as the physical repairs. The techniques used must be constantly improved to minimize traffic delays and safely accomplish the necessary improvements.

An Approach to Concrete Pavement Replacement That Minimizes Disruption of Traffic

R. F. Grimsley and B. G. Morris
Florida Department of Transportation

The Florida Department of Transportation is currently using a method of replacing damaged concrete slabs on the urban Interstate System that results in minimum disruption to traffic flow and maximum safety for maintenance crews. The operation involves using accurate field measurements of damaged slabs to construct forms and pour new slabs in the maintenance yard rather than at the jobsite. These cured slabs are then hauled to the area being repaired and placed on prepared subgrade. Removal of the old slab and setting of the new one are done at night because of the lower traffic volume. As many as four slabs have been replaced during an 8-hour worknight. Slabs are set approximately ½ in. (12 mm) low and are raised to final grade by slab jacking. Joints are then sealed, and the road is ready for traffic. Careful planning and coordination of all crews involved in preparation and installation of slabs are necessary to ensure effectiveness in terms of smooth traffic flow and safety to workmen.

In maintaining today's highways, maintenance engineers are confronted with tremendous traffic volumes. Maintaining production and protecting workmen require that maintenance activities be scheduled when traffic is at its lowest point. Nighttime and weekends are the most opportune times for rerouting or restricting traffic flows.

A recent review of the Interstate System in an urban area in central Florida revealed numerous broken concrete slabs that needed repair or replacement. Until several years ago, slabs at this stage of deterioration were either ignored or leveled with asphaltic concrete. This method was fast and easy but did not prevent further deterioration or present a smooth riding surface to motorists.

A more professional and permanent approach to such maintenance problems involves closing one or more lanes of Interstate roadway to traffic while the deteriorating slabs are removed and replaced with a surface that can bear traffic loads as soon as it is in place.

Today this problem is being solved by precasting roadway slabs in an area remote from the jobsite (Fig. 1). During the hours when traffic volumes are low, the precast slabs are installed with minimum interuption to traffic flow. As many as four broken slabs have been replaced on urban Interstate Systems during the course of an 8-hour worknight. The cost of this method of replacement during 1972 was $35 per square yard.

After a particular slab is located for replacement, it is measured to obtain exact width and length. Interstate slabs are approximately 12 ft (3.66 m) wide by 20 ft (6.10 m) long, but smaller or irregular slabs may be replaced with this method.

Many slabs were cut, skewed, short, or irregular when initially constructed, so a precise measurement is necessary to construct a new slab that will fill the void left by the broken slab. A ¾-in. (19-mm) tolerance is allowed on two ends and one side, but the outside edge of a two-lane facility is adjacent to a paved shoulder and can be fitted snug to the precut shoulder edge. If slabs are replaced in six-lane roadway, a ¾-in. (19-mm) tolerance around the slab is required.

Measurements are given to the casting crews so forms may be set properly for a particular slab. Wooden forms were used for casting slabs for this particular job. However, if this were to be a continuous activity, metal forms would increase production. Slab widths on normal roadway are basically uniform, but lengths were found to vary from 4 to 5 in. (102 to 127 mm). Metal forms are easily adjusted for these varying lengths.

After a bed is prepared for casting a slab, a 6-mil (0.15-mm) Visqueen sheet is placed in the bottom of the form to prevent moisture loss, thus ensuring proper curing. A double mat of reinforcing steel is placed on chairs that are designed to provide necessary clearance for steel placement. These reinforcing bars were Number 5 bars (16 mm); transverse spacing was 6 in. (152 mm) and longitudinal spacing was 15 in. (381 mm). The steel was placed in two layers to compensate for uneven stresses caused by lifting from the bed and placing on an uneven base. This also prevented damage to slab when mud jacking for leveling.

In the casting bed, aluminum tubes were placed vertically in designated locations, making holes through the slab to allow for slab jacking of the slab. Location of these holes is very important, because it is necessary to maintain uniform support under the slab when jacking. Three lines of holes were cast into slabs for this purpose. They were located approximately 18 in. (457 mm) from the ends with 5 ft 8 in. (1.73 m) between holes on the side adjacent to the pavement. The center row was staggered 4 ft 3 in. (1.30 m) from the end with 5 ft 9 in. (1.75 m) between holes. The last row of holes was cast 12 in. (305 mm) from the side with the same distance between holes as the first. Holes usually are spaced a maximum of 6 ft (1.83 m) center to center so that uniform pressure will be exerted on the entire slab.

Six threaded sleeves were placed in the slab to allow lifting from beds and placing in the roadway. These sleeves were located in two rows 3 ft (0.9 m) from the longitudinal side and 5 ft (1.52 m) from the ends with one in the middle of the rows.

Casting required approximately 6 yd^3 (4.6 m^3) of concrete for a slab measuring 8 in. (203 mm) thick by approximately 12 ft (3.6 m) wide and 20 ft (6.1 m) long. Slabs were poured from 3,000-psi (2068-MPa) concrete followed by a vibrating screed and a light broom finish. The slabs were allowed to cure as specified in the department's standard specification.

Slabs were hauled to the approximate job location and stored near an entrance ramp in the area where they were to be installed (Fig. 2). Because the slabs were overwidth for movement on the highway at night, they were transported during daylight hours. Each slab weighs approximately 12 tons (10.8 Mg), so only two were hauled on each transport.

Traffic control installations were handled by a separate unit that preceded the slab installation by several hours. These crews were briefed on the lane closure plan before beginning the operation. In this briefing, they were reminded of proper procedures to follow and their impact on safety. Units consisted of a pickup truck with a roof-mounted sequential arrow light. It entered the Interstate System and raised the roof light as the work area was approached. When it stopped at the work site, the truck's sequential light directed the traffic to the left lane, while another truck and crew set up necessary trailer-mounted sequential lights and A-frame barricades for further traffic control. A sign crew installed detour signs that are turned away from traffic when not in use.

We used the Manual on Uniform Traffic Control Devices for Streets and Highways to

Figure 1.

Figure 2.

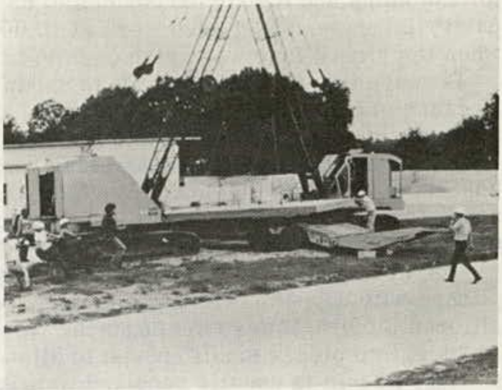


Figure 3.

Figure 4.

Figure 5.

determine the proper detour for the work location. It usually takes a couple of hours for the signs and detour barricades to be erected according to this manual. In this particular case, we started work at 10:00 p.m. and finished at approximately 6:00 a.m., when the closed lane was again open to traffic.

Because the work site was in an urban area, the width of roadbed was restricted. Because of its location in a fill section, all operations were coordinated and planned to allow equipment to pass and to offer maximum safety for men and traveling public.

Joints surrounding the broken slab were sawed with a 24-in. (610-mm) diamond concrete saw blade that cut through the dowel bars and the concrete slab. When slabs are removed, it is important to saw through the joint, or adjacent slabs will be broken or damaged when the slabs to be replaced are broken.

The next operation was to break up the damaged concrete slab. This was accomplished with a ½-ton (450-kg) mobile dragline and a headache ball. The operator positioned the dragline so the headache ball could be dropped on the designated slab and break it into pieces small enough to allow easy removal from the work area. Approximately 45 min is usually required to break a slab into pieces small enough to be loaded into dump trucks by a gradall (Fig. 3). Extreme care in dropping the headache ball is required to prevent damage to adjacent slabs.

After the broken slabs were hauled away, the subgrade was compacted by adding material and releveling. A plate compactor was used to compact the subgrade because the area was too small for other types of compaction equipment. Density under the slabs is important to ensure that no further settlement of the subgrade takes place when a new slab is installed. All slabs were installed approximately ½ in. (12 mm) below final grade. The slabs then were raised to final grade by slabjacking. This ensures uniform support under the new slabs.

The next operation consisted of placing the slabs in the prepared location (Fig. 4). A crane was used to lift the slab from the transport onto the roadway base. Heavy equipment such as the crane should be kept off newly placed slabs until the final grade is established.

Slabjacking is accomplished through holes arranged in the slab when poured. The area under the slab at each pumping point is blown out sufficiently to allow grout to be pumped into this void. The grout forms a lubricant under the slab, causing the slab to float on the subgrade. After all holes have been pumped and the slab is floating on the subgrade, the holes are pumped with a heavier grout to raise the slab (Fig. 5). After three to five piston strokes of the pump in each hole, the slab is normally raised to grade or the operation is repeated.

After the slab has been raised to the proper grade, a pressed fiberboard is placed in the joint. This serves as a filler and a bond breaker. The joint sealer used is a rubberized asphalt sealant that is heated in an oil kettle bath according to the manufacturer's recommendation. The sealer is poured ¼ in. (6 mm) below the top of the slab to prevent damage by traffic. The seal is allowed to cool and the equipment is removed from the jobsite.

The traffic control crew removes the barricades and signs and opens the land to traffic.

In conclusion, precast slabs for maintenance is one method of repairing broken slabs in a heavily traveled urban highway. This method of replacement offers the engineer the quickest and safest approach. The public is least inconvenienced by this method, because the work is accomplished during low-volume hours.

An activity such as this requires a coordinated effort by all involved crews. Each phase has to be planned to ensure maximum efficiency.

Advance Planning for Emergency Handling of Highway Traffic

Elmo C. Meister
California Department of Transportation

Advance planning for emergency detours on a highway system benefits the user and the personnel establishing the detours. Development of such plans requires preparation of maps, determination of critical points, location of detours for short- and long-term closures, review and coordination with involved law enforcement agencies, planning of news releases, establishment of cross-communications systems, and testing and revision of detour plans. Two major incidents proved the value of advance planning. In the case of a large forest fire, the planned detour was used. In another case, the advance planning process aided in improvising a new detour plan under emergency conditions. This occurred when a munitions explosion precluded the use of the planned detour. The handling of these emergencies demonstrated that advance planning of detours is beneficial whether preplanned or field improvised.

Advance planning for handling highway traffic when emergency closures occur is essential to the orderly movement of traffic. Closures can be caused by natural disasters, spectacular accidents, civil disorders, and organized blockades. Adequate advance planning for detours around these closures minimizes their effect on highway traffic and reduces the normal congestion that develops at these incidents.

District 03 of the California Department of Transportation is headquartered in Marysville, about 40 miles north of Sacramento. It includes 11 counties and 29 incorporated cities and is responsible for more than 1,400 miles of state highways. The district is divided into eight maintenance territories, each under the supervision of a maintenance superintendent.

Certain steps were followed for development of an adequate emergency plan. The first step was the appointment of a coordinator for plan development. He was responsible for preparing emergency detour plans for the entire highway system in the district.

Maps of the area were prepared on a large scale for the urban areas and a smaller scale for the rural areas. Critical locations, i.e., structures over large rivers and waterways, complex interchanges in urban areas, freeway-to-freeway separations, mountain passes, and any location that if closed would necessitate lengthy out-of-direction travel or other problems for the highway user, were identified on these maps.

Through the experience of maintenance and traffic engineering personnel, detours were planned around these critical locations, both for short-duration detours (1 to 3 hours) and for extended closures (days, weeks, or longer). The time element has a bearing on the complexity or simplicity of the detour route and the ease of following temporary signing. The longer the closure is, the greater the need is for an easier route to follow. Detours were also planned for short and long segments of freeways.

These plans were reviewed and discussed in detail with local law enforcement personnel (city police or California Highway Patrol) and, in certain situations, local fire department personnel. Input from these meetings is paramount to effective cooperation and coordination. Who does what? Who, in each organization, should be contacted? Who is responsible for providing the needed material and equipment? What types of signs are needed, and how many? Adequate intersection signing and "reassurance signing" (between intersections) are essential. Locations of storage areas for signs, barriers, barricades, and other materials to establish detours can be selected.

Planning news releases is beneficial. Keeping the public informed is essential to effective use of the detour. An informed public will know why the detour is necessary and how long it will last and, therefore, will be less likely to complain about inconveniences or out-of-direction travel. Bringing the district and state public information officers into the picture early relieves the operation personnel of the responsibility for keeping the media informed. News reporters clamor for up-to-date information; they do not like "old news."

Testing the effectiveness of advance planning occurs almost daily in urban areas where accidents or incidents, i.e., multiple-vehicle collisions, tanker-truck fires, overturned propane transporters, spilled cargo, and similar mishaps, usually require detours. Review of how the detour worked after each emergency provides for continuous upgrading of the total plan. The real test comes when nature washes out bridges and also closes secondary roads planned as detours or when a major explosion devastates a large area, including the highway and the planned detour.

The forest fire that closed US-50 is a good example of a test of advance planning for emergency closures. US-50 is mainly a two-way highway and is the direct route from Sacramento to South Lake Tahoe. It provides access to mountain recreation areas and the gambling casinos of Nevada. On August 18, 1973, the California Highway Patrol reported that a forest fire in the Kyburz area, east of Placerville, was out of control and that it might be necessary to close US-50. About 1 hour later, the highway was closed. The 1-hour delay gave the maintenance superintendent of Placerville the opportunity to contact the U.S. Forest Service Fire Control Center and establish communications for support of the closure and frequent updating on the status of the fire. The superintendent alerted maintenance crews of the need for equipment and materials to establish screening points and traffic diversions to selected detours. The transportation district office was notified, and the public information officer was alerted to the situation and the need for dissemination of current information to the news media.

Temporary signs were prepared for the diversion and closure points. Figure 1 shows the general area and the affected routes and the location of the fire.

At Placerville, eastbound traffic was diverted over Calif-49, either north to Interstate 80 at Auburn or south to Calif-88 at Jackson. Signs were placed in eye-catching locations preceding the diversion point to supplement the primary diversion point signing. At South Lake Tahoe, signs directed westbound traffic to Calif-88 and I-80 via Calif-89. Calif-89, -88, and -49 are two-lane highways in foothills or mountain regions, and the geometrics were not of a standard to handle the high-volume traffic that normally occurs on US-50. Allowing through traffic a choice of routes to circumvent the fire area tended to split the volumes into more acceptable numbers.

Next, closure and screening points were set up between the fire area and the diversion points (Fig. 1). These were required to handle local traffic to the small communities between Placerville and South Lake Tahoe. Manning of the points was jointly handled by California Highway Patrol and Department of Transportation maintenance men. Also, effective screening prohibited sightseers from filtering into the fire area and adding to the congestion of men and equipment assembled to fight the fire.

Information received from the screening and diversion points soon indicated that initial signing needed to be supplemented by extra signs between Sacramento and Placerville and in the South Lake Tahoe area east of the point of diversion. This additional signing provided advance notice and gave the motorist more time to decide which route he wanted to use.

This detour plan remained in effect for 9 days, including 2 weekends, which gave us time to observe the operation under both weekday and weekend use. Minor modifications of the detour plan were made for future closures. Throughout the duration, communications with California Highway Patrol and the U.S. Forest Service personnel were maintained by twice daily meetings. In addition to keeping our operational crews informed, it also provided up-to-date information to the media through our public information officer.

Another major test of our planning occurred on April 28, 1973. On that date a munitions-laden freight train caught fire and exploded in the switching yards at Roseville, approximately 20 miles northeast of Sacramento. The yard is located adjacent to I-80, one of the state's primary east-west routes between the San Francisco Bay area and Reno, Nevada. This multilane freeway normally handles an average daily traffic of 45,000 vehicles. Figure 2 shows a map of the general Roseville area.

April 28 was not a normal day. It was a springlike Saturday and the opening day of trout season, and the weekend traffic developed earlier and in larger numbers than usual. A majority of state maintenance men took advantage of their weekend off and the beginning of fishing season and were part of the general exodus to the high country. This contributed to a shortage of available manpower.

Tremendous explosions ripped through the railyard situated less than ¼ mile north of I-80. The first blast occurred at approximately 8:28 a.m., and the resultant detonation and debris from it and subsequent explosions forced the immediate and complete closure of I-80 and many other local streets and highways.

Previously adopted emergency plans and use of prearranged detour roads and traffic shifts were, of necessity, superseded by on-the-spot decisions regarding affected areas and the question of future major explosions of train cargo. Because of the many uncertainties, local residents were evacuated from ever-widening areas. This decision was responsible for establishing flexible traffic controls and communications able to meet the maximum safety requirements of the roadway users. It also provided free unrestricted access and maneuverability for the various emergency forces and their equipment to enter the critical areas.

Soon after the first blast, a coordinated command post was established by the California Highway Patrol approximately 1 mile southwest of the explosion location. This center was used by all the forces cooperating in the emergency situation. Through use of this combined center, it was possible to coordinate all group activities and press releases on an informed and timely basis.

Word of the initial explosion and its potential effect on the state highway was received by department of transportation maintenance supervisors approximately ½ hour after its occurrence. Within another ½ hour, these field supervisors had established direct contact with the California Highway Patrol command post and had informed the district office in Marysville. The district office organized a communications center for the department of transportation, manned by staff personnel, and effectively coordinated activities within the three maintenance territories affected by the explosion. This involved locating manpower and shifting of equipment from various areas to fill the needs established by the field operations. The district center was also responsible for keeping the headquarters office informed and assisted with the dissemination of statewide press information.

Once the decision was reached to close I-80, maintenance crews and the highway patrol coordinated to establish the necessary diversion points and detour routes. Eastbound freeway traffic was diverted at an interchange located about 3 miles west of the explosion site. There, motorists were detoured over a multilane county road (roughly paralleling I-80, approximately 1 to 2 miles to its south) and were directed back to the freeway at an interchange about 2 miles east of the railyard. Shifting and handling of

Figure 1. Map of area affected by Kyburz fire.

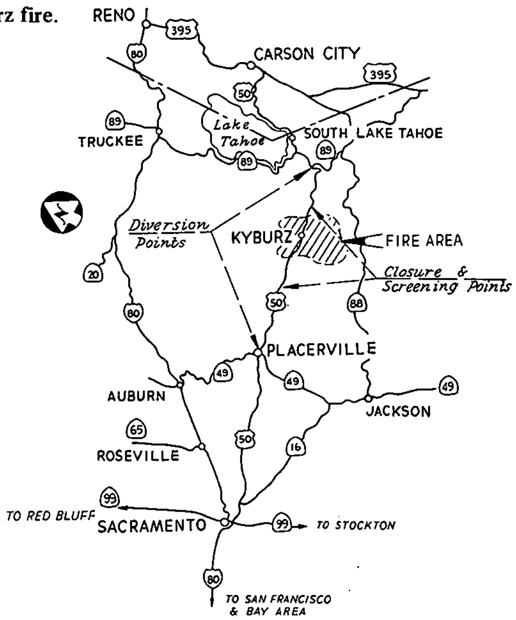
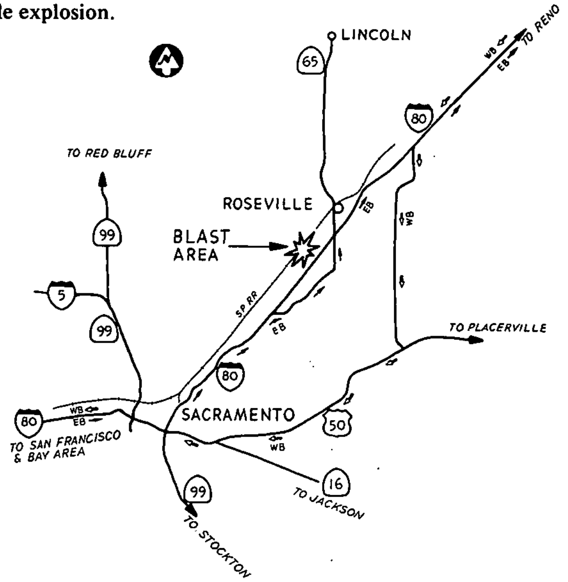


Figure 2. Map of area affected by Roseville explosion.



traffic were accomplished through a combination of control devices including barriers, state trucks with warning lights, temporary signing, law enforcement officers, department of transportation employees, and local volunteers. Basically this detour remained in effect all of the 28th and part of the 29th.

Coincident with the eastbound diversion was the initial rerouting of I-80 westbound traffic at an interchange just east of Roseville. The closure at this location was in effect for only a short while and then lifted, and traffic was able to use the freeway to the point where the detoured eastbound traffic was returning to I-80. Westbound motorists left the freeway and used the same county road that was handling the detouring eastbound traffic. Problems developed because of the numerous crossroad situations, local traffic conditions, and the over-capacity volumes using this county highway.

To alleviate the congestion and conflict between the overloaded, two-directional detour flow, we decided to completely separate the opposing traffic. Consequently a new diversion point was established about 5 miles east of the initial westbound closure location. There, interchange facilities were again used to direct these westbound motorists from the freeway onto a county road system. We used improvised signing and other traffic control devices, liberally supplemented by maintenance employee and law enforcement officer guidance, to direct motorists south over the county road, some 10 to 12 miles to US-50, a multilane freeway. On this major highway, the through motorist could then go directly west to Sacramento and points beyond. Local traffic had a choice of various street patterns leading to the perimeters of the closed section of I-80 and the adjacent restricted community areas.

The splitting of the detour traffic was paramount in alleviating what could have been a severe congestion and delay problem. Although westbound traffic was forced to use rural and somewhat substandard facilities (which at times carried over-capacity volumes), vehicle flow was smooth and there were no significant backups or snarls.

Did the advance planning pay off? Yes and no. Yes, the emergency detours used for the Kyburz fire were preplanned. They were basically the same used each winter when US-50 is occasionally closed by severe storms and heavy snowfall.

No, we did not follow the planned detour for the closure at Roseville because the explosion impacted a greater area than we had anticipated and closed our planned detour as well as the freeway. Yes, it did provide a communication system between law enforcement personnel and maintenance operations people. Yes, detour planning experience proved its value when the maintenance personnel, California Highway Patrol, and sheriff's office improvised other detours over secondary roads not affected by the explosions at the railyard. Despite the fact that the preplanned detour was not followed, the advance planning process did prepare the personnel involved to react swiftly and capably in an emergency. Their training resulted in efficiently establishing detours that satisfactorily carried high-volume Interstate highway traffic without major congestion or long delays. This is the proof that advance planning for emergency handling of highway traffic pays dividends.

Better Management of Traffic Incidents: Scope of the Problem

Conrad L. Dudek
Texas Transportation Institute

When a major incident causes a bottleneck, significant freeway congestion results even though unused capacity may exist on parallel routes within the freeway corridor. The time an incident occurs and how long it blocks a freeway lane are prime factors that influence the amount of congestion and delay that develops. It is shown that, the more frequently incidents occur, the more frequently congestion will result and that, the longer the duration of the incident is, the more likely severe delay is to occur. Incident detection and types of response systems are discussed in relation to the objectives of the freeway system.

Whereas public attention seems to be focused on the energy crisis and the environmental impact of highway construction, highway departments are taking a closer look at operations to make better use of existing facilities. Urban transportation is so complex that it defies meaningful definition. In the broadest sense, urban transportation is the movement of persons, goods, and services within and into and out of the urban area. Any system that provides adequate mobility for the compact concentrations of persons and goods within a relatively small area will necessarily be extremely complex.

In most urban areas, the capacity of the transportation system would be adequate to accommodate all movements required within the urban corridors if the use of space were distributed over time. If the traffic load could be spread out throughout the day, better use of time and space on the streets and highways would be achieved. Most movement is concentrated during the daylight hours and coincides with the starting and ending times of workers. Peaking transportation demand is a characteristic problem for all modes of transportation.

Most intercity traffic is carried through corridors in which multilane freeways serve as the major traffic carrier. The definition of a freeway corridor is different to different people, depending on their perspective. In many metropolitan areas, suburban cities have sprouted adjacent to a major city, thus creating corridors of traffic patterns. In other locations, there is a close proximity of metropolitan areas where the freeway serves as the major traffic carrier although alternate routes are available.

Peak-period congestion occurs daily and thus is quite predictable in both effect and duration. Freeway ramp control systems have proved their effectiveness in reducing recurrent congestion and, consequently, have improved the level of service afforded the motorist. Freeway corridor control systems are under development and are expected to further improve operational characteristics.

The occurrence of an accident or other lane-blocking incident on the freeway reduces the capacity of the freeway section significantly. Freeway incidents occur randomly and result in nonrecurrent congestion. Other events such as maintenance and construction, debris spills, adverse weather or pavement conditions, and natural disasters also result in reduced capacity. When a major incident causes a bottleneck, significant freeway congestion results even though unused capacity may exist on parallel routes within the freeway corridor. Not all incidents result in significant delay; however, each creates queuing on the freeway, which is a serious traffic hazard to uninformed motorists.

Maintenance and construction activities, adverse weather conditions, and natural disasters reduce capacity as well as create safety hazards for uninformed motorists. Some of these events and conditions may warrant complete closure of the freeway.

FREQUENCY AND CHARACTERISTICS OF FREEWAY INCIDENTS

Information on the frequency and characteristics of freeway incidents is documented in several reports (1-6). Studies of a 6-mile section of the Gulf Freeway in Houston, having an ADT of 120,000, revealed that approximately 13 lane-blocking incidents occur per week between 6 a.m. and 7 p.m. (1). On the average, at least one major incident occurs per week from 6:00 to 8:30 a.m. on the inbound lanes of the freeway. Approximately 80 percent of the incidents reduce the directional capacity of the freeway by at least half. High rates of freeway incidents have also been reported by other authors.

The effects of a lane-blocking incident are significant. Goolsby (2) reported that a one-lane blockage reduced the capacity of a three-lane section of freeway by 50 percent, although the physical reduction in usable lanes was only 33 percent (Table 1). Analysis also revealed that the capacity reduction caused by a lane-blocking accident is not significantly different from that due to a stalled vehicle. An accident that blocks two of three lanes reduces the capacity by 79 percent.

When an incident occurs is also important. For example, an incident occurring at the beginning of the peak period will cause more delay than one occurring at the end of the peak period or during the off-peak period. Figure 1 shows the periods of the day that a typical six-lane urban freeway is susceptible to congestion due to an incident. Incidents occur throughout the day on the Gulf Freeway, and most incidents reduce the directional capacity from 5,600 to 2,880 vehicles per hour or less (1). Volume counts on the inbound Gulf Freeway (Fig. 1) show that volumes exceed the incident capacity of 2,880 vph from 6 a.m. to 7 p.m.

Another factor that influences the amount of congestion and delay that develop is the duration of the incident. The longer the duration is, the more severe are the resulting congestion and delay for a given level of demand.

The consequences of incidents are congestion, delay, shock waves in the traffic stream that lead to induced accidents, and other adverse effects. The following hypothetical incident on the inbound Gulf Freeway illustrates some of the relationships involved. It is assumed in Figure 2 that a stalled vehicle requiring police assistance occurs on a lane of the inbound Gulf Freeway at 7 a.m., the beginning of the peak period. The total delay that results is the area between the normal traffic demand curve and the capacity curve. When the stall occurs, the slope of the capacity curve drops, reflecting a reduction in freeway capacity from approximately 5,600 to 2,880 vph. The slope of the capacity curve returns to normal when the stall is removed 18.3 min later (the average duration for a stalled vehicle on the Gulf Freeway). This hypothetical incident would result in 800 vehicle-hours of delay and an average delay per vehicle of approximately 8 min.

These results show that the frequency and duration of incidents occurring on a freeway are primary factors in determining the operating conditions of the freeway. The more frequently incidents occur, the more frequently congestion will result. The longer the duration of the incident is, the more likely severe delay is to occur.

Table 1. Available capacity on inbound Gulf Freeway during different incident conditions (2).

Condition	Number of Incidents	Sample Size	Average Flow Rate (vph)	Capacity Reduction (percent)
Normal flow		312	5,560	–
Stall (one lane blocked)	4	43	2,880	48
Noninjury accident (one lane blocked)	17	167	2,750	50
Accident (two lanes blocked)	6	53	1,150	79
Accident on shoulder	23	254	4,030	26

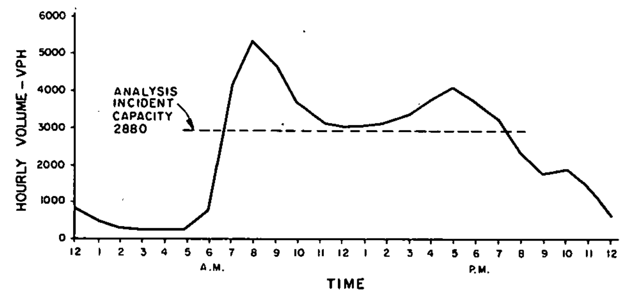


Figure 1. Traffic volumes of inbound Gulf Freeway at Griggs Overpass (1).

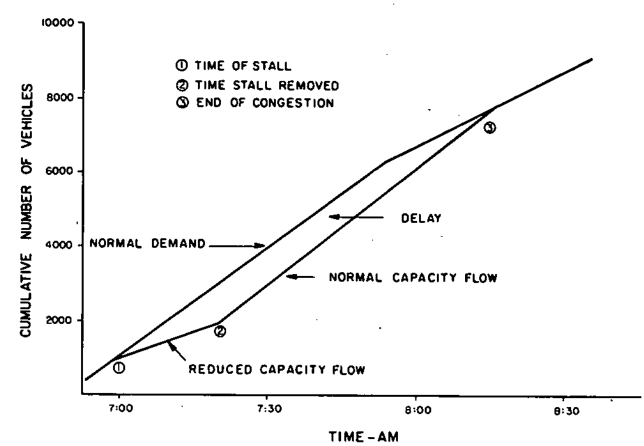


Figure 2. Example of delay caused by a stalled vehicle blocking one lane of inbound Gulf Freeway at 7 a.m.

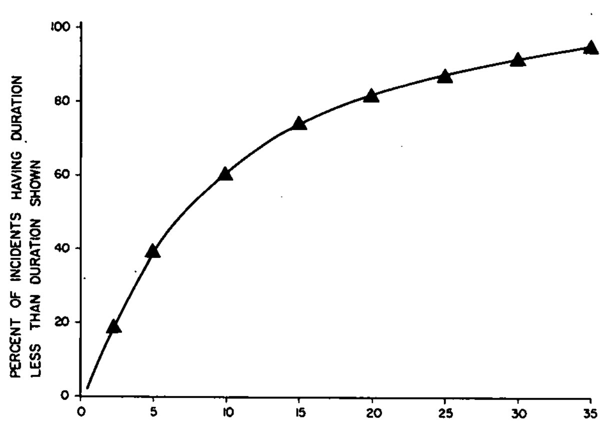


Figure 3. Cumulative distribution of the duration of incidents on Gulf Freeway.

Accidents and stalled-vehicle incidents that require police assistance may exist for a considerable time. Studies conducted by TTI (2) indicated that an average accident requiring police assistance takes 19 min from the moment the accident occurs until it is removed from the freeway. An additional 25 min, on the average, is required to complete the accident investigation. Figure 3 shows the duration of incidents observed on the Gulf Freeway. In earlier studies, Lynch and Keese (7) observed that an average of 45 min was required to remove damaged vehicles from the freeway when emergency vehicles are required.

SOLUTION APPROACH

From a traffic management viewpoint, when an incident occurs on the freeway the vehicles must be removed as quickly as possible, freeway demand must be intercepted before it reaches the reduced capacity caused by the incident, and the demand must be redirected to areas of excess capacity in the freeway corridor. In addition, from a safety standpoint, motorists approaching the queue area resulting from the incident need information.

Incident management consists of the method to detect the incident, offers a means by which the scope and needs are identified, and provides the appropriate response to minimize the adverse effects of the incident. To accomplish these objectives requires corridor surveillance, control, and information.

The surveillance function is required to detect and evaluate the nature of the operating characteristics, to detect any unusual conditions, and to determine the appropriate operational control strategy. The control function provides the response in terms of incident removal, motorist aid, and adjustment to the traffic controllers located at freeway ramps and intersections along alternate routes that will accommodate the short-term changes in traffic patterns. Driver information systems perform a critical role in the successful operation of real-time freeway traffic control systems. The real-time information system provides information to motorists that will enable them to intelligently select and follow the best alternate course of action, from rerouting through the corridor to diverting to another major facility.

Incident Detection

Vehicular incidents can be detected through

1. Electronic surveillance,
2. Closed-circuit television,
3. Aerial surveillance,
4. Emergency call boxes,
5. Emergency telephones,
6. Cooperative motorist aid systems,
7. CB radio, and
8. Patrol vehicles (police, mechanical service, maintenance).

Advantages and disadvantages of each method are discussed by Everall (8) and will not be elaborated on here. It is apparent that some methods provide better detection capabilities; others allow more detailed analysis of the scope and the required assistance. Cost-effectiveness analysis pursuant to the objectives of any proposed system would be necessary to determine the best approach or combination of approaches.

Incident Response

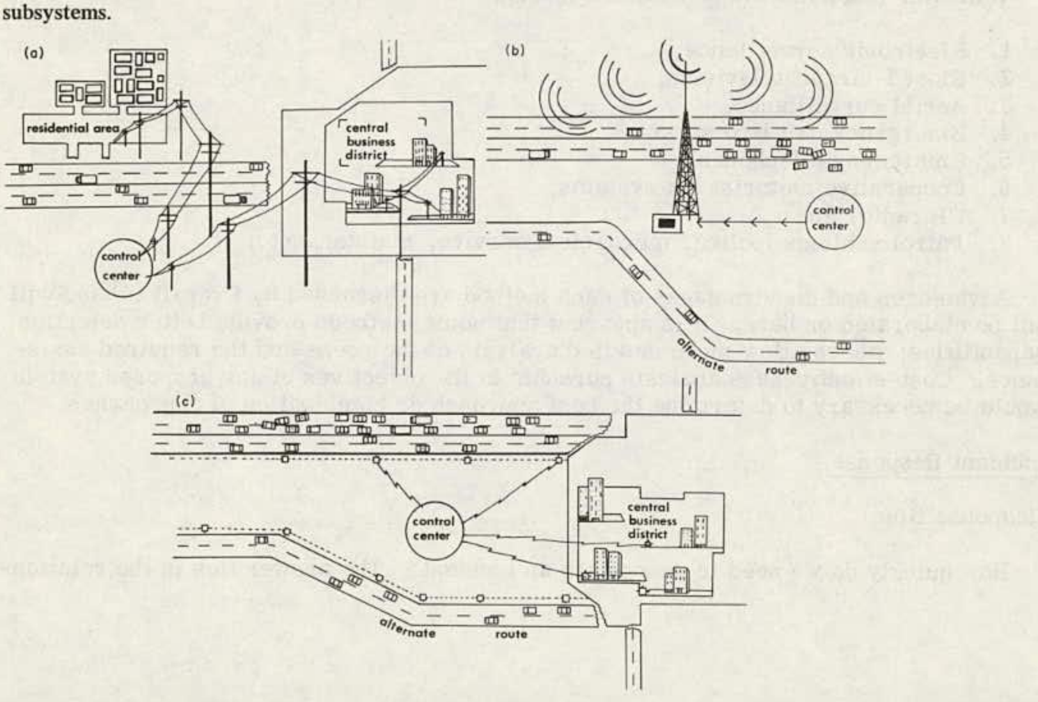
Response time

How quickly do we need to respond to an incident? The answer lies in the relation-

Figure 4. Changeable-message signs in (a) Dallas and (b) Houston.

Figure 5. Safety warning sign on the Gulf Freeway.

Figure 6. Schematics of (a) telephone call-in, (b) commercial radio, and (c) roadside radio subsystems.



ship between required response time and system design. The speed of response is dictated in part by the objectives of the system. If the system is designed to warn approaching motorists of stopped vehicles on the freeway, then the response time must be short.

Response time, of course, includes the time required to detect the incident. It also includes dispatching assistance and removing the involved vehicles. Response time is dictated by the requirements of the system and consequently will affect the cost. A system objective to remove all incidents from freeways during the peak period within 10 min after they occur will cost more than permitting a 20-min response time. The relationship between response time and cost for alternative designs must be determined.

Type of response

Incidents may be serviced by police and highway patrol vehicles, tow trucks, or state-operated maintenance vans. Normally, more than one department of an agency or more than one agency is involved. In other words, successful incident management requires the full cooperation of several government groups. This is difficult to achieve in many cases because of the different priorities within each of the government units involved.

As mentioned earlier, incident response also involves communication with the motorist. The prime methods used to transmit information to the driver are visual and audio. Real-time motorist information displays can be classified according to the relative location to the vehicle at which information is displayed to the driver. These might be classified as external or in-vehicle systems.

Based on a limited amount of human factors input, prototype changeable-message sign systems have been designed and installed in several cities (9). Figure 4 shows the rotating drum signs located along a parallel arterial in the North Central Expressway in Dallas. A prototype warning system used on the Gulf Freeway is shown in Figure 5.

There have been accelerated interest and progress in the development of auditory modes of real-time driver information displays. Audio modes of communication can range from telephone systems for pretrip planning to sophisticated in-vehicle audio systems for a myriad of functions. Conceptual designs and considerations for audio systems (Fig. 6) have been discussed by other researchers (11, 12) and by W. Wolman in a letter to the 1972 HRB Committee on Communications.

Only limited experimentation and application of real-time audio communication other than the commercial radio broadcasts have been implemented in the United States. An example is the low-power radio communication system at the Los Angeles International Airport (13). The California Division of Highways used low-power radio transmitters during and after an earthquake in the early 1970s to advise motorists of conditions along a hazardous route (14). To date, no full-scale field installation, test, and evaluation of these audio modes have been made for highway applications.

SUMMARY

The scope of the incident problem in freeway corridors has been briefly discussed. Approaches to incident management vary in different parts of the United States. The following are a few challenges that need to be addressed so that effective traffic management systems can be implemented and operated.

1. What are the optimal system configurations for incident detection and response?
2. What are the trade-offs between response time and cost?
3. What are the total benefits of freeway patrols, call boxes, closed-circuit television, and so on, and how do we evaluate these on a common basis so that alternatives can be considered from a cost-effectiveness standpoint?
4. What level of reliability can be expected from the various alternatives, and what

maintenance problems and costs are involved?

5. How can government agencies and others coordinate activities and establish priorities for effective incident management?

REFERENCES

1. Messer, C. J., Dudek, C. L., and Loutzenheiser, R. C. A Systems Analysis for a Real-Time Freeway Traffic Information System for the Inbound Gulf Freeway. Texas Transportation Institute, Research Report 139-5, April 1971.
2. Goolsby, M. E. Influence of Incidents on Freeway Quality of Service. Highway Research Record 349, 1971.
3. Goolsby, M. E., and McCasland, W. R. Evaluation of an Emergency Call Box System. Texas Transportation Institute, Research Report 132-1F, December 1969.
4. DeRose, F., Jr. An Analysis of Random Freeway Traffic Accidents and Vehicle Disabilities. Highway Research Record 59, 1964.
5. Kuprijanow, A., Rosenzweig, S., and Warskow, M. A. Motorists' Needs and Services on Interstate Highways. NCHRP Report 64, 1969.
6. Shufflebarger, C. L., and Bergsman, S. W. Shoulder Usage on an Urban Freeway. John C. Lodge Freeway Traffic Surveillance and Control Research Report, Study 417, January 1962.
7. Lynch, F. L., and Keese, C. J. Restoring Freeway Operation After Traffic Accidents. Texas Transportation Institute, Bulletin 28.
8. Everall, P. F. Urban Surveillance and Control: The State of the Art. U.S. Department of Transportation, Federal Highway Administration, November 1972.
9. Changeable Message Signs—A State of the Art. Highway Research Circular 147, September 1973.
10. Dudek, C. L., and Biggs, R. G. Design of a Safety Warning System Prototype for the Gulf Freeway. Texas Transportation Institute, Research Report 165-4, May 1972.
11. Dudek, C. L., and Carvell, J. D., Jr. Feasibility Investigation of Audio Modes for Real-Time Motorist Information in Urban Freeway Corridors. Dallas Freeway Corridor Study, Report RF 953-8, April 1973.
12. Ristenbatt, M. P. A Comprehensive Motorist Vehicle Communication System.
13. Campbell, R. Roadside Radio Smooths Traffic Flow. Popular Science, Vol. 202, April 1973.
14. Roper, D. Traffic Management in Los Angeles. Presented at Advanced Traffic Engineering Short Course, Houston, February 1973.

Incident Surveillance and Control on Chicago-Area Freeways

Joseph M. McDermott
Chicago Area Expressway Surveillance Project

Since 1961, Illinois has had considerable operational experience with surveillance and control of incidents occurring on Chicago-area expressways. In that year, the Illinois radio communications center was joined by two new units, the Emergency Traffic Patrol and the Expressway Surveillance Project, whose duties are to handle expressway incidents and increase expressway efficiency and safety by means of electronic traffic aids. The initiation and subsequent expansion of these three units recognized that urban freeways did not operate by themselves and that attention to continual maintenance and operational responsibilities was needed. This report describes various portions of the Illinois Department of Transportation program for reducing congestion, improving flow, increasing safety, expediting emergency responses, and providing motorist information. The considerable operational experience and refinements to date, as related to handling Chicago-area freeway incidents, are reflected in the contents.

As early as 1920, the importance of emergency traffic vehicles was recognized (1):

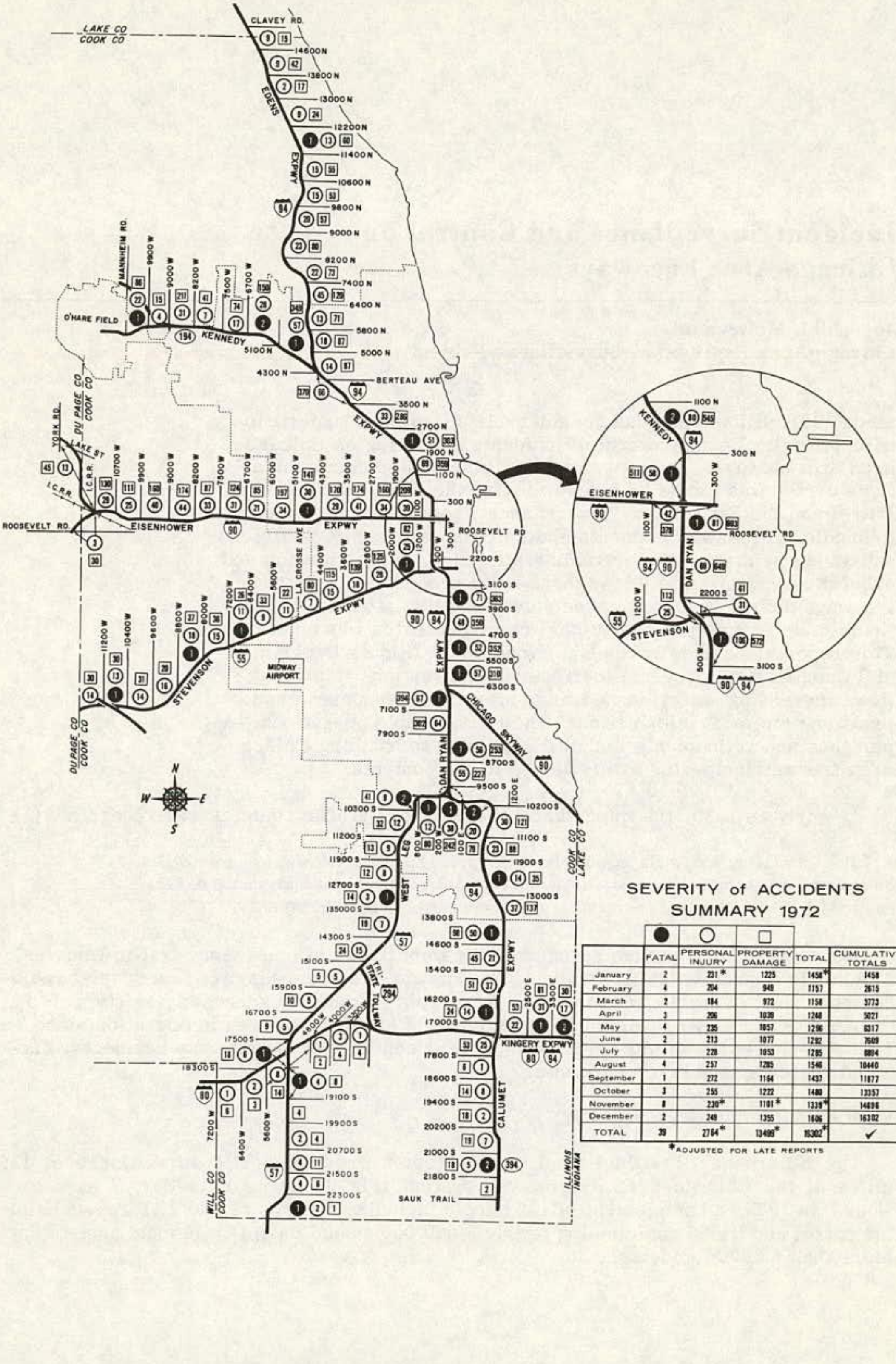
> Often vehicles break down in crowded thoroughfares at busy hours causing expensive delays. Emergency repair vehicles equipped with derricks and other suitable appliances should be kept on hand by all large cities as an economical measure to quickly relieve the trouble.

The state of Illinois also recognizes the importance of emergency traffic vehicles. The key to managing traffic incidents on Chicago-area expressways is a special radio-equipped fleet of trucks, continuously on patrol to promote expressway safety and to expedite traffic flow. This unique public patrol service has been in operation since 1961 and serves as an international model for coping with expressway accidents, disabled vehicles, and other incidents.

EMERGENCY TRAFFIC PATROL

The Emergency Traffic Patrol or minutemen provide mobile surveillance of 135 miles of the Chicago-area expressway system (Fig. 1) 24 hours a day, 7 days a week. In 1972, a complement of 135 people including 88 drivers and 21 foremen helped the patrol and its 56 vehicles log nearly 3,000,000 annual patrol-miles and assisted at more than 68,000 incidents.

Figure 1. 1972 Chicago-area expressway system accidents.



SEVERITY of ACCIDENTS
SUMMARY 1972

	FATAL (●)	PERSONAL INJURY (○)	PROPERTY DAMAGE (□)	TOTAL	CUMULATIVE TOTALS
January	2	231*	1225	1458*	1458
February	4	204	949	1157	2615
March	2	184	972	1158	3773
April	3	206	1039	1248	5021
May	4	235	1057	1296	6317
June	2	213	1077	1292	7609
July	4	228	1053	1285	8894
August	4	257	1285	1546	10440
September	1	272	1164	1437	11877
October	3	255	1222	1480	13357
November	8	230*	1101*	1339*	14696
December	2	249	1355	1606	16302
TOTAL	39	2764*	13499*	16302*	✓

*ADJUSTED FOR LATE REPORTS

The primary objective of the emergency patrol fleet is to keep moving lanes clear of accident vehicles, disablements, and debris. Secondary patrol duties include

1. Assisting at accident scenes by rendering first aid, calling for police, fire, ambulance, or special equipment services, helping to extricate trapped or injured persons, and supplementing police traffic control;
2. Removing accident and nonaccident debris from the roadway by calling for extra clean-up help or special equipment, sanding oil slicks, and removing or assisting with removal of dead animals, large or small dropped items, and spilled loads;
3. Assisting motorists by towing disabled vehicles and abandoned vehicles from hazardous locations, providing gasoline, tire changing aid, and water for overheated radiators, calling for repair service, lending tools for minor repairs, getting other help, and notifying other parties;
4. Establishing emergency traffic detours and controls by placing appropriate temporary traffic cones, barricades, signs, and lights and closing ramps or lanes;
5. Assisting at special expressway maintenance or construction work and traffic routings by protecting workmen placing traffic controls and helping to route presidential motorcades;
6. Reporting traffic information to the Communications Center for distribution to operations personnel and the news media;
7. Reporting faulty state property including damaged signs, fencing, and guardrails, inoperative signals and lighting, pavement defects, and drainage problems;
8. Warning pedestrians to keep off the expressway and notifying enforcement authorities when persons or vehicles do not voluntarily comply with patrol requests; and
9. Providing travel information such as directions and road conditions and map reading assistance to motorists seeking aid.

Services

The patrol vehicles are equipped and the drivers are trained to handle almost any emergency incident likely to occur on Chicago-area expressways, including accidents, disabled vehicles, and small fires. The service is termed emergency because assistance is directed toward actual emergencies and hazardous situations. Enough help is provided to remove or reduce the exposure to high-volume, high-speed traffic. Towing, for example, is only for moving vehicles to shoulders or frontage roads; the motorist must arrange for private towing from there.

The patrol drivers are not mechanics and do not repair disabled vehicles. However, they do provide gasoline, water, air for tires, and the loan of some small tools, which help many disabled motorists get to a service station or garage. All patrol services are free of charge except for gasoline. Upon receipt of 2 gallons of emergency gas, a motorist is presented with an invoice requesting that $2.50 be mailed to the state. More than 75 percent of these charges have been voluntarily paid.

Operations

The patrol fleet of 56 vehicles includes 27 tow, 14 flatbed, and 12 pickup trucks. The basic truck for 41 units has a 1½-ton chassis capable of handling a gross weight of 18,000 pounds (Fig. 2). Equipped with two-way radio, public address system, tow rig, pusher bumpers, rear pintle hooks, and numerous truck and driver equipment items (Tables 1 and 2), each unit represents an expenditure of nearly $9,000. The 1974 budget for patrol personnel and equipment was nearly $1.5 million.

In 1972, the 88 drivers operated over 18 routes on overlapping shifts, which provided a minimum of 25 rush-period patrols. The patrol routes also overlap (Fig. 3) to increase coverage of high-incident sections, such as a 2-mile-long bridge without shoulders on the Dan Ryan Expressway. Foremen patrol the entire system, usually in

Figure 2. Emergency patrol vehicles and operations.

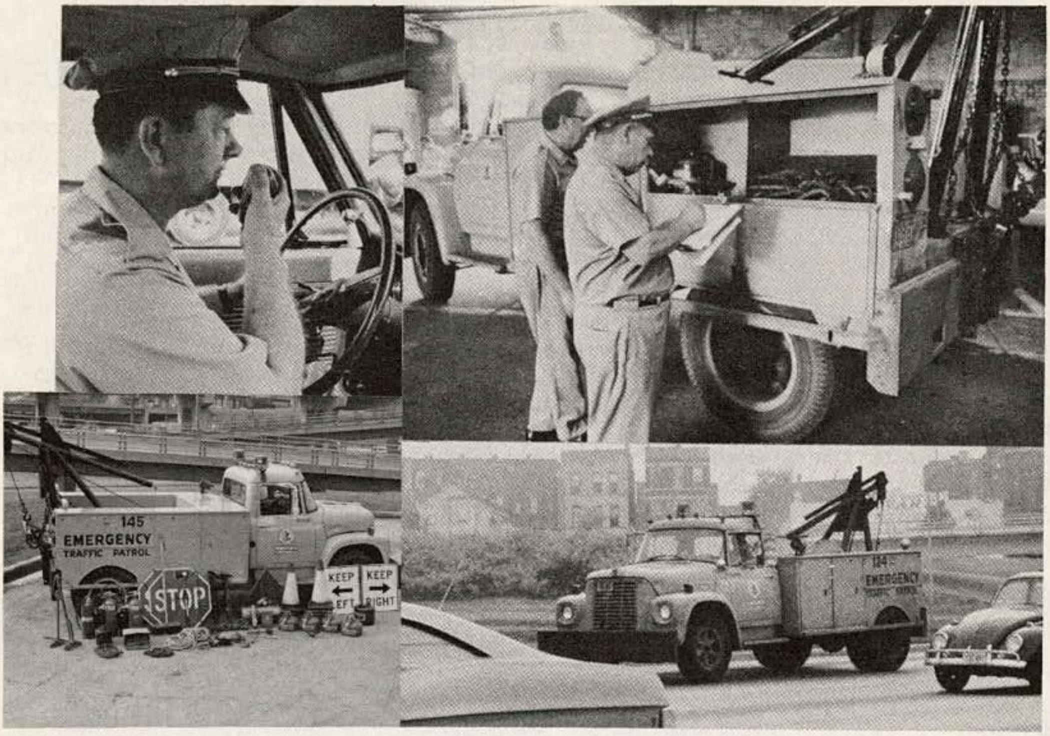


Table 1. Patrol truck equipment.

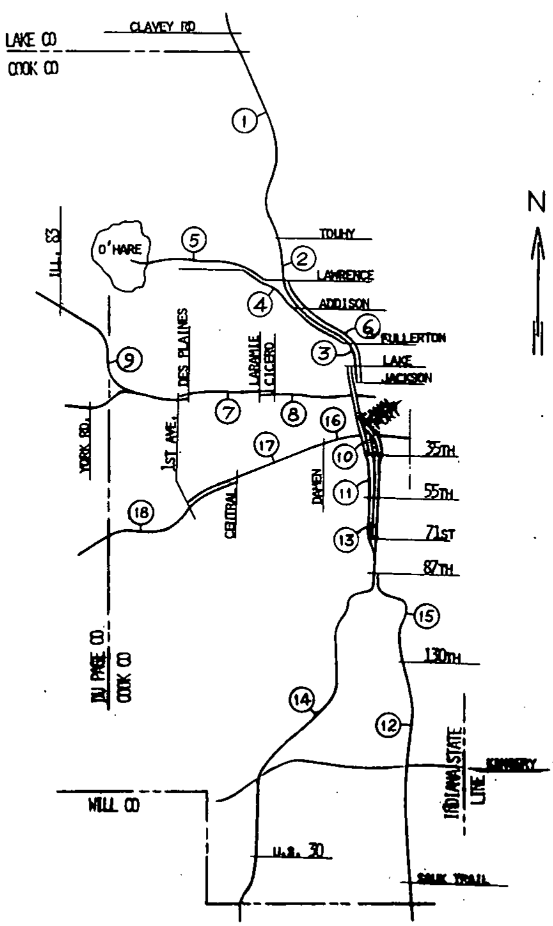
Item	No.	Item	No.
Overhead carrier	1	Snatch blocks	2
Mars lights	2	100 ft of 3/8-in. rope	1
Dietz lights, No. 510	2	50 ft of 5/8-in. rope	1
Spot mirrors	2	1/4-in. rope	1
12-S flasher can	1	3/8-in. tow chain	1
Push-pull switches	2	1/2-in. tow chain	1
Fuse block	1	Long wrench	1
Complete first aid kit	1	Rubber mallet	1
2 3/4-lb purple K powder fire		Tire chocks	1
extinguisher	1	Tripod	1
20-lb CO_2 fire extinguisher	2	Bumper jack	1
Fire axe	1	Traffic signs	
Pry bar	1	KEEP RIGHT	1
5-gal water can (summer)	1	KEEP LEFT	1
2-gal emergency gas cans	4	STOP	2
Air pressure tank	1	Traffic cones	10
Toolbox	1	Red flag	1
Hammer	1	Fuses	36
Common pliers	1	Highway maps	10
Channel lock pliers	1	Whisk broom	1
Wire cutter	1	Paper towel	1
Adjustable 10-in. wrench	1	Shovel	1
Phillips screwdrivers	2	Street broom	1
Screwdrivers	2	Bags of sand	10
Linoleum knife	1	Bags of salt (winter)	10
Electric tape	1	Hand cleaner	1
Wire (stove pipe)	1		

Table 2. Patrol driver equipment.

Item	No.	Item	No.
Utility coats	2	Flashlight	1
Zip-out lining	1	5-in. wand	1
Winter cap	1	Flashlight batteries	2
Garrison cap with 2 covers	1	Clipboard	1
Winter trousers	3	Code cards	2
Winter shirts	3	Badge and shield	1
Summer trousers	5	State shoulder patches	12
Summer shirts	5	Rain jacket	1
Eisenhower jackets	2	Rain pants	1
Neckties	3	Rain cap cove:	1
Safety vest	1		

Note: Uniforms have been replaced with coveralls with reflective safety stripes and hard hats.

Figure 3. Emergency patrol routes (day shift).



pickup trucks, and provide supervision, guidance, and assistance to the patrol drivers. They also help direct special traffic operations, such as changing the coning and barricades for reversing the flow on the Kennedy Expressway.

Driver Training

To handle the various duties and hazards common to urban freeway operations, patrol personnel receive special training in patrol procedures and operational techniques. Periodic classes provide training in basic first aid, fire fighting, state and city police coordination, radio communications, basic traffic control, equipment use and emergency procedures, Chicago Transit Authority operation (rail transit in expressway medians), highway nomenclature and area geography, expressway design and operation essentials, and other related subjects. Completion of a trainee's curriculum requires a month of on-the-job training in a patrol vehicle with an experienced professional.

Assist Characteristics

The patrol reported 68,474 assists in 1972 for the following causes:

Problem	Percent
Vehicle disabilities	77.2
Accidents	14.9
Abandoned vehicles	3.0
Nondisabilities	1.1
Pedestrians	0.4
Fires	0.3
Debris/other	3.1

Vehicle disabilities were of the following problem types:

Problem	Percent
Fuel/mechanical	34.0
Tire	27.3
Out of gas	12.7
Cooling system	10.3
Electrical system	7.4
Unknown/other	8.3

Incident Coordination

Unlike units that must be activated for emergencies, the patrol is always on the job but may or may not be in the right place at the right time. The Illinois Department of Transportation Communications Center serves as the coordination unit for the patrol, handles all incoming reports of incidents, and, through a separate patrol radio frequency, directs the nearest available patrol unit to the scene. The Communications Center maintains radio or direct telephone links with the Chicago police, the state police, Chicago Fire Department, Expressway Surveillance Project, Chicago Traffic Engineering Center, Illinois Tollway Commission, other Illinois DOT vehicles on two additional radio frequencies, and local radio stations operating private traffic helicopters or with traffic hotlines. More than 20,000 radio messages a month are logged by the patrol fleet.

The electronic surveillance network, described later, picks up problems that disrupt expressway traffic flows and that may or may not have been reported by the patrol or other sources. After electronic surveillance identifies troubled expressway sections, the patrol is directed by radio to the problem site to expedite the removal of flow-reducing incidents. There have been cases in which electronic incident detection has resulted in response action beginning 15 to 20 min before the problem was reported through another source. On the other hand, the patrol assists at thousands of incidents that do not noticeably disrupt traffic flows, possibly because the patrol action prevents minor but hazardous situations from developing into more serious problems.

ELECTRONIC SURVEILLANCE AND CONTROL

The Expressway Surveillance Project was established to investigate electronic techniques of achieving more efficiency and safety on existing expressways, particularly in regard to the problem of rush-period traffic congestion. The initial efforts of the project were directed toward installation and use of a pilot detection system for monitoring traffic flow on a short section of the Eisenhower Expressway. In 1962, a total of 25 ultrasonic detectors were installed to monitor the outbound flow along 5 miles of the Eisenhower Expressway. Each detector in the study section was connected via telephone lines to its own analog computer located in the nearby surveillance center. This initial electronic surveillance effort led to the collection of detailed data, which helped to define specific problems and their relations to traffic flow characteristics.

Congestion Causes

There were three general causes for congestion on urban freeways: (a) numerous random incidents, such as accidents, disabled vehicles, and debris on the pavement, that occur at any time at any place; (b) overloading during rush periods (especially at particular entrance ramps or along certain sections where several ramps contributed to the overloading); and (c) highway geometrics that, for one reason or another, create weaknesses in the traffic stream, particularly at high demand levels.

An example of poor geometrics is a lane reduction in which traffic demands approaching the lane reduction exceed the capacity of the downstream section, such that the lane reduction is a physical bottleneck. Another example is an upgrade crossing a river or some physical barrier that requires changing from a depressed to an elevated cross section. Although such upgrades may be 3 percent or less (within design standards), under peak flows with trucks in the traffic stream the geometrics in such sections obviously impose a much lower capacity than adjacent level, tangent sections that may be feeding the traffic.

The problem does not simply involve finding solutions to the three general causes for traffic congestion, primarily because the causes can operate either independently of or in combination with one another. In rush periods, for example, all three causes may compound congestion to the extent that the relative contribution of each cause is difficult to ascertain.

Congestion Solutions

To develop solutions for the congestion problem, however, requires that each cause be approached independently. In the case of poor geometrics, parts of the highway system can be changed or reconstructed to deemphasize the impact of deficient geometrics. A lane reduction, for example, can be corrected by adding an extra lane. Because the expressway surveillance effort has been based on trying to get more efficiency and safety with the existing geometrics, however, attention has been directed primarily toward the other two causes of congestion.

Solving the overloading problem requires techniques that decrease expressway traffic demand or increase expressway capacity or both. The most practical effort thus far has been management of expressway demand by controlling entrance ramp traffic. Experiments on metering entrance ramps, initiated in 1963, have resulted in an operational program that now covers 48 ramps.

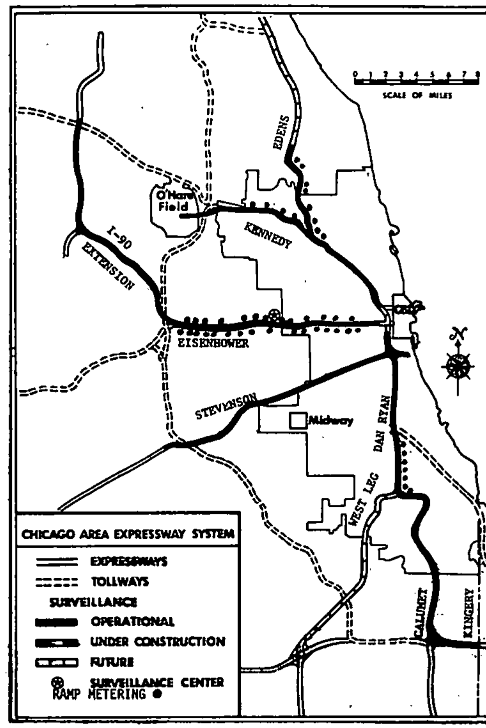
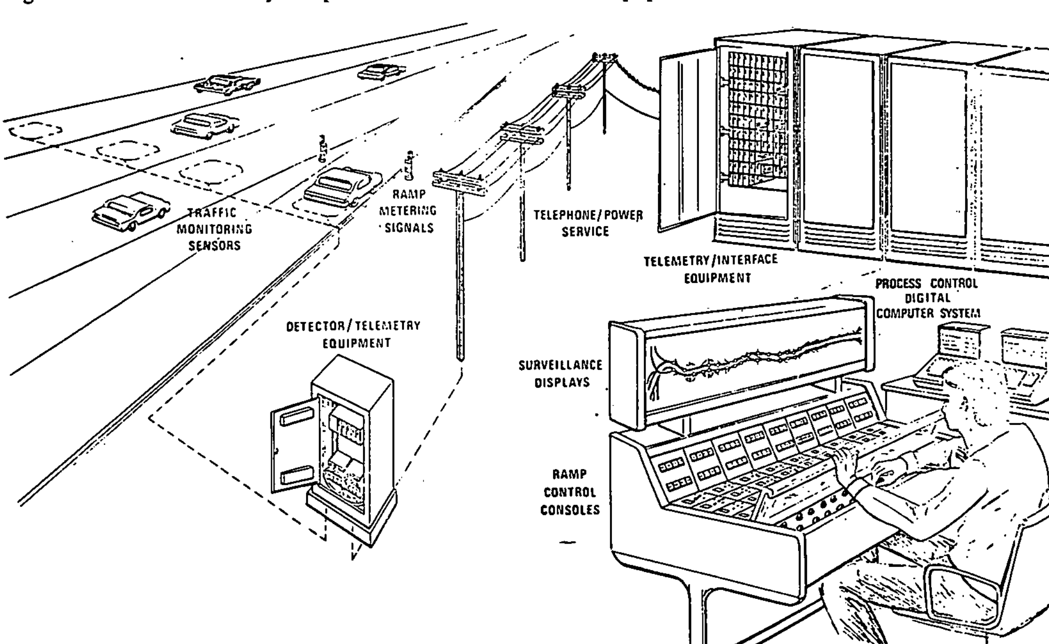


Figure 4. Electronic surveillance and control network.

Figure 5. Detector-telemetry-computer surveillance and control equipment.



Solving the incident problem requires surveillance and service: surveillance to find out that there is a special incident to take care of and service to respond to that incident as quickly as possible with the resources needed to restore normal conditions. Describing the electronic and patrol surveillance and control systems and the capabilities and procedures for detecting and handling operational incidents represents the major portion of this report.

Coverage

From the pioneer experiments with blind detector-telemetry-computer surveillance of traffic flow and the development of ramp metering, the Expressway Surveillance Project has provided the Chicago area with the world's largest computerized freeway surveillance and control network. The 1962 pilot detection system has been expanded to monitor traffic flow at more than 1,300 main-line and ramp points along 90 miles of expressway (Fig. 4). This network handles up to 267,000 vehicles per day at some points and serves approximately 13 million vehicle-miles of travel daily.

Eventually, electronic surveillance will cover all 110 miles of freeway nearest the central city. All instrumented routes are radial with Chicago's CBD and handle extensive commuter trips in both directions during the rush periods. The central computer system for handling this network has been operational since January 1971. Design of a building to house surveillance equipment and staff is currently near completion for a site near the present leased facility.

Instrumentation

Freeway surveillance and control are accomplished through a detector-telemetry-computer network. All the detectors used in the system are presence induction loops embedded in the pavement. The loops are located in each of the main-line traffic lanes at about 3-mile intervals. At approximately ½-mile intervals, one of the center lanes is sampled with a loop. Loops are also provided on all entrance and exit ramps. This detector arrangement gives a surveillance of flow on the main line at about ½-mile intervals and produces a closed subsystem every 3 miles in which all main-line and ramp input and output counts are recorded. Such an arrangement facilitates multipurpose use for surveillance, ramp control, and system evaluation.

The field locations of detectors depend on the availability of phone and power service, which is usually readily available around urban interchange areas. All surveillance and control points in a particular service area are brought to a roadside cabinet through underground interconnect systems. Each roadside cabinet contains detector amplifiers, power supplies, and telemetry equipment for placing detector signals onto leased telephone lines (Fig. 5). By using frequency-division multiplexing techniques, as many as 22 channels can be used for data transmission on each telephone line. Because there may not be 22 detector or other signals for any one service location, multipoint techniques are used to connect adjacent locations to one phone line, so as to maximize use of the telephone line and to minimize leased line costs.

Each detector location has a tone transmitter in the roadside cabinet to encode the detector presence pulse from the detector amplifier on the phone line at a selected frequency. The leased telephone lines (C-1 conditioned) transmit detector signals to the surveillance center, located centrally to minimize communications costs. At the surveillance center, the signals are decoded by tone receivers at the matching frequency for each detector. The tone telemetry equipment in the office decodes and identifies each detector signal and interfaces each pulse into a known position in the surveillance computer.

The digital computer, an industrial process control machine of the type commonly used in steel mills and paper plants, scans the status of each traffic detector 60 times a second. Inasmuch as all detectors are of the presence type, for each scan the computer interrogates the binary status of each detector: Is there a vehicle present or not?

By keeping track of the changes of state from yes to no, the computer records vehicle detection data and calculates volume and lane occupancy for each detection point.

Lane Occupancy

The basic measurement at each surveillance point is lane occupancy, i.e., the percentage of time the detection zone is occupied by a vehicle. The loop detection equipment for measuring lane occupancy also produces lane volume. Although speed is not measured directly, it can be calculated from lane volume and lane occupancy by assuming an average vehicle length for vehicles in the traffic stream (Fig. 6).

Lane occupancy is a most convenient measurement, for it is a summary parameter that includes all the basic aspects of the traffic stream. It considers the volume, speed, and composition (vehicle lengths) of the traffic stream as a whole. Lane occupancy can range from 0, when there is no traffic present, to 100 percent, when there are vehicles continuously in the zone of detection. There should always be some traffic, even at 4:30 a.m., such that the normal operational range is above 0 percent. It is also rare to reach 100 percent lane occupancy, even under stoppage conditions, because there are always gaps between vehicles and usually some movement of the traffic stream.

The basic measurement of lane occupancy gives an indication of traffic stream operations at each detection point. With detectors along each roadway at about ½-mile intervals, sampling the flow at points along the route gives an estimate of overall system operations.

Flow Characteristics

Generalized traffic operations curves show that the optimum peak-period flow condition is 20 percent lane occupancy where traffic speeds near the speed limit coincide with the highest flow rates. Occupancies less than 20 percent indicate flow generally near the speed limit; the corresponding volumes represent demand ranging from zero up to the maximum. This 0 to 20 percent range of flow conditions is referred to as green or free flow.

To sustain the rush-period ideal of 20 percent lane occupancy requires that 80 percent of the traffic stream consist of gaps to keep vehicles moving at high volume and high speeds. Although volume can maintain its maximum throughput, an increase in lane occupancy from 20 to 30 percent causes speed decreases because of (a) fewer and shorter gaps available between vehicles, (b) the increasing difficulty of lane changing, and (c) generally more restrictive flow conditions. These 20 to 30 percent flow conditions are referred to as yellow or impending congestion.

In excess of 30 percent lane occupancy, traffic flow conditions are referred to as red or congested. Speeds continue to degenerate, and volume also decreases. In this red zone, the higher the lane occupancy is, the worse the situation is. High percentages of lane occupancy indicate serious operational problems, such as an accident, a disabled vehicle, or other obstruction to the traffic stream.

Traffic Status Displays

The green-yellow-red zones of operation are used as a convenient on-line expressway surveillance output. The central computer system is used in real time to drive map display panels that indicate through colored lights the green-yellow-red zone of operation for one main-line detector at ½-mile intervals along each expressway direction. A glance at the map displays gives an instant overview of current operations for the entire instrumented expressway system. In off-peak periods, the complete system should operate in the green zone. Any exceptions are clues to operational incidents that require response. In rush periods, a normal pattern of congestion is expected at recurrent bottlenecks. Patterns different from normal indicate operational incidents.

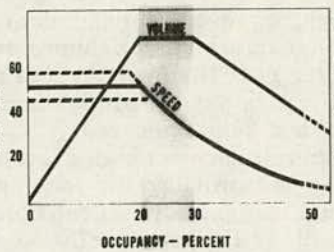
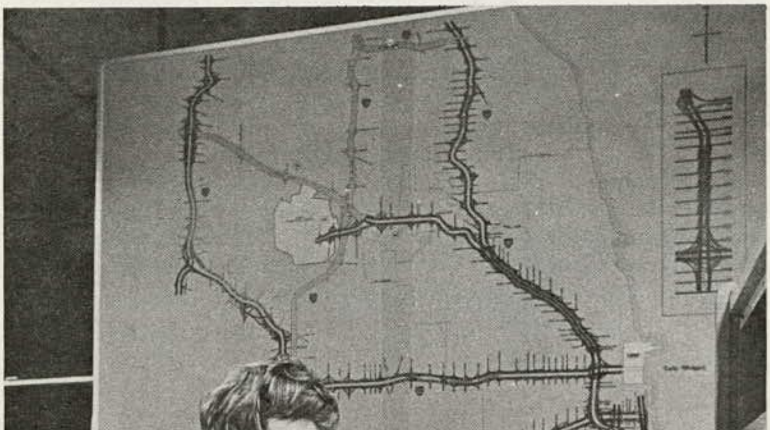


Figure 6. Generalized freeway traffic flow characteristics (one lane).

Figure 7. Map display and ramp control console at surveillance center.



The traffic status displays, however, only summarize the prevailing conditions for main-line traffic. For further information or more detail, the computer system has several peripheral devices for real-time retrieval of the actual traffic flow data. A red condition, for example, can be inspected to determine whether the actual lane occupancy is a 32 percent red or a 74 percent red.

Because the electronic surveillance network grew one expressway at a time, there initially was one map display unit provided for each expressway. Recent expansions consolidated all the individual map panels into one Chicago-area display (Fig. 7). A duplicate area map display allows the green-yellow-red conditions, generated at the suburban Oak Park Surveillance Center, to be repeated via phone line and time division scanning techniques at the Illinois DOT Communications Center in downtown Chicago at Marina City, about 8 miles distant. The Communications Center operates three radio frequencies, coordinates all traffic and maintenance information and operations, including the Emergency Traffic Patrol vehicle fleet, and is staffed 24 hours a day, 7 days a week. Remote duplication of the area map displays at the Communications Center allows the surveillance office to staff only during rush periods and normal working hours, usually each weekday from 5:30 a.m. to 7:00 p.m.

The electronic surveillance displays are the primary sources of expressway traffic condition information disseminated to the public. Both the surveillance center and the Communications Center handle numerous daily telephone requests for traffic information, primarily from commercial radio stations incorporating the information into traffic broadcasts. In March 1974, computer-generated expressway congestion reports were made available to commercial radio stations providing their own remote teleprinters and interconnect with the surveillance computer.

Traffic Control

Traffic flow problems, pinpointed through the electronic surveillance system or reported through other sources, are responded to by a fleet of expressway emergency patrol trucks, operated by the Illinois DOT. These trucks are continuously on patrol and provide mobile surveillance and assistance for all incidents and potentially hazardous situations. When trouble is initially spotted through the electronic surveillance system, the Communications Center dispatcher requests the closest available patrol truck to investigate the site of the traffic flow disruption.

In addition to detection of and response to incidents during rush periods, the direct control of entering traffic demands is effected through expressway ramp metering at 48 entrance ramps on five expressways. In the morning rush period, all 23 inbound and 10 outbound ramps are metered; in the afternoon, all 25 outbound and 8 inbound are metered. The surveillance computer is programmed to operate each ramp metering signal in an automatic one-vehicle-at-a-time mode, which varies the entrance ramp flow rate in response to prevailing traffic conditions on the expressway and ramps. Signal timing is accomplished in the surveillance center, and the commands for ramp signal changes are transmitted to the field via the same phone line feeding detector data to the computer.

Although aimed primarily at recurrent rush-period overloads, ramp metering can provide additional relief during expressway incident situations, particularly where bypass routes are readily available. The amount of relief, however, usually cannot overcome major losses in system capacity.

Computer Capabilities

The process of traffic surveillance, incident response, and ramp metering produces beneficial effects on the traffic stream. Such effects become measurable through changing the traffic flow data, which become a new input to the surveillance system. The computer capability for recording data is the basic source for evaluating the impact of particular control measures on the traffic stream.

The central processor unit has a 32-K memory with a 24-bit word. Most of the data collected through the central processor is transferred to bulk storage on two magnetic disks that have 1,000,000-word capacity each. The primary function of the computer system is to drive the on-line, real-time components, such as the surveillance map displays, ramp metering control rates, and real-time data output or retrieval peripherals. Two CRT video units, two typers, and a line printer allow data to be continuously presented or retrieved through the keyboard from the computer in real time and displayed either on the CRT screens or as hard copy on the typers or the line printer. Most real-time information has a time base of 1 min or less, is used for operational purposes, and is not retained per se unless needed for some particular application.

The surveillance computer system also has a free-time feature that allows data processing and analysis in time periods not required by on-line, real-time functions. The system peripheral devices, particularly a card reader and card punch, are in free-time periods.

Daily Printouts

The usual practice is to store summary data from all detectors for a 24-hour period on the two disk units. Each day three printouts are routinely dumped out on the line printer. The first of these is a printout of traffic counts by hour for each detector in the surveillance system. These data serve as quantitative traffic trend records, as answers to requests for traffic count data from various agencies and researchers, and as a means for locating equipment problems for maintenance. The other two daily printouts are traffic operations performance records for each of the two rush periods.

For the 5 to 10 a.m. rush period, the lane occupancy readings from one main-line expressway detector, at each ½-mile interval, are summarized into 5-min averages. Printouts for each direction for each expressway are similarly retained for the 2 to 7 p.m. rush period. The lane occupancy data on these printouts give a picture of the overall performance for each rush period in terms of the distance versus clock-time pattern of traffic flow, including the effects of any incidents (Fig. 8). Each rush-period printout is further summarized by computing the total minute-miles of congestion, which is a convenient tag for rating the performance each day relative to every other day.

Implementation Philosophy

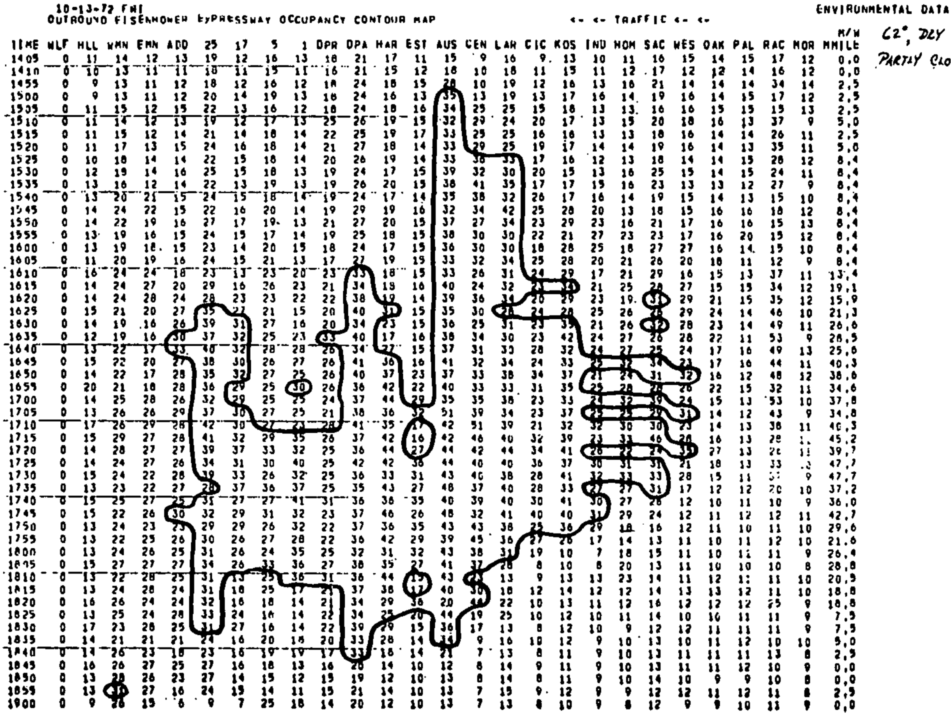
Implementation of freeway surveillance and control in the Chicago area reflects the experience gained since 1961. The surveillance system, featuring the basic loop presence detection, leased phone line interconnect with the central computer, and tone telemetry signal communications, is installed as the initial backbone component. When operational, the surveillance system provides real-time incident detection, which aids the emergency patrol service, and data collection and evaluation capabilities, which help define overloading conditions warranting ramp control or geometric improvements.

Ramp control can be added later to the backbone surveillance system by installing two ramp traffic signals, two additional ramp detectors, buried interconnect, cabinet, and tone equipment at each entrance ramp. Additional freeway and ramp detection can be provided if research demonstrates that increased sophistication is worth the cost. The backbone surveillance system can be used to evaluate the operational effect of ramp control implementation and other electronic traffic aids added on an experimental or operational basis. Examples of other electronic components are changeable-message signs, automatic reversible roadway control, and motorist aid systems.

AUTOMATIC INCIDENT DETECTION

Illinois has had considerable experience with developing and using various large-scale, real-time incident detection strategies. The electronic surveillance system is used to continuously scan various traffic flow parameters, in an attempt to automat-

Figure 8. Typical rush-period lane occupancy printout.



ically identify locations of possible traffic incidents as quickly as possible, with a minimum of false alarms.

Although the traffic status displays of the surveillance system are available for monitoring by operational personnel and further traffic data, both current and prior, can be retrieved for analysis, particularly through use of CRT devices, automatic incident detection is aimed at converting as much manual observation and data checking as possible to computer logic. Such logic analyzes available real-time data to quickly and reliably signal occurrence of a traffic incident. The following general criteria are important to any automatic incident detection scheme:

1. Highest possible detection rate,
2. Fastest possible response time,
3. Lowest possible false alarm rate, and
4. Minimum manual input.

On-Line Experience

Initial automatic incident detection efforts were based on operational experience in spotting traffic incidents from the surveillance map display conditions, particularly sudden significant differences in traffic flow characteristics between adjacent mainline expressway stations. To develop computer logic for detecting incidents required that we determine which basic traffic data (occupancies and volumes of various possible time bases) and features derived from them would most effectively quantify the incident occurrence pattern. The threshold values that, when attained by these features, define the occurrence of a traffic incident could then be optimized. Also considered in initial phases were various checks on data validity and means for evaluating the performance of logic using different time bases, features, and threshold values.

The electronic surveillance network includes nearly 400 main-line incident detection subsystems, i.e., pairs of adjacent sampling detector stations approximately ½ mile apart. Based on the traffic volume and occupancy for each station and the time base variations available, the size of the surveillance network suggests concentration on subsystem upstream and downstream station changes.

One problem with signaling incidents is the wide range of normal operations common to daily traffic flows. Such normal conditions as rush-period overloads, environmental changes, geometric conditions, and faulty surveillance equipment components can mimic the traffic flow characteristics associated with the random traffic incidents requiring detection and response. In addition, there are incidents that self-correct or become removed before response arrives on the scene. In these cases, however, there is no way of knowing at the time of detection, for example, whether a stalled vehicle blocking the pavement will remain until the response arrives.

Primarily by using subsystem station occupancies, 14 variations of basic pattern recognition schemes have been tested and calibrated on-line. Various modifications in time base, features, and threshold values have been implemented. The computer output typer automatically signals possible incidents and prints associated pertinent data. The primary result of on-line, systemwide experimentation with various schemes has been the identification of inputs needed for formulating, evaluating, calibrating, and refining candidate models for operational use.

Logic Evaluation

Although on-line testing has served to define the traffic incident detection framework, it has been impossible to evaluate various schemes because field verification and documentation of all signaled incidents are not possible. Therefore, we plan to use off-line testing of various schemes on a standardized and fully documented data base, including incidents and verified nonincident periods. The off-line efforts will enable comparison of different logic schemes on the same data base as well as optimization of each logic's parameters on the same data.

Effort, in addition to logics, will be made to formulate and test logic schemes based on forecasting and time series analyses. Other specific objectives on the off-line testing will be intensive study of the two major difficulties encountered on the on-line work: electronic equipment malfunctions and highway geometric problems that mimic traffic incident situations. One approach to handling the locations with geometric weaknesses is to determine special threshold values. For the equipment problems, faulty equipment characteristics will be identified to signal maintenance needs and to remove stations from incident consideration.

The experience with on-line testing of incident detection models has pointed out the unreality of most reported off-line tests. One example of the difference between on-line and off-line environments is seen in the acceptance of a 1 percent false alarm level in much of the literature. This rate would produce 17,280 false messages a day for a 20-sec base logic scheme using 400 detector subsystems. Thus, although the more easily documented off-line testing approach will be pursued, contact will be maintained with the on-line environment in which the operational model must perform.

Applications

For the Chicago area, the initial response to expressway incidents is the radio-equipped emergency patrol vehicles. The primary application of automatic incident detection is deployment of these vehicles to reduce response time. Secondary applications include incorporation into ramp metering strategies and driver information techniques such as changeable-message signing and advisory traffic reports.

Currently two interim logic schemes for automatic incident detection are being used on-line. Although problems with regard to their completely automatic operation remain to be solved, these logics provide relatively reliable incident occurrence information, when supplemented with various manual checks.

MOTORIST INFORMATION

Illinois has conducted comprehensive driver attitudinal surveys and has had operational experience with real-time motorist information techniques. One of the requirements for management of traffic incidents is to give drivers information on changing traffic conditions.

Radio Traffic Reports

Since the research surveys pointed out that most commuting drivers listen to radios at home and en route, a great potential for reaching motorists appeared to be through traffic reports on commercial radio stations. Historically, several Chicago-area radio stations have provided public service traffic reports, based either on private ground and helicopter surveillance or on information obtained through the state Communications Center or from various other sources and techniques.

With expansion of the electronic surveillance network, several radio stations requested the current traffic conditions depicted on the traffic status displays, and the Expressway Surveillance Project prepared for directly outputting basic computer-generated traffic information on a remote teleprinter network serving commercial stations. Such communications appeared to have great potential for improving the timeliness, accuracy, and quality of expressway traffic reports, particularly since the electronic surveillance covers all routes in real time and is not limited by the frequency of other coverage methods.

In March 1974, the state notified all Chicago-area radio and TV stations of the availability of a teleprinter service consisting of receive-only, computer-generated traffic reports, as often as each 5 minutes, listing only the significant expressway congestion areas. In effect, this report averages the surveillance map display readings every 5 minutes, on a continuous 24 hour-a-day basis, and prints out the location limits when two or more adjacent, or once removed, locations are congested (red) (Fig. 9). Each radio station provides its own teleprinter, modems, and leased phone lines for hookup with the surveillance computer. Thus far, five of the major AM radio stations are using this service as a source of motorist information reports.

As part of the automatic teleprinter network, the state plans to add two transmit-receive units, one each at the Oak Park Surveillance Center and at the Marina City Communications Center. These units will allow the addition of manual keyboard messages for incidents confirmed by the Emergency Traffic Patrol and for other supplemental data. Each of the state centers would communicate back and forth with each other, particularly with regard to investigating suspected incidents, with selective confirmed information released to the receive-only network.

Off-Freeway Signing

In connection with 1965 ramp control experiments in the Chicago area, the Expressway Surveillance Project installed advisory information signs at four surface street locations in advance of nearby expressway interchanges. The purpose of the signs was to advise expressway-bound motorists of real-time expressway and ramp traffic conditions. When conditions warranted, the motorist could be diverted to parallel alternate routes. At each sign, color-coded information depicted current conditions for the nearest interchange and for the next downstream interchange. In cases where traffic congestion was depicted, the motorist could consider taking an alternate route to the

Figure 9. Typical remote teleprinter traffic congestion report.

```
                    CONGESTION LIMITS 05-03-74 FRI
0600
IB RYAN EXP      ROOS   ->TYLOR*
0610
IB RYAN EXP      87TH  -> 83RD      21ST  -> ROOS
0615
IB RYAN EXP      90TH  -> 83RD      26TH  -> ROOS
0620
IB RYAN EXP      * 95TH -> 83RD
0625
OB KENNEDY       KIMBAL ->ADDSON
OB IKE           CENTRL ->AUSTIN
IB RYAN EXP      * 95TH -> 79TH     29TH  -> 21ST
0630
OB KENNEDY       KIMBAL ->ADDSON
OB IKE           CENTRL ->AUSTIN
IB IKE           W.MANN ->E.MANN
IB RYAN EXP       90TH  -> 87TH     29TH  -> 21ST
0635
OB KENNEDY       KIMBAL ->ADDSON
OB IKE           CENTRL ->AUSTIN
IB IKE           HILSDE ->E.MANN
IB RYAN EXP       90TH  -> 83RD     29TH  -> 21ST
IB RYAN LOC       55TH  -> 51ST
0640
OB KENNEDY       DIVRSY ->SACRAM
OB IKE           LARAME ->AUSTIN
IB IKE           HILSDE ->E.MANN
IB RYAN EXP      * 95TH -> 75TH     29TH  -> ROOS
0645
OB KENNEDY       WESTRN ->DIVRSY    KIMBAL ->ADDSON
OB IKE           LARAME ->AUSTIN
IB IKE           HILSDE ->E.MANN     9TH  ->  1ST
IB STEVENSON     HARLEM ->6900 W    5100 W ->PULASK
IB RYAN EXP      * 95TH -> 83RD     26TH  -> 21ST
IB RYAN LOC       55TH  -> 51ST
0650
OB KENNEDY       BELMNT ->ADDSON
OB IKE           LARAME ->AUSTIN
IB IKE           W.MANN ->E.MANN     9TH  ->  1ST
IB STEVENSON     HARLEM ->6900 W    CENTRL ->PULASK   HOYNE  -> WOOD
IB RYAN EXP      * 95TH -> 79TH
0655
OB KENNEDY       BELMNT ->ADDSON    NAGLE  ->HARLEM
IB KENNEDY       W.CUMB ->E.HAR
IB IKE           HILSDE ->E.MANN
IB STEVENSON     7500 W ->6900 W    CENTRL ->PULASK   HOYNE  -> WOOD
OB RYAN EXP      *TYLOR -> 16TH
IB RYAN EXP      * 95TH -> 71ST     29TH  -> 16TH
IB RYAN LOC       55TH  -> 51ST
0700
OB KENNEDY       SACRAM ->ADDSON    AUSTIN ->HARLEM
IB KENNEDY       E.HAR  ->SAYRE
IB IKE           W.MANN ->E.MANN     9TH  ->  1ST
IB STEVENSON     7500 W ->PULASK    CALIF  -> WOOD
IB RYAN EXP      * 95TH -> 79TH
0705
IB EDENS         WILSON ->MONT *
OB KENNEDY       SACRAM ->ADDSON    FOSTER ->NAGLE
IB KENNEDY       CANFLD ->SAYRE     CICERO ->MONTRE
```

EXAMPLE

At 6:35am, the outbound Kennedy Expressway congested from Kimball Ave. through Addison St., the outbound Eisenhower, Central thru Austin, etc.

next interchange, or at least to the next sign, where additional advisory information was available.

These advisory signs, aimed at repeat users in rush periods, produced marginal results. Although motorists passing the signs responded in a questionnaire survey that they understood and used the signs, actual traffic pattern changes due to information shown on the signs were not perceivable enough to be significant. After 5 years of experimental use, the four signs were deactivated from real-time surveillance computer operation. Three months later, without any interim public inquiries in regard to nonoperational signs, the signs were removed. There also have not been any public inquiries about the signs since their removal.

On-Freeway Signing

In followup driver information research, two attitude surveys (totaling 940 Chicago-area motorists) indicated that drivers feel that freeway traffic information is desirable and that on-freeway changeable-message signs are a preferred means of presentation. Because such signing would be added to the electronic surveillance and control network, Illinois has a research project for testing and evaluating one real-time, changeable-message, on-freeway sign in the Chicago area.

Current efforts consist of monitoring other changeable-message sign projects and preparing signal tests for demonstrating remote prototype sign control via interface with the surveillance computer and through manual override. After signal tests and experimentation are conducted with one actual sign, the results will determine the utility of future traffic information aids of this type.

ACKNOWLEDGMENTS

The research reported on here was conducted under the sponsorship of the state of Illinois and the city of Chicago in cooperation with the U.S. Department of Transportation, Federal Highway Administration. The contents of this report reflect the views of the author, who is responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policy of the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. Eno, W. P. The Science of Highway Traffic Regulation, 1899-1920. Brentano's, April 1920.
2. McDermott, J. M., Kolenko, S. J., and Wojcik, R. J. Chicago Area Expressway Surveillance and Control: An Interim Report. Chicago Area Expressway Surveillance Project, Report 25, March 1975.
3. McLean, C. H. Public Information Problems in Urban Interstate Operation. Institute of Traffic Engineers, Compendium of Technical Papers, 42nd Annual Meeting, 1972.
4. McLean, C. H. Chicago Expressway Traffic Patrol. Presentation to HRB Freeway Operations Committee, July 1970.
5. McLean, C. H. Illinois "Minutemen" Keep Traffic Moving on Expressways. Traffic Safety, January 1964.
6. Dudek, C. L., Messer, C. J., and Nuckles, N. B. Incident Detection on Urban Freeways. Transportation Research Record 495, 1974.
7. Cook, A. R., and Cleveland, D. E. Detection of Freeway Capacity-Reducing Incidents by Traffic Stream Measurements. Transportation Research Record 495, 1974.
8. Hoff, G. C. Development and Evaluation of Experimental Information Signs. Chicago Area Expressway Surveillance Project, Report 18, 1968.
9. Heathington, K. Interim Report: On the Development of a Freeway Driver Information System. Chicago Area Expressway Surveillance Project, August 1969.
10. Hoff, G. C. A Comparison of Selected Traffic Information Devices. Highway Research Record 366, 1971.
11. Heathington, K. W., Worrall, R. D., and Hoff, G. C. An Evaluation of the Priorities Associated With the Provision of Traffic Information in Real Time. Highway Research Record 336, 1970.
12. Heathington, K. W., Worrall, R. D., and Hoff, G. C. An Analysis of Driver Preferences for Alternative Visual Information Displays. Highway Research Record 303, 1970.
13. Heathington, K. W., Worrall, R. D., and Hoff, G. C. Attitudes and Behavior of Drivers Regarding Route Diversion. Highway Research Record 363, 1971.
14. Everall, P. F. Urban Freeway Surveillance and Control: The State of the Art. Federal Highway Administration, U.S. Department of Transportation, June 1973.

Alternate Route Planning: Successful Incident Traffic Management

David H. Roper, Richard J. Murphy, and Robert F. Zimowski
California Department of Transportation

A congestion-causing freeway incident causes the capacity-demand relation to go out of balance. The solutions to such incidents are to restore the temporarily reduced capacity as soon as possible and to limit demand at the incident by diverting traffic. This paper discusses the alternate route planning procedure being developed in southern California. As part of the preplanning effort, information on all segments of the freeways was gathered and mapped. The main elements of the maps are discussed, and a sample map is presented. The two concepts used in implementing alternate route planning are the on-site command post and the incident response team. These are discussed as they relate to incident management.

A commuter is traveling along the freeway on his way to work. About halfway to his destination, traffic comes almost to a complete stop in heavy congestion. At that moment the radio relates the sad story of an overturned truck miles ahead on the freeway. Then comes the simple statement, "Take an alternate route." He drives the freeway twice a day but has no knowledge of the surface streets in the area. Where is an alternate route? Where can he get on it? Where can he get back on the freeway? Will he really save time by using the alternate route? With all these questions to answer, all too often the motorist stays on the freeway and frets his way through the problem area.

Table 1 gives data on the magnitude of nonrecurrent capacity-reducing incidents on the more than 650-mile Los Angeles freeway network. It is estimated that 10 percent of the incidents are major and account for 30 percent of the delay due to incidents.

PREPLANNING

Alternate route planning, being developed in southern California, holds the answer to many of the motorist's questions. When a congestion-causing incident occurs, the capacity-demand relationship is thrown out of balance because of a temporary reduction in capacity. Restoration of full capacity as rapidly as possible is essential; in the meantime, limiting demand at the incident, rerouting the remaining capacity to alternate routes around the incident, and providing advance information of the incident can do much to alleviate the problem. In other words, incident management can yield substantial benefits. And preplanning of alternate routes is the necessary foundation on which to build effective incident management.

In California, the responsibility for management of incidents such as overturned trucks on freeways rests with the California Highway Patrol. In Los Angeles we and the Los Angeles police and traffic department have for some time been working with the California Highway Patrol in an extensive program of planning and implementation of alternate routes.

As a first step in the planning phase, we inventoried all existing surface streets that might serve as alternate routes for every section of freeway in the Los Angeles area. Street widths, curvature, grades, condition of surfacing, intersections, side friction, and turning radii were all recorded. At the same time, we identified missing links that might be constructed to close gaps in desirable alternate routes. We could then select best alternate routes for each possible freeway closure.

In planning, we took the position that we would use whatever capacity remained on the freeway (measured by the number of lanes remaining open) to the maximum extent possible. Physical data collected during the inventory were converted and expressed in terms of capacity. These were incrementally added to remaining freeway capacity (if any) to develop the total capacity needed by implementing the alternate route in varying stages.

In analysis of demand, it is impractical to check demand at each incident before alternate routing is implemented. We gathered demand information for each freeway by time of day. By breaking the day into various segments, we could assign a demand value to each time slice.

We then had the necessary values to analyze the capacity-demand relationship for incidents of varying severity and during varying demands. A matrix relating number of lanes remaining, time of day, and stages of implementation could then be developed.

Coordination with all agencies involved in alternate route planning is of primary importance. It is essential that local agencies having jurisdiction over the surface streets to be used be brought into the planning process as early as possible. Their opinions and knowledge of the facilities complement established route development procedures. We do not want to run heavy traffic past hospitals or schools, through residential areas, or past the mayor's house if we can help it. Local police must be involved in the planning, for they will be involved in directing traffic. Agencies responsible for patrolling and maintaining the freeway must also be committed to the plan to ensure successful implementation and operation. Manpower requirements for directing traffic, barricading, and placing signing must be determined, and agreements on who does what must be reached.

Based on experience in incident management, we decided to put together all necessary information on each segment of freeway in the form of a map. A sample map is shown in Figure 1. Each map contains the following vital elements:

1. Identification of problem location,
2. Location of primary and secondary alternate routes,
3. Tabulation relating lanes remaining in service, time of day, and stages of implementation,
4. Manpower requirements and locations,
5. Required signing,
6. Necessary closures,
7. Responsible parties and telephone numbers, and
8. Special notes unique to this incident area.

To cover more than 475 miles of heavily traveled freeways will require some 2,500 maps. Keeping these maps up to date will be a major activity in itself, but the system cannot work without accurate, current information.

IMPLEMENTATION

Preplanning is an excellent tool, but, if it is not coupled with a means of implemen-

Figure 1. Sample map of alternate routes for incident on Ventura Freeway eastbound.

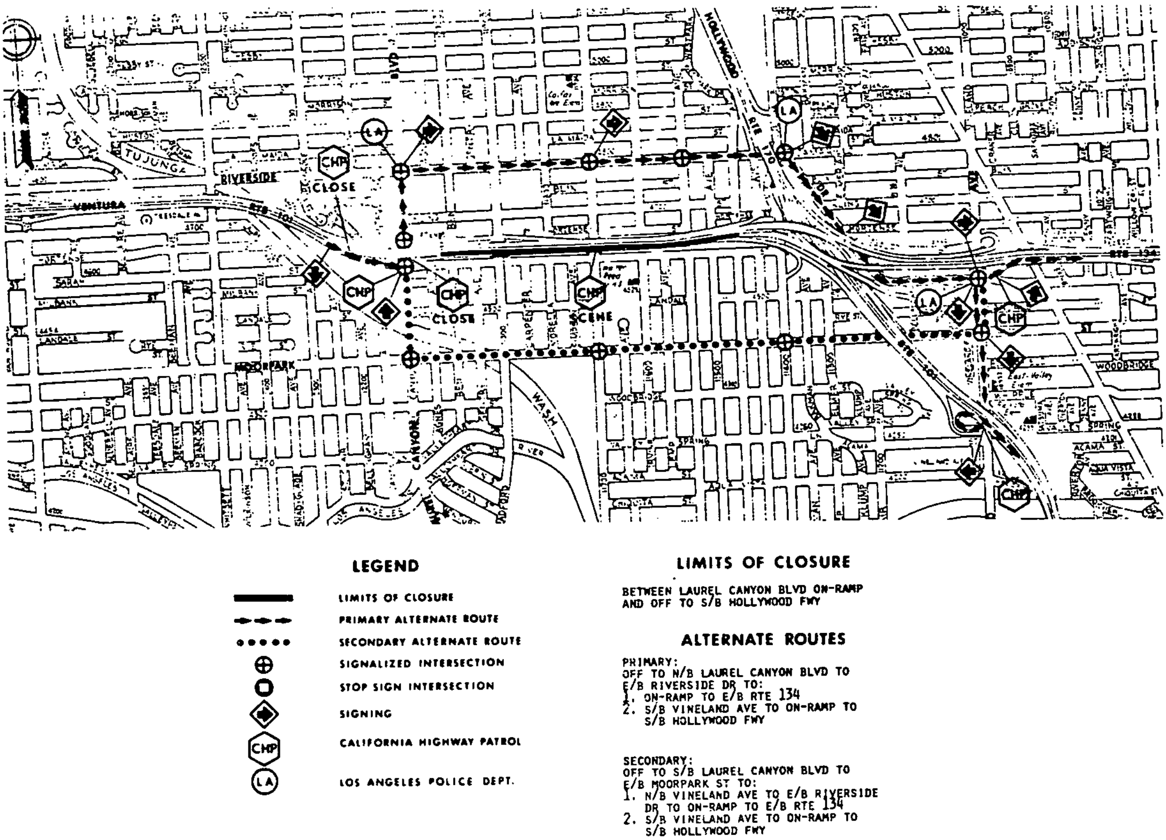


Table 1. Magnitude of delay-causing congestion on Los Angeles freeway network.

Type of Congestion	Delay (1,000 vehicle-hours per year) Number	Percent
Recurrent, weekday peak	9,300	43-57
Nonrecurrent		
Incidents	5,000-10,000	30-47
Weekend	1,400	6-8
Holidays	600	3-4
Other	100	1
Total	16,400-21,400	100

tation and real-time management to provide proper on-the-spot adjustments, results could be counterproductive—sometimes even disastrous. We are meeting these needs through development and use of two operating concepts: the command post and the incident response team.

Command Post

An on-site command post, consisting of representatives from the agencies involved, is essential to effective management of major incidents. Development of this concept is a vital part of our alternate route efforts.

It is unfortunate that the goals of the agencies involved are often in conflict with each other. For example, consider the case of a vehicle fire in which firemen, in placing hoses, have unknowingly closed the primary alternate route. Conflicts such as this must be identified and resolved, and this is where the command post can be most effective.

In the command post concept the question arises, "Who is in charge?" Based on our active participation in incident management, this can best be answered with the question, "For which phase of the overall operation, at what location?" Face-to-face discussion with those responsible for the multitude of actions that must take place brings about a planned effort that best solves all problems and meets joint responsibilities by using joint authorities. Cooperation and coordination are the keys to success. We have been using this approach for more than a year, and it works.

Incident Response Team

Implementation of the plan involves establishing the alternate route and providing the signing, early warning capabilities, and the traffic flow expertise. To meet these needs, we have organized and staffed a response team in our Freeway Operation Branch that goes to the site of incidents estimated to block two or more lanes for 2 or more hours. Team members are traffic engineers and technicians and are available on a 24-hour basis. We use a truck carrying a changeable-message sign that can be strategically located to warn or inform motorists of the situation. Temporary signs, both metal and cloth, are carried in each response team vehicle; these are used to direct traffic along the alternate route and back to the freeway.

The procedures described here have been in use for a relatively short time and much is still to be learned, but, when we get all the pieces in place, the sequence of events will be as follows.

The first man on the scene, usually a traffic officer of the California Highway Patrol, reports the location, analyzes the situation, and estimates the duration of the problem. Communication centers then phone responsible parties as set forth in the plan, and the team assembles at the site and establishes the command post. At that time, a review is made, the appropriate stage of implementation is decided on, the alternate route is physically checked, and action is taken as called for in the plan. One or more changeable-message sign trucks are placed in the most advantageous location with proper message displayed; temporary signing directing motorists around the incident location is placed along the alternate route. Where necessary, traffic officers direct traffic off the freeway and through intersections. Then clearing operations commence. Monitoring and adjustment of the operation plan continue as directed from the command post. Information sources such as helicopters and traffic-monitoring devices, i.e., electronic surveillance with or without closed-circuit TV, can be most helpful to the command post in this phase.

Already some side benefits are emerging from this alternate route program. Knowledge of available routes and contacts with people in the field have been used in planning detours around construction and maintenance activities on the freeways. Many of the same procedures and techniques used in incident management apply to planned closures, and the public can realize similar savings in time, increased safety, and reduction of frustration.

SUMMARY

Preplanning for major freeway incidents by planning alternate routes and providing on-site management of incidents will not solve all congestion problems related to incidents. In many cases it will do little more than make the best of a bad situation. With relatively long-term incidents, it provides the starting point and allows us time to make more extensive arrangements. It can bring some sense of order to what is often a chaotic situation. Providing the whole package, including preplanning, the command post with the atmosphere for overall incident management, the hardware, and traffic expertise, represents a sound approach to solving an all too common problem.

Experience in Handling Freeway Corridor Incidents in Houston

William R. McCasland
Texas Transportation Institute

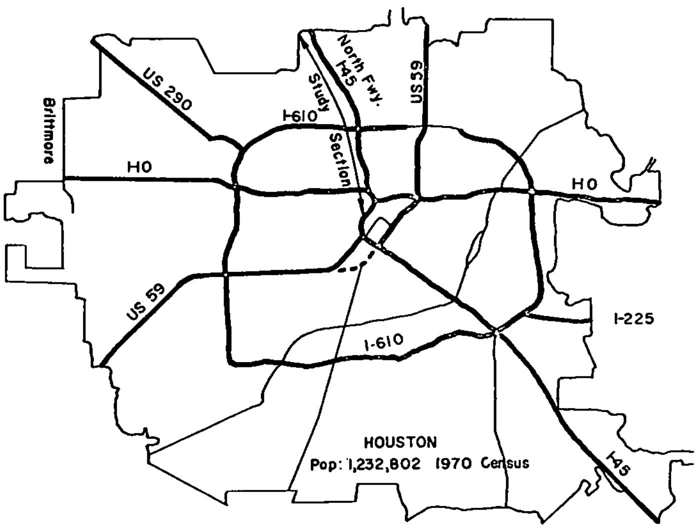
The need to improve the operational efficiency and safety of urban streets and freeways is greater now than ever before. But at the same time urban areas are faced with an increasing problem of traffic congestion caused by capacity-reducing incidents. The incidents are serviced by many organizations in many ways, but few cities have an efficient and coordinated procedure to restore urban roadways to peak efficiency after an incident occurs. The Texas Highway Department, City of Houston, and the Texas Transportation Institute are working toward a solution of this problem. Current practices for the detection and removal of freeway incidents and the results of research studies on response times by police, effects of accidents on traffic operations, and the effectiveness of emergency call-box systems are reviewed. Recommendations are presented for the improvement of the total incident management system in Houston.

Houston has an excellent freeway system. Six radial freeways are connected by an inner freeway loop at the central business district and by Interstate 610 approximately 6 miles from the inner loop (Fig. 1). A belt road that will circle the Houston area approximately 12 miles from the CBD is under construction.

The freeways have the following geometric design characteristics: freeway shoulders on both sides of the roadway, roadways at grade with the surface street system, overpasses at the interchanges, parallel frontage roads with frequent entrance and exit ramps, and a separation between frontage road and freeway that can be traversed by emergency vehicles. The freeways are generally well lighted by highway luminaires and lighting from commercial developments adjacent to the freeways and frontage roads. Only in the immediate area of major interchanges of freeways does the design deviate from these characteristics.

The geometric design of the freeway has a significant impact on incident management. There are only a few miles of urban freeways in Houston on which a disabled motorist could be considered trapped. In most instances, motorists can leave their vehicles and walk a few hundred feet to summon aid. Emergency vehicles can use the shoulders and parallel frontage roads to reach the scene and can cross the separation between freeway lanes and the frontage roads to provide service to stranded motorists.

Figure 1. Map of Houston freeways.



THE PROBLEM

Why are transportation engineers concerned with better management of traffic incidents? First, there is a need to make the best use of existing roadways in the urban area, whether for private automobiles, buses, or car pools. Second, there are better techniques, procedures, and systems available for controlling traffic and providing service to motorists that will keep freeways and arterial streets clear of capacity-reducing incidents. Third, there are problems in implementing these innovations of surveillance, communication, control, and service.

The need for more efficient operation is demonstrated daily when traffic overloads urban streets and freeways during the morning and afternoon peak periods and when the slightest disturbance to traffic causes delays and hazardous conditions. The problems of implementing improvements in incident management are many, but a few of those experienced in Houston are discussed below.

Most urban areas have had increases in traffic and congestion similar to Houston. As the traffic problems have increased, procedures for handling traffic incidents have evolved to meet the situation. Organizations responsible for these procedures have attempted to stay abreast of increasing demands and, for the most part, have succeeded. This success contributes to the problem of implementation. It is difficult to convince administrators that more money should be allocated, more equipment should be purchased, and new organizational responsibilities should be developed to improve what is now a good operation. As one official stated, "All accidents are reported, all vehicles removed from the travel lanes, all trash and debris are picked up in a reasonable time period. What more do we want?" Costs seem to be the primary deterrent even though studies and demonstrations have provided cost-benefit data. Benefits due to the reduction in response time from reasonable to very reasonable are not so apparent as the costs of the systems.

Another problem of implementation is that a full-time system for handling infrequent and nonrecurring freeway incidents has the appearance of being an inefficient operation when it stands ready for long periods of time. When the department heads of the Texas Highway Department were instructed to reduce fuel consumption by 15 percent, one of the first cutbacks was the freeway courtesy patrol, although this service probably reduces the fuel consumption of thousands of vehicles. When the patrol service was compared to other maintenance operations, such as pavement repair, street sweeping, hardware maintenance, landscaping, and trash pickup, its effectiveness was not properly evaluated.

A third problem is that incident management affects governmental and operational agencies, and coordination of their activities is essential. In most situations, there are no questions of which agency is in charge and what actions must be taken to clear an incident. Even in situations involving major capacity reductions and requiring special decisions, the problems are usually time and availability of facilities; but often matters of jurisdiction can complicate the management process. The formation of a multiagency organization to cope with these matters is both a problem and a solution.

Finally, semiautomated and fully automated systems for incident detection and location and for traffic control and driver information are subsystems of large traffic surveillance and control networks. The problems in implementing these larger systems are the same as those listed above but are greater in scope and magnitude.

INCIDENT CLEARANCE PROCEDURES IN HOUSTON

Incidents that cause a reduction in traffic capacity and safety can be grouped into four categories: accidents, disabled vehicles, debris or lost cargo, and environmental conditions. By far, the most important and the most numerous incident is the traffic accident.

Clearing Accidents

Most accidents on urban freeways in Houston are handled in a similar manner. The accident is reported to the police by telephone or radio from a patrol unit. The appropriate police vehicle is dispatched by radio to the scene. Wrecker vehicles monitor the police radio transmissions and respond without being requested. On arrival, the police assess the situation, determine whether additional emergency vehicles are required, tend to the injured, and clear the roadway. If there are no injuries or fatalities, the vehicles are moved from the freeway lane by the drivers or by the wrecker service before the police unit arrives. If there is a spillage of cargo or damage to a highway structure, the police contact the highway department. The police then conduct the accident investigation, and the vehicles and drivers are released from the accident scene.

In 1968, studies were conducted on the Gulf Freeway on the time required to clear an accident scene. A 14-camera closed-circuit television (CCTV) system for detection and location was used to collect data on more than 250 accidents and the following results were obtained (1):

Procedure	Average Time (min)
Detecting and locating incident	1
Police responding after being dispatched	11
Clearing incident from freeway	4
Conducting accident investigation	25
Total	41

During the last 3 years, the research program has been working with the Texas Highway Department and the Houston Police Department to reduce the total time required to clear the roadway and to reduce the total time an incident affects freeway flow. The detection-location time was considered to be a maximum of 1 minute in this study because of the CCTV system. Normal methods of detection may require

30 minutes longer. Response time can be reduced if adequate manpower and communications are available. The time required to clear a roadway may be greatly improved for major incidents by providing the necessary equipment quickly, but most minor incidents are moved as quickly as possible without special assistance. The time for investigation may or may not be reduced, but its effect on freeway traffic can be minimized. The development of this research program is discussed later in this report.

Clearing Disabled Vehicles

When a disabled vehicle blocks a freeway lane, the police respond in the same manner as for an accident. The police assist in removing the vehicle to the shoulder of the roadway and in obtaining service that the motorist might require. The average response time is approximately the same as for accidents, but the clearance time is higher because all vehicles are not operative.

When a disabled vehicle is parked on the shoulder of the freeway, a police unit does not respond unless directed to do so by the dispatcher or when requested by the motorist. The highway department provides a courtesy patrol that stops and offers assistance to vehicles parked on the shoulder.

If a vehicle remains on the shoulder unattended for 48 hours, it is classified as abandoned and a wrecker is called by the police to remove the vehicle to the city pound.

Clearing Debris or Spilled Cargo

The highway department is responsible for keeping the roadway clear of debris. Normal sweeping and trash pickup operations are conducted daily. When unusual debris is reported to the department, a field unit in the area is dispatched. The procedure for clearing the roadway is left to the discretion of the field crew. They may request additional personnel and equipment to provide traffic control and protection for the field crew, but the usual procedure is to use flagmen and cones in light traffic and maintenance vehicles equipped with signs and flashers in moderate to heavy traffic.

When a truck spills its cargo, the police and the highway department are notified. The operator of the truck may obtain help from his company to salvage the cargo, but, if the time required to clear the roadway is critical, the police and highway department provide the traffic control and heavy-duty equipment necessary.

Clearing Ice and Water

Houston experiences flooding and icing several times a year. The Texas Highway Department is responsible for maintaining the roadways under these conditions. Flooding is of short duration, and critical locations are well known. Temporary barricades and signs are erected to direct traffic around these areas.

Icing conditions are also of short duration but usually affect all freeways. If icing is limited to bridges, sand and salt are spread as soon as possible and the roadways are open to traffic. If icing is severe, the freeways are closed to all traffic by barricades placed at the entrance ramps.

RESPONSIBLE ORGANIZATIONS

Police Departments

Houston freeways are under the jurisdiction of the local police agencies. The police departments are responsible for law enforcement, accident investigation, and traffic control around incidents.

Motorcycles

Each of the six freeways is patrolled by two motorcycle officers to provide fast response to a call and to provide visual surveillance of freeway and traffic conditions. These patrolmen assist in the clearance procedure by directing traffic and calling for the appropriate equipment and personnel. They do not conduct accident investigations.

Radio patrol vehicles

The police provide vehicle patrols that cover specific areas of the city. These units are not normally used for traffic control or accident investigation, but respond to incidents in the same manner as the motorcycle patrolmen.

Three-wheeled motorcycles

Three-wheeled motorcycle units are used for point control and limited patrol duty in the CBD. If additional personnel are required to control traffic at a freeway incident, these patrolmen are assigned.

Accident investigators

Accident investigators are assigned to patrol specific areas of the city, but respond to other areas if directed by the central radio dispatcher. There are only 25 accident investigation vehicles on call at any one time. During inclement weather when the accident rate increases, there are not enough accident investigators to respond to each call. Public service announcements advising motorists to report minor accidents to the police station within 24 hours are made by the radio stations.

Wrecker Service

The wrecker service for the clearance of freeway incidents is provided by private companies. The city licensed 100 emergency wreckers that respond to all accidents and another 573 that respond only upon request of the driver or the police. The location of an incident is obtained by the wrecker service in several ways:

1. A driver observes an incident,
2. A call is made by a motorist requesting assistance,
3. Police calls are monitored by the wrecker service, or
4. Police central dispatcher requests assistance of a wrecker service.

The wrecker service usually arrives at the scene of the incident before the police. There are ordinances controlling the activity of wreckers at an incident scene and establishing the rates for services. The procedure that allows emergency wreckers to respond without being called by the police or individuals has both merits and disadvantages. The most important benefit is that the time of response is most assuredly less than that of any other system. The major disadvantage is the disturbance caused by several wreckers waiting at the accident scene for the completion of the investigation. However, with constant monitoring of the wrecker service by the police department and the Houston Wrecker Association, this system will continue to provide good service.

An emergency wrecker can clear vehicles from the freeway before an accident investigator arrives if the accident involves only property damage and the drivers of the vehicles give their consent. There are no charges for this work.

Emergency Service

Houston operates an ambulance service under the supervision of the fire department. Ambulances and fire vehicles are dispatched to freeway incidents through the police and fire dispatchers. Calls for assistance may come from an individual or the police unit responding to the call. The direction and control of these vehicles at the incident are under the supervision of the fire department. Response time for those services is very good. The average response time from the fire station to the accident scene is 5.5 min, and the total time from alarm to arrival at the hospital is 21.6 min. These figures were compiled from data on 28,000 accidents in 1973.

A medical heliport was recently opened in the Houston Medical Center. It has been used to transport patients from one hospital to another over long distances. Helicopters have not been used to respond to an accident location in the Houston area.

Operational Organization

The Texas Highway Department is responsible for the design, maintenance, and operations of the freeway system. Therefore, because freeway incidents impair the operation of the freeway and may also cause damage to the roadway surface or other physical structure, the department must be responsive to incidents.

Roadway conditions

Many of these incidents are handled in a routine manner: the removal of trash and small debris, repair of small pavement failures, replacement of damaged hardware along the roadway, and treatment of roadways for ice and snow conditions. Those incidents that are less predictable and that involve motorists are not handled in a routine way.

Disabled vehicles

The Texas Highway Department provides a limited patrol service that responds to motorists whose vehicles are disabled. This courtesy patrol is free to the motorist, although remuneration is requested on a voluntary basis for the cost of gasoline. The courtesy patrol does not perform enforcement functions such as traffic control or accident investigation, but it does provide protection by signs, flares, and similar devices. The patrol provides limited mechanical assistance, gasoline, oil, and water and assists the motorist in obtaining other necessary services to remove the vehicle from the freeway.

The patrol is under the supervision of the maintenance section of the Texas Highway Department. The patrol follows an established route over 64 miles of freeway and reports incidents by radio to the central office of the department, but the patrol is not dispatched to incidents off the primary route.

In addition, all field personnel of the Texas Highway Department are instructed to render service to motorists on the roadside when it can be done safely. On urban freeways, it is often advisable not to stop unless the service vehicle is equipped with adequate flashers, signs, and other equipment.

Accidents

Most traffic accidents do not require the services of the Texas Highway Department to restore the roadway surface or structure to its potential capacity. There are cases, however, that require special handling such as cargo spill that reduces the coefficient of friction of the pavement, clearance of equipment and vehicles that require heavy-duty wreckers, cranes, or other construction machinery, and the damage of roadway or structures by fire or impact that may reduce the strength of the structure. The department is notified by the police agency when such an event occurs, and the district office dispatches the appropriate personnel and equipment to the scene.

RECENT DEVELOPMENTS IN INCIDENT MANAGEMENT

The procedure for incident management on freeways and arterial streets consists of

1. Detection and location,
2. Response,
3. Clearance, and
4. Report or accident investigation.

Detection and Location of Incident

The time required to detect and locate an incident varies greatly depending on the severity of the incident, location of roadside communications, traffic volume, location of freeway patrols, and many other factors. To reduce this time to a minimum for all incidents would be extremely costly in personnel and equipment, but there are only a few incidents for which cost-effective solutions that would reduce the time by as much as 50 percent are not available.

The usual manner for reporting accidents to the police is by public telephone. The use of a two-way mobile radio by motorists passing an incident scene is increasing and is being coordinated by such organizations as REACT (2). In Houston, other methods for improving accident detection and location time are being investigated.

Emergency call-box system

A few years ago, the Texas Highway Department installed a motorist aid call-box system along 11 miles of I-45 (3). The system design was a four-button radio communication to the central dispatching office of the police department. These four buttons indicated the type of service requested and the location of the call. This system has been removed from the freeway and there are no other roadside motorist aid systems now in operation in Texas, although they are being included in the design of new freeway surveillance and control projects. The failure of this system was not due to the mechanical or electrical operation of the equipment but to the lack of coordination and direction by those agencies responsible for responding to the calls placed on the system. Although the call boxes had separate buttons for fire, service, wrecker, and police, a police patrol unit responded to each call regardless of the type and called for additional service by radio upon his arrival at the scene. There were a large number of false or gone-on-arrival calls received by the police, which prompted the police dispatcher to give a low priority to the call-box system. The reason for the gone-on-arrival calls is not clear, but many can be traced to the length of time required for the response to the call. Another very important factor is that many of these roadside units were placed in close proximity to roadside service centers, such as service stations, shopping centers, and other forms of communications. Still another factor is the charges that were made for services rendered when calls were placed on the call box. These charges were very often much higher than the motorist was willing to pay and, after placing a call often, the motorist decided to seek service on his own. The police were never too enthusiastic about this system and their reluctance to provide faster service to calls placed on this system eventually led to its removal.

Electronic incident detection systems

The research program on the Gulf Freeway Surveillance and Control Project led to the development and testing of incident detection systems. These systems are used to detect shock waves for the operation of warning signs and to alert television monitors of impending congestion and the possibility of accidents or stalled vehicles blocking the roadway. These programs have been successful in detecting shock waves during high and moderate levels of volume (4, 5).

This program is being extended to the study of incidents under low-volume conditions. These programs will be applied on those sections of roadway where disabled vehicles are not easily detected or reached by emergency vehicles.

Also, the study of the detection and computer system has included the system design to automatically detect and locate lane-blocking incidents under high to moderate traffic volumes.

Freeway patrols

The Texas Highway Department's courtesy patrol is being studied to determine whether a more optimum routing and dispatching system can be employed for the same costs, equipment, and personnel.

Television surveillance

Television surveillance is operated on a 6-mile section of one freeway in Houston. The Texas Highway Department has provided observers to man the television system for a 12-hour period from 6:30 a.m. until 6:30 p.m. These observers note any breakdown that significantly affects the traffic flow and safety of the freeway and report by telephone to the police or the highway department's maintenance section the location and type of incident. Under this set of guidelines, incidents located off the travel lanes, such as disabled vehicles on the shoulder or on the outer separation of the freeway, are not usually reported unless they have been involved in an accident or in some other way are distracting to freeway motorists. Television surveillance of course has its limitations, and particularly during periods of light flow the detection of incidents is difficult. Also, the system is not used during hours of darkness, although new developments in CCTV design will enable such use. The department feels that television surveillance is a necessary part of an urban freeway surveillance system and intends to include CCTV as part of its operational design. A microwave television system is being tested for application to remote or temporary installations.

Aerial surveillance

A radio station provides aerial surveillance, but the reports of incidents are not official and are not a part of the highway department or police department. However, it does provide some additional information to motorists, particularly about location of accidents or major incidents. This system is only used during morning and afternoon peak hours and suffers from the same type of visual limitations that television surveillance does.

The Houston area does not have a coordinated system for surveillance except for the Gulf Freeway, which has a system of electronic and video surveillance operated by the highway department and monitored by the police department. The patrol systems, operated independently by the highway department and the police department, provide good surveillance and response to the incident.

Response to Incidents

The second step in clearing incidents from the freeway is response by emergency vehicles. Four years ago, a study was conducted of the time required for police units to arrive at the scene of the accident after receiving a call from the central dispatcher over the radio system. The high variation of response time, from 2 min to more than 30 min, was due to factors such as the location of the accident, availability of police units, condition of traffic, and time of day. On the average, the police responded to incidents on the Gulf Freeway in 11 min. The response time for the wrecker service was considerably shorter, because they stand by the freeways and are able to respond to a police call immediately. A sample of response times during 1974 in this same area showed similar results: an average response time of 11 min. So, although the police have been actively engaged in the operation of the Gulf Freeway surveillance project, police reporting and responding procedures have not changed. At the same time these latest response time studies were made, the staff of the Freeway Surveillance and Control Project measured the response times that were possible by having a vehicle standing by at the control center. The results indicated that the average response time per incident could be reduced by 5 min.

There are approximately 600 accidents per year in the area of the control center. If 300 accidents caused serious congestion, then reducing the time that one or more lanes are blocked is significant. Steps are being taken to have the police department implement this response procedure.

Clearance of Freeway Incidents

Observations have been made on the time required for the police and wrecker units to clear an accident or disabled vehicle from a freeway lane. The distribution varies from 1 min to more than 20 min; approximately 80 percent of the accidents are removed in 4 min or less. Because most of the accidents on our urban freeways are of the minor rear-end type, most vehicles are still operative. The clearance of vehicles from a freeway is of high priority to the police unit at the scene. Therefore, reductions in time for this phase of the clearance procedure are not to be expected. However, vehicles can and should be removed from the freeway before a policeman arrives. If both vehicles are operative and no one is injured, the motorists should move to the shoulder. The Houston Wrecker Association has an agreement with the police department that permits a wrecker to assist in clearance of vehicles before a policeman arrives, if there are no injuries and if the motorists give their consent. There is no charge for the removal to the shoulder. These facts are not widely known by the motoring public, and there are no plans to promote these and other accident procedures with the public.

The time required to clear major accidents involving large numbers of vehicles, heavy trucks, or the spillage of unusual cargo can be improved. It has been the custom to allow companies to provide the necessary equipment to clear their own trucks and cargo if that equipment could be transported to the accident scene in a reasonable period of time. However, during periods of the day when a blockage of the freeway is most critical, this privilege has been restricted. The police and the highway departments have developed an agreement that allows the highway department to provide the necessary heavy-duty equipment to remove trucks or cargo when requested by the police. This agreement was cleared by the Texas Attorney General in the event of legal problems due to vehicle or cargo damage sustained in the removal procedure.

Investigation of Accidents

The time required to complete accident investigation reports was studied in Houston in 1968-69. The time for the investigation of an accident on the freeways ranged from 5 to 70 min and averaged 25 min. The procedure at that time allowed the investigating officer to move the vehicles from the travel lanes to the shoulder of the freeway or to the outer separation and to conduct the investigation at that location. The officer had the prerogative to move the accident vehicles from the scene to a parallel street or to another area in the immediate vicinity if, in his opinion, it was safe to do so. However, most officers preferred to leave the vehicle adjacent to the roadway. The city ordinance on the responsibility of the motorist involved in an accident states that, if the accident results only in vehicle damage, the person shall immediately stop the vehicle at the scene of the accident, or as close thereto as possible, and shall remain at the scene of such accident until he has fulfilled the requirements of exchanging information and reporting to the police officer. It also states that every such stop shall be made without obstructing traffic more than is necessary, a statement that is not understood by the public. In the event that the accident results in injury or fatalities, the vehicle should remain in place long enough to have the occupants cared for and to determine the state and cause of the accident. The highway department has worked with the police department during the past several years to encourage the removal of all accident and emergency vehicles, such as wreckers and police units, from the roadway and from the field of view of the freeway when the accident involves only property damage. The highway department has constructed an Accident Investigation Site (AIS) system on 6 miles of the Gulf Freeway (6). This system consists of designated parking areas—some specially constructed paved aprons under overpasses, others designated curb parking on city street systems in the immediate area of the freeway—where the accident investigation is conducted. The results of removing an accident scene from the field of view of the motorist are obvious to laymen and traffic engineers. The accident scene is a distraction that causes a reduction in speed, which results in a reduction in flow rates past the accident scene, even though all lanes are open. This reduction is 25 to 35 percent of the capacity of the roadway for a six-lane freeway, in the direction of the accident. Often, similar reductions in speeds and flow rates are observed in the opposing lanes of travel.

The benefits of removal of the accident from the shoulder and outer separation of the freeway to a site out of view of the freeway motorists can be calculated in terms of total delay to freeway traffic. Other benefits of reduced congestion on freeway lanes, such as operating costs and the reduction of secondary accidents, can be estimated. The costs of installing such an AIS system, which consists of signing, small special-duty parking aprons, and lighting of the parking area, are low. In a 1-year study of the system on the Gulf Freeway, the benefit-cost ratio was calculated to be 28 to 1, even though only the reduction in delay to the freeway in the direction of the accident was used in the analysis and only 65 percent of all accidents were removed from view of the freeway. It is estimated that 85 percent of the accidents can be moved from the freeway for conduct of the investigation.

A training and public relations film is being prepared on the AIS system and the responsibilities of the public in the event of an accident.

RECOMMENDATIONS

The following recommendations are made to improve the incident management system in Houston.

1. Establish a central communications center for the purpose of maintaining a current-status board for roadway conditions, maintenance and construction activities, and traffic incidents on the urban freeways. This facility should be expanded to include major arterials and suburban freeways as time and resources permit. The center should be directed by an organization that has the responsibility and authority to control all activities on the urban roadway system to maintain safe and effective traffic operations.
2. Improve the frequency of patrols by coordinating the police and highway department activities.
3. Improve the response time of patrols to an incident by permitting the highway department patrol to be dispatched, increasing the number of patrol units during peak periods, and assigning special police units for the purpose of responding to incidents during peak traffic periods.
4. Develop contingency plans for diverting traffic from the urban freeways to the frontage roads and surface street system.
5. Expand the AIS concept to all urban freeways and major arterials.

6. Develop more detailed guidelines for the clearance of incidents that require special equipment and personnel.

These six recommendations are basically people-oriented, and the capital outlay is small. Many of these activities are being conducted daily but simply require coordination with other activities to be more effective. There are many other recommendations that can be made: Develop an electronic and radio surveillance system to reduce the time of detection, increase the size and number of freeway patrol units to reduce the time of response and clearance of incidents, and reform the accident investigation procedure to shorten the time that police personnel are involved. All of these things will be done in time, but emphasis should be placed on those things that can be done now within the constraints of time, money, and organizational responsibilities.

SUMMARY

Management of freeway incidents in Houston may be described as unstructured and loosely coordinated, but responsive and effective. All agencies, acting somewhat independently, provide good service, but service can be achieved if a higher degree of coordination can be developed, if more personnel and facilities can be dedicated to this important function, and if proven new and experimental systems of surveillance and control are made operational.

REFERENCES

1. Goolsby, M. E. Accident Reporting and Clearance Procedures on the Gulf Freeway. Texas Transportation Institute, Research Report 139-1, September 1969.
2. Trabold, W. G., and Reese, G. H. Performance of Volunteer Monitors Using Citizens Band Radio for a Highway Communications Service. Transportation Research Record 495, 1974.
3. Goolsby, M. E., and McCasland, W. R. Evaluation of an Emergency Call Box System. Texas Transportation Institute, Research Report 132-1F, December 1969.
4. Dudek, C. L., Messer, C. J., and Friebele, J. D. Investigation of Lane Occupancy as a Control Variable for a Safety Warning System for Urban Freeways. Texas Transportation Institute, Research Report 165-6, March 1973.
5. Messer, C. J., and Dudek, C. L. Development of a Model for Predicting Travel Time on an Urban Freeway. Texas Transportation Institute, Research Report 165-8, January 1974.
6. Pittman, M. A., and Loutzenheiser, R. C. A Study of Accident Investigation Sites on the Gulf Freeway. Texas Transportation Institute, Research Report 165-1, September 1972.

System Challenges of Traffic Incident Management

H. Nathan Yagoda and Bernard Adler
Computran Systems Corporation

Because of the restrictions and limitations confronting efforts to construct new urban freeways, it has become increasingly important to efficiently operate and manage those facilities that exist. Incident management is one element of freeway operation that is under intensive research and that appears to offer significant opportunity for improving the quality and quantity of service provided. This paper defines the various elements of the incident management system. The intent is to identify system considerations that are significant in the design and implementation of an incident management system including within it the elements of cost, response time, performance characteristics, and trade-off analyses.

An incident detection and response system can be separated into three parts: detection, verification, and response. Figure 1 shows these elements along with the cause, the incident occurrence, and the response, the operational recovery.

One of the objectives of any incident management system is to minimize the time between the start of an incident and the completion of a recovery. As a result, the detection, verification, and response elements must be designed to produce a rapid recovery process. The alternatives available to the designer in the definition and specification of each of these elements are described subsequently.

However, for all configurations, the combined delay associated with detection, verification, and response must be considered in the design process. Speedy detection and sluggish response or sluggish detection and speedy response are equally undesirable.

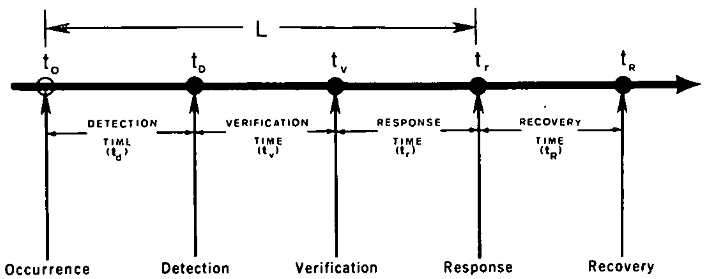
DETECTION ALTERNATIVES

On various U.S. urban freeways a variety of detection alternatives have been used as part of the incident management process (1). In some instances, the detection mechanism contains within it some elements of the verification process as well as elements of the response process. In general, however, detection is an independent function and is accomplished through various mixes of instrumentation, automation, and manpower.

Among the most common detection elements currently used (2, 3) are

1. Motorist call box-telephone systems,
2. Cooperative motorist alarm systems, e.g., FLASH,

Figure 1. Time components of a traffic incident.



3. Patrol-observer systems, e.g., police or rescue,
4. Stationary observer-television surveillance systems, and
5. Electronic surveillance systems.

Motorist Call Systems

One of the first incident detection systems used motorist call boxes (4, 5, 6, 7). With this system, a motorist experiencing an incident proceeds to the nearest call box and informs the operating agency of the nature of the incident being experienced either by selecting a button with a precoded message or by using voice communications or both. The time delay inherent in this type of operation is primarily due to the delay associated with the motorist's determining that an incident has occurred, determining that the proper action involves using the call box, locating the nearest call box, and proceeding to that location to inform the operating agency.

In view of present practices in call box placement, the detection delay in this type of system is often quite significant. In addition, motorist call boxes inherently result in a large percentage of undetected incidents. Many occurrences go unreported because, for example, the incident does not immobilize any one motorist or the motorist attempts to remedy the problem without assistance.

Cooperative Motorist Alarm System

Incident detection that relies on motorist cooperation has been the topic of experimental research for rural freeways and is one of the mechanisms available on toll facilities. One example of this concept is the FLASH system in which passing motorists signal the operating agency by flashing their automobile headlights at an optoelectronic detection system (8).

Another example is verbal communication between motorists and toll collectors at barrier toll stations. In both of these instances, incident reports are received from passing motorists willing to provide cooperation. As a result, a travel time delay from the incident to the monitoring station and the need for multiple reports of a particular occurrence to minimize the false alarm rate contribute to the detection lag time inherent in these systems.

Patrol-Observer System

A variety of patrol-observer systems have been implemented on urban freeways, and in some instances they provide verification and response as well as detection (9). For example, police patrol cars that circulate with the traffic stream provide a commonly used detection mechanism. In these instances, the police patrol detects the incident and verifies the nature and extent of response and the appropriate rescue services. Generally, these vehicles provide little or no assistance for the majority of the incidents requiring aid.

Patrolling service and rescue vehicles provide a similar, but slightly expanded, detection mechanism. In these instances, light-duty service vehicles provide detection, verification, and, in incidents involving minor mechanical failures, response services.

Finally, in high-density traffic on urban freeways, helicopter-borne incident observers have been used for incident detection. Police helicopters are generally used for this purpose and provide both detection and verification.

Stationary Observers and Television Surveillance

Conceptually, the use of stationary observers in strategically located positions is an attractive mechanism for incident detection. However, when two or more observers are required for any given facility, it becomes economically more attractive to use a single observer at a control center and two or more remote-controlled television cameras (10).

Experience with this type of operation indicates that the ability of an observer to detect an incident during a prolonged tour of duty rapidly diminishes. As a result, significant delays are often experienced because the viewer watching the traffic flow does not comprehend that an incident has occurred.

Electronic Surveillance

With the advent of inexpensive, relatively reliable electronic detection equipment and the development of inexpensive digital computer systems, automated surveillance of traffic streams became feasible. With this mechanism, incidents are detected by logically evaluating the variations in flow characteristics (11).

The spacing of detectors along the roadway directly affects incident detection time. Furthermore, because the measured parameters are subject to random variations, data must be averaged in the processing algorithm to ensure that the false alarm rate is acceptably low. Thus, there is an inherent relationship between the delay in the data processing logic and the false alarm rate of such systems.

VERIFICATION

Subsequent to the detection of an incident, the nature and extent of rescue services needed to remedy the problem must be verified. The mechanisms available for this purpose depend on the detection scheme and the competence and reliability of the detection elements. The more commonly used verification processes are

1. Verbal communications by the motorist,
2. Dispatching of land patrol and rescue vehicles,
3. Dispatching of helicopters, and
4. Use of remote-controlled television.

Verbal Verification by the Motorist

As indicated earlier, the verification of an incident and the type of response that is appropriate to remedy the problem are often integrally related to the detection mechanism. For example, in a motorist call box system with voice communications, the dispatcher for the operating agency can interrogate the motorist reporting the incident to determine the character and severity of an incident for which response is requested. Similarly, patrol vehicles provide experienced and knowledgeable evaluators at the scene from whom the dispatcher can determine with reasonable reliability the nature and extent of additional required rescue services.

Vehicle Dispatches

In cases where incident detection does not provide for voice or visual communications, the dispatcher must send an observer to verify incident occurrence and the type of response that is appropriate. For this purpose, two approaches have been used.

The first approach involves dispatching an observer capable of reaching the scene quickly but incapable of providing significant assistance. Typical of this approach is the use of generally available police vehicles, helicopters (12), and the like. The alternative involves dispatching a special-purpose service vehicle that can arrive at the scene reasonably quickly and, in addition, provide the service appropriate for a large proportion of incidents. In this alternative, the verification process is often accomplished by an observer with the capacity to provide remedial assistance for many of the occurrences. In those instances for which additional response is required, the dispatcher is informed and appropriate further action is taken.

Television Surveillance

Under suitable conditions, it is often economically desirable to provide the rescue service dispatcher with the capability of verifying the nature and extent of an incident through the use of remote-controlled television. Using this mechanism, the dispatcher can quickly evaluate the apparent characteristics of the incident to determine the appropriate response. The use of television for this purpose is particularly attractive when electronic surveillance and automatic incident detection are provided, for both approaches use a limited staff at a central control facility.

RESPONSE MECHANISMS

The response capabilities available in an incident management system can generally be separated into three major categories. Class I includes service vehicles capable of providing minimum mechanical assistance, as well as vehicle removal services when the affected automobile remains capable of movement but lacks the power or fuel. When an automobile sustains severe physical damage as a result of collision, overturning, major fire, or even major mechanical failure, generally a tow vehicle capable of clearing the wreckage must be dispatched. These vehicles are aggregated into class II. Futhermore, certain incidents may require the dispatch of selected special-purpose rescue vehicles. Within class III are ambulances, fire fighting apparatus, and sanitation service vehicles.

The particular response methodology used on any specific urban freeway is a design option available to the operating agency. However, the alternatives available for the incident response process somewhat depend on the detection and verification methodologies used. For example, when an incident is verified by a patrol vehicle dispatched to the location at which an incident has been detected, class I rescue services are available. Furthermore, it is frequently considered good practice to provide police patrol vehicles at any location to which class III special-purpose rescue vehicles are sent.

Class I Patrol Vehicles

Patrol vehicles dispatched to incident locations have a wide range of service capabilities. Perhaps most limited of the vehicles in this class are ordinary police cruisers manned by police officers. These vehicles may provide no services other than the initiation of an emergency warning signal to improve highway safety by reducing the probability that additional incidents will result. Furthermore, even the best equipped service patrol vehicle is limited to very minor mechanical repairs. Thus, at best, there is approximately a 28 percent probability that either a class II or class III rescue vehicle will also be required for a typical highway incident.

Class II Tow Vehicles

All incidents in which a vehicle sustains significant mechanical damage and is rendered inoperable require the services of a tow vehicle. Generally, an ordinary tow vehicle can remove passenger vehicles and small trucks. However, there are many incidents in urban areas in which the disabled vehicle is a large truck or bus. In these instances, it is necessary to dispatch heavy-duty tow vehicles with capacity significantly beyond that of ordinary towing trucks.

The primary function of a tow vehicle is to remove a damaged vehicle from the highway. However, as part of this process, it is often necessary to remove wreckage from the areas adjacent to the main roadway or to right a vehicle that has been overturned. In these instances, the towing capacity required may also significantly exceed that of class II vehicles.

The tow truck in class II is restricted to the common size of tow vehicle. Heavy-duty tow trucks suitable for removal of buses, tractor trailers, and the like are generally not available and are included in class III vehicles.

Class III Special Vehicles

Class III vehicles include fire fighting apparatus, ambulances, heavy-duty tow trucks, and sanitation vehicles. These vehicles are generally not required in the more common incidents experienced on an urban freeway. However, estimates based on typical urban freeway characteristics indicate that ambulances and fire fighting vehicles are required approximately 2 percent of the time.

INCIDENT CHARACTERISTICS

To determine the response capabilities that are appropriate for the operation of an incident management system requires that the incident characteristics that are likely to occur be identified. Obviously, incident histories in a particular freeway system will depend on many factors including the geometrics, the nature and characteristics of the users, and volumes and speeds. Although it is not possible to develop a complete characterization that will apply to all freeways, there have been numerous attempts to characterize the relationships between selected identifiable freeway parameters and incident statistics. One typical summary of incident characteristics is given in Table 1 (13). This summary represents a reasonable starting point in the analysis of urban freeway incident problems for use in design of a traffic incident management system.

Analysis of information given in Table 1 reveals that, in approximately 53 percent of the incidents, the problem that occurs is easily solved and the incident is very likely to clear without the aid of external services. Of the remainder, approximately 19 percent involve incidents that require minimal services generally within the capability of the average motorist and certainly well within the capabilities of even the most basic patrol vehicle equipped to provide simple remedial services.

Table 1. Incident characteristics.

Category of Stop	Probability of Occurrence	Vehicle Class Needed
Miscellaneous or nonemergency	0.34	I, II, or III
Tire trouble	0.17	I, II, or III
Mechanical problem	0.17	I, II, or III
Information	0.18	I, II, or III
Gas, oil, and water	0.11	II or III
Medical problem	0.01	III
Fire	0.01	III

Approximately 26 percent of the incidents that occur involve failures that are beyond the repair capabilities of the average motorist but that are easily solved by a trained service mechanic with a reasonably well-equipped service vehicle. The remaining 2 percent involve events and failures that are beyond the repair capabilities of the most advanced technician and mobile vehicle. In these instances, the damaged vehicles must be removed to return the freeway to normal operations.

Within this class of incidents, a wide variety of remedial services are required, including ambulance services, heavy-duty tow trucks, fire fighting apparatus, and sanitation, emergency, and maintenance vehicles.

The types of incidents that require minimal assistance and that are resolved in the shortest periods of time have the highest frequency of occurrence. Similarly, the most severe incidents that require the most significant assistance occur with the lowest frequency. As a result, it is necessary to combine the incident rate with the incident detection, verification, and response time to develop a complete characterization of the traffic incident management problem.

RESPONSE TIME ANALYSIS

A response time analysis can be performed by considering the total detection, verification, and response delay, as well as the time necessary to restore the freeway to normal operations. Alternatively, this analysis can be restricted to the elapsed time between completion of incident verification and initiation of the recovery process during which freeway operations are restored to normal.

Because the total cost to the user is directly related to the elapsed time between the occurrence of an incident and the restoration of normal freeway operations, it is preferable, but more difficult, to consider the former quantity. This quantity can be computed by

$$\hat{T}_e = \sum_{L=1}^{N} P_i(\nu, o)[T_{di}(\nu, o) + T_{vi}(\nu, o) + T_{ri}(\nu, o) + T_{Ri}(\nu, o)] \qquad (1)$$

where

$\hat{T}_e$ = expected elapsed time between the occurrence of an incident and the restoration of normal operations,

P_i = probability of a type i incident, and

T_{di}, T_{vi}, T_{ri}, T_{Ri} = detection, verification, response, and recovery times associated with an incident of the ith type.

Furthermore, all of these parameters are functionally dependent on several variables including volume, speed, and average trip length.

As a first-level analysis effort, it is possible to assume a set of constant parameter values and to conduct an evaluation for a prescribed probability distribution of incidents, and a given set of associated delay times. When this is done, it is possible to develop either a statistical description of delay or a single value of the expected user delay. Similarly, it is possible to investigate any reasonable delay quantities on the detection, verification, and response alternatives available to the system designer. In this way, the cause-effect relationship can be developed between the expenditure of funds for specific elements of the traffic incident management system and the net expected delay in the incident response system.

USER COST ANALYSIS

A second and perhaps more meaningful characterization of the effectiveness of a traffic incident management system is the user cost incurred as a result of incidents. One method of evaluating this cost involves computing the expected value of user cost incurred as a result of incidents on the freeway. An analytic expression for this purpose is

$$\widehat{TC} = \sum_{i=1}^{N} P_i C_{Di} \tag{2}$$

where the expected user total cost $\widehat{TC}$ is defined as a function that is dependent on the type of incident i and the delay associated with the occurrence of the ith event.

Furthermore, it is well known that the cost of an ith type of incident directly depends on the time of occurrence. For example, a major collision at the beginning of the rush hour is significantly more costly than the same type of collision at the end of the rush hour. One method of accounting for this direct dependence on the type of occurrence is

$$TC = \alpha \sum_{i=1}^{N} K_i (t - T_i)^2 U_{-1}(t - T_i) \tag{3}$$

where

TC = total cost,
α = cost per unit time,
K_i = function of the changes in capacity-demand,
T_i = time at which capacity-demand changes, and
$U_{-1}(t - T_i)$ = unit step function.

Equation 3 is derived on the assumption that the cost is directly related to the delay experienced by affected motorists. The delay experienced by motorists is quadratically related to the duration of the incident.

A simple analytic model was constructed to identify the costs associated with the occurrence of an incident and the savings available through the proper management control system. In this idealized model, the peak period was characterized by a 2-hour rush period in which the average demand increased from 5,000 to 7,000 vehicles per hour and then dropped back to 5,000 vph (Fig. 2). The actual short-term demand was assumed to vary about the average by ±1,000 vph. The urban highway section under consideration is a four-lane limited-access road designed to the latest standards, and a capacity of 8,000 vph was assumed. Thus, under normal conditions, the peak-hour demand, which varied randomly between 6,000 and 8,000 vph, was assumed to flow along the highway at an acceptable level of service. The occurrence of an incident was modeled as a reduction in the capacity of the highway. The incident was assumed to occur at T_i min after the onset of the rush period and was assumed to persist for a period of L min thereafter as shown in Figure 2. The incident was assumed to reduce the capacity of the highway to 5,000 vph (i.e., one lane was blocked and a gaper effect resulted).

The change in flow, until dissipation, is mathematically described as

$$\Delta f = (D_2 - C_2)\, U_{-1}(t - t_i) - (C_1 - C_2)\, U_{-1}(t - T_i - L) - (D_2 - D_1)\, U_{-1}(T - T_e) \tag{4}$$

where

C_1 = basic roadway capacity,
C_2 = reduced capacity due to an incident,
D_1 = pre-rush-hour demand,
D_2 = peak demand,
L = length of time incident remains on roadway,
T_i = time after onset of rush period that incident began, and
T_e = time after onset of rush period that incident ended.

Subject to these assumptions, the flow onto the section upstream of the incident exceeded the flow past the incident, and upstream congestion was assumed to occur. A queue of unsatisfied demand resulted (Fig. 3). Associated with this queue is a delay cost representing vehicle-minutes of delay experienced in the congested area. The accrual of this is shown graphically in Figure 4.

From the flow equation (Eq. 4) the queue is derived by accounting for the change in flow for each unit of time. Thus, flow Q is described as

$$Q = \int_{-\infty}^{t} \Delta f \, dt \tag{5}$$

The cost of delay associated with this incident accumulates as the product of the number of vehicles in the upstream queue and the time wasted by these vehicles increase. As a result, the user cost associated with the delayed queue as shown in Figure 4 is directly related to the area under this curve. Hence, the total computation of user delay encompasses the interval from T_i, the occurrence of an incident, through T_e, the time at which freeway operations have completely recovered.

The total cost of delay is given by

$$TC = \int_{-\infty}^{t} \alpha \, Q dt \tag{6}$$

Thus, for this example the total cost is expressed as

$$TC = \alpha \left[(D_2 - C_2) \, U_{-3}(t - T_i) - (C_1 - C_2) \, U_{-3}(t - T_i - L) - (D_2 - D_1) \, U_{-3}(T - T_e) \right] \tag{7}$$

As can be seen from the figures, the shape of the queue and total cost curves depend directly on the shape of the demand growth, the capacity curve in Figure 2, and the instant at which the incident is removed. However, generally the queue increases linearly with the duration of the incident, and, therefore, the cost associated with a given incident varies quadratically with the duration of the incident.

Typically, for a 45-min incident, using the above demand characteristics gives an accrued delay of approximately 1,250 vehicle-hours. Based on an average cost figure of \$3.00 to \$4.00 per hour of delay for the vehicles and passengers and the occurrence of 2,000 peak-period incidents per year, the annual cost to the user ranges from \$7.5 to \$10 million. With traffic incident management, it is estimated that response time can be reduced by one-third to one-half.

Given that the accrued cost varies as the square of the response time, these time savings reduce user cost by one-fourth to four-ninths of the amount before an incident management system. This corresponds to an approximate yearly savings of between \$4.1 and \$7.5 million.

As can be seen from this example (14), the nature of the user cost analysis can be investigated with reasonable ease if certain simplifying assumptions are made. Furthermore, a reasonably reliable estimate of the expected user cost can be computed based on logical assumptions regarding the probabilities of incident occurrence.

SUMMARY

A traffic incident management system is an integral part of any urban freeway system. However, in many prior implementation projects, the design of this critically important element has been left to chance and, in many instances, is evaluated as a patchwork solution developed by the operating agency during the final stages of construction or after the highway has been opened. A simple evaluation of incident statistics indicates that urban freeways with ADTs of up to 200,000 vehicles experience as many as 50 in-

Figure 2. Model for characterization of peak-period demand and capacity.

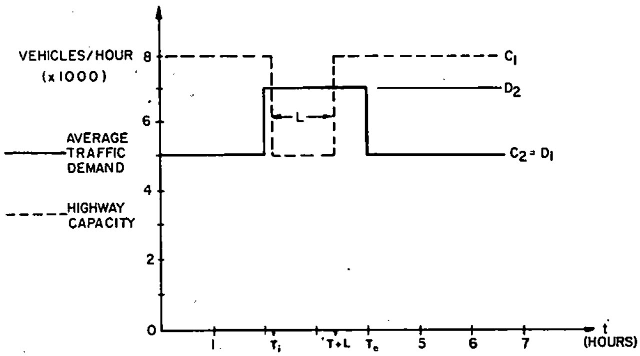


Figure 3. Queue due to unsatisfied demand.

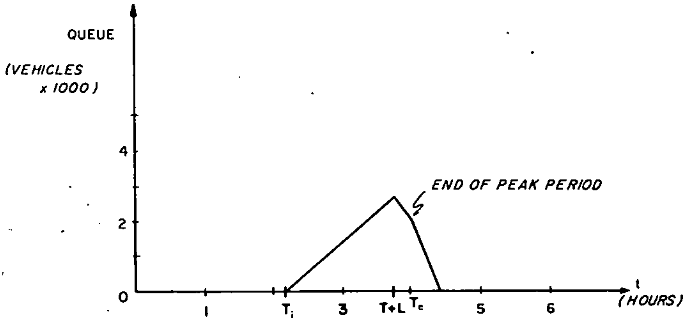
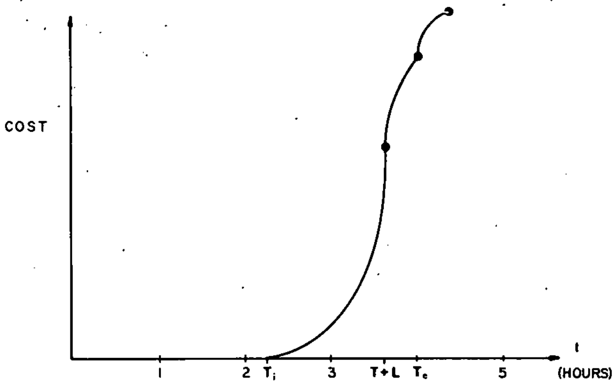


Figure 4. Accrual of cost versus the amount of motorist delay.



cidents per mile per day. Based on this frequency of events, incident management is perhaps the most important single element in maintaining facility production. In view of this, it is extremely important that an incident management system be designed for each urban freeway so as to maximize the benefit-cost ratio for that facility.

There are many challenges to accomplishment of this objective including the following.

1. Analytic procedures must be developed to identify and evaluate the design parameters that affect the performance of an incident management system. Included here are the design relationships that characterize the detection, verification, and response subsystems.
2. Subsystem performance characteristics must be analyzed so as to identify the benefits and disadvantages of the alternative subsystem configurations.
3. Trade-off procedures must be developed to evaluate alternative designs involving man-machine systems for incident management. Consideration must be given to capital and operating costs and equipment reliability, maintainability, and availability throughout the design life of the system.
4. Subsystem components to be used to improve the effectiveness and relieve the burden placed on the operating agency responsible for the freeway facility must be analyzed.

The present situation indicates that the construction of new urban freeways has become extremely difficult, if not impossible. As a result, the emphasis has been shifted to the development of procedures and techniques through which more effective use of available facilities can be achieved. In view of this reemphasis, it is extremely likely that efforts during the next few years will result in the development of sufficiently greater insight into the mechanics of designing and implementing improved traffic incident management systems for urban freeways.

REFERENCES

1. Everall, P. F. Urban Freeway Surveillance and Control. Federal Highway Administration, 1973.
2. Motorist Aid Systems. NCHRP Synthesis of Highway Practice 7, 1971.
3. Mammano, F. J., and Surti, J. Emergency Call Systems for Distressed Motorists. American Highways, Vol. 47, No. 2, April 1968, pp. 3-8.
4. Emery, R. H. Evaluation of Motorist Aid Systems. Federal Highway Administration, 1971.
5. Goolsby, M. E., and McCasland, W. R. Use of an Emergency Call-Box System on an Urban Freeway. Highway Research Record 358, 1971, pp. 1-7.
6. McDermott, J. M. I-80 Motorist Aid System in Illinois. Presented to HRB Committee on Motorist Services, 1974.
7. Green, D. A., and Bregman, H. A. The Northway Emergency Telephone System, Final Report. New York State Department of Transportation, 1970.
8. Adler, B., et al. FLASH—A Cooperative Motorist Aid System. AIL, A Division of Cutler-Hammer, Inc., New York, 1971.
9. Forster, H., Kuprijanow, A., and Pogust, F. Patrol Effectiveness in the Detection of Stopped Vehicles. Presented at HRB 44th Annual Meeting 1965.
10. Goolsby, M. E. Accident Reporting and Clearance Procedures on the Gulf Freeway. Texas Transportation Institute, Research Report 139-1, 1969.
11. Foote, R. S. Traffic Control at the Lincoln Tunnel. Port Authority of New York and New Jersey, 1973.
12. Helicopters to Patrol Freeways Seeking Accident Causing Debris. U.S. Department of Transportation News, 1960.
13. Traffic Surveillance and Control in the West Side Highway Corridor. Computran Systems Corporation, Hackensack, N.J., April 1974.

Work Schedule Changes to Reduce Peak Transportation Demand

Brendan W. O'Malley
Port Authority of New York and New Jersey

This paper discusses a program to reduce transportation congestion by encouraging voluntary work schedule changes in New York-New Jersey central business districts. More than 220,000 workers in more than 400 organizations in Manhattan are participating by staggering their work hours by at least a half hour earlier or later than the customary work schedule of 9 a.m. to 5 p.m. First started in lower Manhattan in 1970, the program has been expanded to include midtown Manhattan and is now under study for Newark, New Jersey. The results of surveys of transportation systems, worker attitudes, work schedules, and productivity are included. The survey conclusions in almost every case were positive and have been used in an ongoing program to expand participation. Elements of the marketing effort are discussed. Also discussed is the recent UNTA technical study grant that will enable the sponsors to fully document the program and provide instructions that can be used by other central business districts considering implementation of staggered work hours to help reduce transportation congestion.

On April 1, 1970, the Port Authority of New York and New Jersey in cooperation with the Downtown-Lower Manhattan Association (D-LMA) initiated a staggered work hours program (SWHP) that is expected to affect more than 2 million workers in the constricted central business districts of the New York-New Jersey region by easing their work trips. Corollary advantages of full participation will be improved efficiency of business operations, reduction of vertical travel time in buildings and building lobby congestion, employee punctuality, better morale, and possibly less crowded lunch periods in restaurants.

After the completion of a successful experimental program conducted solely among headquarters staff of the port authority, the program was implemented in downtown Manhattan (south of Canal Street). In 1972, the program was expanded to midtown Manhattan through the support of major civic and government organizations, and studies have been initiated to determine whether it should be expanded to Newark.

The goal of the program is to ease the problems in commuting to and from work by persuading organizations to alter work schedules from peak commuting times. If participation can be secured on a sufficiently broad scale, future capital requirements for transportation facilities may be decreased.

As the chairman of the port authority noted:

Staggering working hours is definitely a program that could lead to savings on the part of the public transportation operators throughout the country, if it is participated in fully enough. In highly congested central business districts, a major portion of the capital cost for transportation facilities is based on the demand which these facilities are expected to handle during peak travel periods. If this demand could be permanently altered downward by work-schedule changes, we could alter downward the physical requirements.

ROOTS OF STAGGERED HOURS PROGRAM

For many years prior to the SWHP, interest had been shown in staggering work hours to help reduce transportation congestion in the Manhattan CBD. A number of plans were developed and implemented on a small scale during World War II, but were abandoned at the war's end. In the early 1960s, an exhaustive study was commissioned by the City of New York. However, the recommended plan for staggered work hours was never inplemented. In 1970 the port authority and the D-LMA decided to implement the SWHP in lower Manhattan.

Downtown lower Manhattan, the area between the Battery and Canal Street, has undergone tremendous changes. Old inefficient office buildings have been replaced by taller, more modern ones. The growth of office space in Manhattan has been tremendous.

Midtown Manhattan has congestion problems similar to those downtown, but midtown has a greater area, more employment, and greater problems. More than 1½ million people work in midtown, and almost all subway, rail, and bus lines converge on the area.

To cope with the growth in employment, the New York City Transit Authority and the Port Authority Trans-Hudson (PATH) Corporation have provided expanded transportation facilities. However, these improved transit facilities could be used more efficiently if the highly peaked demand for transportation were spread out. Before SWHP was implemented in downtown Manhattan, some port authority studies had indicated that an effective program involving the shifting of a substantial number of workers from the traditional 9 to 5 work schedule offered a promising avenue for relieving transportation congestion.

After the downtown program was under way and its success had been demonstrated, the program was extended to midtown Manhattan. Late in 1972 the Midtown Staggered Work Hours Task Force was created to cosponsor the midtown phase along with the port authority.

In 1974 the Greater Newark Chamber of Commerce requested the port authority's assistance in conducting studies in Newark and its environs to determine the feasibility of staggering work hours there. The studies will cover more than 85,000 workers in downtown Newark. Experience with Newark is expected to be particularly worthwhile because Newark probably is more typical of other U.S. cities from the standpoint of transit and automobile access during commuting hours. The employees are not so concentrated as in uptown or downtown Manhattan and are probably more dependent on transit than are the employees in Manhattan.

In 1973 the U.S. Department of Transportation via the Tri-State Regional Planning Commission awarded a $200,000 technical study grant to investigate and advance the SWHP. Under the grant a number of pioneering areas are being fully documented as the Manhattan SWHP leads the way in the nation. Federal funding is a boost to the program, in terms of financial support for expanding the efforts and the recognition of the Manhattan program as a pioneer in this concept.

One of the most important but time-consuming aspects of the program procedure, and the key to its success, has been the task of convincing business and industry to change from their usual 9 to 5 work schedules. The port authority tackled this problem by securing the assistance of prominent businessmen and creating cooperative task forces composed of civic and trade organizations and public agencies.

The first step followed in the program is to conduct a work schedule survey of all businesses in the area to determine work hours, the number of employees on each schedule, and their mode of transportation and area of residence. If this survey shows that work schedules are unusually concentrated at certain starting or quitting times and that this pattern can be directly related to transportation congestion, individual firms are requested to change their hours of work. The port authority, as project director, maintains a careful record of all contacts and coordinates follow-up with the individual firms requested to participate. This method holds costs to a minimum, and the backing of the prominent civic and business organizations lends force to the requests for participation.

Effects on Transportation

The prime focus of attention has been to determine the effect of staggered work hours on transportation patterns. The goal is to adjust work schedules so that all transportation systems serving Manhattan will have less congestion, which now is largely caused by everyone commuting at the same time.

All transportation professionals are familiar with the twice-daily peaks that occur in American cities. The thrust of staggered work hours is not to accept this demand pattern as a given. When a peaking problem exists, as opposed to a capacity problem, there are ways in which demand can be altered to better use existing facilities. This is the thrust of current national emphasis on low-capital measures to reduce urban transportation congestion. Staggered hours ranks high on the list of such measures.

Before and after surveys conducted as part of the downtown Manhattan effort have shown dramatic reductions in peaking. The more comprehensive surveys in midtown Manhattan have yet to reflect much change because participation there is still low.

As early as 1972, changing demand patterns in three of the busiest subway stations in lower Manhattan and at PATH's World Trade Center terminal provided clear indications that the program was reducing congestion. Passenger counts taken before and during the project indicate that peaking at these locations has been significantly reduced.

Surveys taken in 1970 and 1972 revealed a substantial and continuing reduction in congestion of 26 percent in the peak 15 minutes at three of the busiest transit authority subway stations. In the peak from 9:00 to 9:15 a.m., passenger travel has been reduced from 17,658 to 13,074. In the earlier 8:30 to 8:45 a.m. period, when greater platform and stairway capacity is available because of lighter crowds, passenger travel increased at these three stations from 12,024 before the program to 14,864 at the present time, a 24 percent increase (Fig. 1).

At PATH's World Trade Center terminal in lower Manhattan, similar congestion reduction was determined by comparing 1970, 1971, and 1972 surveys. Passenger counts during the 5:00 to 5:15 p.m. peak are down 18 percent since the program began: from 7,500 to 6,224. During the 4:30 to 4:45 p.m. period, when additional train capacity is available, passenger travel has increased by more than 50 percent from 3,100 to 4,750 (Fig. 2). On PATH, in fact, the demand pattern in the afternoon is almost level for almost 45 minutes—a considerable improvement over the pattern that existed several years ago.

The project sponsors have been gratified by these changes inasmuch as the principal aim of the project was to provide improved levels of service on public transit systems. Greater use of previously underutilized train and station capacity in what were off-peak periods provides this improvement to all subway and PATH passengers.

When congestion has been reduced that much at all rail transit stations in lower Manhattan, as is possible, many thousands of commuters will be affected. This is significant, for about 85 percent of lower Manhattan's workers ride rail transportation to work. It is expected that congestion reduction will continue as additional firms implement staggered hours.

Surveys to assess the impact of the program on transportation in midtown Manhattan are more comprehensive, although we have yet to see large reductions in congestion because fewer than 100,000 of the 1.5 million workers in the area have shifted work hours.

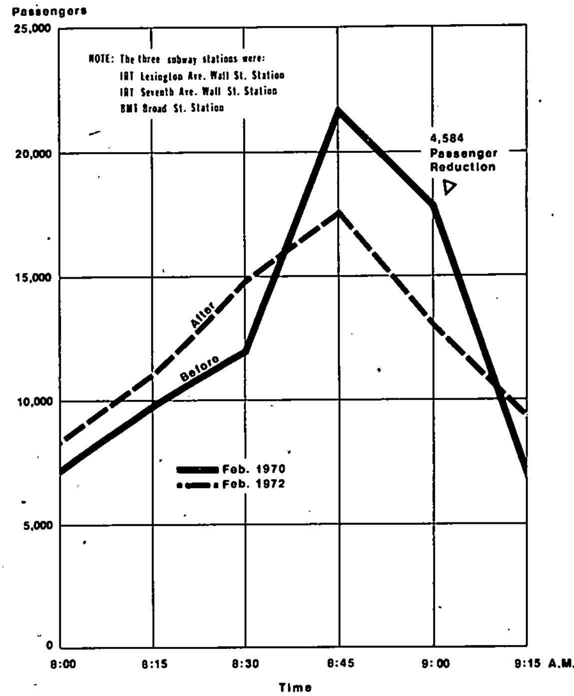


Figure 1. Passenger counts at three major downtown Manhattan subway stations before and after the SWHP.

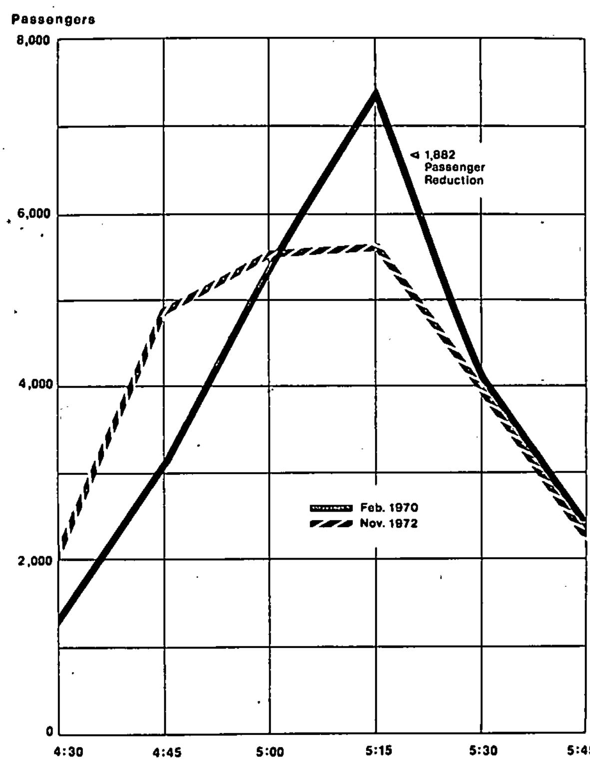


Figure 2. Evening peak passenger counts at PATH World Trade Center station.

Further, midtown differs from lower Manhattan, and the differences complicate the effect of staggered hours on transportation. Midtown is much larger, its employment level is greater, and its complex transportation network intermingles commuter, transit, pedestrian, bus, and vehicular operations. In addition, there is a large overlap of transportation between the separated CBDs of midtown and lower Manhattan. This is being carefully examined in order to avoid nullifying the achievement in lower Manhattan with an "out-of-phase" midtown plan.

A major strategy in midtown has been to focus on the larger firms and to base work schedule recommendations on the number of employees. Those with under 700 staff received recommendations keyed to their postal zip code. Firms with more than 700 employees received recommendations tailored to their work schedule data as submitted to the midtown task force. From a transportation standpoint, the enlistment of a major firm can mean instant and significant congestion relief in the immediate area. For example, when New York Life Insurance's 3,500 employees shifted from 9:00 to 4:30 to 8:30 to 4:00, there was an immediate diversion of over 1,000 people on the IRT and BMT lines serving the 28th Street stations.

We have conducted numerous traffic and passenger surveys to determine the peaking phenomenon on various modes serving midtown. These include rail transit, commuter rail, transit bus, commuter bus, pedestrian, and automobile.

For the surveys, data are gathered on at least 2 weekdays, usually Tuesdays and Wednesdays (Thursday evening is a shopping night in Manhattan and may present unusual patterns). We take 5-minute intervals in almost all surveys, especially in areas of highly peaked demand such as at transit stations.

Queuing at escalators

Commuters arriving for a 9 a.m. start experience heavy backups on escalators in Grand Central Station. On escalators from the IRT Flushing Line exiting to Third Avenue, for example, an unbroken queue forms for 30 minutes until 9:00. Although more than 250 persons wait before 9:00 a.m., the queue sharply reduces to no waiting just minutes later after 9:05 a.m. (Fig. 3).

Subway station counts

Counts of subway riders entering the IND 47-50 Street station at Rockefeller Center reflect the intense 5:00 p.m. quitting time; 75 percent more passengers enter between 5:00 and 5:15 p.m. than in the 15-minute intervals immediately before and immediately after (7,757 versus 4,423 and 4,544). Subway train operations on several midtown lines were found to become strained immediately after 5:00 p.m. when concentrated passenger arrivals contribute to large increases in station dwell times. For example, for IRT No. 7 trains at Grand Central terminal bound for Queens dwell times quickly rise by 50 percent after 5:00 p.m. from an average of 27 seconds per train from 4:30 to 5:00 p.m. to 41 seconds from 5:00 to 5:30. Heavier train loadings and frequent door-holding result in longer travel times and delays to following trains, similar to back-ups on highways.

Subway travel times

Surveys taken at checkpoints on various subway lines traversing the Manhattan CBD show a marked decrease in operating speed during the buildup of the 9:00 a.m. and 5:00 p.m. peaks, particularly in the afternoon.

Automobile volumes at Lincoln and Queens-Midtown Tunnels

A series of traffic counts over the past several years at the Lincoln and Queens-Midtown Tunnels, which connect the Manhattan CBD with New Jersey and Queens, shows that there is very little peaking. Instead, the essentially flat demand curve at both facilities during the peak periods suggests a capacity problem in which the demand is tempered by the ability of the tunnels to accommodate the traffic.

Transit bus

Counts taken on a representative sample of north-south, crosstown, and combination north-south and crosstown bus routes in midtown suggest that buses also are caught up in peak traffic delays or are subject to passenger overloading due to the 9 and 5 habit.

Work schedule and transportation operations relationship

The crux of the analyses we are performing, largely under the technical study grant, is the correlation between work schedules and transportation operations. Obviously, one is related to the other in a CBD, but to what degree and to what sensitivity?

In a first attempt to combine all our data in a single case study, we looked at the peaking problem on the IND E&F subway line to Queens during the afternoon peak. Ridership on this high-capacity route is very heavy, and the equipment is modern and runs well. In a comparison of the quitting times in the zip code area surrounding the Lexington Avenue station, the maximum load point, the pattern of quitting times almost matched the arrivals at the station (Fig. 4). Further, because these arrivals peak sharply after 5:00 p.m., the dwell time correspondingly increases, for more people hold doors on crowded trains. This leads to delaying other trains and slowing train speeds through midtown.

It can be theorized, therefore, that spreading out quitting times in this area would reduce delays to train operations. We are going to great lengths to quantify these results so that clear benefits may be seen from our staggered hours effort.

Transit Service Adjustments

The large-scale shifting of workers to staggered work hours has naturally affected the demand on the area's transit systems, and many have already made adjustments in their schedules. To pinpoint any problems that do occur, transportation survey forms (Fig. 5), the so-called "complaint" questionnaire, are distributed to each participating firm for distribution to employees who report transportation difficulties only as a result of changing their hours.

The approach of following up on any problems is more effective than trying to predict and effectuate service changes prior to (and assuming large participation in) staggered work hours. There are several reasons for this.

1. There is usually no way to accurately predict the level of participation in staggered hours, the work schedule change, transportation mode, residential distribution, or transportation impact from these variables until it in fact happens.
2. When problems do surface, however, the anonymous mail-back questionnaire enables the project staff to pinpoint specific transportation problems, and present them to the appropriate transportation operator, who can make suitable service changes.

Initially, in lower Manhattan, the rate of return of the questionnaires was moderate; about 2,000 were returned. After analysis, D-LMA and port authority staff suggested to area commuter rail and transit system operators certain schedule changes to better serve project participants. As a result, PATH has several times added extra trains to its evening peak service from the World Trade Center terminal to Newark. The New York City Transit Authority improved its IND-BMT E and RR services; the Erie Lackawanna Railroad provided improved service on several of its branches; the Central Railroad of New Jersey, though bankrupt, began running a new main-line train; and the Penn Central adjusted schedules or added cars to its trains, or both, on several sub-

Figure 3. Passenger queuing at Grand Central Station escalators during and after the morning peak.

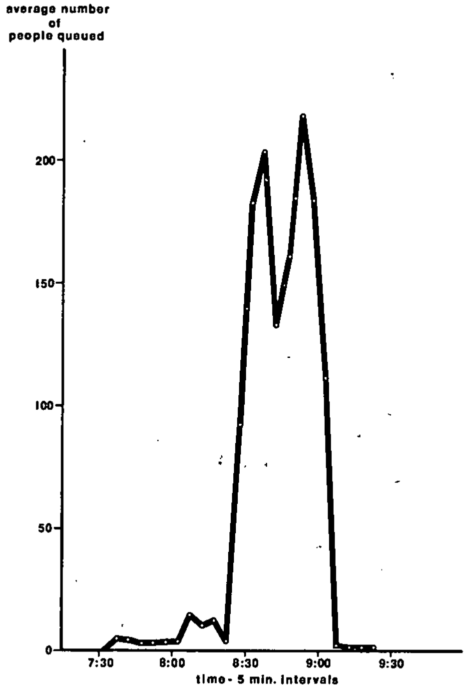


Figure 4. Case study of the peaking problem.

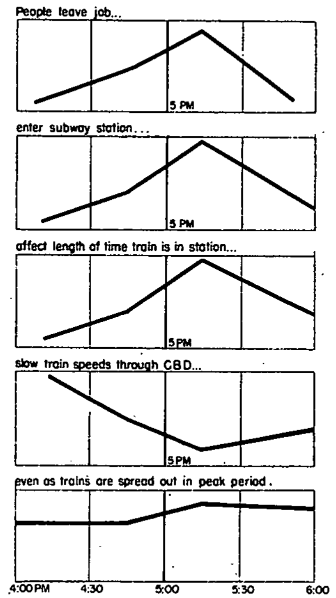
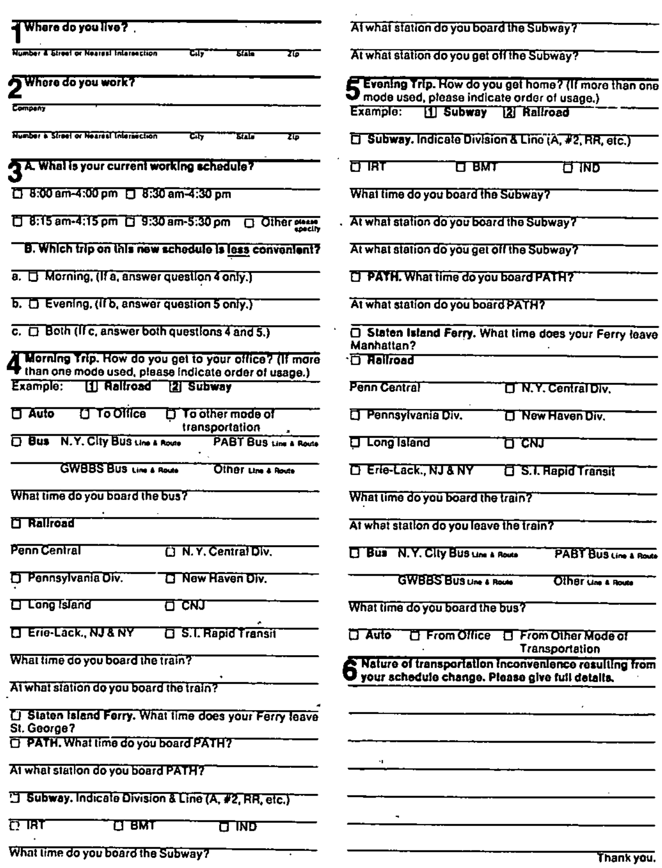


Figure 5. Transportation survey forms ("complaint" questionnaire) distributed to midtown Manhattan participants in SWHP.



1 Where do you live?

Number & Street or Nearest Intersection — City — State — Zip

2 Where do you work?

Company

Number & Street or Nearest Intersection — City — State — Zip

3 A. What is your current working schedule?

☐ 8:00 am-4:00 pm ☐ 8:30 am-4:30 pm

☐ 8:15 am-4:15 pm ☐ 9:30 am-5:30 pm ☐ Other please specify

B. Which trip on this new schedule is less convenient?

a. ☐ Morning. (If a, answer question 4 only.)

b. ☐ Evening. (If b, answer question 5 only.)

c. ☐ Both (If c, answer both questions 4 and 5.)

4 Morning Trip. How do you get to your office? (If more than one mode used, please indicate order of usage.)

Example: [1] Railroad [2] Subway

☐ Auto ☐ To Office ☐ To other mode of transportation

☐ Bus N.Y. City Bus Line & Route PABT Bus Line & Route

GWBBS Bus Line & Route Other Line & Route

What time do you board the bus?

☐ Railroad

Penn Central ☐ N.Y. Central Div.

☐ Pennsylvania Div. ☐ New Haven Div.

☐ Long Island ☐ CNJ

☐ Erie-Lack., NJ & NY ☐ S.I. Rapid Transit

What time do you board the train?

At what station do you board the train?

☐ Staten Island Ferry. What time does your Ferry leave St. George?

☐ PATH. What time do you board PATH?

At what station do you board PATH?

☐ Subway. Indicate Division & Line (A, #2, RR, etc.)

☐ IRT ☐ BMT ☐ IND

What time do you board the Subway?

At what station do you board the Subway?

At what station do you get off the Subway?

5 Evening Trip. How do you get home? (If more than one mode used, please indicate order of usage.)

Example: [1] Subway [2] Railroad

☐ Subway. Indicate Division & Line (A, #2, RR, etc.)

☐ IRT ☐ BMT ☐ IND

What time do you board the Subway?

At what station do you board the Subway?

At what station do you get off the Subway?

☐ PATH. What time do you board PATH?

At what station do you board PATH?

☐ Staten Island Ferry. What time does your Ferry leave Manhattan?

☐ Railroad

Penn Central ☐ N.Y. Central Div.

☐ Pennsylvania Div. ☐ New Haven Div.

☐ Long Island ☐ CNJ

☐ Erie-Lack., NJ & NY ☐ S.I. Rapid Transit

What time do you board the train?

At what station do you leave the train?

☐ Bus N.Y. City Bus Line & Route PABT Bus Line & Route

GWBBS Bus Line & Route Other Line & Route

What time do you board the bus?

☐ Auto ☐ From Office ☐ From Other Mode of Transportation

6 Nature of transportation inconvenience resulting from your schedule change. Please give full details.

Thank you.

urban divisions. Discussions continue with transportation operators on other changes as they become necessary, and all have given their assurances of continuing cooperation in the project.

This follow-up effort, it is felt, gives tremendous credibility to the New York program in that, once a firm has adopted a different work schedule, there is still contact with the project staff. However, the tremendous breadth of transportation services offered in the New York-New Jersey region is certainly more pervasive than in most areas, so the luxury of waiting to hear from the "squeaky wheels" regarding problems incurred is reasonable. Other urban areas may not be able to do this, especially if there is an obvious lack of public transit schedules for periods to which people's work schedules are being shifted. It appears certain, however, that attempting to make major changes in transit schedules before shifts in work schedules occur is a fruitless exercise, especially if there is little success in enlisting significant participation in staggered work hours.

Elevator and Lobby Congestion

It was anticipated that the adoption of staggered arrivals and departures in Manhattan firms would have a beneficial effect on elevator operations in buildings. Studies conducted in the Chase Manhattan, Federal Reserve, and Morgan Guaranty Trust office buildings in lower Manhattan indicated conclusively that staggered work hours reduced elevator waiting times for workers in these firms. This also contributed to the increased employee punctuality reported by many firms.

The studies indicated that, prior to the SWHP, employees in the buildings surveyed experienced delays that averaged more than 2½ minutes during peak periods. This waiting time was reduced by more than half after the project began, with delays averaging only about 1 minute. Another service improvement at the Morgan Guaranty Trust building indicated by the survey was that only 278 persons had to wait for elevators during the program as compared to 673 on a typical weekday before the program. Moreover, the maximum waiting time declined from 6 minutes to 2 minutes. Before the new hours, as many as 112 persons waited in the lobby just prior to 9:00 a.m., whereas after staggering only 37 people, at most, waited for elevator service (Fig. 6).

Three observations are particularly important with regard to elevators.

1. Although some organizations may be apprehensive about staggering work hours to reduce congestion on transportation systems, they will readily identify with elevator congestion in their buildings, especially when it is associated with punctuality problems. This is a strong selling point.
2. Building managers always insist that their elevators provide excellent service and are very sensitive to this issue; therefore, elevator congestion data should be used cautiously.
3. Whereas it may be logical to assume that spreading out the peak load on elevators will reduce the number needed in a given building, the minimum number required is fixed by fire regulations for mandatory quick and complete building evacuation.

Transportation Tardy Study

The transportation operators in the New York-New Jersey region also cooperated in a transportation tardiness study, a detailed examination of the probability and magnitude of train delays during various parts of the morning peak period.

After most participants in lower Manhattan shifted to an earlier schedule, a number of firms reported increased employee punctuality. This was apparently due to fewer and less severe transportation delays in the morning peak period. The tardiness study investigated the relative reliability of rail transportation systems during the morning peak period. For instance, would a person working earlier than a 9 to 5 schedule be less likely to incur a transit or commuter rail delay in going to work? This is of

Figure 6. Passengers waiting for elevators in Morgan Guaranty Trust Bank building before and after SWHP.

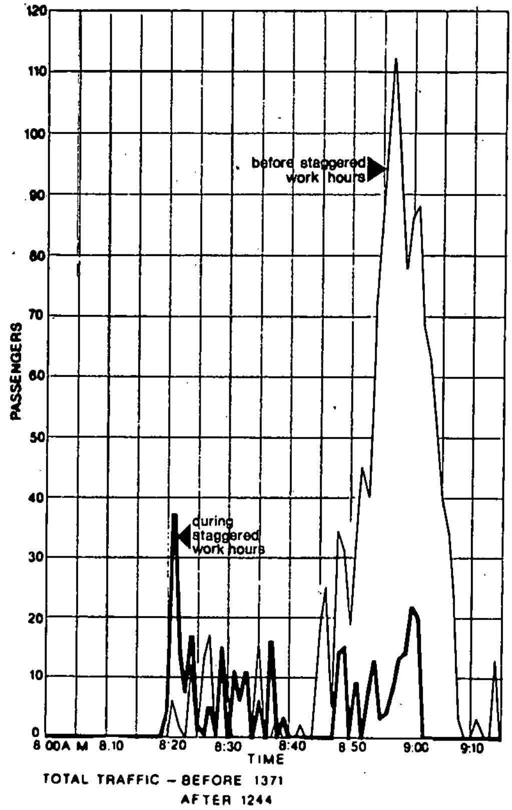
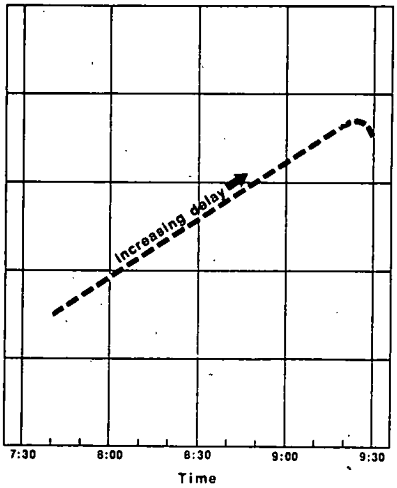


Figure 7. Rail transit delay pattern during morning peak.



obvious importance to the SWHP because, if patterns of train delay were found to build up during the morning peak period, employees who started work at 8:30 could be expected to be punctual more often than those who started at 9:00.

The tardiness study was limited to commuter and transit rail systems, inasmuch as 85 percent of lower Manhattan employees use the rail mode for a significant portion of their work trip. With the cooperation of the rail transportation operators in the region, train "on-time" arrival data were collected for each morning peak-period train for 13 randomly selected days in 1970. Rail systems surveyed included all New York-New Jersey area commuter lines at their inbound terminals, several New York City Transit Authority subway lines at key stations, and the PATH system. For the purpose of the study, a train was considered delayed if it arrived at its checkpoint terminal or station at least 5 minutes behind schedule.

In the tardiness study, significant relationships were found on many of the commuter and transit rail systems studied, which indicated an increasing pattern of train delays as the morning peak period progressed. These patterns represented either an increasing likelihood (or probability) of train delay or an actual increase in train delay, or both, during the major portion of the peak period (Fig. 7).

The tardiness study also revealed that Manhattan employees using the rail systems encountered greater and more frequent delays if they started work at 9:00 than if they started at 8:30. Rail delays were even less before an 8:30 start. Findings for the transit systems studied indicated more than a 25 percent greater likelihood of train delay and over a 40 percent increase in the average length of delay time for a 9:00 start versus an 8:30 start.

The commuter railroad systems serving Manhattan exhibited similar relationships for the 8:30 and 9:00 a.m. starting times. The chance of being delayed is two-thirds greater for the later starting time, whereas the average delay is 50 percent greater for the 9:00 a.m. start.

The difference in length of time delay means, in other words, that Manhattan employees using rail systems in the morning saved on the average more than 1 hour of commuting time every month if they started work at 8:30 a.m. rather than 9:00 a.m. Commuters would save additional travel time if they started before 8:30 since delays before 8:30 are even lower.

Rail transit service for the systems surveyed is almost as frequent for an 8:30 start time as it is for a 9:00 start in the New York area. Between 8:10 and 8:30 a.m., 2,369 trains are scheduled, whereas 2,427 are scheduled between 8:40 and 9:00 a.m.

Work Schedule Survey

Before we can judge whether adherence to certain work schedules contributes significantly to the transportation congestion being experienced, we must determine the prevailing work schedule practices followed by organizations in the area under study.

The port authority has refined the work schedule survey by conducting separate surveys in lower Manhattan, midtown Manhattan, and downtown Newark and among Manhattan advertising agencies and consulting engineers. Each is briefly discussed below. A sample of the work schedule survey form used is shown as Figure 8.

Downtown Manhattan

In 1970, a work schedule survey questionnaire was prepared and distributed by the president of the D-LMA to all its member firms. The mode of transportation used by employees was not asked, for it was feared this might inhibit the preparation of the information and reduce the response rate. Experience with the midtown and Newark surveys suggests this may not be a problem. Approximately 70 percent of the D-LMA membership, 113 firms employing about 136,000 individuals, responded to the questionnaire, and a high concentration of the 9 to 5 schedule was evident: 66 percent

Figure 8. Work schedule survey distributed to midtown firms.

Please complete one questionnaire for your principal location and one for each subsidiary location, if any. Do not complete a questionnaire for subsidiary locations in which a relatively small number of employees (less than 50) are located, such as bank branches, small stores, etc. Please only include locations which are between 14th Street and 59th Street in Manhattan, and immediately adjacent areas.

1. Name of Company: ______________________________

2. Address: ______________________________

3. Number of employees working at this location in Midtown Manhattan between 14th and 59th Streets or immediately adjacent areas: ______________

4. Do all of your employees work on the same time schedule?

   Yes: ______________ Schedule is: ________ a.m. to ________ p.m.

   No: ______________ Go to Question #5

5. If the answer to #4 is "No", how many different schedules does your firm work? ______________________

   Please list approximate number of employees on each schedule:

   (a) From ______ a.m. to ______ p.m. No. of employees: ________

   (b) From ______ a.m. to ______ p.m. No. of employees: ________

   (c) From ______ a.m. to ______ p.m. No. of employees: ________

   (d) From ______ a.m. to ______ p.m. No. of employees: ________

6. Would you estimate how many of your employees live in: *

   New York City

   Manhattan ____________ New Jersey ____________

   Bronx ____________ Rockland & Orange ____________

   Queens ____________ Long Island ____________

   Brooklyn ____________ Westchester, Putnam & Dutchess ____________

   Staten Island ____________ Connecticut ____________

7. If such information is readily available, would you estimate how many of your employees use the following modes of transportation for the major portion of their commuting trip? Please do not complete this question if it would require an arduous and costly firm-wide survey.

	Approximate Number of Employees
Subway	____________
Commuter Railroad into Grand Central Station	____________
Commuter Railroad into Penn Station, New York	____________
PATH System into 33rd Street Station	____________
Bus into Port Authority Bus Terminal	____________
Automobile into Manhattan Central Business District	____________
Other modes please specify	____________

____

If your firm maintains automated records of employee residences by zip code, such zip code summaries would be quite helpful.

were scheduled to start at 9 a.m., and 64 percent quit work at 5 p.m. The results of the survey revealed that the lower Manhattan firms had not adopted a staggered work hours system to any considerable extent even though they had been urged to do so by the D-LMA as early as 1961.

Midtown Manhattan

The midtown work schedule survey, conducted in 1972, gathered information on the work schedules, residences, and travel habits of almost 300,000 workers from 1,450 firms. This represents about one-fifth of total midtown Manhattan employment. The results indicated that about 54 percent of the midtown employees were scheduled to start at 9 a.m. and quit at 5 p.m. In contrast, only 15 percent began at 8:30 a.m., the next most preferred starting time.

The midtown work schedule survey was the most comprehensive of those performed. Questionnaires were sent to the 26 member business, civic, and trade organizations of the Midtown Task Force on Staggered Work Hours, an organization specially set up for the purpose of communicating with the business community.

Of the 2,800 questionnaires mailed, 1,450 were returned. Of these, 1,192 qualified on the basis of location. These returns represented a sample size of about 300,000 midtown employees, about 20 percent of the approximately 1.5 million persons employed in midtown Manhattan. Further validity of the sample was sought by comparing the results of the survey to the journey-to-work study conducted by the New York/New Jersey Transportation Agency in 1961.

Comparison of the peaking phenomenon in similar surveys taken a decade earlier indicated that peaking had apparently increased. In all probability this was caused by the decline in manufacturing jobs in the Manhattan CBD and an increase in office and clerical jobs, with relatively more dependence on shorter workdays and peak-hour schedules. Since the future of Manhattan will undoubtedly be toward expansion of the office function, the peaking phenomenon can be expected to worsen unless countermeasures are taken.

Advertising Agencies and Consulting Engineers

Separate work schedule surveys were conducted of Manhattan advertising agencies and consulting engineers to ensure complete understanding of their work scheduling practices. In addition, these two industries were not well represented in the 26 midtown task force associations.

Conducting the surveys was considered worthwhile because it revealed the variety of scheduling practices that exist from industry to industry. It also provided a means to initiate communications with many firms not already contacted.

Newark, New Jersey

A work schedule survey was conducted in Newark in 1974 to determine work scheduling practices by downtown Newark organizations and whether a staggered work hours program is desirable to reduce peak transportation demands.

Employee Response (Attitude)

Although the port authority's prime interest in staggered work hours was to determine whether it would relieve transportation congestion, we realized that the program would never succeed if the people involved reacted negatively to revised work hours. To study this aspect, the port authority and the D-LMA engaged Dr. Derek Phillips, associate professor of sociology at New York University. The necessary data were obtained by distributing questionnaires to the employees of 27 of the companies participating in the downtown project. Employment at these firms ranged from some 5,000 or more to 400 or fewer. The survey returns included ample representation from firms in the four major segments of the lower Manhattan business community: banking, insurance, investment, and corporation headquarters groups.

The study found that about 85 percent of the employees sampled had a favorable overall reaction to the project. With regard to organizational efficiency, the results emphasize that, in all the industry groups, the changed hours had very few negative effects on efficiency. In fact, some organizations reported positive gains in work effectiveness.

The positive way in which the SWHP is viewed was apparent in answers to the question, On the new schedule, does the workday seem to you to be longer, shorter, or about the same? The results showed that, even though people were working the same number of hours, three times as many felt the day was shorter under the new schedules. Moreover, although nothing was changed but their work schedule, four times as many were more satisfied with their jobs than were less satisfied.

Because the SWHP is aimed at relieving the peak-hour congestion that workers in lower Manhattan face in commuting to and from work, it was of interest to find out how the participants felt about their daily commuting. Almost 50 percent of those who changed their work schedules reported that they were more satisfied with their trips to and from work. Only 9.8 percent were less satisfied. These findings were highly significant, for they indicated that the improvements in transportation are very strongly perceived by those on staggered hours.

Those whose schedules changed and those whose did not also were asked to evaluate elevator congestion during the staggered hours project. Again, the responses were surprisingly positive: 45 percent reported less elevator congestion, whereas 50 percent reported no change.

Responses to questions to determine whether home lives were disrupted in any way by the new schedules, i.e., whether the new schedules caused inconvenience, show that, although certain changes did occur, they were viewed positively by the participants.

From the enthusiastic reactions of the participants in both government and industry and the positive effects recorded thus far on transportation facilities, the staggered hours project in lower Manhattan can be termed a success. Similar surveys conducted by companies in Manhattan have confirmed both the direction and the degree of the findings.

Management Response

Regardless of how employees feel about staggering work hours, such a program cannot be implemented unless management is convinced that the efficiency of operations will not suffer. Phillips investigated this aspect and reported that

1. Six times as many supervisors reported gains in productivity under the new hours as reported losses and
2. The punctuality of employees increased.

In summary, all surveys have indicated that the changed hours had very few negative effects.

One of the important specifics on the question of work efficiency was whether shifted starting times affected internal or external communications among participating firms. A substantial majority of unit heads surveyed reported that no severe communications problems resulted from the changed hours. About 15 percent cited some impact, but evidently the problems were not sufficient to cause a drop in efficiency.

In discussions with company representatives before the experiment, many commented on current problems of employee punctuality and asked that this area be studied. For this reason, unit heads have been specifically asked about punctuality. They reported that staggered hours appear to have a beneficial effect. Compared to previous experience, almost 80 percent of the supervisors said their employees were arriving on time or earlier under the new schedules. Only 11.6 percent reported that they were arriving later. This generally agrees with the findings of the tardiness study discussed earlier.

It is also worthwhile to note the individual experience of major firms in adjusting to new work schedules. For example, the New York Life Insurance Company reported, "It is gratifying to be able to report that service and productivity had been very adequately maintained during the two months of experimental activity on staggered work hours." After adopting a permanent staggered work hours program, the Bristol Myers Company, Union Carbide, Inc., Westvaco, and Sears, Roebuck, Inc., made similar comments.

Firms that communicate regularly with western time zones often express concern about keeping in touch on earlier staggered schedules. We therefore canvassed several major New York corporations that have successfully shifted to staggered hours, and none has found this to be a serious problem. There are several reasons for this; the major reason is simply that the new work hours required only a small adjustment in communications habits.

General Program

Although the approach to individual companies and government organizations to change their work hours has remained relatively similar since the earliest efforts, there has been constant improvement in the sales technique as additional studies provide success stories and enable us to develop additional sales and promotional tools. It is also important to note that efforts to implement the SWHP in the New York-New Jersey region were developed from scratch and few guidelines were available.

Generally, the procedure has been as follows. After sponsoring organizations have been secured, as many companies as possible are surveyed to determine their work hours. Subsequently personalized requests to consider SWHP are made to company officers. In many instances, questionnaires are prepared for distribution to the company's employees to assist the company in deciding whether to experiment with staggered work hours. Basically, these questionnaires request employees' opinions on their work schedule preferences and the effects of a change of hours on their home lives and their journey to work. Having a company agree to survey its staff usually ensures eventual participation, for employee interest is invariably overwhelmingly favorable. After the change has been made to a different work schedule, questionnaires are distributed to determine employees' attitudes and actual effects on transportation and living habits.

The following sequential steps are usually taken in an organization considering the adoption of staggered hours:

1. Discussion with company officials and presentation of the results of the SWHP in other companies, usually via an association of which the firm is a member;
2. A work schedule survey of the hours actually worked and the age, sex, marital status, place of residence, and mode of travel of the employee;
3. A subsequent survey of the employee's attitude toward the shift in work hours generally tabulated by department, age, sex, and so on; and
4. A summary report on the results of the staggered work hours experiment.

Concurrently, when a sufficient number of employees have shifted their hours, studies are made of the congestion on local streets and at transportation facilities as well as in building lobbies. Management is also questioned on data regarding the efficiency of employees on new schedules, the rate of absenteeism and tardiness, and any operating problems such as faulty communications with out-of-town offices or customers that might have occurred.

All of the material received is used in personal solicitation with firms that have not yet joined the program, and every effort is made to circulate literature and newspaper articles that will encourage further participation.

Specific Case Histories

It is illustrative to describe briefly the case histories of a number of Manhattan organizations that have adopted staggered work hours in the last 4 years. Additional information on any of the situations may be obtained from the author.

1. Union Carbide switched all 3,500 headquarters staff from 9 to 5 to 8:30 to 4:30 after a survey showed great preference for earlier hours. A later survey indicated that about 85 percent like earlier hours and that management was very pleased.
2. Bristol-Myers adopted a program similar to Union Carbide's although it took more than a year to convince management to make a staff survey, which showed 84 percent preferring earlier hours. The company switched to 8:30 to 4:30. It has now adopted a floating day, which allows staff to come in between 8:00 and 9:00 a.m. each day and leave the corresponding number of hours later.
3. Westvaco management was immediately enthused about staggered hours and conducted a survey of staff showing 75 percent like earlier hours. Westvaco shifted within a month to 8:30 to 4:30, and the after survey again showed that 86 percent favored staggered hours.
4. New York Life Insurance was on 9:00 to 4:30 schedule, and it took more than a year of persuasion, analyses, and surveys—mainly harping on their congested 9:00 a.m. start time—to convince the firm to try an 8:30 to 4:00 schedule for its 4,000 staff. Because of more free evening time and better commuting, 87 percent of employees like the new schedule. N.Y. Life adopted the earlier hours in the summer, which is an excellent time to make such an experimental change. Since the change, N.Y. Life has received much publicity from participation and has enjoyed the exposure.
5. Celanese Corporation shifted 1,000 staff to 8:30 to 4:30 after 2 years of effort. Critical were direct contact to the president of the corporation by staggered hours chairman Andrew Heiskell and also the company's recent move to consolidate staff.
6. Sears, Roebuck, the nation's largest retailer, shifted from 9 to 5 to 8:30 to 4:30 and 8:45 to 4:45 after a preference survey conducted by port authority staff. Much publicity was accorded its participation as the 400th firm to join the program.
7. McGraw-Hill switched from 9 to 5 a few weeks after a breakfast meeting sponsored by Rockefeller Center. Most of its staff adopted 8:30 to 4:30 schedules, although about 25 percent chose 9:30 to 5:30 hours.
8. Montgomery Ward adopted the floating day concept after moving from 9 to 5 (similar to Bristol-Myers).
9. Lever Brothers, after moving to 8:30 to 4:30, shaved 15 minutes off the lunch period to go to a 4:15 p.m. quitting time.

Promotion

As previously mentioned, no avenue of potential promotion should be overlooked if staggered work hours implementation and significant reductions in traffic congestion are to be achieved. Although these activities may seem out of the realm of transportation professionals, they are essential to enlist the broad participation needed to make staggered hours successful. Some examples of far-reaching promotional efforts are discussed below.

Real estate management firms and their leasing agents have been quite cooperative in the SWHP. Their properties are more attractive when lobbies are uncrowded and elevator service rapid. There may, however, be some resistance to staggered hours, for they may feel that future prospects will consider the program as necessary because of an inadequate physical plant. Appropriate real estate contacts can provide lists of new tenancies for staggered hours solicitation. In addition, spreading hours should promote sales in a building's ancillary facilities, such as stores and restaurants.

Minimum congestion in stores provides ease in shopping and frequently more space for new lines of merchandise.

All media should be contacted to the fullest extent possible. Newspapers should be used to run news stories and the resultant fact sheets distributed to prospective program participants. Under no circumstances should any opportunity for TV and radio interviews be ignored, and the television stations in particular should be advised that early quitting times may provide them with a larger night prime time audience. Public service announcements can be developed for TV and radio to take advantage of the nonprofit nature of the program. These can often be done by advertising agencies on an out-of-pocket cost basis. A complete program of press releases for all newsworthy items is essential.

Banks should be urged to schedule their retail hours for early opening to accommodate program participants. Where banks do adjust their hours, tie-in promotions should be encouraged.

The trucking industry can be advised of the benefits of earlier deliveries, particularly if participants have an early opening and a late closing. Several trucking firms have requested the names of firms on staggered hours to facilitate earlier deliveries.

A formal program including a newsletter and up-to-date mailing lists should be maintained. Periodic reports should be issued to participants and prospective participants. Participating companies should be urged to include in "help wanted" advertising the fact that they are on preferred hour schedules.

Wherever possible, restaurants should be encouraged to provide "early-bird" breakfasts, say to 8:30 a.m., and cocktail lounges can provide early "happy hours," say to 4:30 p.m.

In addition, the SWHP should be run in a strictly businesslike manner with attractive stationery, business cards, and descriptive hand-out materials. Brochures keyed to staggered work hours can be distributed with telephone and department store bills to encourage increased participation and to further public knowledge of the program. Business associations should be encouraged to report in their house organs particular items of staggered work hours interest.

A very successful event calling attention to the Manhattan program was held on May 7, 1974, at the MTA's new 57th Street subway station. A ceremony there highlighted the installation of some 12,000 staggered work hours cards and posters (Fig. 9) that depict crowded conditions in a cartoon. On May 7, New York City public and corporate officials plus several hundred men and women from participating firms simulated with good spirit and much enthusiasm the crowded conditions shown on the poster. The event attracted wide coverage by TV and radio stations and newspapers.

FLEXIBLE HOURS, 4-DAY WEEK, AND OTHER CONCEPTS

There are other work schedule concepts that may relieve traffic congestion in a central business district without major investment in physical facilities. Notable is the use of flexible work hours or a shortened workweek. Flexible work hour scheduling, also called flextime, Gleitzeit, and gliding work hours, is spreading rapidly in Europe and is now being tried by many American firms. Basically, this concept permits employees to set their own starting and quitting times around a "core" time when all must be at their work station.

Not only does the flexible hours concept have the advantage of reducing traffic congestion, but also it has been found to be personally desirable. Employers find that more time is actually spent working by staff. It is notable, however, that the flexible hours concept is more difficult to adapt to assembly line operations, particularly where groups are unionized or where overtime payments are a factor.

The port authority plans to experiment with flexible hours. This experiment will probably include about 1,000 professional and clerical personnel and, as with the original port authority SWHP experiment, will provide in-depth documentation on the effects of flexible hours on transportation, efficiency, and worker attitudes.

Figure 9. Card used to advertise the staggered work hours program.

The concept of a shortened workweek, usually 4 days, has also been receiving a great deal of attention. Under this concept, total work hours are compressed into fewer days or the total number of hours worked is reduced. The trend has been based on efforts both to give employees another full day of leisure and to conserve energy and reduce transportation congestion.

FEDERAL TECHNICAL STUDY GRANT

In 1973, the port authority was awarded a technical study grant of $133,000 from the U.S. Department of Transportation to study the SWHP in midtown Manhattan. The grant, made by the Urban Mass Transportation Administration, is being administered by the Tri-State Regional Planning Commission.

In making the grant, UMTA asserted its support of efforts by Manhattan firms to relieve transportation congestion through voluntary staggering of work hours. Support was expressed for such low-capital projects as a way in which local communities can do something now about urban transportation congestion.

The port authority's responsibility will include the preparation of a technical report detailing the nature of objectives of the project, the work performed, and an analysis of the results with concomitant recommendations. Also included will be a manual containing criteria for determining the feasibility of staggered work hours in any CBD. The manual will discuss ways and means of developing alternate work schedules and an evaluation of the many variations of staggered work hours in terms of their effect on transportation peaking. This will include split shifts within an individual company, flexible hours, voluntary versus employer-prescribed hours, and the 4-day workweek.

CONCLUSION

At the risk of being repetitious, it should be emphasized again that the success of any staggered work hours program is dependent on constant documentation, promotion, and sales effort. The gains in employee comfort and the reduction in transport congestion from SWHP have been amply demonstrated by many studies, including those reported in this paper. Implementation, however, has generally not achieved satisfactory levels of participation in many areas because sales effort and promotional programs have not been stressed. As noted, the backing of influential businessmen and civic leaders is a prerequisite to a successful program. Obviously, these leaders must first be convinced of the value of the program to their own organizations, to their personnel, and to the community as a whole. After this, success can only come through constant unremitting effort.

Pricing and Transportation Capacity

Herbert Mohring
University of Minnesota

Both consumer behavior in taking trips and the rules for optimizing the facilities on which these trips are taken are similar to the corresponding phenomena of relevance to any other commodity. In traveling, consumers act as if the price they pay for a trip is its direct money cost plus the time required times the value of travel time. Using travel facilities efficiently requires that consumers be charged tolls that force them to take into account the costs their trips impose on other travelers. Optimizing travel facilities requires that the sum of the direct resource costs and the value of consumer-supplied travel time be minimized. Failure to charge appropriate tolls yields welfare losses that are of substantial size on heavily congested facilities. Facilities that operate under the constraint that efficient tolls cannot be charged are developed for optimizing facilities. The implications of these rules for expressway travel are roughly quantified.

The basic purpose of this paper is to inquire into the role of price in the use and development of transportation facilities. More specifically, it seeks to answer the following questions:

1. What is the nature of the price that governs consumer decisions on the number, destination, and timing of trips?
2. What is the nature of the price that must be charged households if existing transportation facilities are to be efficiently used?
3. What price should govern social decisions regarding the expansion or contraction of transport facilities when the prices charged consumers are and are not those required for the efficient use of these facilities?
4. Finally, what costs does society incur as a result of its failure to price these facilities efficiently?

The prices of relevance to household travel decisions and to social decisions regarding the efficient use and expansion of transportation facilities are quite similar to those for any other commodities. This is true despite the fact that transportation is, in a way, qualitatively different from most of these other commodities. Decisions (whether efficient or inefficient) about the nature and pricing of transportation facilities have effects that pervade the economy. To explore these effects fully requires a truly monumental study. This analysis is therefore restricted to identifying the quantitative implications of inefficient pricing of highways from two very simple models.

HOUSEHOLD TRAVEL DECISIONS AND THE NATURE OF EFFICIENT PRICING AND INVESTMENT PROCEDURES

In benefit-cost analyses, time savings have long been regarded as a benefit of highway improvements. If an improvement cuts an hour from the time required for a truck round trip between here and there, a trucking company saves 1 hour's truck driver wages for each round trip it provides. It clearly seems appropriate to regard this saving as a benefit of the improvement.

Although the legitimacy of doing so is still occasionally questioned, the same sort of procedure is also commonly used in dealing with the benefits consumers derive from travel time savings. That is, if an improvement reduces the time required for a particular type of trip, this time saving multiplied by the number of trips of that type that are taken multiplied by some more or less arbitrarily selected value of private passenger vehicle travel time is normally recorded as a benefit of the improvement.

Two fundamental questions about this prevailing procedure for dealing with consumer travel time savings deserve answers: First, does the representative consumer value time savings in the same manner as the truck operator? That is, is the way in which he values these savings equivalent to multiplying the number of hours saved by an appropriate hourly value? Second, even if this is the way he takes them into account, might it not be appropriate to attach a different value to these savings in striving to allocate society's scarce resources efficiently? For example, an argument once commonly heard was that time saved in private passenger vehicle travel should be assigned a zero value inasmuch as this travel is not directly involved in any productive activity. Might not this or related views be valid?

As it turns out, the answers to these two broad questions are respectively yes and no. That is, utility-maximizing consumers do value travel time savings in essentially the same fashion as do truck operators. Furthermore, analysis of transportation improvements should take these consumer valuations into account. The validity of these assertions is demonstrated by analyzing two simple problems. First, consumer i, one of n individuals in a society, is assumed to derive utility from consuming a general-purpose commodity ("everything else") and a specific commodity—what happens there (be it work or recreation) when he takes bus trips from here to there and back. However, the consumer derives disutility from the time he spends taking these trips. He is subject to a budget constraint

$$I^i = a^i + Fb^i$$

where

I^i = his fixed dollar income,
a^i = his everything else consumption level (priced at \$1 per unit),
b^i = his bus trip consumption level, and
F = fare charged per bus trip.

The analysis demonstrates that, in maximizing his utility under these circumstances, the consumer acts as if the price he pays for a bus trip, P^i, equals $F + v^i t$ where t is the time required for a trip and v^i is his value of travel time, the amount of everything else he would be willing, at the margin, to give up in return for saving an hour's travel time. (This problem can easily be reinterpreted to account for any other transportation mode. For example, b^i could be interpreted as the number of automobile trips the consumer takes between here and there and F as the sum of his vehicle operating costs per trip and the highway user charges imposed on him. X, defined later as the number of bus-hours of service provided on the here-there route, would then become a measure of the capacity of the here-there highway.)

The second simple problem involves assuming that society wishes to maximize social welfare, a function $W(u^1, \ldots, u^n)$ of the utility levels, u^i, of its individual members. In doing so, it is subject to a constraint on total output that can be written

$$R = A + CX$$

where

A = total amount of everything else produced for all consumers,
X = number of bus-hours of service devoted to providing here-there trips, and
C = cost of an hour's bus services, i.e., the number of units of everything else given up to obtain these services.

The time required for a here-there round trip is a function $t(B, X)$ of X and the total number of trips taken, B. The analysis demonstrates two basic conditions for this maximization problem. First, the fare for a bus trip must be set equal to the cost that that trip imposes on other travelers

$$\Sigma b^i v^i \partial t/\partial B = BVt_B$$

the total number of trips taken ($B = \Sigma b^i$) times the weighted (by number of trips taken) average value of travel time ($V = \Sigma b^i v^i/\Sigma b^i$) times the change in the time required for a trip resulting from an additional trip being taken. Second, given that this efficient or benefit-maximizing fare is charged, maximization of social welfare requires that the quantity of bus service provided be that that minimizes $CX + BVt$, the dollar cost of bus service plus the value consumers attach to the time they spend traveling. This is essentially the same cost-minimization problem that would be faced by a truck operator in determining the total number of trucks to buy and the number of driver hours to hire in providing truck service between here and there (5).

In the first problem, consumer i's goal is to maximize his utility, $u^i(a^i, b^i, \tau^i)$, subject to the budget constraint, $I^i + a^i + Fb^i$, where $\tau^i = b^i t$ = total travel time. (Alternatively, the consumer's utility could be regarded as depending on time per trip, t, rather than total travel time, bt. Although the algebra would be more complicated, the conclusion reached would be unchanged.) Setting up the Lagrangian expression

$$z^i = u^i(a^i, b^i, \tau^i) + \lambda^i(I^i - a^i - Fb^i)$$

and differentiating with respect to a^i and b^i yield

$$z^i_a = u^i_a - \lambda^i = 0 \qquad (1)$$

$$z^i_b = u^i_b + u^i_\tau t + u^i_\tau bt_B - \lambda^i F = 0 \qquad (2)$$

as first-order conditions for utility maximization where subscripts refer to partial derivatives. It seems reasonable to suppose that consumer i does not take into account the effect his trips have on his own travel time. If so, $u^i_\tau bt_B$ can be ignored. Dividing Eq. 2 by Eq. 1 then yields

$$u^i_b/u^i_a + t\, u^i_\tau/u^i_a = F \qquad (3)$$

In Eq. 3, u^i_τ/u^i_a is the ratio of the marginal disutility of travel time to the marginal utility of dollars in dollars per hour. It therefore seems reasonable to substitute for this ratio $-v^i$, the money cost consumer i attaches to his travel time. Doing so changes Eq. 3 to

$$u^i_b/u^i_a = u^i_b/\lambda^i = F + v^i t \qquad (4)$$

This relationship says that consumer i will set the ratio of the marginal utility of bus trips to that of dollars equal to the fare plus the time cost of a trip.

In regard to society's welfare maximization problem, the economy is assumed to be subject to the constraint

$$R = A + CX$$

where

R = weekly flow of services from the available stock of resources,
A = weekly consumption of everything else, Σa^i, and
C = resource service cost of providing a bus-hour of services.

Implicit in this formulation is the assumption that the cost of a bus-hour is independent of both the carrying capacity of the bus and the speed at which it travels. Although clearly unrealistic, this assumption is not so bad as it may seem. An increase in bus capacity results in a less-than-proportionate increase in its capital, fuel, and related costs. More important, driver wages and fringe benefits account for about 70 percent of the total costs of typical urban bus operation. The cost of a driver-hour is independent of the size of the bus he operates. This implicit assumption has two consequences for the analysis. First, there is always room aboard a bus for an additional passenger. That is, no one must ever wait for another bus to come because the first bus passing his stop is loaded to its capacity. Second, the bus company incurs no direct cost as the result of serving an additional passenger. The additional passenger affects the system only by increasing the travel time costs incurred by other passengers. Modifying the analysis to allow for bus capacity constraints and costs imposed on the bus company by additional passengers would fundamentally not alter the conclusions reached.

Setting up the Lagrangian expression

$$Z = W(u^1, \ldots, u^n) + \eta(R - A - CX) \tag{5}$$

and differentiating with respect to a^i and b^i yield

$$W_i u_a^i - \eta = W_i \lambda^i - \eta = 0 \tag{6}$$

$$w_i(u_b^i + u_T^i t) + \Sigma W_j u_T^j b^j t_B = 0 \tag{7}$$

as first-order conditions where $W_i = \partial W/\partial u^i$ = the "marginal welfare weight" attached to individual i. The second equality in Eq. 6 follows from Eq. 1. By appropriately substituting Eqs. 3, 4, and 5, Eq. 7 can be shown to reduce to

$$\eta(F - BVt_B) = 0 \tag{8}$$

where V is the weighted (by number of trips taken) average value of travel time, $\Sigma b^i v^i / \Sigma b^i$. The Lagrangian multiplier, η, can be interpreted as the welfare gain resulting from a one-unit increase in available resource services. It is presumably positive. Hence, Eq. 8 says that, if welfare is to be maximized, the fare per trip must equal the difference between the marginal and the average time costs of a trip. That is, the fare must equal the additional time costs resulting from an additional trip minus those time costs incurred by the traveler himself.

Differentiating Eq. 5 with respect to X and making substitutions similar to those that led from Eq. 7 to Eq. 8 yield

$$-\eta(VBt_x + C) = 0 \tag{9}$$

This is the same expression that would result from selecting that value of X required to minimize VBt(B, X) + CX. This sum is the total time and dollar costs of B trips if an hour of consumer travel time is assigned its weighted average value.

This simple bus problem suggests that in transportation activities $F + v^1t$, the direct money cost of a trip plus the money equivalent of the time it requires, plays the same role in the consumer decision process as does the price of commodities usually dealt with in textbook discussions of consumer behavior. Furthermore, this simple problem suggests that it is efficient to take consumer travel time values, v^1, directly into account both in setting prices for transportation facilities and in designing them.

Expressed differently, the basic qualitative difference between transportation activities and the commodities that are commonly dealt with in textbook discussions is that, in transportation, consumers play a producing as well as a consuming role. In taking trips (whether by automobile, bus, or air), consumers both use transportation facilities and provide inputs—their own time, in particular—vital to the production process. The analysis indicates that, in the design and pricing of transportation facilities, it is appropriate to take these consumer-supplied inputs into account in precisely the same fashion as are inputs that are purchased on the open market for more conventional production processes.

COST OF INEFFICIENT USE OF EXISTING HIGHWAY CAPACITY

The following attributes of highway travel seem well established: The time required to complete a trip on a given street or highway increases with the number of trips travelers attempt to make on it. Starting from a low level, an increase in attempted trips is associated with an increase in the number of trips actually completed. However, once the capacity of the highway is reached, a further increase in attempted trips results in a reduced number of completed trips. In this range of highway utilization rates, the volume-capacity ratio is inversely related to trip demand. Indeed, as attempted trips continue to increase beyond highway capacity, the volume-capacity ratio approaches zero. That is, attained travel speeds become smaller and smaller, implying that the time required to complete a trip becomes indefinitely great.

The simplest functional relationship that yields these attributes can be written

$$N - 1 + a^2/4b = aS - bS^2$$

where

N = volume-capacity ratio = the ratio of completed trips T to highway capacity K and
S = speed.

[Adding $-1 + a^2/4b$ to the right of this expression is necessary to ensure that the maximum traffic flow occurs at a volume-capacity ratio of one. No empirical studies of speed-volume-capacity relationships have relied on this functional form. However, assigning reasonable values to a and b yields speed-volume relationships similar to those that have been reported in the literature (4).] If vehicle speed at a volume-capacity ratio of zero is twice that at a volume-capacity ratio of one, this relationship reduces to the simpler form

$$N = 4c(S - cS^2) \tag{10}$$

Equation 10 can be solved for

$$t = 2c/[1 \pm (1 - N)^{1/2}] \tag{11}$$

where t (=1/S) is the time required to travel 1 mile. In this relationship, the plus and minus signs hold when attempted trips are respectively above and below capacity levels. On an expressway, 60- and 30-mph speeds are reasonable round numbers to use for

Table 1. Travel times and marginal cost tolls per mile for various volume-capacity ratios.

Travel	Volume-Capacity Ratio	Average Time Per Mile (minutes) Below Capacity	Above Capacity	Time Increase for All Travelers, Below Capacity (minutes)	Marginal Cost Toll (cents/mile)
Expressway	0.0	1.00	∞	1.00	0.0
	0.1	1.03	38.97	1.05	0.1
	0.2	1.06	18.94	1.12	0.3
	0.3	1.09	12.24	1.20	0.5
	0.4	1.13	8.87	1.29	0.8
	0.5	1.17	6.83	1.41	1.2
	0.6	1.23	5.43	1.58	1.8
	0.7	1.29	4.42	1.83	2.7
	0.8	1.38	3.62	2.24	4.3
	0.9	1.52	2.92	3.16	7.7
	0.95	1.63	2.58	4.47	14.2
	1.0	2.00	2.00	∞	∞
City streets	0.0	2.40	∞	2.40	0.0
	0.1	2.47	93.53	2.52	0.2
	0.2	2.54	45.46	2.69	0.7
	0.3	2.62	29.38	2.88	1.2
	0.4	2.71	21.29	3.10	1.9
	0.5	2.81	16.39	3.38	2.9
	0.6	2.95	13.03	3.79	4.3
	0.7	3.10	10.61	4.39	6.5
	0.8	3.31	8.69	5.38	10.3
	0.9	3.65	7.01	7.58	18.5
	0.95	3.91	6.19	10.73	34.1
	1.0	4.80	4.80	∞	∞

Figure 1.

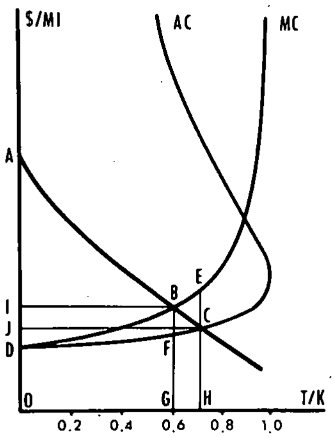


Figure 2.

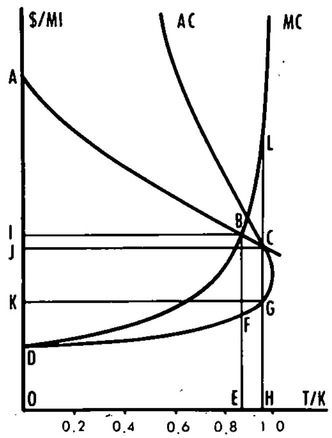
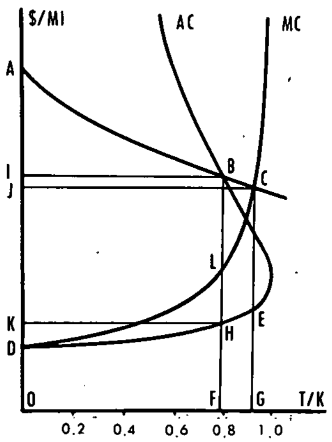


Figure 3.



travel rates respectively equal to zero and to highway capacity. Corresponding reasonable numbers for city streets might be 25 and 12.5 mph. With speed expressed in miles per minute, these values imply c equals 1 for expressways and 2.4 for city streets.

Multiplying Eq. 11 by T, the total number of trips taken, yields the total travel time required for these trips. Differentiation of tT with respect to T then yields marginal travel time per trip, the additional travel time resulting from an additional trip being taken. When demand is less than capacity, this expression can be written

$$\partial(tT)\partial T - t = N/\{[1 + (1 - N)^{1/2}]^2 (1 - N)^{1/2}\} \quad (12)$$

In Eq. 12, t is the travel time input, per mile, of the marginal traveler himself and the right side is the number of minutes by which his 1-mile trip increases the trip time of all other travelers. If the effect of volume-capacity ratios on vehicle operating costs can be ignored, the analysis presented earlier indicates that efficiency would require charging each traveler (each traveler is, after all, the marginal traveler for the purpose at hand) a toll equal to the right side of Eq. 12 times an appropriate weighted average value of travel time. For alternative volume-capacity ratios, data given in Table 1 rely on Eq. 11 to determine the travel times per mile associated with efficient and inefficient use of expressways and city streets, inefficiency being involved when demand exceeds capacity. Table 1 also uses Eq. 12 to determine marginal travel time per mile and, if the appropriate value of travel time is \$3 an hour (1, 2), marginal cost tolls when the highway is operating efficiently.

The toll actually charged for highway use in the United States is user charges such as gasoline and other excise taxes. These currently appear to average about 1 to 1.5 cents per vehicle mile. This is a level appropriate for volume-capacity ratios on the order of 0.5 for expressways and 0.25 for city streets if the assumptions underlying Table 1 can be believed. However, for the purpose at hand, i.e., determining the social costs involved in not charging efficient tolls, it is useful to ignore any direct cash outlays by consumers for travel including those for vehicle operation. That is, it is useful to suppose that the only current cost of highway trips is the time they consume.

A commonly used measure of the gross consumer benefit derived from a commodity is the area under the demand schedule for it, the sum over all units consumed of the maximum amount that a consumer would be willing to pay for each unit. A commodity's net consumer benefit or, more commonly, its consumers' surplus is its gross benefit minus the amount that consumers actually pay for it. Given some restrictive assumptions, the efficiency conditions discussed earlier can be interpreted as requiring maximization of the difference between the gross consumer benefit and the cost of producing a commodity.

Figures 1, 2, and 3 show data from Table 1 and three alternative demand relationships plotted. Assume that the demand schedule for trips along a particular highway is ABC in Figure 1. The efficiency conditions would be satisfied at a trip consumption rate of OG, this being the rate at which the price some consumer would be willing to pay for a trip equals its marginal time cost (again, vehicle operating costs are being ignored). Consumers would take trips at this rate if they face a price of BG per mile. Because OG trips are associated with a time cost of FG per mile, this price could be effected by imposing a toll of BF per mile. The gross consumer benefit associated with OG trips is the area OABG. The net consumer benefit is area ABI, the difference between OABG and the sum of the time costs incurred by travelers and their toll payments.

Although the tolls travelers pay are clearly a cost to them of taking trips, the tolls are not the cost these trips impose on society as a whole. Rather, toll payments involve a transfer of income, of purchasing power, from travelers to whatever agency is responsible for collecting tolls. Payments such as these are commonly referred to as rents or producers' surpluses. In the case at hand, this rent can be interpreted as a reward to society for its provision of scarce, valuable road services. Payment of this reward serves two important purposes. First, it is essential if these road services are to be used efficiently. Second, it enables the reduction of other taxes collected for public purposes.

Now suppose that tolls are no longer imposed. The trip consumption rate would then increase to OH, the level at which the amount a consumer would be willing to pay for a trip equals its time cost to him. As a result of this increase in trip output, gross consumer benefits would change from OABG to OACH, an increase of GBCH in these benefits. However, these additional gross benefits come about at the expense of additional travel time costs aggregating to GBEH. (If OG trips are taken, total travel time costs equal FG dollars per trip mile times OG trips. FG times OG can be shown to equal the area, ODBG, under the marginal time cost of trips schedule in Figure 1. Similarly, if OH trips are taken, total travel time costs equal CH times OH. This quantity can be shown to equal the area ODEH in Figure 1. GBEH is the difference between ODEH and ODBG.) Thus, the cost increase that results from the additional GH trips exceeds the gross benefits generated by these trips by an amount equal to BEC. This difference between the increases in gross benefits and travel time costs can be interpreted as follows: In the demand situation depicted in Figure 1, elimination of the toll has the direct effect of making travelers better off. The increase from FG to CH in travel time per mile is more than offset by elimination of the toll. As a result, the price travelers pay for trips falls by IJ per mile. They therefore receive additional consumers' surplus benefits, which equal area IBCJ. At the same time, however, the government agency involved loses a source of revenue. Area BEC is the amount by which the taxes that would have to be levied to replace the toll revenues lost exceed the benefit consumers derive from not having to pay tolls.

Consider, now, demand schedule ABC in Figure 2. Efficiency would dictate that OE trips be taken. This trip output level could be effected by imposing a toll of BF. With such a toll in effect, total benefits from using the highway would equal ABD, which can be divided into a consumers' surplus of ABI plus a rent (equals total toll collections, BF times OE) of IBD.

In the absence of tolls, the system would be in equilibrium (if at all) at C in Figure 2. This equilibrium is inefficient in two senses. First, as with the equilibrium at C in Figure 1, an inefficiently large number of trips would be produced. A loss of BLC (which corresponds to BEC in Figure 1) in potential benefits results from this excessive output rate. Second, equilibrium at C involves producing OH trips at a cost per trip HC that is substantially greater than the minimum cost HG, at which this same quantity of trips could be produced. The total loss associated with the no-toll equilibrium in Figure 2, then, is area JCGK plus area BLC. As in Figure 1, this area equals the amount by which the taxes that would have to be levied to replace the toll revenues generated by equilibrium at B exceed the additional consumers' surplus benefits, IBCJ, derived from not having to pay tolls.

Finally, for demand schedule ABC shown in Figure 3, efficiency would dictate equilibrium at C with a toll of CE. In this equilibrium, total benefits would equal the area ACD. Elimination of tolls would lead to equilibrium at B. A pair of inefficiencies comparable to those of Figure 2 is involved in this equilibrium:

1. Producing an output different from that at which price equals marginal cost leads to a loss of BCL in potential benefits.
2. Inefficiency results from the loss of IBHK associated with producing OF trips at a time cost of BF per trip rather than the minimum possible cost HF.

Unlike the situations in Figures 1 and 2, however, the equilibrium trip price in the absence of tolls is higher in Figure 3 than that that would eventuate with tolls in effect. Thus, if the demand for trips along a stretch of highway is so great that an equilibrium eventuates like that at B in Figure 3, everyone involved could immediately be made better off by the imposition of tolls: The government agency would reap a formerly unharvested source of revenue, and travelers would consume more trips at a lower

Table 2. Percentage loss due to failure to charge marginal cost tolls as a fraction of travel expenditures when demand is inversely proportional to price.

Volume-Capacity Ratio				
Efficient	With No Toll	Loss 1[a]	Loss 2[b]	Total
10	10.3	0.04	–	0.04
20	21.1	0.17	–	0.17
30	32.6	0.45	–	0.45
40	45.0	0.98	–	0.98
50	58.2	1.95	–	1.95
60	72.4	3.77	–	3.77
70	87.0	7.52	–	7.52
80	98.9	17.06	–	17.06
90	82.1	1.68	59.45	61.13

[a]From producing inefficient number of trips.
[b]From producing trips in backward bending portion of average time cost schedule.

total cost per trip.

Just how consequential are the losses suggested by Figures 1 through 3? Some idea of the orders of magnitude that might be involved can be obtained by (a) assuming a specific functional form for the demand relationship, (b) assigning the alternative values to the parameter in this relationship necessary for it to intersect the marginal time cost schedule at different volume-capacity ratios of interest, and (c) determining the point at which this demand schedule intersects the average time cost schedule. With this information, the total benefits derived from use of the highway with and without tolls given alternative positions of the demand schedule could be computed. The difference between these two benefit levels divided by, say, the sum of the time and money expenditures on trips that would result if tolls were charged would then give a measure of the proportionate consequence of inefficient pricing.

Calculations of this sort are given in Table 2. The demand relationship underlying these calculations can be written PT = k. (This demand relationship is used for two reasons. The first is computational simplicity. The second is the fact that price changes that are not offset by changes in income must, on the average, result in offsetting proportional changes in consumption.) According to this relationship, total consumer expenditures on travel are independent of the price paid for a trip. An increase in trip price would result in an exactly offsetting reduction in the number of trips taken, thereby leaving total expenditures unchanged.

The absolute loss per mile traveled as a result of inefficient pricing is greater at any given volume-capacity ratio for high values of travel time than for low values and greater on city streets than on expressways. However, the nature of the demand and travel time-volume-capacity relationship underlying Table 2 is such that the relative loss measure tabulated depends only on the efficient volume-capacity ratio. It is independent of both the value of travel time and the specific type of trip under examination. Table 2 indicates that, if the demand for street or expressway services is so low that volume-capacity ratios of 40 percent or less would result when marginal cost tolls are imposed, complete elimination of these tolls would lead to neither significant increases in volume-capacity ratios nor appreciable welfare losses. As for the latter, if the efficient volume-capacity ratio is 40 percent or less, the ratio of the loss from eliminating tolls shown in Figure 1 to the trip expenditures that would result under efficient pricing is less than 1 percent. As efficient volume-capacity ratios increase above 50 percent, however, the effect of toll elimination on both volume-capacity ratios and losses becomes more and more significant. Indeed, if efficient pricing would result in a 90 percent volume-capacity ratio, the sum of the two losses shown in Figure 3 would amount to more than 60 percent of the total expenditures on trips. Such losses may well be common in major urban areas during morning and afternoon rush hours.

OPTIMIZATION OF HIGHWAY CAPACITY THROUGH INEFFICIENT PRICING

The model developed earlier dealt with a short-run situation. It asked, what are the social costs of inefficiently using a highway of fixed capacity? Given time, the capacity of any highway can be altered, which would partially offset the inefficiencies discussed. Now we will examine the rules for optimizing highway capacity when efficient prices cannot be used to ration highway services and determine rough orders of magnitude for the losses that result from inefficient pricing under these circumstances.

As stated previously, a common measure of the gross benefit consumers derive from taking trips is the sum over all trips taken of the maximum amount they would be willing to pay for each trip they take. If, for simplicity, each traveler is assumed to place the same value on his travel time, the inverse demand function involved in this summation can be written as

$$P = F + Vt(T/K) = g(T) \tag{13}$$

where

V = average value of travel time,
F = toll charged for a trip,
$g(T)$ = inverse demand function, and
t = travel time per trip, a function of T, the total number of trips taken, and K, the highway's capacity.

The net benefit derived from using the highway is the gross benefit to consumers minus the total cost of providing its services. In the long run, this total cost includes both the time costs borne by consumers (again, vehicle operating costs are ignored) and the costs of providing highway capacity borne by whatever highway authority may be involved. Suppose, albeit unrealistically, that the cost of providing the services of a unit of highway capacity is independent of the size of the highway. Then the net benefits afforded by the highway can be written as

$$B = \int_0^T g(\tau)d\tau - TVt(T/K) - P_K K \tag{14}$$

The integral in this expression is the area under the trip demand schedule, the gross consumer benefit. $TVt(T/K)$ is the value of travel time inputs, and $P_K K$ is the cost per time period of providing the capital invested in the highway.

A sensible objective would seem to be to establish those values of T and K that maximize Eq. 14. Differentiating with respect to T and K and rearranging terms yield

$$g(T) = F + Vt = Vt + TV\partial t/\partial T \tag{15}$$

$$TV\partial t/\partial K + P_K = 0 \tag{16}$$

Equation 15 will be satisfied if the toll per trip F is set equal to $TV\partial t/\partial T$. Equation 16 can be interpreted as saying that arterial capacity should be increased to the point where the cost, P_K, of the last unit of capacity added equals the saving in travel time afforded by that unit, $-TV\partial t/\partial K$. Further manipulation would reveal that, if Eqs. 15 and 16 hold simultaneously, total toll collections would just cover total capital costs. If construction of the highway involves increasing returns to scale (i.e., if a doubling of its capacity would result in less than a doubling of its total costs), however, simultaneous satisfaction of these two equations would result in toll revenues that fall short of total capital costs.

Suppose that it is impossible to charge the toll implied by Eq. 15 but possible to charge some other toll F_o that may be greater or smaller than the one that would maximize benefits. With this toll in effect, the number of trips taken will be determined by

$$G(K, T) = T - f[F_o + Vt(T/K)] = 0 \tag{17}$$

A reasonable objective would seem to be to select K and T to maximize net benefits subject to this constraint. That is, the objective is that of maximizing

$$B^* = B + \lambda[G(K, T)] \tag{18}$$

where B is given by Eq. 14. Setting the derivatives of Eq. 18 with respect to K and T equal to zero and rearranging terms yield

$$F_o - TVt_T + \lambda(1 - f'Vt_T) = 0 \tag{19}$$

$$TVt_K + P_K + \lambda f'Vt_K = 0 \tag{20}$$

Dividing Eq. 19 by Eq. 20 and again rearranging terms yield

$$P_K + TVt_K = (F_o - TVt_T)[f'Vt_K/(1 - f'Vt_T)] \tag{21}$$

The left side of Eq. 21 is the difference between the price of an additional unit of capacity and the value of the travel time inputs that would be saved by adding that additional unit. In Eq. 21, ($F_o - TVt_T$) is the difference between the arbitrarily specified toll and the increase in the travel time costs of other travelers that would result from an additional trip. In the second term, f' is the change in trip consumption rates that would result from a change in the full price of a trip and t_T and t_K are the change in time per trip that would result respectively from an additional trip and from an additional unit of capacity. The signs of f', t_K, and t_T are negative, negative, and positive. As a result, the second term has a positive sign. This means that the entire right side of Eq. 21 will be negative if F_o is less than TVt_T, which seems to be the prevailing state of affairs on urban arterial streets and expressways, at least during morning and afternoon rush hours. If F_o is less than TVt_T, Eq. 21 indicates that (constrained) efficiency requires establishing a capacity level such that P_K is less than $-TVt_K$, i.e., such that the price of the last unit of capacity added is less than the value of the travel time inputs saved by that unit.

In an attempt to provide a verbal rationalization for this finding, suppose that the toll component of the full price of a trip is, for whatever the reasons, less than the cost that trip imposes on other travelers. Also suppose that the road has been designed so as to minimize the sum of the cost of providing road capacity and the travel time costs incurred by those who use the road. The sort of loss described in discussion of Figure 1 would result from charging an inefficiently low toll. Contracting the capacity allocated to the road would increase travel costs by more than it would reduce capacity costs. At the same time, however, doing so would increase the travel time component of the price of a trip and thereby reduce the number of trips taken and, in turn, the size of the loss resulting from the inefficiently low money component of the price of a trip.

The prescription implied by Eq. 21 can be stated briefly. If the toll levied on trips is constrained to be less than that required to equate their price and their marginal costs, (constrained) efficiency requires building what is, in a sense, an inefficiently small road, i.e., a road that has a capacity less than that required to minimize the sum of the time and capacity costs associated with the number of trips being taken. Essentially the same reasoning applies when a too-high toll is charged. Then, constrained efficiency would dictate building an inefficiently large road—one for which the saving in travel time costs resulting from the last unit of capacity added to the highway is less than the cost of that unit of capacity.

It would be of interest to quantify both the capacity adjustments and the remaining inefficiencies that would result from application of Eq. 21. Unfortunately, the cost of an additional unit of city street capacity varies so much from situation to situation that this sort of quantification is impossible. Information enabling rough quantification of the implications in Eq. 21 for expressway travel is available, however.

Rough data developed by Meyer, Kain, and Wohl (3, pp. 200-211) can be interpreted as indicating that the annual cost at 6 percent interest of the captial invested in constructing 1 mile of an L lane urban freeway is approximately

$$Z = 11{,}200 + 2{,}500D + (7{,}200 + 300D)L$$

where D is the net residential density of the area through which the freeway passes, i.e., population per acre of land actually used for residential purposes. In addition, annual right-of-way costs are on the order of 0.005DZ. In the United States, net residential densities in urban areas range between about 10 and 300.

That a substantial proportion of both construction and right-of-way costs are independent of the number of lanes provided reflects the fact that a freeway includes paved shoulders on both sides of each set of traffic lanes, a median strip between opposing traffic streams, and buffer strips between the freeway and adjacent land. As a result, of the land area used by four- and eight-lane freeways, only about 33 and 50 percent respectively is devoted to freeway lanes themselves.

The capacity of a four- to eight-lane expressway is about 1,800 vehicles per lane-hour. Using this number to convert annual cost data into average and marginal costs per vehicle mile of capacity yields the data given in Table 3. These data suggest using 0.1 and 3 cents as lower and upper bounds in the illustrative calculations.

Suppose that the travel time-volume-capacity relationship described holds and that the average value of travel time is $3 per hour. Inserting these values into Eq. 16 yields a benefit-maximizing volume-capacity ratio of 24.6 percent for a highway with 0.1 cent per vehicle mile capacity costs and 78 percent for one with a 3-cent capacity cost. From Eq. 15, the corresponding efficient tolls are 0.41 and 3.85 cents per mile and the corresponding trip prices are 5.76 and 10.66 cents per mile.

As was suggested earlier, the prevailing user charge for freeway travel in the United States appears to lie in the 1- to 1.5-cent range. Suppose that the demand for trips takes the form TP = k where P is given by Eq. 13. Inserting a 1.25-cent toll, this demand relationship, and the values described in the preceding paragraph into Eq. 21 yields 23.4 percent as the constrained optimum volume-capacity ratio for the 0.1-cent capacity cost highway and 83 percent for that with a 3-cent capacity cost. The trip prices associated with these volume-capacity ratios and a user charge of 1.25 cents are respectively 6.58 and 8.33 cents. They would lead respectively to 12.5 percent fewer trips on the low-cost highway and 27.9 percent more trips on the high-cost highway than would be optimum if efficient prices could be charged.

Relative loss measures of the sort given in Table 2 can be constructed for these two examples by (a) evaluating Eq. 14 for optimum and constrained optimum trip consumption and volume-capacity levels and (b) dividing the difference between the resulting net benefit measures by the highway expenditure levels implied by the demand equation, TP = k. The results are 0.87 and 3.79 percent for the low- and high-cost highways respectively.

As is common in empirical research into the costs of inefficient pricing and investment policies, these numbers are less than earthshaking. In partial response to the obvious question, "Why bother?", two points are worth making: First, although

Table 3. Average and marginal capacity costs.

Net Residential Density	Marginal Capacity Cost (cents)	Average Capacity Cost (cents)		
		4 Lanes	6 Lanes	8 Lanes
10	0.07	0.13	0.11	0.10
100	0.35	0.98	0.77	0.66
200	0.85	2.47	1.93	1.66
300	1.54	4.56	3.55	3.05

3.79 is not a large percentage, 3.79 percent of the total time and money costs of urban trips is a large sum of dollars. Second, the constrained efficiency rules imply a grossly inequitable allocation of the burden of financing highways. To repeat, an efficiently designed and priced road would yield toll revenues just sufficient to cover its total capital costs if these costs are proportional to its capacity. However, the constrained optimum low-cost road would generate user charges equal to 2.91 times its capital cost. For the high-cost road, on the other hand, user charges would equal only 34.3 percent of capital costs. Thus, constrained optimization would, in effect, call for the users of rural roads to provide heavy subsidies to their urban counterparts. Elimination of these subsidies by limiting the costs of roads in a particular geographical area to user revenues generated in that area, for example, would lead to far greater aggregate losses than those that would result from following the rules developed here.

ACKNOWLEDGMENT

I am indebted to Donald Dewees, Richard Klem, and Alan Walters for helpful comments.

REFERENCES

1. Beesley, M. The Value of Time Spent in Traveling: Some New Evidence. Economica, Vol. 32, (NS), 1965.
2. Lisco, T. The Value of Commuters' Travel Time: A Study in Urban Transportation. University of Chicago, PhD dissertation, 1967.
3. Meyer, J. R., Kain, J. F., and Wohl, M. The Urban Transportation Problem. Harvard University Press, 1965.
4. Mohring, H. Relation Between Optimum Congestion Tolls and Present Highway User Charges. Highway Research Record 47, 1964, pp. 1-14.
5. Strotz, R. Urban Transportation Parables. In The Public Economy of Urban Communities (Margolis, J., ed.), Resources for the Future, Washington, 1965.

Changes in Population Distribution and Automobile Ownership and Implications for Urban Transportation

Irving Hoch
Resources for the Future, Inc.

Emergent long-term trends in population and automobile ownership are examined, and some implications for urban transportation are drawn. Population trends include (a) a strong national trend toward zero population growth, at a pace well above demographers' expectations; (b) an inverse relation between growth and urban size, where the largest metropolitan areas exhibit little or even negative growth; and (c) a general shift in population from higher to lower density areas, with associated lower congestion, and some reversal of migration flows from nonmetropolitan to metropolitan places. Statistical analysis of growth rates shows two key relations: high rates of growth in three subregions (Pacific Southwest, Florida, and middle-sized Texas areas) and an inverse relation between size and growth for both the high-growth and low-growth areas of the country. Despite the continuing shift of population from central cities to suburbs, available evidence suggests some current diminution of suburban sprawl and a likely buildup of population densities in urban areas, a development that can be related to a corresponding movement toward saturation of automobile ownership. Data on intraurban density relations and rates of automobile ownership are presented in support of these predictions. The discerned trends in population and automobile ownership are likely to reduce traffic congestion and hence the needed highway investment and improve the viability of public transit in the large urban areas investing in such systems.

The apparent slowdown in population growth and the population shifts from more congested to less congested areas should relieve traffic congestion and reduce transportation requirements. The movement toward saturation of automobile ownership implies a ceiling on highway requirements but may involve higher urban densities, particularly in less congested areas. Higher densities, however, could make public transit more viable.

Nationally, there is a strong trend toward zero population growth (ZPG); recent birth rates are below replacement levels. At the same time, there has been very slow or even negative growth in large metropolitan areas, the areas of greatest congestion. Between 1970 and 1973, five of the 15 largest metropolitan areas experienced declines in population, and, of the rest, only Washington, D.C., grew as much as 3 percent, the national growth rate. In light of these trends, it may not be too long before the concern is with population shortages.

Not only has there been a shift from the more congested to less congested areas, but also, at the regional level, there is a long-term shift from the northeast and north central regions to the south and west. Three of the fastest growing regions are Florida, middle-sized Texas areas, and the Pacific Southwest, exclusive of Los Angeles and San Francisco. Not only are large areas growing slowly, but also there is a statistically significant inverse relation between metropolitan size and growth rate. The relative shift toward nonurban areas is not the major reversal of historic migration streams that some see it as, for much nonurban growth occurs just outside metropolitan boundaries.

These movements from more congested to less congested places are no mere coincidence. I see long-term equilibrating mechanisms at work, wherein people are moving to places that yield higher real income or general well-being (not necessarily measured in economics). A major component of the higher real income is traceable to less congestion and corresponding decreases in driving and time costs, air pollution, and noise. Admittedly, the destination areas become more congested and polluted with the inflow of migrants, but the system as a whole gains. The benefits at the origin, and to people moving from the origin, should outweigh the costs to those at the destination.

There is some empirical evidence supporting the movement toward saturation of automobile ownership and an end to urban sprawl, but the thesis here also depends on the argument explaining intraurban distribution of population. Population density within urban areas shows steady decline from the center. This distance-decay relation can be explained by the access-space trade-off or, put another way, by the trading of travel time for rent. Changes in this density relation can be tied to (a) population change and (b) the extent of the highway network and car ownership.

The larger the city is, the greater the density at any distance from the center is. As an urban area grows, land values go up, causing more intensive land use and greater density at all points. Congestion effects near the center inhibit growth there and encourage growth near the periphery. As reliance on the automobile increases, i.e., as roads and automobile ownership increase, density near the center decreases and population spreads out farther (or sprawl occurs). For example, the newer cities in the west, Texas, and Florida were built with the automobile in mind and are much less dense than older cities in the east. In recent history the impact of population change has been dominant over changes in the highway network and car ownership, but the latter are likely to become much stronger in the future, as areas approach automobile saturation. Thus, changes in the density relation for Baltimore show continued dominance of the second factor from 1960 to 1970 and from 1950 to 1960, whereas five western cities show the "sprawl" shift dominant only from 1960 to 1970 and some signs of a building up of density from 1950 to 1960. In 1970, Baltimore's density pattern roughly paralleled that of Los Angeles 2 decades earlier. Density data on a number of cities in specific time periods support the pattern of change discerned here. Additional support comes from analyses of automobile ownership data. Over time, the percentage of multiple-car families as a function of income has increased, which indicates suburbanization. This increase reached a peak in 1963-1966 and then slowed substantially, and it can be hypothesized that the pace of suburbanization moved in the same fashion. Reduction of air and noise pollution level should counteract some of the inhibition of growth near the urban center and higher energy prices should reinforce the pattern, making cities more dense and further limiting sprawl.

Analyses were made of automobile ownership rates; the equations developed can be used to forecast car ownership for specific years. Such forecasts would indicate how fast we are approaching "ultimate" saturation levels. In 1971, there were 1.15 cars per household, and available data indicate an ultimate level of 1.75 cars per household, an increase of 52 percent. Given alternative ZPG forecasts, ultimate levels for total car ownership can be projected as 2.0 to 2.3 times present levels, which seems hardly overwhelming.

With slow population growth and a movement toward saturation of automobile ownership, there should be a slowdown in demand for new highways. Density increases, particularly in fast-growing urban areas, should make public transit more viable, though probably not economic in the sense of covering costs.

Almost all of these trends are beneficial, for they should reduce the pressure for solutions to transportation problems considerably.

DEMOGRAPHIC TRENDS

Total Population

Since 1960, annual population growth in the United States has shown a steady decline, from about 3 million in 1960 to half that in 1970. There was some reversal of the trend in the period 1968-1970, but the decline was reestablished thereafter (Fig. 1).

If we mechanically fit a trend line to the data in Figure 1 by least squares, we reach ZPG in short order. A linear fit yields a ZPG date of 1989, whereas a logarithmic function, with a bit more variance explained, moves the date to 1993. The respective equations are

$$\Delta P = 2{,}425.8 - 101.5\ T \qquad \bar{R}^2 = 0.88,\ \text{t-ratio} = 10.5 \tag{1}$$

and

$$\Delta P = 30{,}543.7 - 15{,}515.8\ \text{LOG}\ T^* \qquad \bar{R}^2 = 0.89,\ \text{t-ratio} = 10.6 \tag{2}$$

where

ΔP = annual population change,
T = YEAR-1965, and
T* = YEAR-1900.

Maximum population is only 226 million based on a January 1974 population of 211 million. The mechanical fit is a limiting case, of course. Because of the large number of women of child-bearing age, a consequence of the post-World War II baby boom, it is plausible that the downtrend will taper off.

Recognizing the trend, the Bureau of the Census has substantially reduced its population projections (2). Yet there are indications that the bureau has not gone far enough. In previous projections, the average number of births per woman upon completion of child-bearing was alternatively set at 3.1, 2.8, 2.5, and 2.1 to yield projections respectively labeled the B, C, D, and E series. Late in 1972, the bureau dropped the B series and added an F series, which assumed 1.8 births per woman. For the year 2000, the midpoint population projection for the initial set of alternatives was 296 million; for the revised series, the midpoint projection is 275 million. The E series corresponds to eventual ZPG, 2.1 births per woman. If that level holds indefinitely, population will be 264 million in the year 2000 and will stabilize at 320 million around the year 2025. But actual experience in 1972 and 1973 (Table 1) is best represented by the F series.

The Bureau of the Census (1) made the following growth predictions (in millions) for June 30, 1972, to July 1, 1974:

Series	Growth
C	4.5
D	4.2
E	3.3
F	3.0

In the F series, projected annual growth first increases to 1.8 million and then declines to around 1 million by the year 2000 at a total population of 250 million; a peak population of 270 million is reached by 2020. Thus, the F series decline is not nearly so drastic as that obtained by straight-line projections; nonetheless, it represents a considerably reduced rate of growth relative to recent expectations.

Given oral contraceptives and abortion on request, a marked reversal of the trend seems unlikely. In retrospect, the furor over ZPG in the U.S. context seems somewhat surprising in the light of the downtrend shown in Figure 1. Perhaps both ideology (ZPG enthusiasm) and mores (changing life-styles) follow technology (improved birth control).

The deceleration in population growth has led to overcapacity in pediatrics and elementary education, and similar problems are conceivable in transportation capacity. However, the considerable lag between birth and trip-making should provide enough lead time to adjust transportation planning to changed circumstances. It must be added that the reduced birthrate may allow more women to enter the labor force and thus increase trip-making for a considerable period. The increased number of working women has increased the number of journeys to work, and for many families the additional income has made a second car feasible. It has been suggested, in fact, that working women are the major cause of the last decade's increased energy demand (3).

Population Distribution

Concern about population distribution generally focuses on urban growth and the possibility that cities are too large. Here again, trends appear to be preempting the problems. The trends include (a) little growth or even declines in population in the largest metropolitan areas; (b) an inverse relation between growth and size; (c) a continuation of shifts from northeastern and north central urban areas to those in the south and west; (d) continuation of the shift from central cities to suburbs; but (e) some apparent reversal of migration flows from nonmetropolitan to metropolitan areas.

The census estimate that, between March 1970 and March 1973, there was a net flow of around a million people from SMSAs to places outside of SMSAs has been viewed by some as a major turning point (4). The statistic is a bit misleading because it disregards the location pattern of 2.4 million people who moved to the United States from abroad. Most of these people settled in metropolitan areas, which countered the internal migration shift somewhat. Detailed information on these migration flows is given in Table 2, which shows where the population in 1970 was located in 1973. The table does not include births since March 1970.

The percentage distribution of the population in the two periods was as follows:

Location	March 1970	March 1973
Central cities	32.20	30.21
Remainder of SMSA	36.69	38.34
Outside of SMSA	31.11	31.45
	100	100

The increase for the nonmetropolitan percentage contrasts with a previous long-term decline. Thus, between 1960 and 1970, the nonmetropolitan population as a percentage of the total declined from 33 to 31 percent (6).

If we accept the estimated March 1970 and March 1973 percentage distributions as given and apply them to estimated total resident populations for the respective periods, we obtain the following estimated distributions of number of people (in thousands):

Figure 1. Annual U.S. population growth from 1960 to 1973 and fitted trend line.

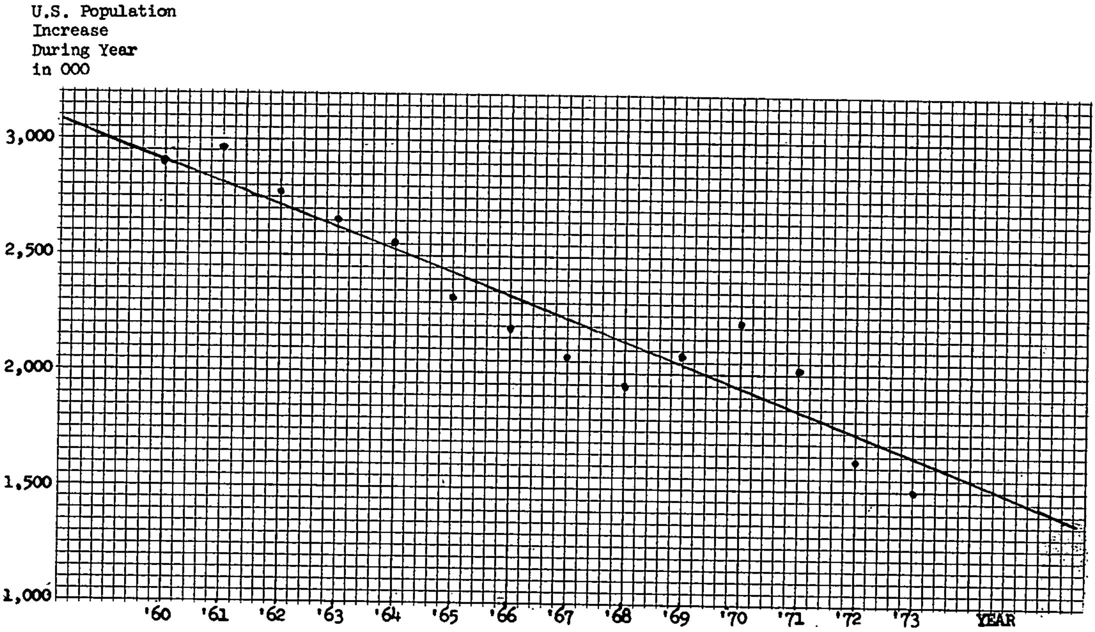


Table 1. Annual U.S. population.

Year	Population Growth[a] (in thousands)	Year	Population Growth[a] (in thousands)
1960	2901	1967	2072
1961	2955	1968	1952
1962	2771	1969	2089
1963	2655	1970	2223
1964	2555	1971	2017
1965	2316	1972	1628
1966	2197	1973	1499

[a]Includes armed forces overseas.

Table 2. Distribution of the U.S. population (in thousands) in 1970 and 1973.

				Population in March 1973			Net Flows	
Location	Population in March 1970	Movers	Nonmovers	Central City	Remainder of SMSA	Outside SMSA	Excluding Movers From Abroad	Including Movers From Abroad
Central city	62,333	25,197	37,136	52,927	7,278	2,128	-4,124	-3,112
Remainder of SMSA	71,042	22,922	48,120	3,673	64,704	2,665	3,209	4,114
Outside SMSA	60,225	20,525	39,700	1,609	2,269	56,347	915	1,431
Outside U.S.	2,433	2,433		1,012	905	516		
Total in 1973	196,033	71,077	124,956	59,221	75,156	61,656	0	2,433

Location	March 1970	March 1973	Ratio 1973/1970
Central cities	65,394	63,239	0.967
Remainder of SMSA	74,513	80,257	1.077
Outside of SMSA	63,180	65,834	1.042
	203,087	209,330	1.031

(This distribution accounts for births since 1970.) The central cities declined in total population by 3.3 percent, whereas the remainder of SMSAs grew faster than the remainder of the country (7.7 versus 4.2 percent). These percentages contrast with respective percentage growth between 1960 and 1970 of 5.3, 28.2, and 6.5 for the three locales. The increase in central city population in the 1960-1970 period may well reflect annexations rather than growth within city limits as defined in 1960. (Cities that annexed no territory showed a small population decline in the period.) Hence, the decline in central city population can be viewed as a continuation of earlier trends. However, the relative shift to nonmetropolitan areas seems substantial; if it is not a reversal of earlier trends in population growth, it does represent a considerable shift in relative strength of those trends.

Much of the nonmetropolitan growth probably involves a form of urbanization, and reflects (a) growth in the "far suburbs" around existing metropolitan centers but outside the county lines defining SMSAs, (b) development of industrial complexes at intersections of major highways at some distance from metropolitan areas, and (c) expansion of service industry employment in rural counties (7, 8). In all cases, improved highway transport is probably a major underlying cause for the shift.

Beale (9) argues that there has been a narrowing of life-style differences so that "urban areas are not as urbane as they were" and "rural areas are not as rustic" and that this is an additional reason why more people remain in nonmetropolitan areas. Although occurring more slowly, the shift in pattern can be expected for the migration of minority groups in the future.

In addition to these explanations for nonmetropolitan growth, the recently improved terms of trade for agricultural production, with attendant high prices and profits, should bring higher wages and increases in farm employment if the changed market conditions persist. (Admittedly, the long-term substitution of nonhuman for human inputs can limit growth in agricultural employment, but at least some city-to-farm reverse migration now seems plausible.)

Some perspective on regional migration flows can be obtained from Table 3, which gives recent interstate migration flows (people moving from one state to another). Central cities in all regions had net outflows of migrants between March 1970 and March 1973, although the change was most pronounced in the northeast and north central regions and the net losses in the south and west were only limited. A similar pattern occurred for the remainder of the metropolitan areas; the northeast and north central had net outflows and the south and west net inflows of migrants. For nonmetropolitan areas, however, only the northeast had net outflows. The northeast had the lowest ratio of inmigrants to outmigrants in every locale, central cities, remainder of metropolitan areas, and nonmetropolitan areas. The highest ratio for both central cities and nonmetropolitan areas occurred in the west, though the ratio for the south was almost as large. For the remainder of SMSAs, the ratio for the south was well above that of the other regions, including the west. For regional aggregates, the order was south, west, north central, and northeast for both the ratio and the total net immigration. It is plausible that the same pattern, though less pronounced, holds for total population changes.

Some information is available on recent population changes for individual metropolitan areas (10). From April 1970 to July 1972, the population of all SMSAs combined grew 2.18 percent, whereas nonmetropolitan areas grew 2.27 percent. It is noteworthy that growth in the 15 largest SMSAs (populations over 2 million) averaged less than

1 percent, and five of the group registered small declines (New York City, Los Angeles, St. Louis, Pittsburgh, and Cleveland). In contrast, 52 of 235 smaller areas grew by more than 5 percent.

Individual area statistics were analyzed in detail by forming the ratio of 1972 to 1970 population for each SMSA and relating the ratio to population size and region. Initially, six population size classes were set up, and inspection of the data led to the classification of three fast-growing subregions:

1. Pacific Southwest desert and shore consisting of Arizona, Nevada, and California SMSAs below 2 million population and excluding the California central valley SMSAs, which typically had somewhat lower growth rates;
2. All Florida SMSAs except Titusville-Cocoa-Melbourne, which lost population because of cutbacks in the space program; and
3. Texas SMSAs ranging in 1970 from 140,000 to 2 million population.

There were 13 SMSAs in each grouping. Average population ratios for these classifications are given as Table 4. Aside from the high-growth subregions, it seems that growth declines as size increases, and there is the suggestion of such a pattern for the high-growth subregions as well, although it is less pronounced. The inverse relation between size and growth was confirmed by statistical test. Regression analysis was used to relate the growth ratio to population size and region; dummy variables were used for the major census regions and for the three fast-growing subregions. Before the subregions were introduced, both south and west were statistically significant. After the subregions were introduced, however, only the west retained significance. Thus, accounting for the high-growth subregions caused the south as a whole to drop out as a significant variable; i.e., high growth in the south seems specific to Florida and middle-sized Texas SMSAs. When brought into the equation, both the south and northeast had small positive effects, relative to the north central region as base. Most of the SMSAs in the south are in the small size classes, whereas the opposite holds for the northeast, where a greater proportion of SMSAs are in the larger size classes than holds nationally. Hence, some of the differential growth between regions appears to be attributable to size effect, since 1970 population size was statistically significant with negative coefficient.

The statistical results obtained are given in Table 5. The basic growth ratio of 1.02195 is adjusted by population size; thus, for a population of 0.5 million, we obtain 1.02195 - 0.5 (0.00349) = 1.02020, the estimated growth rate. This value applies for SMSAs outside the fast-growing regions. If we want the fitted ratio for Florida, we obtain 1.02020 + 0.05658 = 1.07678; the value for the Pacific Southwest is obtained by adding the coefficients for both that region and the west, i.e., 1.02020 + 0.01899 + 0.02973 = 1.06892.

The results can be summarized as follows:

1. With increasing size, there is an increasing drag on growth, and
2. Rapid growth tends to be localized in specific subregions.

Hypotheses explaining these results can be based on a number of presumptive causative factors by drawing on outside evidence. It seems plausible that one of the major causes is an equilibrating process involving a shift from more to less congested places.

Part of the population shift involves the migration of retired persons to pleasant places (the sunny coasts) or to places with a lower cost of living; because the cost of living is higher in the north and increases with size (11), people on fixed incomes have higher real incomes in the south and their real income increases the smaller the locale in which they live is. If we assume that similar migration patterns also hold for workers, real "net" wages are higher in areas with the highest rate of growth. It seems plausible that level of highway service is a factor in both the regional and the

Table 3. Migration flows of population, March 1970 to March 1973.

Locale and Region	Number of Migrants in Thousands: From Origin March 1970	To Destination March 1973	Net Flows	Ratio In/Out
Central cities				
Northeast	869	314	-555	0.361
North central	1,179	589	-590	0.500
South	1,324	1,202	-122	0.908
West	899	889	-10	0.989
	4,271	2,994	-1,277	0.701
Remainder of SMSAs				
Northeast	1,067	886	-181	0.830
North central	1,100	1,016	-84	0.923
South	993	1,856	863	1.869
West	924	1,240	316	1.342
	4,084	4,998	914	1.224
Outside SMSAs				
Northeast	507	470	-37	0.927
North central	853	1,007	154	1.181
South	1,468	1,559	91	1.062
West	702	857	155	1.221
	3,530	3,893	363	1.103
Regional totals				
Northeast	2,443	1,670	-773	0.684
North central	3,132	2,612	-520	0.834
South	3,785	4,617	832	1.220
West	2,525	2,986	461	1.183
	11,885	11,885	0	1.000

Table 4. SMSA population growth by size.

Population Class[a] (in thousands)	Average Ratio[b] of 1972 to 1970 Population: SMSAs in High-Growth Subregions	Other SMSAs	Distribution of SMSAs: In High-Growth Subregions	Other
<150	1.0790	1.0228	9	59
150 to 250	1.0668	1.0224	8	39
250 to 500	1.0641	1.0197	11	52
500 to 1,000	1.0664	1.0190	4	34
1,000 to 2,000	1.0615	1.0158	7	12
>2,000	—	1.0073	0	15

[a] 1970 population.
[b] Essentially the same as the weighted average ratio.

Table 5. Regression equation results for ratio of 1972 to 1970 population as dependent variable.

Independent Variable	Coefficient	t Ratio
Constant	1.02195	506.965
Population in 1970		
(in millions)	-0.00349	2.367
West	0.01899	3.664
Pacific Southwest	0.02973	3.540
Florida	0.05658	7.974
Texas, middle-size	0.01899	5.267

Note: $\bar{R}^2 = 0.3397$.

Table 6. Average growth of all places with 1960 population of 10,000 or more by distance to nearest SMSA and accessibility to Interstate highways.

Distance From SMSA (miles)	Accessibility	Places: Number	Percent
<50	High	585	27.6
	Average	399	16.7
	Low	41	15.5
	None	24	14.6
50 to 100	High	36	15.0
	Average	103	13.8
	Low	52	5.6
	None	24	2.2
100 to 150	High	15	15.4
	Average	114	15.1
	Low	98	10.0
	None	111	6.7
>150	High	3	14.7
	Average	55	17.2
	Low	22	4.2
	None	64	7.1

urban growth patterns. The larger the urban area is, the greater the cost of operating a car (e.g., parking costs, insurance, and accident costs not covered by insurance), the longer the journey to work, and the greater the traffic congestion and pollution. Such costs may help limit the growth of large areas. Further, the newer urban areas of the south and west typically have much better highway networks than those of the northeast and north central regions.

The thesis that level of highway service is an important factor in urban growth is supported by data given in Table 6, which relates 1960-1970 growth for places of at least 10,000 population to access to Interstate highways and proximity to an SMSA. Caution must be exercised in interpreting the future, however, because it is possible that some of the causation runs the other way, i.e., that fast growth generates highway construction, or that both fast growth and highway construction are caused by other factors. For example, areas with comparative economic advantage may have that advantage because they are newer, and, because they are newer, their land use pattern was influenced by the automobile and, hence, is of lower density and has lower costs for highway construction. Nevertheless, it is plausible that at least some of the causal relation flows from highway access to growth. With increased access, a number of costs must decline, which increases real income.

Intraurban Patterns

Although evidence shows a continued shift of population from central cities to suburbs, other evidence suggests some current diminution of suburban sprawl and a likely future buildup of population densities in urban areas. Increasing densities can be tied to a corresponding movement toward saturation of automobile ownership (as an asymptotic limit). The density buildup can be expected first in the west, given present high rates of automobile ownership in that region, whereas sprawl should continue for some time in areas of relatively low automobile ownership rates, the northeast, in particular. Evidence on automobile ownership rates is more suggestive than definitive, but the items of evidence are consistent with each other and with economic theory applied to urban land use.

That theory is concerned with the trade-off of space for access (or of rent for travel costs), which accounts for the distribution of urban population per unit of land area, which is approximated by

$$D = Ae^{-bK} \tag{3}$$

where

D = population density,
K = distance from the center in any direction ($K \geq 0$), and
A and b = parameters.

With $K = 0$, $D = A$, or density at the center, b is termed the density gradient, which measures the relative rate of decline from the center (14, 15). Shifts in the density relation can be related to (a) population growth, (b) crowding effects, including congestion, air pollution, and noise, and (c) transportation improvements. As population increases in an area, other things being equal, we can expect increased demand for land at all points, a general bidding up of land values, and a corresponding increase in population density. If there were no crowding effects, we could hypothesize a constant percentage increase everywhere; however, crowding effects are likely to be strongest the higher the density is, tending to inhibit growth near the center and to increase it at the periphery of the urban area. Finally, as transportation improvements make outer areas more accessible to the center, we can expect a shift of population outward and corresponding declines in the values of both A and b in Eq. 3. Over time, higher per capita incomes reinforce both the crowding and transportation improvement effects through increased automobile ownership and shifts to single-family housing.

Figure 2 shows the argument in diagram form (15, p. 95). The density-distance relation is plotted in terms of the log of density, which yields a linear relation of the form

$$\log D = \log A - bK \tag{4}$$

Line 1 is the initial pattern, and line 2 represents the effect of a once-and-for-all transportation improvement, which implies lowered density at the center and higher density at the periphery. The intercept, log A, declines, as does the absolute value of b; because the slope is negative, its upper bound is zero. Line 3 exhibits the effect of population growth and assumes crowding effects; if the increase were the same percentage at every distance, in the logs this would be the equivalent of adding a constant to line 2. However, with crowding effects, which inhibit growth near the center, the increase is less than proportionate near the center and greater than proportionate at the periphery. In reality, of course, transportation improvement and population growth occur in a continuous process over time. But there is some basis for hypothesizing that the shift from 1 to 2 was predominant in the recent past, whereas that from 2 to 3 is likely to be strongest in the future.

Some evidence is available to support these conclusions. Muth estimated the density function for 46 U.S. cities for 1950 (15, p. 142). By regressing the logs of Muth's parameter estimates on log of urbanized area population and regional dummy variables, we obtain results supporting Figure 2. Putting the equations in antilog form yields the following estimates:

$$A = 6.03\ (1.71)^{NE}\ (0.73)^{W}\ P^{0.208} \qquad R^2 = 0.26 \tag{5}$$

$$b = 4.54\ (1.49)^{NE}\ (0.90)^{W}\ P^{-0.424} \qquad R^2 = 0.34 \tag{6}$$

where P is urbanized area population and NE and W are regional dummy variables that take on values of 0 for observations outside the given region and 1 for observations within the region. (The south and north central regions are excluded.) In the latter regions, A equal 6.03 times the population effect, $P^{0.208}$. In the northeast, A equals (6.03) (1.71) or 10.31 times the population effect; and in the west, A equals (6.03) (0.73) or 4.40 times the population effect, reflecting regional variations in urban density pattern, which fit with the shift from line 1 to line 2 in Figure 2. With population increases, there is some increase in A as well as a decrease in the absolute value of b, which shows that density increases at the center as well as at the periphery, given population growth. However, the population exponent in Eq. 5 has a low value, and the hypothesis that it equals one is rejected, showing that population increases are more than proportionate at the periphery and less than proportionate near the center.

Barr (16) obtained individual city estimates of A and b for 1960 by using essentially the same sample of census tracts that Muth had used for 1950. He obtained estimates for only 30 cities because of massive redefinition of census tracts in the remaining cases. Between 1950 and 1960 both A and b decreased substantially, presumably reflecting the growth in intraurban highways and attendant suburbanization. Comparing Barr's estimates to Muth's shows that the 1960 value of A declined to 0.65 and the value of b declined to 0.63 of the corresponding 1950 value, on the average. These changes correspond to the shift from line 1 to line 2 in Figure 2. Regression results using the Barr estimates as data parallel those for the 1950 cases:

$$A = 2.89\ (1.92)^{NE}\ (0.63)^{W}\ P^{0.236} \qquad R^2 = 0.21 \tag{7}$$

$$b = 11.80\ (1.41)^{NE}\ (0.69)^{W}\ P^{-0.616} \qquad R^2 = 0.34 \tag{8}$$

Figure 2. Relation of urban population density to distance from urban center.

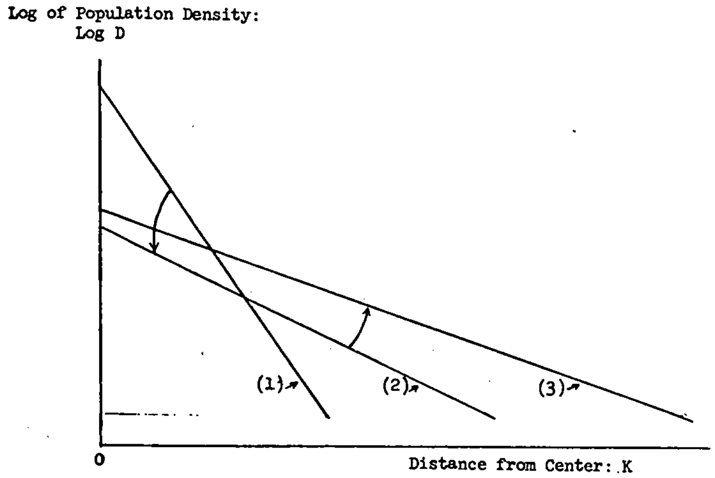


Table 7. Density equation results.

Estimate	Year	Baltimore	Kansas City	Denver	Los Angeles	Riverside	San Bernardino
Log A (intercept)	1950	4.226	—[a]	3.599	3.517	—[a]	0.961
	1960	3.556	1.737	2.220	3.127	1.136	1.211
	1970	3.102	1.866	2.192	3.033	1.300	1.198
b (coefficient)	1950	0.649	—[a]	0.706	0.246	—[a]	0.303
	1960	0.387	0.174	0.148	0.169	0.230	0.228
	1970	0.258	0.155	0.143	0.148	0.162	0.191
Change in log A	1950 to 1960	-0.670	—[a]	-1.379	-0.390	—[a]	0.250
	1960 to 1970	-0.454	0.130	-0.028	-0.094	0.164	-0.013
Change in b	1950 to 1960	-0.262	—[a]	-0.558	-0.077	—[a]	-0.075
	1960 to 1970	-0.129	-0.019	-0.005	-0.021	-0.068	-0.037
$\bar{R}^2$		0.681	0.436	0.389	0.645	0.671	0.535
N		112	76	81	219	76	90

[a]Not available.

Table 8. Percentage of automobile-owning households by income and residence.

No. of Cars	Place of Residence	Annual Family Income <$5,000	$5,000-7,499	$7,500-9,999	$10,000-14,999	>$15,000	All Income Groups
1 or more	Central cities of 12 largest SMSAs	22	45	83	83	78	58
	Central cities of other SMSAs	43	78	85	95	90	70
	Suburban areas of 12 largest SMSAs	53	93	93	98	95	89
	Suburban areas of other SMSAs	63	96	96	98	99	90
	Adjacent areas of SMSA	68	95	98	99	98	86
	Outlying areas of SMSA	51	92	97	99	97	76
	All places	52	85	93	96	95	79
2 or more	Central cities of 12 largest SMSAs	1	4	7	22	22	10
	Central cities of other SMSAs	3	17	25	47	51	21
	Suburban areas of 12 largest SMSAs	5	27	29	38	66	37
	Suburban areas of other SMSAs	12	11	30	55	74	36
	Adjacent areas of SMSA	8	24	33	43	67	26
	Outlying areas of SMSA	5	24	39	40	59	21
	All places	5	19	28	41	62	26

The hypothesis that the exponent of P equals 1 in Eq. 7 is again rejected.

Table 7 gives fitted density equation results for six SMSAs in 1950, 1960, and 1970 (17). In some cases, residential density is low near the city center because the land use is other than residential. Therefore, observations of less than 1 mile were omitted for Baltimore and Denver and less than 2 miles for Kansas City and Los Angeles. Coefficient estimates were appreciably affected only for Kansas City, but $\bar{R}^2$ improved in all other cases.

In general, the decline in log A and in absolute value of b is much less pronounced between 1960 and 1970 than it was between 1950 and 1960. In the later period, only Baltimore showed a change consistent with a strong shift of the line 1 to line 2 variety. In all of the other areas, that shift seems much less pronounced; log A is relatively stable in three of the areas and even shows a slight increase in the other two, whereas the value of b shows only a small decline. Thus, some support emerges for the hypothesis that the impact of road-building and automobile ownership on density is still considerable in the east, but is greatly attenuated in the west. It seems worth noting that Baltimore's density relation for 1970 roughly corresponds to that of Los Angeles about 15 years earlier. (In 1970, Baltimore's intercept is roughly that of Los Angeles in 1960, and its slope corresponds to that of Los Angeles in 1950.)

TRENDS IN AUTOMOBILE OWNERSHIP

The regional pattern of automobile ownership (18) parallels the urban density pattern by region, as demonstrated by the following data on percentage of households owning cars in 1971:

Region	At Least One Car	Two Cars	Three or More Cars
Northeast	73.1	21.3	4.5
North central	83.6	26.3	5.7
South	79.2	24.9	4.1
West	85.9	28.1	5.2
U.S.	80.0	25.0	4.8

Besides density, income is a major factor explaining automobile ownership. The joint effects of the two factors (19) are indicated by the data given in Table 8.

Maximum or "ultimate" car ownership levels can be estimated as follows. Available evidence indicates that the ultimate percentage of Americans owning one or more cars is 97 and the percentage owning three or more cars is 15. Given the data of Table 8, it is possible to estimate "ultimate" percentages for two-car households under some fairly reasonable assumptions. Assume that the percentage of car ownership for the highest income class is very close to the ultimate percentage for each locale, and project an ultimate population distribution by locale. (An income of \$15,000 in 1968 corresponds to \$19,200 at 1973 prices.) This procedure provides the data given in Table 9.

With the 1970 population distribution, ultimate U.S. ownership of two or more cars is 60.8 percent of all households; under the ultimate distribution that figure rises slightly to 62.7 percent. We can now obtain total cars per household for 1971 and for the ultimate period by noting that each category involves a one-car increment (assuming three cars and over is very close to 3.0). We find:

Cars	1971	Ultimate
First	0.80	0.97
Second	0.30	0.63
Third	0.05	0.15
	1.15	1.75

Table 9. Estimation of ultimate automobile ownership levels.

Area	Location	Maximum Percentage With 2 or More Cars	1970 Population Distribution	Ultimate Population Distribution
Central cities	12 largest SMSAs	25	11.1	8
	Other SMSAs	55	20.3	17
Suburbs	12 largest SMSAs	67	14.1	13
	Other SMSAs	75	23.5	27
	Adjacent areas	68	15.0	17
	Outlying areas	60	16.0	18

Table 10. Percentage of households owning automobiles by locale.

Car Ownership	Residence	1964	1965	1966	1967	1968	1969	1970	1971
1 or more	Central cities								
	12 largest SMSAs	54	57	56	54	58	55	62	61
	Other SMSAs	80	77	77	73	70	67	66	78
	Suburban areas								
	12 largest SMSAs	83	90	86	88	89	91	91	91
	Other SMSAs	88	88	92	87	90	88	89	90
	Adjacent areas to SMSA	84	83	85	86	86	86	86	89
	Outlying areas to SMSA	75	76	75	76	76	77	83	82
	All areas	78	79	79	78	79	79	82	83
2 or more	Central cities								
	12 largest SMSAs	14	10	11	9	10	11	18	15
	Other SMSAs	24	22	24	22	21	22	27	30
	Suburban areas								
	12 largest SMSAs	25	39	32	36	37	39	46	41
	Other SMSAs	33	33	37	34	36	39	35	34
	Adjacent areas to SMSA	24	26	28	27	26	29	29	29
	Outlying areas to SMSA	15	18	18	20	21	20	20	18
	All areas	22	24	25	25	26	27	28	28

Table 11. Percentage of households owning automobiles by income class.

Car Ownership	Income (dollars)	1964	1965	1966	1967	1968	1969	1970	1971
1 or more	<1,000	32	27	24	25	27	32	25	13
	1,000-2,000	33	43	31	38	38	39	41	34
	2,000-3,000	70	56	54	53	48	46	50	52
	3,000-4,000	72	68	67	63	63	54	60	61
	4,000-5,000	72	76	76	76	73	68	70	71
	5,000-6,000	86	82	84	82	81	78	75	83
	6,000-7,500	87	88	89	86	89	88	86	89
	7,500-10,000	94	94	93	93	93	93	92	89
	10,000-15,000	98	97	96	95	96	95	96	95
	>15,000	93	94	95	93	95	97	96	97
2 or more	<1,000	3	2	3	6	0[a]	1	3	0[a]
	1,000-2,000	2	2	3	2	2	4	1	0[a]
	2,000-3,000	8	6	3	5	1	1	7	0[a]
	3,000-4,000	11	12	6	10	10	8	6	9
	4,000-5,000	12	12	11	14	13	7	9	10
	5,000-6,000	19	17	16	15	17	15	9	12
	6,000-7,500	19	21	21	19	21	18	15	19
	7,500-10,000	34	32	30	29	28	31	26	26
	10,000-15,000	46	47	46	45	41	44	41	40
	>15,000	57	57	60	62	62	61	60	54

[a]Less than 0.5 percent.

The second-car increment in 1971 is obtained by adding the fraction for those with exactly two cars (0.25) to the fraction for three or more (0.05), thus accounting for the latter group's second car. However, the proper second-car fraction in the ultimate case equals that for those owning two or more cars.

It follows that the ultimate ownership rate is about 52 percent above that for 1971 (1.75/1.15 = 1.52). If we accept the census bureau's F series projection of a maximum population of 270 million, the ultimate growth in population is about 28 percent of the present level. The population and car ownership projections combined yield a projected ultimate number of cars equal to around twice the present level ($1.52 \times 1.28 = 1.95$). The E series projection of the ZPG level is 320 million or 1.52 times present population. Under this projection, ultimate car ownership is about 2.3 times present levels ($1.52 \times 1.52 = 2.31$). Inasmuch as these are estimates of ultimate levels, intuitively they do not seem very large. Some may well be encouraged by these magnitudes.

Time series data are not readily available on car ownership by both income and locale; most series relate ownership percentage to only one of these factors, so that changes in the other factor are not controlled. Thus, an apparent trend in one set of percentages may well reflect changes in the uncontrolled factor. Table 10 gives car ownership percentages by locale, whereas Table 11 gives those percentages by income. The percentages are based on surveys with a sample size of 2,500 (the 1971 survey had only 1,300 observations). Hence, sampling variability can cause distortions also. In Table 10, some upward trend is manifested in all locales, though the change is small for suburban areas since 1965 because of their high initial percentages. It is likely that much of the upward trend can be explained by increases in real income.

Data given in Table 11 appear remarkable for an apparent absence of trend, though increased suburbanization would cause some upward shift in car ownership. The absence of trend is more apparent than real because income data are in current dollars, and (as we are all aware) there has been marked inflation in recent years. The data given in Table 11 seemed worth detailed analysis for the information they might yield on suburbanization, assuming that most of the discerned shifts in relationships could be attributed to that cause. Data of the same form were available for the years 1950 through 1965 from a study by Lansing and Hendricks (20), and the two series were merged so the period 1950 through 1971 was covered. The income measures for the earlier period were in constant dollars on a 1957-1959 base, in contrast to the current dollar measures of the later series. We used consumer price level deflators to convert all data to 1973 dollars. Car ownership was related to income for each subsample separately, and patterns of year effects were determined. Then the subsamples were pooled, and separate regression relations were estimated for percentages owning one car and two or more cars, based on the patterns of year effects. For percentage owning one or more cars, observations were omitted for incomes above $12,500 (1973 prices) because this seemed to be the upper bound for the ownership-income relation. Both intercept and slope shifts were investigated, but the final form of equation accepted was based on statistical significance (22). For percentage owning one or more cars, the equation had the following form:

$$(\%\ 1 + \text{car}) = (a + b_1 B_1) + (c\$ + dT\$)$$

where

$\$$ = income in thousands of dollars,
T = year minus 1950,
b_1 = the coefficient of a dummy variable,
B_1 = representing a set of years in a grouping, and
$T\$$ = the product of T and $\$$.

The equation form specifies both intercept and slope shifts over time; intercept shift is accounted for by b_i and slope shift by dT. The numerical results obtained are given in Table 12.

Significant intercept shifts occurred only in the earliest years of the series (covering 2-year periods from 1950 through 1955), but a significant slope shift occurred over the entire period, 1950-1971. A linear equation for a given year can be obtained directly from Table 12. Thus, for 1950, we have (% 1 + car) = 27.36 + 5.14 $; and for 1970, (% 1 + car) = 32.80 + 6.16 $. The fitted equations for 1950, 1955, and 1970 are shown in Figure 3. It seems that the upward shift in ownership has decelerated since the early 1950s, which is hardly surprising, given the high level attained some time ago.

For percentage owning two or more cars, the equation has this form:

$$(\%\ 2\ \text{car} +) = (a) + (b_i B_i \$ + c\$ + dT\$)$$

In this equation, the intercept is stable over time and there are two kinds of slope shifters: dT shifts the slope in a regular fashion, whereas $b_i B_i$ generates an irregular shift specific to a particular grouping of years. Numerical results are given in Table 13.

Fitted equations for specific years can again be obtained by appropriate multiplications and additions. Thus, for 1950, the equation is (% 2 car +) = 2.697 + 1.003 $, and, for 1970, the equation is (% 2 car +) = 2.697 + 2.904 $. [In 1970, the slope is equal to 1.003 + 0.0294 T + 1.313 (1) or 1.003 + (20)(0.0294) + 1.313.] Fitted equations for some specific years are shown as Figure 4. Data given in Table 13 and Figure 4 seem to indicate an initial accelerating upward shift in the relationship, followed by a decelerating shift in recent years. If we divide the increment in the slope shift ($B_i$$) for each period by the number of years covered and add the common shift (T$), we find the following estimated annual shifts from Table 4:

Year	Shift
1950-52	0.0294
1953-54	0.0925
1955-58	0.0897
1959-62	0.0840
1963-66	0.1401
1967-68	0.1294
1969-71	0.0577

There is a peaking in the 1963-66 period and a substantial falling off in the most recent period. If we interpret the shift as indicative of the suburbanization process (or sprawl), the slowdown in the shift may be taken as additional evidence that sprawl may be approaching a limit and that future urban densities are likely to increase.

It may be noted that the data given in Tables 4 and 5 might be extended to future years; then, given forcasts of the income distributions for those years, forecasts of car ownership for specific future years could be derived. Such forecasts would suggest how fast car ownership is approaching the "ultimate" levels presented earlier.

Those forecasts might be affected somewhat by increased energy prices, and by air pollution and noise regulation. Higher gasoline prices will make long automobile trips more expensive and location at the urban periphery less attractive. Cost per mile driven and the purchase price of automobiles have also increased because of air pollution control devices. All of those cost increases may slow down automobile purchases, although the major effect might well be a shift to smaller cars. Air pollution and noise abatement should lead to some countering of the flattening effect on intraurban population distribution caused by such externalities. Insofar as such environmental changes make large cities more attractive than small ones, there may also be some reversal of the relative shift of population from large places. All of these trends will reinforce the shift toward higher population densities.

CONCLUSION

Some inferences can be drawn, and a number of hypotheses formulated, on the basis of the trends examined in this paper.

The problems of growth at the national level and in terms of the burgeoning of "too" large urban areas appear to have been grossly overstated in recent years. At the national level, in fact, it is likely that lack of growth will eventually be seen as a problem, as will attendant labor "shortages" (23). At the urban area level, the inverse relation between growth rate and size and the little or no growth of the largest areas agree with the hypothesis that there are eventual diseconomies of scale, including increased congestion and pollution effects. In response, people can and do move to more pleasant and generally smaller places.

The slowdown in total growth and in growth of large areas and the shift of population to less congested places should imply a considerable slowdown in high construction needs. The U.S. Department of Transportation has estimated a need in 1990 for approximately 18,000 miles of additional freeways and expressways within urbanized areas, compared to about 8,000 miles in 1968 (24). That estimate may be too high if earlier growth patterns were used in its development.

The discerned shifts in population can be viewed as part of an equilibrating process, improving congestion problems in dense areas and making them somewhat worse in places that are currently less dense and les congested. This pattern is likely to be intensified if we are, in fact, moving toward saturation of automobile ownership. Available evidence lends support to the hypotheses that suburban sprawl and, concomitantly, increases in automobile ownership are likely to continue at near their former pace only in the high-density areas of the east, while the rest of the country, led by the west, enters an era of building up of densities at distances from the urban center and much slower growth in automobile ownership. If developed far enough, the pattern of increasing urban densities may make public transit operations more viable in a number of areas. They may not become profitable, but losses ought to be reduced. Given a commitment to extensive investment in public transit facilities, in any event, some density increases can thus be viewed in optimistic fashion.

It might be added that public transit investment will probably be an additional causal factor yielding higher urban densities. In general, of course, causation runs both ways for transportation and land use. Higher gasoline prices and air pollution regulations affecting the cost of automobile operation should also inhibit suburban sprawl. Parenthetically, air pollution regulation initially involved some planned use of pricing to ration automobile use and, hence, reduce pollution, but that development now seems less likely.

The patterns of population redistribution and the possible coming saturation of automobile ownership can be viewed as benign, even hopeful, developments that will tend to reduce traffic congestion and perhaps to reduce needed highway investment and to make public transit a more economic investment.

Although market-like processes can be discerned in trends, that does not vitiate the need for planning, benefit-cost analyses, and policy devices, including pricing. The use of such devices might speed up or even reverse the trends.

Nevertheless, present trends seem to mean a considerable reduction of the pressure for solutions to transportation problems. It is hoped that the time gained will be used constructively.

REFERENCES

1. Current Population Reports, Population Estimates and Projections. U.S. Bureau of the Census, Series P-25, No. 513, Feb. 1974; No. 523, June 1974.

Table 12. Regression equation results for percentage owning one or more cars.

Variable	Coefficient	t-Ratio
Constant	32.8016	28.413
Intercept shifting dummy, B_i		
1950-1951	-5.4439	2.388
1952-1953	-3.6551	1.788
1954-1955	-1.8163	0.900[a]
\$(thousands of dollars)	5.1376	24.772
T\$	0.0510	3.049

Note: $\bar{R}^2 = 0.897$, number of observations = 186.

[a] Included because omission caused other dummies to lose significance.

Table 13. Regression equation results for percentage owning two or more cars.

Variable	Coefficient	t-Ratio
Constant	-2.6966	9.936
\$	1.0030	25.437
T\$	0.0294	2.322
Slope shifting dummies, (B_i\$)		
1950-1952	0[a]	—[b]
1953-1954	0.1262	2.144
1955-1958	0.3672	4.538
1959-1962	0.5854	4.601
1963-1966	1.0282	5.947
1967-1968	1.2282	5.696
1969-1971	1.3130	5.356

Note: $\bar{R}^2 = 0.972$, N = 278.

[a] Set equal to zero because one dummy is omitted to avoid colinearity.

[b] Not applicable.

Figure 3. Estimated percentage of families owning at least one car as a function of income.

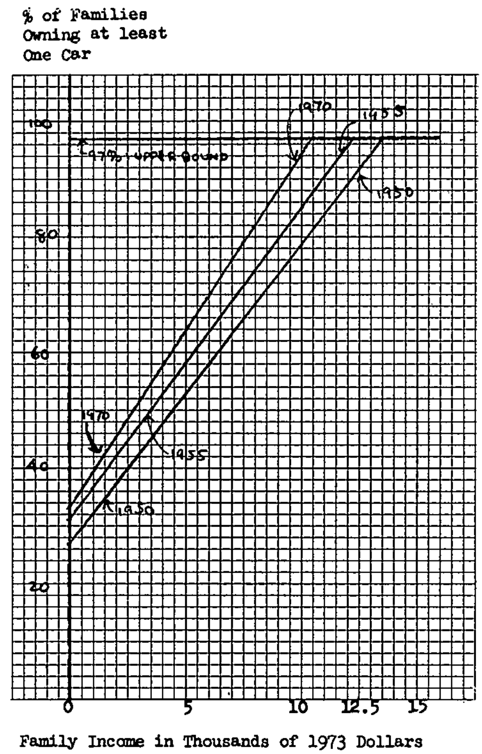
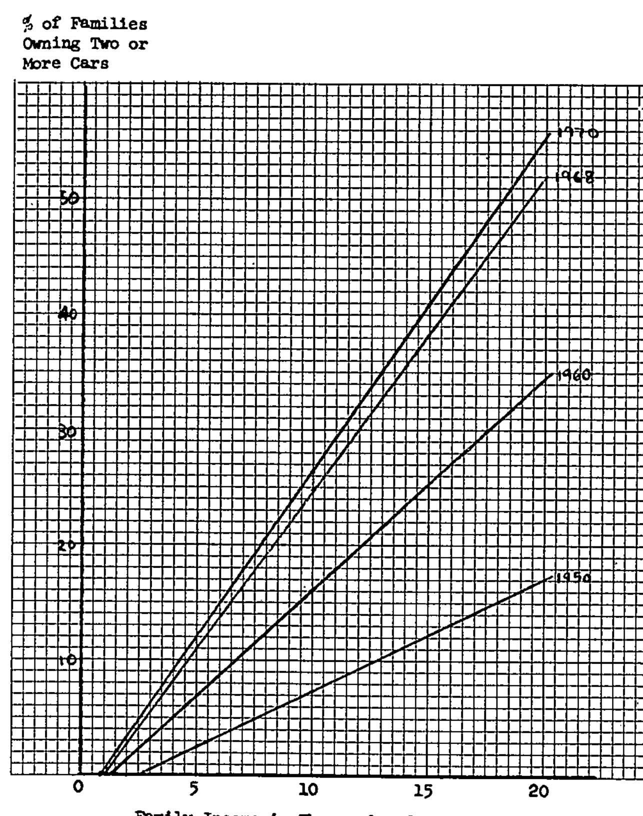


Figure 4. Estimated percentage of families owning two or more cars as a function of income.



2. Projections of the Population of the United States, by Age and Sex: 1972 to 2020. U.S. Bureau of the Census, Series P-25, No. 493, Dec. 1972.
3. Fisher, J. In New Scientist, Feb. 28, 1974.
4. Back to the Small Town. Newsweek, Jan. 21, 1974, p. 53.
5. Mobility of the Population of the United States, March 1970 to March 1973. U.S. Bureau of the Census, Series P-20, No. 256, Nov. 1973.
6. General Demographic Trends for Metropolitan Areas, 1960 to 1970. In Census of Population and Housing: 1970, U.S. Bureau of the Census, Final Rept. PHC(2)-1.
7. Mayer, H. M. Changing Urban Structure and Its Implications for Terminals and Pickup and Delivery Problems in Metropolitan Areas. HRB Spec. Rept. 120, 1971, pp. 110-120.
8. Pickard, J. U.S. Metropolitan Growth and Expansion, 1970-2000, With Population Projections. U.S. Commission on Population Growth and the American Future, Research Reports, Vol. 5, 1973.
9. Beale, C. L. Migration Patterns of Minorities in the United States. American Journal of Agricultural Economics, Dec. 1973.
10. Estimates of the Population of Metropolitan Areas, 1971 and 1972. In Current Population Reports, Population Estimates and Projections, U.S. Bureau of the Census, Series P-25, No. 505, Sept. 1973.
11. Hoch, I. Interurban Differences in the Quality of Life. In Transport and the Urban Environment (Rothenberg, J., and Heggie, I. G., eds.), Macmillan, London, 1974.
12. Taeuber, C. M. Measuring Patterns of Urban Growth. Oak Ridge National Laboratory, ORNL-UR-104, p. 24.
13. Sturgis, R. B. Selected Factors Influencing City Growth 1960-1970. Oak Ridge National Laboratory, ORNL-UR-109.
14. Clark, C. Population Growth and Land Use. St. Martin's Press, New York, 1968.
15. Muth, R. Cities and Housing. University of Chicago Press, 1969.
16. Barr, J. L. Transportation Costs, Rent and Intra Urban Location. Department of Economics, Washington Univ., 1970, pp. 44-45.
17. Land Use Change and Environmental Quality in Urban Areas. Earth Satellite Corp., 1973.
18. 1972 Automobile Facts and Figures. Automobile Manufacturers' Association, Inc., Detroit, 1972.
19. Katona, G., et al. 1968 Survey of Consumer Finances. Institute for Social Research, Univ. of Michigan, 1969.
20. Lansing, J. B., and Hendricks, G. Automobile Ownership and Residential Density. Institute for Social Research, Univ. of Michigan, 1967.
21. Mandell, L., et al. Surveys of Consumers, 1971-1972. Institute for Social Research, Univ. of Michigan, 1973.
22. Hoch, I. Automobile Registration Forecast. Chicago Area Transportation Study, 1959.
23. Power, J. Europe's Defenseless Immigrant Worker. New York Times, March 23, 1974, p. 31.
24. Reilly, W. K., ed. The Use of Land. Thomas Y. Crowell Co., New York, 1973.

Recent Observations on the Effect of Gasoline Price and Supply

Eugene J. Lessieu and Arnold Karvasarsky
Port Authority of New York and New Jersey

This paper addresses the economic effects of the gasoline shortage of 1974 and the ensuing gasoline price increases. Further, the investigation covers what is commonly thought of as two separate markets: the work or peak-hour travel market and the off-peak or nonwork travel market. Observations were made at the six port authority vehicle crossings: the Holland and Lincoln Tunnels and the George Washington Bridge, which provide access across the Hudson River between New Jersey and Manhattan, and the Bayonne and Goethals Bridges and Outerbridge Crossing which connect New Jersey and Staten Island. Traffic observations in the 6 months of 1974 showed that (a) traffic volumes declined much less in the second quarter of 1974 than in the first quarter; (b) weekend traffic was down much more than weekday traffic in the first quarter, whereas in the second quarter there was only a slight difference in weekday and week end declines; (c) the decline in weekday peak and off-peak traffic differed little in the first and second quarters; (d) the number of passengers per vehicle decreased in weekday off-peak hours and increased during some peak hours; (e) cross elasticity between automobile and transit modes was small but not zero; and (f) longer trips decreased more than short trips.

Many things influence traffic growth, including changes in socioeconomic variables (population, employment, and income), travel systems (new highways or transit systems), and relative travel costs. In an economic sense, travel is a "normal good" that demonstrates an inverse relationship between price and quantity and a direct relationship between quantity and income.

Many forces helped to maintain the high rate of growth in traffic volumes (3½ to 4 percent) during the past 10½ years. Pricing is only one explanatory variable and perhaps only a small part of the impetus to long-term growth. It could be argued then that a relative price increase in the cost of driving would have only a small effect on the growth of automobile traffic if all the other contributing forces remained constant.

In the long run, many of the other forces change, and pricing as a single force is difficult to isolate. During the first half of 1974, the short run, it may be inferred that pricing was the major cause of any changes in traffic growth.

Observations of short-run changes in traffic growth, however, must consider short-term influences. On a year-to-year comparison, weather proved to be an important variable. Minor changes caused by construction, strikes, and service disruptions in competitive modes must be taken into account as must holidays. Many of these influences were present in the raw data that were used for this paper, and attempts to adjust for them were made.

REACTION TO GASOLINE PRICE INCREASE

During the first quarter of 1974, reductions in volume due to gasoline price increases were masked by the overwhelming effects of gasoline shortages. As a result, the pattern that emerged in the first quarter showed heavy weekend losses (especially on Sundays) compared to weekdays. Following the lifting of the oil embargo in March, gas station lines virtually disappeared and with them the major problems of gasoline availability. Higher prices, however, remained as a deterrent to traffic growth.

Measurement of the effects of rising prices while all other variables are held constant can be accomplished by cross-sectional analysis, in which differential costs are typically related to differential distances, or time series analysis with cost comparisons between two time periods. In a modeling effort the cost coefficient that emerges, if it is free of the effects of the other socioeconomic variables considered, may be easily converted into an elasticity coefficient. Economists define elasticity as a measurement of the percentage of change in quantity (in this case, the number of trips per household) for a given percentage of change in price (total cost of trip). It is then a simple task to interpolate the total cost elasticities so as to consider the individual components of the cost, e.g., gasoline prices, tolls, parking costs, on the quantity of trips taken. For example, before the recent price increase, gasoline expenditure was estimated at about 20 percent of the total trip cost, the remainder being other operating and maintenance costs, parking fees, and tolls. Therefore, an elasticity coefficient of -1.0 for total cost would translate to -0.2 for gasoline alone.

Differing regions might be expected to show different elasticity coefficients, largely because the magnitude of the coefficient reflects the number of available substitutes for the good in question, including the obvious choice of doing without. Nevertheless, empirical studies around the country have yielded a remarkably stable gasoline price coefficient, ranging from -0.1 to -0.5. In this range, the effect is said to be inelastic, such that a 10 percent increase in gasoline price would yield only a 1 to 5 percent decrease in trips consumed. In cross-sectional studies done several years ago, the elasticities that emerged were related to trip purpose. The elasticity coefficient of the work trip was higher than that of the nonwork, leisure trip. The elasticity coefficients used were total cost coefficients and were estimated at -1.2 for the work trip and -0.7 for the nonwork trip. The effects of prices on work and nonwork trips translated into a 12 percent decrease in work trips and a 7 percent decrease in leisure trips for each 10 percent rise in total trip cost. Although it seems unusual that work trips are more responsive to price changes than nonwork trips, it should be noted that, in the New York-New Jersey metropolitan region, the availability of rail and bus transit, especially during peak periods, provides greater ease of substitution than in many other regions. For nonwork trips, most often made in other than the weekday peak period, transit alternates are not so readily available, and, as a result, the traveler's reaction to a price increase may be to change the destination or leisure activity to avoid the higher cost. This may take the form of substituting shorter trips for longer ones, e.g., a trip to a nearby beach rather than a camping trip.

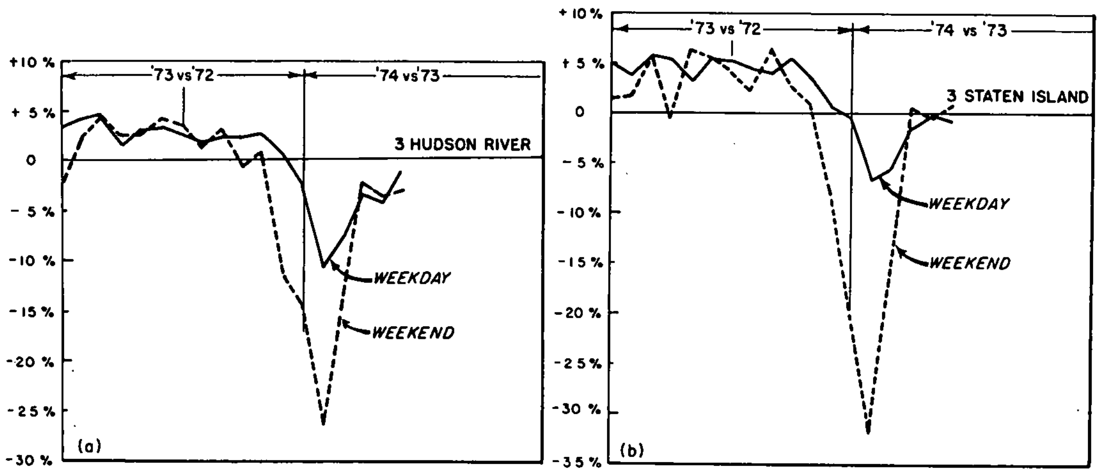
As of June 1974, gasoline prices had increased in the New York-New Jersey region by 35 to 40 percent. Feeding this range of price increases into the price elasticity coefficients yields an expected decline on the order of 6 to 7 percent in total trips. We may infer from the data given in Table 1 that, after the gasoline shortage in the first quarter of the year, April, May, and June traffic volumes fell short of 1973 volumes by about 2.8 percent. When the expected 3 to 4 percent growth from 1973 to 1974, which would have been anticipated in the absence of energy constraints (Fig. 1), is added to this, the resulting decrease in volume is about 6 to 7 percent, which the price effects lead us to expect.

Table 1. Adjusted decreases from 1973 to 1974 at six port authority crossings.

Month	Percentage Decrease		
	Weekday	Weekend	Monthly
January	3.9	14.2	6.4
February	10.4	27.7	15.1
March	7.3	13.6	9.0
April	3.0	2.0	2.9
May	3.5	3.4	3.4
June	2.4	1.3	2.1

Note: Erratic factors such as unusual weather, holidays, and strikes have been compensated for.

Figure 1. Adjusted monthly growth rates for 1973 and 1974 for (a) Hudson River crossings and (b) Staten Island crossings.



It should be noted that this 6 to 7 percent loss is measured in vehicle trips and measured from expected traffic. The estimated distribution of this traffic loss is discussed later. In brief, whereas some of the traffic loss is from trips that have been discontinued, some is accounted for by carpooling and the remainder by shifts to public transit.

PEAK VEHICULAR TRAFFIC CHANGES

Observations of peak traffic volumes both during the height of the gas shortage and in the second quarter of 1974 showed no difference between the change in peak traffic and the change in weekday traffic. The peaking pattern has remained similar. It seems that the effects of the gas price increase, or overall price increase, on peak-hour work traffic and on off-peak nonwork traffic have been similar.

Before we address the magnitude of the changes observed in the peak, some relationships must be established between peak travel and work travel: Are the two terms really synonymous?

Figure 2 shows automobile volumes by time of day and trip purpose. The data for this figure are from a 1972 port authority continuous-sample origin-destination survey. Figure 2 shows eastbound traffic only: New Jersey residents going to work in the a.m. peak and New York residents returning from work in the p.m. peak. The business trips bulk up in the middle of the day; other trips continue throughout the day into the night. It is apparent that not all the peak travel (particularly the p.m. peak) is work travel and conversely not all the work travel occurs in the peak periods. Data given in

Figure 2. 1972 average weekday trip purposes for (a) Hudson River crossings and (b) Staten Island crossings.

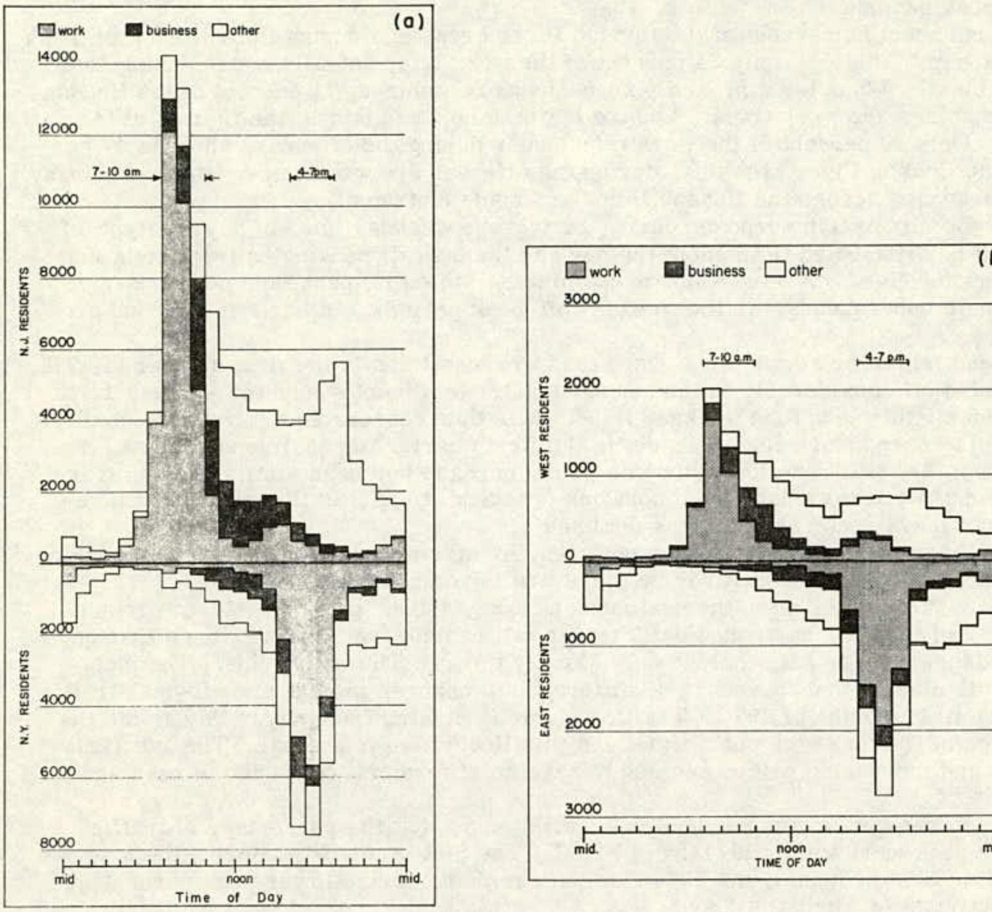


Table 2. Percentage distribution of eastbound trips by trip purpose.

Time Period	Hudson River Crossings		Staten Island Crossings	
	Work	Nonwork	Work	Nonwork
Weekday				
7 to 10 a.m.	83.7	16.3	82.0	18.0
4 to 7 p.m.	55.9	44.1	65.4	34.6
7 to 10 and 4 to 7	70.8	29.2	71.4	28.6
Off-peak	43.9	56.1	38.4	61.6
24 hours	56.6	43.4	52.9	46.1
Weekend				
Saturday	19.5	80.5	25.0	75.0
Sunday	8.9	91.5	9.0	91.0

Table 3. Percentage distribution of weekday eastbound trips by time period.

Time Period	Hudson River Crossings		Staten Island Crossings	
	Work	Nonwork	Work	Nonwork
7 to 10 a.m.	33.6	9.5	24.6	6.2
4 to 7 p.m.	19.5	22.4	34.2	20.8
Off-peak	46.9	68.1	30.2	73.0
24 hours	100.0	100.0	100.0	100.0

Tables 2 and 3 show the percentage distribution of work and nonwork trips during peak and off-peak periods.

Although about 84 percent of the Hudson River crossings during the 7 to 10 a.m. peak are work trips (Table 2), only 34 percent of the work trips actually occur during that peak (Table 3). When the a.m. and p.m. peaks are combined, 71 percent of the Hudson River crossings are work trips. For the Staten Island crossings, the figures differ slightly. Only 30 percent of the work trips occur during the off-peak, whereas 47 percent of the Hudson River crossings during the off-peak are work trips. However, many of the work trips across the Hudson River are made by transit.

Given the mix of trip purposes during an average weekday, loss of any one type of trip could be distributed throughout the day and the overall peaking pattern might not change much. However, work trips predominate in the a.m. peak, and nonwork trips predominate on weekends. In the weekday off-peak periods, neither trip purpose predominates.

Weekend trips were down much more than weekday trips in the first quarter of 1974. This leveled off considerably in the second quarter to a point where the weekend trips were down slightly less than weekday trips. It is thus reasonable to attribute the difference in weekend and weekday levels in the first quarter to gasoline shortages. It would seem then that reaction to the gas price increase has been similar among travelers who make work (weekday) and nonwork (weekend) trips. Is this phenomenon reflected in weekday peak and off-peak declines?

From observation so far, we have not seen any major change in the weekday peaking pattern on either the Hudson River or the Staten Island crossings.

The data in Figure 3 show the eastbound weekday 4-hour peak (6 to 10 a.m.) as a percentage of the total eastbound daily traffic (all vehicle types) on the three Hudson River crossings. The data include only Monday through Thursday trips. The plot shows quite clearly that there was no difference in changes in peak and off-peak traffic during the first months of 1974. Traffic volume fluctuated from day to day as did the percentage of traffic in the peak hours, but this fluctuation is normal. The gasoline shortage and increasing prices seemed to have no differential effect in the peak and off-peak.

The second-quarter comparisons in Figure 3 show that the percentage of traffic during the peak was lower in 1974 than in 1973. The plot is distorted during the 3 weeks surrounding Easter Sunday, but 1974 continues to demonstrate lower peak percentages throughout most of April and May.

It would appear from this figure that the pricing effect was greater in the peak than in the off-peak; however, this is not the case. The June 1973 and 1974 comparison is similar to the first-quarter comparisons, i.e., no apparent difference in effect on peak and off-peak traffic. The traffic pattern shift that shows up in April and May is due to the PATH strike in 1973 when more peak trips were diverted to automobiles than off-peak traffic, thus raising the peak percentages to a higher level than would occur at this time of the year.

Comparison of 1973 and 1974 traffic on the Staten Island crossings also shows no basic difference between the change in peak traffic and the change in off-peak traffic. This is true even though there was a significant shift of traffic when a new section of highway opened in the period between the two counts.

It could be concluded then that the relatively higher work orientation in off-peak trips has some effect on the fact that the peak and off-peak patterns did not change significantly with the increase in gasoline price, even though other changes may have taken place in work trips and nonwork trips.

It should be noted, of course, that these traffic shifts are small in all of the second-quarter observations. Further, an assumption that trips were lost in the work or nonwork categories cannot be proved without some field surveys on trip purpose. There are, however, additional data from which further insight into the peak and off-peak, work and nonwork situation can be derived. These data are from a survey of passengers per vehicle (p/v) made in April 1974.

Figure 3. Peak-period (6 to 10 a.m.) eastbound traffic as percentage of 24-hour traffic for Hudson River crossings.

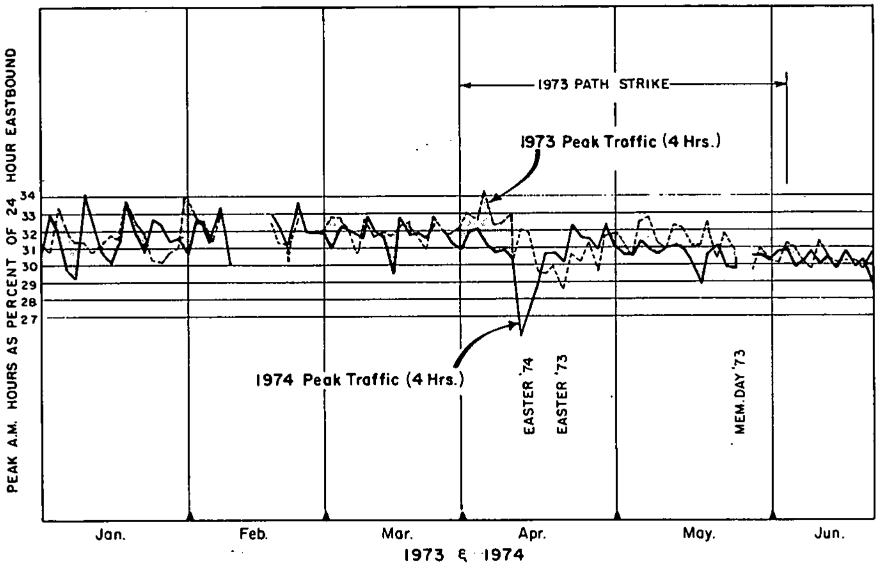


Figure 4. Passengers per car for Hudson River crossings.

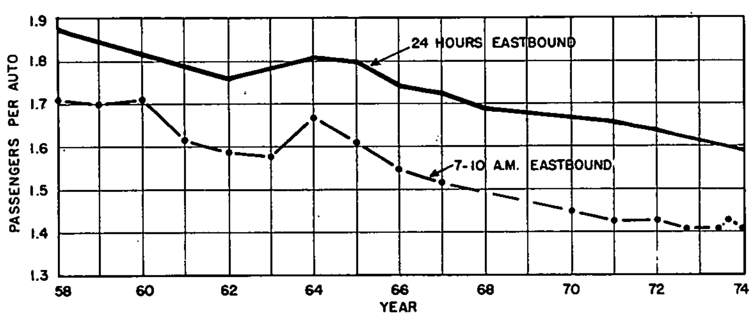
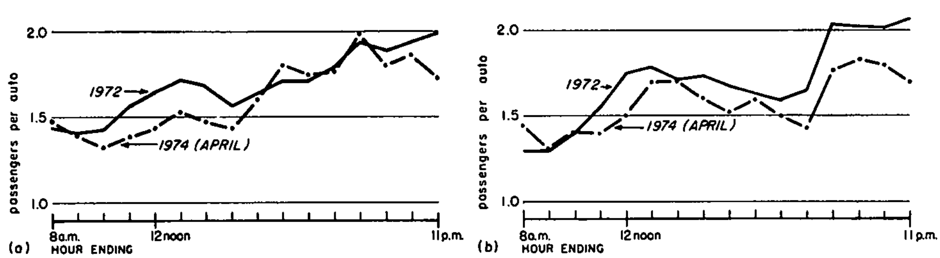


Figure 5. Passengers per car for (a) Hudson River crossings and (b) Staten Island crossings.



PASSENGERS PER AUTO

One of the methods suggested to reduce gasoline consumption during the height of the gas shortage was car pooling. Many efforts were and are being made to encourage car pools. Surveys of p/v rates were made at all six port authority facilities in April 1974 to determine whether the gas price increase or promotion of car pools effected car pooling.

Car pooling, in theory, can have a tremendous impact on peak vehicle demand reduction, on gasoline use, and not the least of all on the cost of using an automobile. The reduction in per-person cost due to 2, 3, or 4 people sharing the cost is dramatic. Car pooling is defined here as a group making a work trip. It is most prevalent where origins and destinations are dense enough to make the establishment of a car pool convenient, and it is done to reduce costs without sacrificing too much convenience.

Data collected over the years in continuous origin-destination surveys have allowed us to keep track of changes in the p/v index. The overall p/v index has been declining for many years (Fig. 4), probably because of increased affluence, increased automobile ownership, and dispersal of residences and jobs.

Data derived from the 1972 continuous origin-destination survey reveal some findings that tend to validate the statement on dispersal of activity causing lower p/v rates. First, work trips for west of Hudson residents show an increase in p/v in the a.m. peak. Also, the p/v rate of work trips at the three Hudson River crossings is greater than at the Staten Island facilities. An additional breakdown for the Hudson River crossing destined for the Manhattan CBD showed even higher p/v rates in this time period. Thus, the convenience of time and location does appear to foster car pooling. For residents east of the Hudson, p/v is higher in the return from work trip. Actually, on the three Hudson River facilities the p/v rate for the New Yorkers returning from New Jersey work places (the reverse commuters) is significantly higher than for the typical journey to work movement. This is probably caused by a greater car pooling effort on the part of the former group for the purpose of cost savings. This is in general a lower income group, traveling to common destinations at large New Jersey factories. Public transportation is less satisfactory, which leaves car pooling as the most viable means of reducing costs.

Three surveys were taken prior to this study: December 1973, January 1974, and the end of April 1974. The first two were taken in the 7 to 10 a.m. peak period only eastbound, and the last one included part of the remainder of the day, to 11 p.m. The results of these surveys indicate little change in p/v rate during the peak, but significant decline in the off-peak p/v rate. Figure 5 shows a comparison of p/v rates for 1972 and April 1974.

In addition to classifying the p/v data by time of day, we classified the data according to whether tolls were paid by cash or ticket. The toll for automobiles is $1 round trip (collected eastbound only). Reduced tickets are available at 50 percent discount for 20 nontransferable tickets good for 30 days and at 20 percent discount for 12 tickets good for 2 years. The following observations were made:

1. Use of discount tickets (mostly for work trips) decreased slightly more than did the number of automobile users;
2. The p/v rate of discount ticket users was up in all time periods, especially in the p.m. peak; and
3. The p/v rate of cash toll payers was down noticeably, especially during off-peak periods.

At the Hudson River crossings, the work trip volume is down as much as or more than the leisure trip, verifying that the elasticity of the work trip in this market is higher than that of the nonwork trip. The fact that the p/v rate of the discount ticket user is up may be a sign of increased car pooling. The fact that it is up more in the p.m. peak indicates greater car pooling in that market. In the a.m. peak, a greater shift to public transit by the one-occupant car user is a reasonable expectation.

In addition, using discount tickets for tolls has decreased for work trips because fewer work trips are made by automobile and discount tickets have a time limit. Some of the substitute trips are possibly made by public transit. Such a shift would result in a lower overall p/v rate among those paying cash tolls.

It also appears that the greatest loss in the nonwork category is in the family-friends type of trip. These trips have historically had the highest p/v rate, and the loss of these trips in the off-peak hours explains the significant drop in the p/v of cash toll payers in the off-peak hours. It is also possible that the loss of some of these trips during peak periods is one reason why the peak p/v rate has not changed as much as the off-peak, i.e., loss of high p/v nonwork trips is counteracted by low p/v work trips.

One difference in the observations at the Hudson River and Staten Island crossings should be explained. The a.m. peak p/v rate changed significantly more on the Staten Island facilities than on the Hudson River ones. Because the market served by the Staten Island facilities has little or no public transit service, the tendency toward car pooling in an effort to reduce cost in the work trip category was greater.

The fact that the p.m. peak traffic now shows a higher p/v rate at the Hudson River facilities and a lower p/v rate at the Staten Island crossings might be explained by the exceptionally high p/v rate for work trips in that period on the Hudson River facilities and therefore a smaller difference in p/v between work and nonwork trips.

SHIFT TO PUBLIC TRANSIT

It has been shown that in the second quarter peak traffic volumes on the Hudson River crossings declined about 3 percent or roughly 4,000 to 5,000 passengers for both peaks. Based on analysis of p/v changes, it is estimated that some of the travelers, particularly for work trips, shifted to public transit.

Attempts have been made to trace this shift, but the volume is small, the public transit volume is large, and the task was difficult. The transit passenger volume crossing the Hudson River in the peak is about three times as high as the automobile passenger volume. A loss of 3 percent from automobiles then means only a 1 percent increase in public transit volumes. The special counts that would reveal a shift are expensive and cannot be made very often. As a result, a 1 percent change derived from isolated counts cannot really be recognized. Isolated counts have indicated some small increases on the public transit system in the peak period, and it is not unrealistic to attribute it to a shift from the automobile.

OTHER HIGHWAY FACILITIES IN NEW JERSEY METROPOLITAN AREA

Because traffic volumes at port authority facilities changed in certain ways during the gas shortage of the first quarter of 1974 and in other ways in the second quarter, traffic data were collected on other specific highway types throughout the area to determine whether the findings were consistent and to test ideas not readily apparent from port authority facility data.

First, we observed the dramatic decline in Sunday traffic in the first quarter and much less in the second quarter. Second, based on port authority data but not origin-destination information, we made the assumption that local trips declined less than longer trips. To check these observations, we investigated permanent count data from several highway locations in New Jersey. Different types of highways were selected to illustrate the findings.

In almost all cases, observations were consistent with findings on port authority facilities. The hypothesis that longer trips were affected to a greater extent than shorter trips was supported by interchange data from the New Jersey Turnpike. Equally supportive were the vehicle counts taken on express and competing local roadways in the region, such as I-80 approaching the George Washington Bridge.

CONCLUSIONS

In the progress on this study, observations of total traffic volumes, weekday and weekend volumes, peak traffic volumes, and automobile occupancy at various times of the day were made. Comparisons of the changes observed in these traffic data provided sufficient evidence to estimate the magnitude of the effect on the driving public of the gasoline price increases in the first half of 1974.

First, the substantial reduction in traffic volume during the first quarter of 1974 can be attributed for the most part to the gasoline shortage, not the price. When gasoline was generally available in the second quarter, traffic recovered. The recovery was not complete, and the declines in traffic volumes from a year ago can be attributed mostly to rising gasoline prices.

From previous studies, total cost elasticities were estimated in the range of -0.7 to -1.2. However, gasoline cost only accounts for about 20 percent of the total cost of driving; therefore, an expected decline in traffic from a 35 percent gas price increase would be about 7 percent. Traffic declines of 2 to 3 percent from last year were observed that, when added to a normal growth of 3 to 4 percent, result in a decline of 6 to 7 percent, as expected.

Observations of peak and off-peak weekday traffic indicated no change in the peaking pattern. Both peak and off-peak traffic reacted to the gasoline price increase. The mixture of trip purposes in both the peak and off-peak hours tends to soften any differential reaction to rising prices that may have occurred among different trip purposes. Nevertheless, there is evidence that both work and leisure trips were affected by the price increase and even some evidence that work trips were affected more than nonwork trips.

From the data available, it appears that peak-period work travelers reacted to the price increase by car pooling and by switching to public transit. Off-peak leisure travelers seem to have reduced the number of trips or made shorter trips.

Regarding peak travel demand reduction, it is apparent that pricing does have some effect. In this case, because the peak travel market is strongly work trip oriented, person trips were not reduced, but the automobile trips were. In large measure, as many trips are still being made, but some automobile trips are now being made on public transit or in car pools.

Legislation for Increased Highway Safety

W. A. Goodwin
University of Tennessee and
J. B. Phillips, Jr.
Tennessee Office of Urban and Federal Affairs

The Federal-Aid Highway Act of 1956 authorized positive action on the problem of highway safety. The National Traffic and Motor Vehicle Safety Act and the Highway Safety Act, both passed in 1966, gave highway safety national priority. The Highway Safety Act includes 18 safety standards that, if fulfilled, will increase the effectiveness of the total highway transportation system. Other federal legislation, including title II of the Federal-Aid Highway Act of 1973 (known as the Highway Safety Act of 1973), is discussed. Among the goals of these pieces of safety legislation are to standardize and simplify operation of the system, which in turn will increase its utility.

Public awareness and concern for highway safety date back to the early 1920s when Secretary of Commerce Herbert Hoover called the first National Conference on Street and Highway Safety. The conference was aimed at furthering means to save life and prevent accidents and to make highway travel safer for both pedestrians and passengers. The death rate at that time was 17 per 100 million vehicle-miles traveled.

In 1946 the first President's Highway Safety Conference was convened. In 1954, the White House Conference on Highway Safety was convened, and out of this conference the President's Committee for Traffic Safety was formed.

It was not until the 1956 Federal-Aid Highway Act that positive action was authorized. This act directed the Secretary of Commerce to undertake a comprehensive investigation of the entire subject of highway safety and to report the findings. The findings and recommendations were published in 1959 (1).

Among the recommendations of this study was the establishment in 1960 of an interdepartmental highway safety board chaired by the Secretary of Commerce and including the Departments of Health, Education, and Welfare, Defense, and Labor and the Interstate Commerce Commission, General Services Administration, and Post Office Department.

This board formed a working committee. The board in 1965 published a report on federal policy and programs for highway safety (2). By 1965, the number of deaths per 100 million vehicle-miles traveled had dropped to about 5.8.

Subsequent to the Commerce study, several important pieces of legislation were enacted, including standards for seat belts, hydraulic brake fluid, and the Baldwin amendment to the 1965 Federal-Aid Highway Act. This amendment clearly established leadership and coordination for a national highway safety effort. It noted among other things that "after December 31, 1967, each state should have a highway safety program... designed to reduce traffic deaths, injuries, and property damage...."

In late 1965, President Johnson commented to the public about traffic deaths and noted that in 1966 he planned to recommend steps to mount a major campaign against such senseless and terrible loss of life.

Such highway accident statistics for 1965 as 49,000 fatalities, 1.5 million disabling injuries, and $8.5 billion in property damage gave the impetus to Congress to promulgate the 1966 acts. When the landmark Highway Safety and National Traffic and Motor Vehicle Safety Acts of 1966 were signed, safety on the nation's highways became a national priority.

Although street use was not among the issues that led to the highway safety acts, several of the 18 highway safety standards that have been promulgated under section 402 of the 1966 act have a direct bearing on the issue; an indirect effect is noticeable in others if all requirements are fulfilled.

Increased effectiveness of the total highway transportation system will occur when accidents that cause delays and confusion among system users have been reduced. In addition, the safety requirements standardize and, to some extent, simplify the operation of the system, thereby increasing its utility. Among the standards that have a direct bearing on street use are the following:

Standard No.	Issue
301	Periodic motor vehicle inspection
306	Codes and laws
609	Identification and surveillance of accident locations
310	Traffic records
612	Highway design, construction, and maintenance
613	Traffic engineering services
316	Debris, hazard control, and cleanup

Implementation of the requirements of many of these standards may necessitate legislative action on the part of the states. For example, Tennessee must conduct periodic motor vehicle inspection (PMVI) by June 1975, or the state will lose federal funds.

PMVI may be justified on the basis that some accidents are caused by faulty vehicle components; studies reveal that about 6 percent of all fatal accidents may have been caused primarily by a faulty vehicle component. The primary goal of PMVI is to reduce crashes resulting from improperly maintained or repaired vehicles.

State legislation concerning PMVI predates the Highway Safety Act by many years; Massachusetts initiated a motor vehicle inspection program in 1926. Eleven states have enacted legislation since the passage of the 1966 act, and today 31 states and the District of Columbia require vehicle inspection on a statewide basis. Obviously vehicles that are well maintained have a greater chance of being operated on roads and streets with minimum downtime and interference with traffic flow.

Making maximum use of existing highway transportation facilities requires that their users be governed by specific rules and understand what conduct is necessary and permitted for safe travel. Where the traffic laws of a state conflict or are inconsistent with those of other states, drivers crossing state borders may find themselves in situations where adherence to the traffic statutes of their home states results in traffic violations and accidents. To establish a substantial degree of uniformity among the states' traffic regulations, standard 306 was issued in 1967. It requires that each state develop and implement a program to achieve uniformity of traffic codes and laws throughout the state.

The Uniform Vehicle Code sets forth rules that should govern the behavior of drivers and pedestrians as they interact on public roads. The National Committee on Uniform Traffic Laws and Ordinances states that, if the public is to understand, remember, and observe these rules in moving from state to state, the rules should be exactly the same, word for word, in every state. Such uniformity makes it easier for police officers, judges, and traffic engineers to exercise their duties in a fair and equitable manner.

Standard 609 covers the identification and surveillance of accident locations. It requires that each state maintain a current inventory of accident locations and that it organize a system for identifying and correcting particularly high accident locations. Probably the most important standard, as related to the management of the states' overall highway safety effort, is the one setting forth the requirements for traffic records. Standard 310 requires that information on vehicles, drivers, and accidents be systematically entered into a computer-based data system for rapid entry and referral.

Standard 310 has probably been the greatest stumbling block for the states because of the lack of guidance and coordination at the federal level. Many states began in ignorance setting up computer-based systems, and those that did begin did so with little or no guidance in making the systems compatible among themselves. A nationwide traffic record system with each state having a subsystem could have tremendous benefit in the determination of causes and the identification of remedies as well as efficient allocation of resources. FHWA standard 612 has a direct impact on street and highway use. It covers highway design features that lead to accident prevention such as breakaway utility poles and postcrash activities including routing for emergency vehicles. FHWA standard 613 encourages and requires identification of needs and reporting of deficiencies. Another significant requirement of this standard is the upgrading of existing traffic control devices. Several states have completed inventories of all traffic control devices and have devised a system for keeping the inventory current. Information contained in the system will aid in the selection and installation of the needed traffic control devices.

Highway safety standard 316 was issued in 1968 and requires that each state have a debris, hazard control, and cleanup system. Each state, in cooperation with its governmental units and political subdivisions, is to ensure rapid, orderly, and safe removal from the highway of wreckage, spillage, and debris resulting from highway incidents. Such a system will also reduce the likelihood of secondary and chain-reaction collisions. The standard calls for an operational procedure to enable rescue and salvage equipment and personnel to arrive on the scene of accidents and incidents with minimum delay. The following description illustrates how standard 316 can improve street use and capacity during incidents:

> A heavily loaded truck crashed into a bridge railing recently at mid-afternoon in Washington, D.C. The driver was tossed 50 feet to the river below while the truck remained above, blocking traffic and causing a huge traffic jam on the bridge that extended considerably beyond both bridge approaches. At one time traffic in almost one-fourth of the city was involved in the jam. The truck was so heavy police cranes had to be threaded—and threaded seems to be the right word—through the knotted traffic to the scene. Overall, over 3 dozen policemen spent almost 3 hours untangling the traffic confusion.

The actual and potential impact on street capacity and use of the 1966 Highway Safety Act and its resulting standards is recognizable and constitutes an improvement to the transportation system. Although saving lives is the most important element of the nationwide highway safety goals, the disheartening element is that the federal government has not felt a mandate to adequately fund highway safety activities. The most important section of the act is section 402, which establishes the states' partnership with the federal government in implementing the 18 standards and reducing deaths, injuries, and accidents. Even the most uninformed clearly recognize that a greater impact on saving lives and reducing costs to motorists could be achieved by appropriate funding. For example, the authorization for section 402 for fiscal year 74 was approximately $100 million, whereas the appropriations were only $80 million. The direct federal aid to states and communities under section 402 totaled $470,779,000 from fiscal 1967 through fiscal 1974. These funding levels might not be questioned except that the 1968 study to determine the cost of carrying out the Highway Safety Act of 1966 estimated that it would require $8.3 billion in federal funds for fiscal 1974 and 1975. This study, commonly known as the needs study or the 207 study, took a critical look at the funding requirements for the 10-year period for 1967 to 1976. Inasmuch as the study was released in October 1968, the estimate is probably conservative.

A commendable feature of the 1966 act is the requirement that each state develop and submit for review and approval a 5-year comprehensive plan as a planning guide to program development. In 1971, planning and budgeting procedures were further refined when the NHTSA required the states to initiate an annual work program that outlined program plans on a fiscal year basis.

The comprehensive plans and fiscal year work programs submitted by the states provide a realistic estimate of funds needed to fully implement the highway safety standards. In the absence of appropriate funding levels from the federal government, the states are reluctant to pass needed legislation. For example, a number of states are dragging their feet on initiating statewide PMVI and, as a consequence, will be forced into acting when the Secretary of Transportation invokes the penalty clause in the act. This clause essentially permits the withholding of federal funds. Some believe that adequate funding by the federal government would lead to greater cooperation by the states in their passage of needed legislation.

The most important federal legislation on highway safety passed by Congress since the 1966 act is title II of the Federal-Aid Highway Act of 1973, which is frequently referred to as the Highway Safety Act of 1973.

In addition to section 402 and section 403, the 1973 act sets forth specific programs that Congress deems appropriate and needed. These programs are identified as separate sections of title II.

Section 203 requires that each state conduct and systematically maintain a continuing survey of all highways to identify those railroad crossings that may require separation, relocation, or protective devices and to establish and implement a schedule of projects to correct the deficiencies. Specific reporting requirements were included in the legislation, and the first report was due in September 1974; thereafter, they are due annually. An important aspect of this section is the availability of funds to local governments for matching state funds for the improvement of railroad crossings.

Section 205 amends the U.S. Code by adding a pavement marking demonstration program to provide greater vehicle and pedestrian safety. Under this section, work can be performed on any highway whether or not it is on the federal-aid system. The only exclusion is the Interstate System. This, of course, means that roads and streets are included and their use can be improved by clearly visible markings that conform to FHWA standards.

Section 209 requires the states to make a continuing survey of all highways to identify and correct high-hazard locations that may constitute a danger to vehicles and pedestrians. For fiscal year 74, $50 million was appropriated out of the Highway Trust Fund. Section 209 has specific reporting requirements including an assessment of the cost of, and safety benefits derived from, the means and methods used to mitigate or eliminate hazards and also report on the previous and subsequent accident experience at locations where projects are undertaken.

Section 210 authorizes $25 million in funds for use solely on projects to eliminate roadside obstacles on any federal-aid road except Interstate. This section also authorizes the replacement of existing sign and light supports that are not designed to break away upon impact.

FHWA published instructions for conducting surveys required by sections 203, 209, and 210. These instructions make use of the rail-highway crossing inventories that were begun in early 1973. The information in these inventories, coupled with on-site inspection, should allow the states to advance some of their grade crossing improvement projects to a higher priority for funding. Many states already have continuing engineering surveys of high-hazard locations on the federal-aid system, and these procedures can be expanded to apply to all highways. Section 210 suggests using a windshield survey of statistically selected sections of highways to determine the number and types of hazardous obstacles—for example, utility poles that are within 30 feet of the edge of the traveled way, except those installed in protected locations. The FHWA instructions also suggest that this survey use or be supplemented by other roadway data that may be on file with the state and local governments.

Although the 1973 act has other sections such as educational programming and driver education evaluation that have an impact on highway safety and are not discussed here, there are two more sections that are unique and have a more direct bearing on improved use of existing transportation facilities.

Section 214 authorizes expenditure of federal-aid highway funds for the construction of bikeways and pedestrian walkways outside the normal highway right-of-way along federal-aid highways. The new program provides for the construction of bicycle and pedestrian facilities on a 70 percent federal and 30 percent state funding basis. The maximum annual expenditure during any fiscal year is $40 million, and there is a $2 million limit for an individual state. Another feature of this section is that the funds may be used to acquire additional right-of-way to assist in the construction of facilities to serve traffic that would have normally used the federal-aid route. The states are being encouraged to take advantage of this opportunity to use federal-aid funds because of the anticipated environmental, recreational, and safety benefits.

Section 230 is probably the most exciting element of the 1973 act as it contains the Federal-Aid Safer Roads Demonstration Program. This program applies to all public roads or segments that are not on the federal-aid system and that need improvements to correct safety hazards selected or designated by the state. This section is another first, for it authorizes federal-aid funds for safety improvements on non-federal-aid roads. The section requires that each state provide a report identifying and setting a priority on projects for improvement of highway marking and signing, elimination of roadside obstacles, elimination of hazards at railroad-highway grade crossings, and correction of high-hazard locations that have been identified by accident reporting and traffic records systems. FHWA has suggested that these projects may be selected from the states listings developed through the safety improvement program and TOPICS studies.

Tennessee implemented this section through its Department of Transportation and the Office of the Governor's Highway Safety Coordinator who met with the state's nine development districts and are providing funds for local surveys of needs on non-federal-aid roads.

The priority of these project needs will be based on local government evaluation and will be reviewed by the state DOT and Highway Safety Office for statewide coordination and allocation of funds. Activities are under way in Tennessee to meet the requirements of this section.

Section 402 on state and community programs has a new provision that authorizes the Secretary of Transportation to offer incentive grants—up to 25 percent of a state's annual grant apportionment—to states that enact mandatory seat belt use laws or make significant progress in reducing traffic fatality rates. Other provisions written into section 402 include teaching bicycle safety in driver education programs.

The most important change in section 403 on research and development is the safety needs study. By January 10, 1976, Congress is to receive a report on the evaluation of ongoing federal programs and is to provide a basis for further authorizations.

In 1972, the Motor Vehicle Information and Cost Savings Act was passed. This act has four titles.

1. Title I authorizes the promulgation of bumper standards.
2. Title II requires that a comprehensive study and investigation of the methods for determining certain characteristics of passenger vehicles relating to damage, crashworthiness, and repair be conducted.
3. Title III requires the establishment of diagnostic inspection demonstration proj-

ects to determine repair costs for vehicles when inspected in accordance with the vehicle-in-use standards. Also, information on the costs for maintenance and repair of emission control systems will be collected. The first demonstration project is under way. The act specifies not less than five nor more than 10 projects.

4. Title IV relates to odometer requirements and establishes the reliability of odometers and safeguards against their alteration or resetting.

SUMMARY

It has taken a long time for the Congress and the public to recognize the importance of a nationally organized highway safety program to save lives and reduce injuries and property damage.

The 1966 act was a major piece of legislation in this respect. Had appropriate federal funding and adequate guidance from the federal level been provided and had sincere dedicated response been forthcoming from state and local governments, much more could have been accomplished in the past 8 years. This has not been the case; and, as a consequence, a rejuvenation is in order. The 1973 Highway Safety Act goes a long way toward putting highway safety back on the track of meeting its 1966 goals. We know what needs to be done, and how to do it, so why can't we!

REFERENCES

1. The Federal Role in Highway Safety. U.S. Government Printing Office, March 3, 1959.
2. Federal Policy and Programs for Highway Safety. Interdepartmental Highway Safety Board, March 1965.
3. Comprehensive Plan for Highway Safety, State of Tennessee. Office of Urban and Federal Affairs, Highway Safety Planning Division, 1973.

Making Better Use of Existing Facilities Through Highway Safety Improvements

Roy E. Jorgensen
Roy Jorgensen Associates, Inc.

This paper describes work done on NCHRP Project 17-2A. Part of the project goal was to develop a users manual giving a step by step approach to establishing safety improvement programs. This paper summarizes some of the findings and management issues and outlines the system incorporated in the users manual.

Making improvements that reduce safety hazards is obviously one of the ways in which we can make better use of existing highway facilities. But what, specifically, do we mean by highway safety improvements?

First, we mean identifying hazardous locations, places where the relationship between physical conditions and operating characteristics of traffic creates a special hazard. It may be an intersection with a variety of or no traffic control devices. It may be a stretch of highway with soft shoulders or a narrow bridge or a blind sight distance or a slippery pavement, i.e., places on the highway that may lead drivers to wrong decisions, or compound the mistakes of drivers.

Second, we mean determining what can be done by the highway agency to reduce the hazard at specific locations. What alternatives should be considered? Should we consider signs, signals, or channelization at problem intersections or problem areas? Widening the bridge, flashing lights, and reflectorization may be alternatives at a narrow bridge.

Third, we mean evaluating alternative safety improvements as the basis for establishing a specific project improvement and the development of short-term and long-term highway safety improvements. How much of an improvement can be expected, and what will it cost? How in the short term can we get the maximum payoff in safety? How can we obtain adequate long-run financing of safety improvements? And how do we merge the desire of early payoff with the maximizing of safety in the long term?

For many years highway agencies have been making improvements specifically designated and programmed for safety. Many agencies have earmarked funds for so-called spot improvements for safety. The federal government has promoted spot improvements for 10 years. The 1973 Highway Safety Act establishes programs oriented toward specific types of improvements such as striping and railroad grade crossing protection.

Many agencies have reported striking accomplishments through their safety improvement programs.

In spite of all the attention that has been given to safety improvements, there are no well-defined scope and objective and planning and evaluating process to ensure attainment of the objective. To meet the need for a coordinated system and to give guidance and impetus to the highway safety improvement programs, the National Cooperative Highway Research Program established a project on methods for evaluating highway safety improvements. The project included development of a users manual that provides a step by step approach to establishment of safety improvement programs and evaluating accomplishments.

Carrying out the project included

1. Examination of the state of the art,
2. Development of a coordinated system for evaluation of needs and programming and control of improvements,
3. Examination of management issues related to the system, and
4. Development of a users manual.

In this brief presentation, we are summarizing some of the findings and management issues and briefly outlining the system incorporated in the users manual.

FINDINGS

In conducting the project, we determined that most agencies use accident data to identify hazardous locations, but there is need for more accurate data and better use of data. The users manual presents several basic procedures for identifying hazardous locations.

Only one agency has an operational system that evaluates needs, considers alternative improvements, programs improvements, and evaluates postimprovement accomplishments. The users manual outlines a management system that can serve as a guide to all highway agencies.

Agencies with low-volume road systems have difficulty using statistical techniques for establishing hazards. The users manual provides alternatives.

Most agencies now have inadequate record data on which to evaluate the potential for accident reduction. Before and after evaluations need to be systematized, and the results of these evaluations should be used to continually improve the data base on which forecasts may be made. The users manual establishes these evaluations as a basic part of the coordinated system.

Most agencies have safety improvement programs based on a more or less arbitrary level of financing. Existing funding levels are not necessarily based on knowledge of the problem size. How much of the safety problem is soluble by highway safety improvements at different levels of funding needs to be determined. Most agencies have not taken the time to determine how much funding is needed to solve the problem. They have tended to make the most of relatively small but protected earmarked funds instead of preparing the facts needed to sell legislators on longer solution-oriented programs.

No agency has put the safety improvement program into open competition with regular construction programs for construction dollars. The users manual can help in several ways to increase the justification for highway safety improvement funding. First, the program evaluation method provides the facts needed to justify highway safety improvements as a program. Second, the evaluation of alternative improvements can be applied to comparisons of safety improvements and regular improvements.

Existing systems are almost totally based on analyses of accident data, and so is the users manual. However, many new approaches are being developed, including the following.

1. Conflicts analysis and other similar field observations may soon provide much quicker ways of identifying hazards and evaluating results shortly after improvements have been installed. In a period of hours, two men can collect conflicts data, which may tell the same story that we now hear after waiting 1 year for accident statistics.
2. Operations research techniques such as economic models can be used to select better programs. Alabama recently developed a dynamic programming model to

maximize benefits from its annual highway improvement program budget.

Many innovative people are developing these new methods. It is very encouraging from the methods side. Establishment of the system outlined in the users manual will ensure that the new methods are directed toward achieving the highway safety goals.

THE SYSTEM

The system outlined in the users manual consists of six major steps:

1. Identifying hazardous locations,
2. Identifying problem causes and selecting possible alternative improvements,
3. Evaluating the alternative improvements,
4. Programming and implementing the improvements,
5. Evaluating implemented improvements, and
6. Evaluating the highway safety improvement program.

Note that there are three distinct evaluations in this system. They relate to the following questions:

1. What is the potential problem-solving value of an alternative improvement, and how much is it likely to cost?
2. What was the value of benefits actually obtained from an improvement?
3. What is the actual value of benefits from the overall program?

The first two steps in the system narrow the field of candidate locations and attempt to determine why each location is hazardous and how the hazard can be eliminated. In addition to the proven tools such as collision diagrams, the users manual presents guidelines for using methods such as multidisciplinary investigation teams and fault tree analysis to get to the cause and effect relationships leading to the identified accident experience.

In developing methods for the third step, we recognized that different economic analyses are used according to the point of view of the agency. For example, budget-oriented engineers generally think in terms of getting the most benefit obtainable from each annual budget. Thus, they use the familiar benefit-cost ratio as a measure of economy. On the other hand, economists are oriented toward solving the entire problem. They look at the big picture and use net benefit as a measure of economy.

These two approaches are in conflict. And the economist is correct in theory. In an environment with an undefined problem scope and apparently inadequate funds, should engineers continue to maximize benefits from the available funds? The response to this question in the users manual is no. The proposed system approach is directed toward a reconciliation of the two viewpoints.

The fourth step is the implementation of the improvements. The users manual provides guidance in (a) setting objectives and policies for the highway safety improvement program, (b) coordinating highway safety improvements with other projects, (c) assigning responsibilities for implementation, (d) formulating the program, (e) budgeting for the program, and (f) scheduling individual projects.

It is difficult to overstate the need to express objectives and agency policy. For this reason, a sample policy statement is presented in the users manual.

The last two steps in the system use conventional statistical analysis to determine the degree of success of individual projects and the program as a whole. The reports from these two steps are a major input to refining the other steps in the system. An example reporting system was designed for the users manual. It consists of

1. Hazardous location identification worksheet,
2. Probable accident cause analysis worksheet,

3. Potential improvement identification worksheet,
4. Improvement analysis worksheet,
5. Improvement evaluation worksheet,
6. Program priority listing worksheet, and
7. Improvement effect worksheet.

These reports contain the basic information needed to begin continuous evaluation and analysis. Suggested forms for these worksheets are shown in Figures 1 to 7. They should be modified as more refined procedures are developed and data requirements change.

CONCLUSIONS

Achieving greater utility from highway safety improvements will depend on more consistently applying the right improvements to the right problems, implementing increased numbers of improvements proven by evaluation to be effective, and altering the highway safety improvement program to meet broader goals and reflect evaluation findings. The highway safety improvement evaluation system is the way to accomplish these things. And it will work if top management takes the following steps:

1. Commit the organization to solve the problem of unsafe conditions on the highway system through systematic evaluation;

Figure 1.

Highway Safety Program
Documentation Record
FORM 101

Location Identification Code ________
Date ________
Prepared by ________

HAZARDOUS LOCATION IDENTIFICATION WORKSHEET
(One for each hazardous location)

LOCATION: ________________________________

Intersection ☐ Spot ☐ Section ☐ Section length ________

CATEGORY: Rural ☐ Urban ☐ 2-lane ☐ 4-lane Undivided ☐ Divided ☐ Freeway ☐

DESCRIPTION: Sketch on back of sheet ☐ See collision diagram ☐ No sketch or collision diagram drawn ☐

ACCIDENT EXPERIENCE:

Year						Total
From (date)						
To (date)						
AADT						Ave=
Exposure						
Total accidents						
Fatal accidents						
No. fatalities						
Injury accidents						
PDO accidents						
Accident rate						

SYSTEMWIDE AVE. RATE FOR CATEGORY: ________ Acc/MV

CRITICAL RATE FOR LOCATION: ________________

INDEX OF SEVERITY: ________________

COMMENTS: ________________

Figure 2.

Highway Safety Program
Documentation Record
FORM 102

Location Identification Code ________
Date ________
Prepared by ________

PROBABLE ACCIDENT CAUSE ANALYSIS WORKSHEET
(One for each hazardous location)

LOCATION: ____________________

COLLISION DIAGRAM: ☐ Attached ☐ Not drawn

ACCIDENT CHARACTERISTICS:

Type of Accident	Fatal	Injury	P.D.O.	Total	Percent
	No. of Accidents				
Head on					%
Rear end					%
Right angle					%
Side swipe					%
Fixed object					%
Overturned					%
					%
					%
Total					100%
Percent	%	%	%	100%	

CONDITIONS: No. of accidents

Time of day - 6:00 am - Noon ____ 6:00 pm - Midnight ____
Noon - 6:00 pm ____ Midnight - 6:00 am ____
Light conditions - Day ____ Night ____
Surface conditions - Dry ____ Wet ____ Snow or ice ____
Weather - Cloudy ____ Clear ____ Rain ____ Snow ____ Other ____
Other - ____________________

PROBABLE CONTRIBUTORY CAUSES: ____________________

Figure 3.

Highway Safety Program
Documentation Record
FORM 103

Location Identification Code ________
Date ________
Prepared by ________

POTENTIAL IMPROVEMENT IDENTIFICATION WORKSHEET
(One for each hazardous location)

LOCATION: ____________________

APPLICABLE IMPROVEMENTS:

IMPROVEMENT	EXPECTED RESULTS

SPECIAL COMMENTS: ____________________

Figure 4.

Highway Safety Program
Documentation Record
FORM 104

Location Identification Code ________
Date ________
Prepared by ________

IMPROVEMENT ANALYSIS WORKSHEET
(One for each potential improvement)

LOCATION: ____
____

IMPROVEMENT CODE: ________
IMPROVEMENT DESCRIPTION: ________________________
____

ESTIMATED SERVICE LIFE ______ YRS.
CURRENT 19 ___ AADT ________ EST. 19 ___ AADT ________
☐ Constant
☐ Increasing by _____ % annually
☐ Increasing by _____ VPD annually

ESTIMATED ACCIDENT REDUCTION:

By Types:				By Severity:	
Left turn	___ %	Skidding	___ %	Fatal	___ %
Head on	___ %	Wet pavement	___ %	Injuries	___ %
Rear end	___ %	Night	___ %	PDO	___ %
Right angle	___ %	RR crossing	___ %	Total	___ %
Side swipe	___ %	________	___ %		
Fixed object	___ %	________	___ %	All Accidents	
Lost control	___ %	________	___ %	___ %	

EQUIVALENT UNIFORM ANNUAL BENEFITS:
From accident reduction $ ________
From secondard benefits $ ________
Total $ ________

ESTIMATED COSTS:
Initial implementation $ ________
Annual operation and maintenance $ ________
Terminal value $ ________
Equivalent uniform annual cost $ ________

NET ANNUAL BENEFIT $ ________ BENEFIT/COST RATIO ________

SPECIAL COMMENTS: ____
____

Figure 5.

Highway Safety Program
Documentation Record
FORM 105

Location Identification Code ________
Date ________
Prepared by ________

IMPROVEMENT EVALUATION WORKSHEET
(One for each hazardous location)

LOCATION: ____
____

SUMMARY OF EVALUATIONS:

Code					
Description					
Initial cost	$	$	$	$	$
Annual cost	$	$	$	$	$
Salvage value	$	$	$	$	$
Service life	yrs	yrs	yrs	yrs	yrs
Equiv. uniform cost	$	$	$	$	$
Equiv. uniform benefits	$	$	$	$	$
Net annual benefits	$	$	$	$	$
Benefit/cost ratio					

REJECTED IMPROVEMENTS (and explanation): ________________
____
____

ELIGIBLE IMPROVEMENTS AND RANKING:

Code	Description	B/C Ratio

COMMENTS: ____
____

Figure 6.

Highway Safety Program
Documentation Record
FORM 107

Location Identification Code ____________
Date ____________
Prepared by ____________

IMPROVEMENT EFFECT WORKSHEET
(One for each hazardous location)

LOCATION: ____

IMPROVEMENT: ____

DATE IMPLEMENTED: ____________ SOURCE OF FUNDING: ____________

INITIAL COST: $ ____________ ANNUAL COST: $ ____________

BEFORE ACCIDENT EXPERIENCE: See Hazardous Location Identification Worksheet (Form 101)

AFTER ACCIDENT EXPERIENCE:

Year						Total
From (date)						
To (date)						
AADT						Ave=
Exposure						
Total accidents						
Fatal accidents						
No. fatalities						
Injury accidents						
PDO accidents						
Accident rate						Acc/MV

BEFORE ACCIDENT RATE: ____ Acc/MV AFTER ACCIDENT RATE: ____ Acc/MV

PERCENT ACCIDENT REDUCTION: 100 [____ - ____ / ____] = ______ %

ADJUSTED NUMBER OF BEFORE ACCIDENTS: ______ [( )( ) / ( )( )] = ______ Acc

SIGNIFICANT ACCIDENT REDUCTION AT ______ % LEVEL OF CONFIDENCE:

☐ POISSON DISTRIBUTION TEST ☐ POISSON COMPARISON OF MEANS TEST

COMMENTS ____

Figure 7.

Highway Safety Program
Documentation Record
FORM 106

Prepared by ____________
Date ____________

PROGRAM PRIORITY LISTING WORKSHEET
Year ______

Location		Improvement		Estimated Cost	B/C Ratio
Code	Description	Code	Description		

2. Organize personnel, procedures, and organizational units to implement the system;
3. Schedule implementation of the system; and
4. Coordinate development and operation of the system to ensure cooperation from all involved organization units.

The highway system is not likely to fade away soon as a major mode of transportation. Witness to this is the recent panic at the nation's gas stations when motorists found themselves being deprived of power to use their highways. But the recent gasoline shortage and subsequent reactions, such as lower speed limits and reduced traffic volumes, also showed us that the safety picture can change—all the more reason for evaluating and understanding the effects of highway safety improvements and organizing to make better use of existing facilities through such improvements.

Highway Safety Program: A Status Report

R. Clarke Bennett
Federal Highway Administration

The highway safety program consists of improvements to the driving environment that promote safety. The National Emphasis Program, which addresses the need for and methods to achieve highway safety, emphasizes (a) accurate identification of accident locations, (b) development of traffic engineering capability, (c) development of skid accident reduction programs, (d) use of uniform regulatory and warning signs, and (e) development of safer pedestrian crossings. This paper relates the states' activities in these areas.

The National Highway Safety Program has existed for more than 7 years. During that time there has been considerable effort by the states and the federal government to reduce the accident toll. We have made progress; both the fatality rate and the total number of fatalities have decreased nationally. Still there has been no dramatic breakthrough in solving the highway safety problem. This paper has two purposes: to report the present status of the states' efforts to implement their highway safety programs and to identify future trends in program activities.

The highway safety program consists of both capital and operational improvements to the driving environment that will correct hazards and promote safety. Highway-related standards have been established in four categories: identification and surveillance of accident locations; highway design, construction, and maintenance; traffic engineering services; and accommodation of pedestrians.

The majority of the activity associated with these program standards involves improvements to the highway system that are not so visible to the general public.

Even physical improvements, which are sometimes costly and require extensive planning and analysis before implementation often go unnoticed by the general public. However, they represent a vital segment of the total safety program. These activities provide the basic data needed to determine safety needs and effect improvements. Roadway and operational improvements have the immediate and long-lasting effect of reducing the number and severity of accidents. The effectiveness of these improvements can be measured directly.

To give a detailed status report of the more than 40 elements of the highway-related safety program standards would require too much space. We can, however, indicate states' efforts to implement programs in special areas of emphasis and point out trends in developing programs.

In the early years of the highway safety program, the states concentrated on identifying their needs and developing the resources needed to produce effective improvements. These efforts were necessary because, in many cases, the basic data did not exist and the 1966 Highway Safety Act specifically prohibited the use of highway safety funds for construction of highway improvements.

The states also identified some specific national problem areas. First, most states had at least a partial data base for identifying accident locations on the federal-aid system of highways. This was a requirement for safety improvement programs on the federal-aid system instituted by the Bureau of Public Roads in 1962. But highway safety applies to all roads, and accurate identification of accident locations off the Interstate and federal-aid primary systems was far from adequate. The states needed to improve their capability in this area. Because of limited resources, local jurisdictions had difficulty meeting the uniform safety standard requirements.

There was also an immediate need for uniform signing on local streets and highways. As a result, Congress indicated that the intent of the 1966 Highway Safety Act had been misinterpreted and that safety funds should be used for accident reference systems and regulatory and warning sign improvements off the federal-aid system. Previously, these activities had been considered ineligible under the 1966 Highway Safety Act.

FHWA's program goals at the national level are (a) to develop accident data and manpower to implement effective countermeasures, (b) to encourage the immediate implementation of proven high-payoff improvements, and (c) to provide the necessary resources to initiate an effective improvement program for all highways. The first steps toward these goals were to analyze national needs and to emphasize those program features oriented to the national goals. The National Emphasis Program, issued as part of the FHWA Highway Safety Program Management Guide in 1972, initiated the program. This program

1. Provides for the basic capability to identify problem areas and develop corrective measures for all streets and highways;
2. Emphasizes features of the highway standards that produce the greatest reductions in traffic deaths, injuries, and property damage in the shortest possible time; and
3. Establishes target dates for completing various elements of the emphasis program.

Some states already have implemented portions of the emphasis program and have moved on to other projects.

The five elements of the emphasis program furnish the necessary bases on which to develop the specific accident countermeasures needed and to provide professional and technical manpower to implement the program. The priorities also focus on identified nationwide highway safety problems.

The first element is accurate identification of accident locations. To develop effective highway accident countermeasures, jurisdictions must have accurate information on where accidents are occurring. Currently, many jurisdictions, particularly the smaller local agencies, do not have this basic resource. The emphasis program sets a target date of 1975 for a nationwide capability to accurately identify accident locations and to establish appropriate accident reference files.

The second is traffic engineering capability. One of the major hurdles to implementing FHWA safety standards has been the lack of professional manpower to analyze accident data and to develop, install, and evaluate accident countermeasure programs. This is especially true for many small local jurisdictions. Therefore, a high priority for the national safety program is the development of the traffic engineering capability to implement the standards. The emphasis program sets a target date of 1976 for developing needed expertise in all cities and counties throughout the United States.

The third element of the emphasis program is skid accident reduction programs. Skidding accidents contribute significantly to the nation's accident toll. About 20 percent of all accidents occur on wet pavements, and that pavement slipperiness causes these accidents has been clearly established by accident studies and highlighted in congressional hearings. The identification and correction of locations with an incidence of skidding accidents can lead to a dramatic reduction in wet weather crashes. Frequently, a corollary benefit is a reduction in dry weather accidents. The emphasis program sets 1976 as a target date for all states to develop such a program.

The fourth effort is uniform regulatory and warning signs. Standard 13 requires that all traffic control devices conform to the Manual on Uniform Traffic Control Devices. The 1971 revised manual requires a nationwide effort toward conformity as quickly as possible. In addition, there must be an extensive program to ensure public understanding of the new devices when the signing and marking system is improved. Uniformity in application and design and a public understanding of the new devices are essential to safe traffic operations. The emphasis program urges states and their local subdivisions to achieve conformity with the new manual by 1975.

The fifth is pedestrian crossing programs. This portion of the program concentrates on identifying hazardous pedestrian locations and applying traffic engineering measures and good highway design to develop a systematic plan for providing better pedestrian protection. Its importance is obvious, for pedestrians are the victims of one-fifth of highway fatalities.

Beginning with the 1973 Annual Work Program, states were expected to plan and program activities in these five areas. The states have responded well. Analysis of the states' 1974-78 comprehensive plan and their annual work programs indicates the following:

1. Only two states have not yet started work on a program to accurately identify accidents on all road systems. A majority of states already have this capability on the state highway system and are developing location systems for nonstate highways.
2. Forty-four states now have a program to increase the traffic engineering capability of medium-sized cities and counties. Thirty-seven states are working on a program to provide a traffic engineering capability to cities with populations between 25,000 and 50,000.
3. Thirty-six states are now performing a skid inventory on at least their state highway system, and 29 of these states have also started on a corrective program. All but two states are at least in the planning stage of the program.
4. All states are developing a plan to bring their traffic control devices into conformance. Already 14 states have started to replace nonconforming signs at the local level. Another 28 states will be starting shortly.
5. The area that has been the least active is the pedestrian safety program. Only 21 states now have a specific program to identify hazardous pedestrian crossings or to implement improvements. More work will be done in this area after the states begin operation of a basic accident identification system.

Clearly, the states are making substantial progress toward developing basic data needed for a systematic safety improvement program. This does not mean that the states are delaying needed safety improvements before a complete inventory is established. Last year the states spent almost $250 million on safety improvement projects on the federal-aid system alone. However, a better data base will assist the states in planning safety improvements statewide.

The emphasis program will provide the basis for the next step in the safety program: priority implementation. As reflected by the 1973 Highway Safety Act, the safety program is quickly moving into this phase. Congress, impressed by the effectiveness of many highway-related accident countermeasures, authorized a number of specific safety improvement programs and $975 million to assist in their implementation during the next 3 years. These specific programs are identified under five sections of the act.

1. Section 203 calls for the elimination of hazards at railroad-highway crossings on the federal-aid system.
2. Section 205 amends Chapter 1, Title 23, U.S.C., to establish a pavement marking program for both federal-aid and non-federal-aid highways.

3. Section 209 amends Chapter 1, Title 23, U.S.C., to provide for the correction of hazards at specific locations on the federal-aid system.

4. Section 210 amends Chapter 1, Title 23, U.S. C., to authorize the states to develop projects to reduce hazards caused by roadside obstacles on the federal-aid system.

5. Secticn 230 adds Section 405 to Chapter 4, Title 23, U.S.C., to authorize states to develop improvements for the elimination or correction of safety hazards on those highways not on any federal-aid system. (This section is significant because, for the first time, federal-aid funds can be used for safety improvements off the federal-aid system.)

Federal-aid funds will cover most of the cost for these programs. Section 205 will be completely federally funded. All other sections will be 90 percent federally funded. It should also be noted that, in every case, projects on the Interstate System have been excluded.

Among other unique features of the 1973 Highway Safety Act is a requirement for statewide comprehensive engineering surveys on all roads. Information from these surveys will be used to establish improvement priorities on the basis of potential payoff regardless of administrative control of the highway.

The 1973 legislation fills a critical gap that once existed under the safety program. Now the results of high accident location studies, roadside inventories, and other surveys conducted under the safety program can be implemented with the support of federal funds on a statewide basis.

In many ways the highway safety program is only just getting started. We are now entering the action phase; we have laid the basic groundwork and have launched a systematic attack on highway hazards. In the past, our efforts have had a relatively small impact on the nation's accident toll. But the application of corrective measures to the highway safety problem during this phase of the program should result in quick, significant, and long-lasting reductions in traffic deaths and injuries.

The Interstate System, the largest construction effort ever undertaken, is nearing completion. Its most outstanding characteristic is its low fatality rate. A major problem remains: We must make all existing highways safer. Finding the solution to this problem is our challenge for the future.

Manpower Allocation and Countermeasure Evaluation

Gregory L. Goodson
Arizona Department of Public Safety

As a complement to the theme of making better use of transportation facilities, this paper discusses making better use of existing resources to serve those transportation facilities. The resources addressed are manpower and budget allocations. Projects of the Arizona Department of Public Safety are presented.

Analogous to the theme of improved use of transportation facilities is the improved use of existing resources to serve these transportation facilities. The perspective of this paper is, Do not be afraid to innovate.

Traffic law enforcement agencies today are being confronted with increasing demands for service without related increases in manpower and budget allocations to cope with these demands. A new problem has created further complications: decreased fuel availability and increased fuel costs that exceed the funds appropriated for that purpose. Therefore, to cope with the needs for service, law enforcement agencies must seek more efficient and more effective use of the resources at their disposal.

Law enforcement administrators should seek innovative, perhaps unprecedented alternatives and be willing to experiment with alternatives. What is being suggested can be explained by a review of operational experiments of the Arizona Department of Public Safety. The department has a reputation for researching better ways to use both manpower and equipment to provide more and better service to the public. An early example of this was the 1956 Nogales Highway Project. The Nogales Highway, infamously known as Camino de la Muerte (highway of death), is 60 miles of US-89 from Tucson to Nogales. It had an unenviable record of disproportionately reducing the motorist population. The problem lay not so much in treacherous roadway as it did in treacherous drivers.

Highway patrolmen (now part of the department), in studying the accident data, felt they had an answer to the problem. To demonstrate their ideas, they borrowed officers from districts throughout the state for periods of temporary assignment on the Nogales Highway. Special emphasis patrols concentrated attention on the factors that had proved to be predominant in accident causation. Roadblocks were established along the highway so that all drivers could be stopped at strategic times and strategic places.

In the year preceding the project, the Nogales Highway had claimed 35 lives in 26 fatal accidents. In the year after the project, one fatal accident claimed three lives. Injury accidents were reduced by 50 percent and property-damage-only accidents by 29 percent. Continuing modifications in manpower assignment and countermeasure techniques based on project experience have kept accident rates so low that many people today have forgotten that the Nogales Highway was once the highway of death.

US-66 PROJECT

In summer of 1962 the Arizona Highway Patrol embarked on a similar project on the 381.5 miles of US-66 traversing northern Arizona. This highway varies from approximately 480 to 7,000 ft above sea level. Inasmuch as this is basically a corridor route through the state, during the summer months 80 to 85 percent of the traffic is composed of passenger vehicles with origins and destinations outside of the state.

Accident experience indicated that 45.8 percent of the accidents were noncollision, i.e., running off the roadway, overturning on the roadway, or similar incidents. Inattention and fatigue were predominant contributing factors. Ninety patrol officers, with the aid and cooperation of other public and private agencies, began a four-phase service and fact-finding project that included (a) an in-depth accident investigation, (b) an origin and destination survey, (c) a medical investigation, and (d) an adequate enforcement program.

Surveys indicated that most cars traveled from 301 to 1,000 miles between overnight stops; the mean distance between bed rest was 625 miles. The mean trip length was 1,800 miles, and the average driver was 36.9 years old. Further refinement of data, of little significance for this discussion, was made.

Enforcement efforts concentrated on the most significant accident-causing factors at the most frequent times and locations of accidents. During July and August officers made 20,030 written contacts. In addition, officers assisted 8,001 motorists in need of mechanical or other aid. During these 2 months in the prior year, 218 persons had been injured and 31 killed on this stretch of highway. During the project period, only 100 injuries and 12 deaths were recorded, reductions of 55 and 61 percent respectively.

JOINT ENGINEERING AND ENFORCEMENT PROJECT

As a followup to Arizona's US-66 project, the U.S. Bureau of Public Roads in the summer of 1964 coordinated a similar project in the seven states traversed by US-66 from Los Angeles, California, to Joliet, Illinois. This project primarily involved a special study of the causes and characteristics of single-vehicle accidents and an analysis of the types and volume of police services for motorists and the time spent by officers on such services. Supplementary were studies of police manpower, vehicle speeds, median crossover (for enforcement and service purposes), and traffic information and control signing needs.

Of the 850 single-vehicle accidents investigated, in 781 or 92 percent of these at least one wheel left the pavement or area of normal travel and about four-fifths of these occurred on straight highway sections. Falling asleep was a common factor. Of some significance, perhaps particularly today, was the fact that small cars (under 2,000 lb) were involved 3.5 times as often as standard automobiles, and cars towing trailers were involved 4.5 times more than those without trailers.

Other data developed in both US-66 studies provided information to aid in determining future manpower needs, service needs, and ideas for new accident countermeasures. In addition, the need for engineering improvements to aid both the motorist and those providing service to the motorist was determined.

IMPACT

For 3 years, manpower and countermeasure efforts were concentrated in areas of the most need throughout the state. Certain sections of highway were in practice designated as respond-only areas. That is, no active patrol was conducted on those portions of highway, and officers entered the area only in response to accidents or other calls for service. Careful records were maintained of officers' hours, times they were away from their assigned beats, and their reasons for leaving their beats. These data were used to illustrate to the department and subsequently to legislative leaders where and how additional manpower needed to be assigned.

The initial manpower increases allocated as a result of these efforts were assigned for purposes of further refinements of the already-begun manpower and countermeasure research. Practical limitations on recruiting, testing, and training prohibited putting the desired number of officers to work all at the same time; however, as a demonstration project those initially hired were assigned to bring specific patrol districts up to desired strength. For some continuity and control the southern corridor route was selected as the project area. This is made up of Interstate 8 from the California boundary to its junction with Interstate 10 near Casa Grande, then continuing on Interstate 10 to the New Mexico boundary. This new project, begun in January 1973 and still under way, is called the interstate model patrol and accident control technique (IMPACT).

Along with the manpower increases, district commanders were given extensive latitude in the use of personnel, equipment, and accident countermeasure techniques. The issuance of tickets or warnings was of no concern. The primary criterion was measured in terms of accident reduction.

In the Yuma district, which contains approximately 150 miles of I-8 plus frontage roads and connecting highways, innovation precluded precedent in the IMPACT project.

In the Yuma district, which has many miles of relatively straight, level roadway and considerable distance between major population centers, a large number of fatigue-related accidents occurred. Before IMPACT, officers would frequently stop drivers because of some form of fatigue-related improper driving and often issued traffic tickets or warnings for the specific violations committed.

It was hypothesized that issuing tickets sometimes motivated the drivers to stop and rest but, more often, aroused them temporarily or even angered them, and they continued until they again became drowsy and possibly became involved in an accident farther down the roadway. Officers were counseled to approach the problem more toward the prevention goal than toward enforcement. The goal was to motivate the fatigued driver to stop and rest by whatever means necessary. We preferred that the officer not write a ticket, except as a last resort when drivers flagrantly disregarded the hazardous potential of their condition. Final statistical analyses have not been made, but observations by officers involved tend to indicate that this philosophy is producing some positive effects.

A second break with precedent involved which roadways the officers patrolled. Although having statewide jurisdiction, highway patrol officers in Arizona were routinely responsible for patrolling only the state and federal highway system. The Yuma area has a unique network of roadways—crossing and paralleling and zigzagging state and federal highways that access and interconnect numerous well-traveled roadways in the county and city road systems. This area also has its share of problems with motorists driving under the influence of alcohol. A local origin-destination survey was conducted among drinking drivers involved in accidents or stopped for violations on the highways. From this information, we determined most probable routes of travel of problem drinkers en route to the state highway system, and officers were directed to perform active patrol and enforcement on those routes, even if off the state system, in an effort to counter the overall accident problem on the system. Because of manpower limitations, officers were not actually assigned to these county or city roads, but they performed selective enforcement while traversing selected routes between points on the state system. The efforts shown resulted in increases in DWI arrests and some reduction of DWI accidents.

OTHER CONCERNS

In coping with accidents or incidents, law enforcement agencies must consider the goals of both rendering effective and timely aid and restoring normal flow of traffic or providing temporary alternative facilities. Situations vary from the need for assisting an ill or disabled motorist on a lonely stretch of highway to the more major concern of finding food, lodging, or medical aid for hundreds of motorists temporarily stranded by natural disasters such as a sudden, unexpected severe snowstorm.

Officers may routinely patrol in nonpatrol vehicles such as four-wheel drive pickups, snow scooters, snow cats, airplanes, helicopters, and even fully equipped ground ambulances. Even these, however, may not be sufficient for all emergencies, and agencies should make some contingency plans for potential emergencies.

Contingency planning should include the development of a list of resources available from other government and private agencies and development of a cooperative working relationship with key personnel in those agencies. Conflict can be avoided at a critical time by agreeing on duties and responsibilities during an emergency in advance. Clarification of responsibilities may be made through cooperative agreements or may even involve legislation. What is important is that the essential planning be accomplished before the emergency occurs.

However comprehensive the planning and however sophisticated your resources, situations that you have not been specifically prepared for will still occur. When this happens, evaluate the problem, assess the resources available, and innovate.

REFERENCES

1. Median Crossover Study. Arizona Highway Patrol, January 1965.
2. Enforcement Project Route US-66: Manpower Study. Arizona Highway Patrol, 1964.
3. Final Report on U.S. Route 66 Across Northern Arizona. Arizona Transportation and Traffic Institute, University of Arizona, Report 9, June 1964.
4. Fannin, P. Operation 66. Police Chief, IACP, June 1964.
5. Summary of Results of Spot Survey of Selected Law Enforcement Agencies. U.S. Department of Justice, Energy Report 1, March 1, 1974.
6. Report of Operation 66: Joint Engineering Enforcement Project. U.S. Department of Commerce, November 1966.
7. Wallace, J. S. Arizona Takes Pulse of Famed Artery. Traffic Safety, National Safety Council, October 1963.

The Jacksonville Emergency Medical System: A Model for the Seventies

John M. Waters
Jacksonville Department of Public Safety

The Jacksonville, Florida, emergency medical system, one of the most sophisticated systems in the nation, has contributed significantly to the reduction of the death rate of automobile accident victims. The reduction in the death rate of trauma victims has been 38 percent. Training of the emergency medical technicians is described, and data on death rates and survival rates of heart attack victims are presented. The Jacksonville emergency medical system has been a model that many cities and states are seeking to duplicate.

Jacksonville, Florida, with a population of 532,000, has developed an emergency medical system that is considered by many experts to be the finest in the country. Yet, less than 4 years ago the ambulance situation was chaotic; numerous underfinanced private companies and funeral homes competed for business and offered a very poor quality of service. As a result of a campaign by the news media, rising public opinion, and the inauguration of a new consolidated county-city government, the situation has been corrected. This was made possible by the cooperation of government and organized medicine, stiff regulatory provisions for private ambulance operations, and a take-over by the Fire Department of all emergency ambulance service.

Nearly all calls for emergency medical service come into the Fire Department Emergency Operation Center from either home phones or special emergency phones located on street corners. This center has direct phone lines and radio communications with all hospital emergency departments in the city and with the switchboard of the County Medical Society. Punched time cards are kept on each run as well as 24-hour tape records of all communications. In 1969-70, nine rescue squads were distributed throughout the city so that the average running time to the scene was 4.2 min. Over the past 2 years as the population has expanded toward the city limits, the average running time has risen to more than 7 min. However, the first trained help arrives at the scene in an average of 4 to 5 min.

The ambulances have a roomy box type of body mounted on a 1-ton chassis and are fully equipped as mobile emergency rooms. At 60,000-mile intervals, the body is lifted off, the worn chassis is replaced by a new one, and we have a "new" ambulance. Cost for a 10-year period is approximately half that of a limousine type of ambulance.

Five of the 11 rescue ambulances are based at hospitals to afford in-hospital training to the rescue personnel and to provide better ambulance-hospital coordination. The crews are rotated at periodic intervals so that each man spends an average of 5 months each year based at the hospitals. Cooperation between the ambulance and hospital emergency staffs is extremely close.

Training of emergency medical technicians (EMTs) is extensive and starts with the first-aid course given all recruits at the Fire School. In addition to the 20-hour American Red Cross course, training for all recruits includes the 80-hour Department of Transportation Emergency Medical Care course. Advanced lectures in cardiology are then given by physicians. More than 20 EMTs have participated in the program. During in-hospital training, EMTs spend hundreds of on-duty hours in hospital emergency departments. Practical work includes observing childbirths and autopsies, giving EKGs, drawing blood, and working as a part of the cardiac resuscitation team. On-duty formalized exercises and frequent written tests are required, as are periodic lecture series after hours by physicians. Twelve EMTs are enrolled at Florida Junior College in a 61-semester-hour associate arts program in emergency medical technology, one of the first of its kind in the nation. One result of the extensive training program has been the ability of the city to obtain malpractice insurance for all EMTs in the amount of $5,000,000 for an annual total premium of only $7,500.

The medical training program is supervised by the Chief Fire Surgeon who is a prominent cardiologist and who serves at a pay of $1 per year. He is assisted by 10 assistant fire surgeons, all $1 a year volunteers, who act as advisors to the squads, help in the instruction, and monitor their performance to ensure quality control. Each fire surgeon has a two-way fire department radio in his private car and is on call to answer serious emergencies whenever he is on the streets. The dedication of these men and women reflects the outlook of our entire medical community on emergency care.

The most rewarding results have been in the reduction of the automobile accident death rate. After the EMT arrives at the scene, he first checks the victim's airway and stops any bleeding before he attempts to move the victim. If serious injury is suspected, an intravenous blood expander such as lactated Ringer's solution is given to combat shock. Medical advice is instantly available from the hospitals by direct ambulance-hospital radio communications.

The EMT at the scene, with the radioed advice of the hospital physician, makes diagnoses and renders treatment, not merely first aid. The procedure has the full blessing of the great majority of our physicians and is sanctioned by an amendment to the Florida Medical Practices Act passed by the legislature in 1970 at the urging of the Florida Medical Association. That such an advance has occurred in less than 5 years is a tribute both to the quality of emergency care rendered by the rescue branch and to the farsightedness of our medical professionals. Whereas IVs can be initiated by the EMT, the administration of drugs must be prescribed by radio or phone by a physician.

After the victim's condition is stabilized, great emphasis is placed on proper extrication and use of backboards. Patients are kept on these backboards until after they are X-rayed. Spare backboards are provided at the hospitals so that the ambulance can quickly return to service. Less than 5 percent of the runs from the scene of the accident to the hospital are done at high speed with light and siren. En route the patient is given antishock therapy and monitored on an oscilloscope if required. The hospital is warned ahead of time of the nature, vital signs, and the estimated time of arrival of any serious case. In severe cases the patient can be left wired to the oscilloscope until a physician examines him and approves his transfer to the emergency room. A form detailing the treatment of the case is filled out on each patient.

Because of the advance notice given the emergency departments and the fact that four of them have emergency practice groups present at all times (a fifth is a large teaching hospital with full extra-duty manning), little delay is encountered in beginning treatment at the hospital on serious casualties.

The results in reducing the death rate in automobile trauma victims have been spectacular. Before this program was initiated, medical authorities predicted that the provision of rapid and effective treatment on scene could reduce automobile trauma deaths by 20 percent. How has the Jacksonville system done?

In 1968, when we initiated the program, there were 15,846 accidents involving approximately 8,669 injuries and 139 deaths. In 1971, the number of accidents yearly increased to an estimated 22,500 (based on 11 months of data) involving an estimated 12,380 injuries, but the deaths dropped to 117! The save rate of injured victims was 99 percent. Deaths per thousand accidents were reduced from 8.27 to only 5.2: a reduction of 38 percent or nearly twice that predicted by medical experts. In contrast, the number of automobile deaths in the state of Florida rose 7.8 percent in 1971.

Much of this reduction has to be attributed to our advanced emergency medical system. Of course, post-1967 cars, with new safety features such as collapsible steering columns, breakout windshields, and padded interiors, have had their effect as has a stepped-up highway law enforcement system; yet in the surrounding seven rural counties, where new cars are also involved, the death rate per thousand accidents was 23.4 in 1970—four times greater than Jacksonville's present rate. These counties lack sophisticated emergency care systems.

The practice of on-scene diagnosis and advanced treatment has had another marked effect. Prior to 1968, nearly all victims were placed in an ambulance and rushed at high speed to a hospital, where many turned out to be suffering from only minor injuries. In 1971, the rescue units responded to 18,204 calls for help. Yet, they transported only 8,427 to hospitals; the rest either required no treatment, were treated at the scene and released, or were allowed to go to a doctor or hospital by automobile. The reduction in load on the emergency departments and rescue branch is significant.

Whereas 117 people died in 1971 in Jacksonville in automobile accidents, more than 1,600 died of heart disease. Nationally, a large percentage of these deaths occur outside the hospitals. The overall death rate on initial myocardial infarction is approximately 25 percent; but, for patients treated in the intensive care units of hospitals, the rate drops to approximately 18 percent. It is shocking that 60 percent of those who die of a heart attack in the United States die outside the hospital without a doctor in attendance. To increase the rate of survival outside the hospital, our rescue ambulances have been fitted with special cardiac equipment and rescue crewmen have been given special training in treatment of cardiac cases, including defibrillation, cardiopulmonary resuscitation, and, when prescribed by a physician, the administration of drugs. Two-way radio communications can be maintained with a physician in the emergency room, and all of the ambulances are equipped to telemeter EKGs by radio. On all suspected heart attack cases, two ambulances (or an ambulance and a fire engine) are dispatched to provide the manpower necessary to carry out life-sustaining procedures.

Unfortunately, statistics on heart attacks are not so complete as those on automobile trauma, but we are now preparing a study of results for 1971. However, a brief overview of the suspected heart cases for January 1972 is enlightening. It must be emphasized that the final diagnostic results have not been received from the hospitals and that the data are for those persons who had symptoms that were strongly suggestive of cardiovascular distress.

In January 1972, 348 dispatches were made to calls from citizens who described symptoms of breathing difficulty, chest pains, and the like. On 165 of these calls, symptoms were so suggestive that two units were dispatched, the second unit usually being an engine or ladder company that is trained and equipped to help the rescue unit. Transportation and treatment were required for 158 calls. Of these, 113 had suggestive cardiovascular symptoms, and the other 45 were suffering from respiratory distress or various other illnesses.

Of those with heart symptoms, six were dead before arrival of the first unit, and resuscitation was not attempted. The other 107 were treated and transported. Eleven of these suffered heart or breathing arrest while in the hands of the rescue unit and were given cardiopulmonary resuscitation or drugs. Three of these 11 died before arrival at the emergency department, whereas eight who had suffered arrest or arrythmia were delivered viable. Of the 107 suggestive heart patients transported, 97.2 percent were delivered viable to the emergency department! We must caution the reader that past experience has shown that perhaps half of these in the final analysis will prove not to be heart patients; even so, the safe delivery rate of true heart patients would be over 94 percent.

The introduction of the special intensive care rescue units has led one prominent thoracic surgeon to remark, "Jacksonville is the safest place in the United States to have a heart attack."

As the citizens of the city became aware of the immediate availability of help when they were in distress, the calls and load steadily increased and, consequently, so did the time required by the rescue unit on arrival at the scene. The most obvious solution was to add more rescue units. However, the expansion of the city also demanded that we increase the number of engine companies, and four more ladder companies were recommended by the underwriters. Such an augmentation was beyond our financial capability. We came up with what we feel is an innovative solution.

Ladder companies are primarily used to provide special rescue and extrication equipment, lighting and ventilation gear, and manpower. The long aerial ladders are used in only 2 percent of all calls, especially in suburban areas. A pumper, if it arrives promptly, can usually extinguish a fire with the 500 gallons of water carried in its integral tank without having to hook onto a hydrant. An accident or heart victim, if he is treated promptly at the scene, can wait the additional 5 minutes or so until a rescue ambulance arrives.

We, therefore, converted four excess pumpers as triple threat units, manned by three men each, and designated them as quick response squads (QRS). By removing the bulky large hose, we provided space to stow all the special rescue equipment carried by ladder companies. Racks were made for 35- and 14-ft aluminum ladders, which are adequate for most suburban fires. Five hundred feet of $1\frac{1}{2}$-in. preconnected hose is carried for fire fighting, utilizing the water carried. Also, the QRS vehicles were given a full complement of first-aid equipment and were painted a distinctive lemon-yellow color. The men for these units were given a special 7-week course at the fire school and 120 hours of medical training. These units can act as an engine company, a ladder truck, or a rescue unit. Although they cannot transport patients, they can initiate treatment and often have the patient ready to go when the rescue ambulance arrives. The total conversion cost for each unit was less than $3,000. Manpower was obtained by reducing the manning of the urban companies, which we compensated for by adding one more engine company to each assignment, resulting in more manpower actually available at a fire. The assignment of four QRS units to the suburbs has eased the load on the rescue units and has enhanced their ability to extricate and treat victims.

Thirteen rescue-oriented units were thus available to aid citizens in distress, far in excess of the resources usually available in a city of our population. However, we were aware of the fact that, in a city of 840 square miles, we still had time and distance problems. We turned our attention to the fire fighting companies (combat), of which we had 49 in 41 stations, strategically distributed throughout the city. Why not use these, which were often nearest an accident, to render aid until a rescue unit arrived? Firemen have historically been underutilized, yet they must be ready 24 hours a day in case of an alarm. In 1971, we adopted a formal policy of dispatching a combat company along with a rescue squad on all serious medical cases. Trained in comprehensive first aid and carrying first-aid kits and oxygen, they can initiate treatment until a rescue unit arrives with even more sophisticated equipment and EMTs, who can transport the victim if required.

In the first 6 months of 1973, nearly half of all responses by combat companies involved medical emergencies rather than fire. In a number of cases, their prompt arrival definitely saved lives. In taking this step, we have placed the full resources of our Fire Department and its thousand men at the disposal of the citizens who are stricken, rather than just the 120 plus men in rescue and the QRS units. It is a practice that is here to stay and one that every citizen in the nation should consider.

A nominal charge of $22.50 is made by the rescue branch when a patient is transported; no charge is made if he is treated on scene and released. The charge is more a deterrent to needless calls than a revenue source. The cost to the citizens of Jacksonville for this superb rescue service is $1.56 yearly per capita in tax dollars. It is one tax expenditure that brings no public criticism. In nearly 6 years of operation, only 10 written complaints have been made about the service.

Despite the great strides we have made, there are still problems. Here, as nationally, we are continually faced with delays of 1 to 6 hours after the onset of symptoms before the heart attack victim seeks help. All too often it is then too late. The best team in the world can do nothing for a man who has been dead 10 minutes.

Our rescue service has probably been more publicized than any in the nation, and we have mailed emergency telephone number stickers to all citizens. Yet, in January 1972, 23 percent of the people seeking emergency medical rescue service called the police; the fire control center, which handles rescue, got the message secondhand. Furthermore, 10 percent of fire alarm calls were actually made to the police department.

We are deeply concerned that a 1971 study of heart attack patients admitted to five of our city hospitals showed that 67 percent of them drove to the hospital by automobile instead of calling the rescue branch—an extremely dangerous practice that costs a number of lives. Yet the American Heart Association advises in case of a heart attack first contacting your doctor, and, if unable to do that, going to the hospital immediately. This may be practical advice in some areas where ambulance service is poor or where doctors can be reached immediately, but it is not in Jacksonville and we are advising our citizens to first call rescue. The rescue branch and Fire Department respond immediately and notify the patient's doctor if requested.

Because of our highly effective rescue system and the network of high-speed highways throughout the city, we have few occasions to use medical evacuation helicopters within the city limits. Three large naval air stations have seven helicopters available and are most cooperative in providing them to us on short notice for evacuating patients from the surrounding rural counties. The Army National Guard unit has more than 15 modern helicopters available and will provide them when crews are available during weekend drills. Heliports are available at three major hospitals in the city.

Under our agreement with the Navy, our rescue branch provides the EMTs and medical equipment and the Navy provides the helicopter and crew. In 1971, a number of successful medical evacuations were made following a request from one of our surrounding rural counties. A critical test came in February 1971, when a tremendous explosion at a Thiokol plant in southeast Georgia killed or injured more than 80 people. In an isolated area, with only three doctors in the county, outside aid was the only hope. The Jacksonville Fire Control Center took control of the operation and dispatched ground rescue units, four Navy helicopters, a police helicopter, and numerous doctors and EMTs. Patients were treated at the scene and loaded onto helicopters for rapid transportation to Jacksonville hospitals, where surgical teams were mobilized and waiting. Twenty-seven critical patients were brought in by land and air, and 26 survived. When the last patient arrived less than 2 hours after the first alert, the northside hospital complex still had surgical and bed capabilities to handle 80 more patients.

The major urban centers with their great hospitals must extend help to neighboring rural areas, and the helicopter is an ideal transportation mode for such medical evacuations. Because of the tremendous operating expense of large helicopters capable of carrying prone patients, they must be provided by the military rather than the states and cities. Utilizing the communication, control, and medical facilities of a major city system, combined with a military airlift capability, we have the know-how today to quickly set up an areawide model. During the past year, over 70 groups from other cities and states have visited Jacksonville to observe our system, and its components have been copied and implemented in many other areas. Our capability is being extended into the surrounding rural areas as a model of an urban-rural area EMS system, and we have $3 million of federal funding for such a project. The system will be fully operational by 1974.

In a message to Congress in 1971, the president called for a new approach to providing care to those who live in remote rural areas. We in the Jacksonville area are ready with a new approach, and we are confident that it will continue to stimulate other areas in a common effort to avert the appalling and needless loss of life suffered yearly in this country because of a grossly inadequate emergency medical care system.

Sponsorship of This Special Report

THE Transportation Research Board is an agency of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's purpose is to stimulate research concerning the nature and performance of transportation systems, to disseminate information that the research produces, and to encourage the application of appropriate research findings. The Board's program is carried out by more than 150 committees and task forces composed of more than 1,800 administrators, engineers, social scientists, and educators who serve without compensation. The program is supported by state transportation and highway departments, the U.S. Department of Transportation, and other organizations interested in the development of transportation.

The Transportation Research Board operates within the Commission on Sociotechnical Systems of the National Research Council. The Council was organized in 1916 at the request of President Woodrow Wilson as an agency of the National Academy of Sciences to enable the broad community of scientists and engineers to associate their efforts with those of the Academy membership. Members of the Council are appointed by the president of the Academy and are drawn from academic, industrial, and governmental organizations throughout the United States.

The National Academy of Sciences was established by a congressional act of incorporation signed by President Abraham Lincoln on March 3, 1863, to further science and its use for the general welfare by bringing together the most qualified individuals to deal with scientific and technological problems of broad significance. It is a private, honorary organization of more than 1,000 scientists elected on the basis of outstanding contributions to knowledge and is supported by private and public funds. Under the terms of its congressional charter, the Academy is called upon to act as an official–yet independent–advisor to the federal government in any matter of science and technology, although it is not a government agency and its activities are not limited to those on behalf of the government.

To share in the task of furthering science and engineering and of advising the federal government, the National Academy of Engineering was established on December 5, 1964, under the authority of the act of incorporation of the National Academy of Sciences. Its advisory activities are closely coordinated with those of the National Academy of Sciences, but it is independent and autonomous in its organization and election of members.