HIGHWAY RESEARCH BOARD
Bulletin 76

Origin and Destination Surveys
METHODS AND COSTS

National Academy of Sciences—
National Research Council
publication 269

HIGHWAY RESEARCH BOARD

1953

2101 Constitution Avenue, Washington 25, D. C.

HIGHWAY RESEARCH BOARD
Bulletin 76

Origin and Destination Surveys

METHODS AND COSTS

PRESENTED AT THE

Thirty-Second Annual Meeting

January 13-16, 1953

1953
Washington, D.C.

TRAFFIC AND OPERATIONS DEPARTMENT

Wilbur S. Smith, Chairman
Associate Director, Bureau of Highway Traffic
Yale University

COMMITTEE ON ORIGIN AND DESTINATION SURVEYS

D. Grant Mickle, Chairman
Director, Traffic Engineering Division,
Automotive Safety Foundation

John T. Lynch, Secretary
Chief, Highway Planning Survey Section
Highway Transport Research Branch, Bureau of Public Roads

Wroe Alderson, Alderson and Sessions, Marketing and Management Counsel, Philadelphia, Pennsylvania

Maynard A. Bebee, Principal Civil Engineer, Bureau of Highway Planning, New York Department of Public Works

John B. Boddie, Jr., Office of Business Economics, U. S. Department of Commerce

Lloyd M. Braff, Traffic Engineer, Dallas, Texas

James S. Burch, Engineer of Statistics and Planning, North Carolina State Highway and Public Works Commission

Nathan Cherniack, Economist, The Port of New York Authority

Gordon K. Gravelle, Transportation Engineer, Automotive Safety Foundation

Stuart Hawley[1], Director, Travel Department, The Texas Company, New York, New York

Howard Ilgner, Traffic Engineer and Superintendent, Bureau of Electrical Services, Milwaukee, Wisconsin

Roy E. Jorgensen, Engineering Counsel, National Highway Users Conference

Harry W. Lochner, H. W. Lochner and Company, Chicago, Illinois

C. H. Makeever, Ohio Turnpike Commission

Burton Marye, Jr., Assistant Chief Engineer, Virginia Department of Highways

C. A. Rothrock, State Planning Engineer, State Road Commission of West Virginia

I. S. Shattuck, Planning and Traffic Consultant, Wayzata, Minnesota

Lawrence S. Waterbury, General Partner, Parsons, Brinckerhoff, Hall, and Macdonald, New York, New York

[1]Deceased April, 1953

Contents

Dallas Traffic-Survey Methods and Cost Analysis

H. P. Stockton, Jr. 1

Postcard Method of Obtaining Origin and Destination of Traffic and Comparison with Roadside-Interview Method

Robley Winfrey 10

Sampling Methods for Roadside Interviewing

Irwin Miller, P. E. Irick, H. L. Michael, R. M. Brown 31

Sampling Procedures for Roadside Interviews in Origin-and-Destination Traffic Surveys

Robley Winfrey and Robert J. Hansen 53

Dallas Traffic-Survey Methods and Cost Analysis

H. P. STOCKTON, JR., Texas Highway Planning Survey
Texas Highway Department

The rapid growth in population of the metropolitan area of Dallas, with the corresponding growth in the traffic of the area, has resulted in the need for information to guide planning for immediate relief and future development. In 1950 and 1951 a comprehensive traffic survey embracing both external and internal traffic surveys of the 215-sq.-mi. metropolitan area supplemented by external origin-and-destination surveys at six outlying incorporated cities, and a parking survey of the central business district of Dallas was undertaken by the city and the state in cooperation with the Bureau of Public Roads. Current land-use data were obtained while selecting the internal sample.

Provision was made for recording detailed costs expressed in work units applicable to future projects of this nature. A brief description of the methods used in the survey with recorded unit and total cost data are expressed.

Application of the findings in the solution of the traffic problem are outlined. Land-use data found a place in planning for future traffic development.

● DURING the latter part of 1950 and early 1951, comprehensive traffic and parking surveys were made in the Dallas metropolitan area. In January 1951 a progress report on these surveys was presented to the Highway Research Board by D. K. Shepard. This paper is a final report, treating of the methods employed in the surveys and devoting major attention to the cost analysis of the project with the thought of furnishing data which may be used for planning and estimating similar undertakings.

Dallas, like many other cities, emerged from the wartime suspension of major street and highway improvements with the realization that drastic and immediate action was advisable to provide for the formation of a new and modern master plan.

By 1950, Dallas had exceeded both the population and vehicle registrations predicted for 1970. The population had increased from 295,000 persons in 1940 to over 430,000 persons in 1950. The area of the city proper had been extended from 50 sq. mi. in 1944 to 122 sq.mi. in 1950, with other annexations planned to bring the area to more than 150 sq. mi. in 1951.

To be able to formulate a new and modern master plan providing for the existing traffic volumes and to be able to plan wisely for future growth, it was evident that current and comprehensive traffic data were needed. The selection of an east-west expressway route was being studied, and this fact added further impetus to the demand for the data. With these demands in mind, the City of Dallas requested and received the cooperation of Dallas County, the Texas Highway Department, and the Bureau of Public Roads in a traffic survey of the entire metropolitan area.

In preliminary conferences, representatives of the Bureau of Public Roads, Texas Highway Department, and the County and City of Dallas outlined the scope of the survey. The extent of the survey was limited to the estimated cost of $100,000, and all available cost data were brought into use to outline a project which would provide the greatest value within this limit. The need of more extensive cost data was apparent, and it is the intent of this paper to place the information gained in form for such use.

The metropolitan area selected included an area of 215 sq. mi. with a population of 517,035. The area included nine incorporated cities which are completely surrounded by the Dallas incorporation.

In addition to this coverage of the metropolitan area, external origin-destination surveys were conducted at six outlying incorporated cities to obtain data for a county-wide, major thoroughfare plan to supplement the Dallas thoroughfare plan.

The population covered in the survey was 517,035 persons in the internal survey and 20,066 persons in the outlying communities, or a total of 537,101 persons - 88 percent of the total population of Dallas County.

The administration of the survey was assigned to the Traffic Section of the Highway Planning Survey, which organization furnished personnel to serve as project manager; supervisors of internal, external, and parking surveys; and chief accountant.

All other personnel were employed locally or borrowed from various city departments. For various phases of the operations, the city traffic engineer and city plan engineer assigned personnel to the survey; however, the unit cost records given in this report reflect this expense based on regular rates of pay for each item of work accomplished. Such personnel as the city furnished free of cost to the project are also reflected in the unit costs at the rates of pay charged for similar work.

The rates of pay for personnel were as follows: $1 per hour for interviewers, samplers, office clerks and coders; $1.25 to $1.50 per hour for crew leaders; supervisors for each phase of the survey were paid on a monthly rate at $225 to $250 per month. In general all personnel, other than supervisory, were obtained from the Texas Employment Commission. One exception to this was the manner of obtaining home interviewers. Contact was made with district census director, who furnished a list of names of the census enumerators in the census which had just been completed. Excellent personnel were obtained by this method.

A detailed accounting system was established to be able to determine actual costs and man-hour requirements for each unit of all phases of the survey. Time slips showing actual man-hours used on each phase of the work were submitted daily by each supervisor. These costs, along with material and supply costs, were posted to each operation daily using a control ledger system and accounts journal. This not only kept an accurate record of costs and man-hours for final reports, but made possible periodic summaries of accumulated costs for each unit of operation. The costs reflected include administration and administrative overhead of the field office and field operation. They do not include administrative costs of the higher echelon and should be so interpreted. Excellent weather conditions throughout the period of the field survey served to hold costs to a minimum.

Office space was available for 60 people, and this space was used to the maximum during the peak period of operations.

THE SURVEY

The study was divided into three phases: (1) the internal survey, consisting of (a) sampling, with related land use, (b) home interviews, (c) truck interviews, and (d) taxi interviews: (2) the external survey; and (3) the parking survey.

TABLE I

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

COST REPORT SUMMARY - MAY 1, 1952

	Total Man Hours	Total Project Costs
ADMINISTRATION	7192	$ 11574 22
EXTERNAL SURVEY	14323	17183 82
INTERNAL SURVEY	24089	27714 17
O & D REPORT	3419	6920 62
DALLAS COUNTY O & D	3392	4347 89
PARKING SURVEY	23189	29585 45
GRAND TOTAL	75605	$ 97326 17

INTERNAL SURVEY

Sampling

A 5-percent sample of dwelling units was used for the entire metropolitan area. In selecting the sample, it was found that the Sanborn maps, in addition to being obsolete, covered a very small portion of the area to be surveyed. The City Plan Department had available land-use maps drawn to scale 1 in. to 400 ft. which had been maintained approximately up to date. It was determined that the most-accurate way of obtaining a sample was to use these land-use maps as a guide and actually travel in the field, checking the land-use maps against actual land use, and at the same time, bringing the maps up to date by notations relative to new occupancy or abandonment of dwelling units or other structures in the area. The samples were selected by two-man teams in automobiles. The driver called the type of occupancy of each dwelling to his recorder, who made the proper notations on the map. The driver kept count of the total dwelling units and called the house number of the sample to his

recorder, who recorded the addresses of the samples on a form provided for this purpose. In this manner, the sample was selected in every instance from a field inspection, and the City Plan Department, at the same time, obtained data for an up-to-date land-use map not only of Dallas but of the entire metropolitan area.

sample selection required 3,786 man-hours at a total cost of $4,509.80. The unit cost of sample selection was 52 cents per sample for field work and 17 cents for mapping and supplies, or a total of 69 cents per sample.

It is believed that the value of the current land-use information for both present

TABLE 2

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

ADMINISTRATIVE COST DATA

ADMINISTRATION	Date Started	Date Completed 2/	Total Man Hours	Total Costs
OFFICE OPERATIONS				
Supervision	10-2-50	3-7-51	1617	$ 4564.88
Accounting	10-2-50	12-1-51	1100	1343.45
Clerical	10-2-50	3-7-51	1118	1246.78
Typing	10-22-50	3-1-51	500	500.00
Drafting	11-12-50	3-3-51	237	265.41
Office Supplies			-	603.15
Postage & Freight			-	108.92
Office Equipment Rental			-	420.00
Total Office Operations			4572	9052.59
OFFICE BUILDING MAINTENANCE				
Labor 1/	10-3-50	3-7-51	2620	1904.30
Supplies			-	113.77
Utilities			-	503.56
Total Office Bldg. Maintenance			2620	2521.63
TOTAL ADMINISTRATION			7192	11574.22

1/ Night Watchmen and Janitor Service.

2/ Dallas Field Office only.

For this phase of the work the City Plan Department and Traffic Control Department furnished the original personnel to train new men as they were employed. At the peak of this activity seven crews were in the field. It was found that careful selection of personnel for this work was vitally important and required men with good educational background and a thorough knowledge of the city.

The preparation of numerous base maps necessary for this work required 1,081 man-hours at a cost of $1,425.01. This cost included map reproduction and supplies. In addition, the actual field work of

and future planning will be well worth the additional cost, for with these data it will be possible to expand trip information by areas as these areas develop.

In the Truck Study a 10-percent sample of all trucks registered and operated within the metropolitan area was selected from the state's motor-vehicle-registration records. A total of 2,519 truck samples required 286 man-hours at a unit cost of 12 cents per sample.

A 20-percent sample of all taxis operating in the area was selected from taxi-company records. This required only 10 man-hours and cost 13 cents per sample.

Interviewing

Trip information involving all modes of transportation for each member of the household over 5 yr. of age was obtained at each dwelling unit selected for the sample. Supervisors with crews of 10 interviewers were assigned to census tracts and each tract was completed, except for call backs, before moving into a new area. Call backs were assigned to one interviewer with a car.

The use of census tracts as subareas for dwelling-unit interviews proved well worthwhile for checking purposes. As soon as a tract was completed, summary data were expanded and compared with pertinent census data.

Interviewers were schooled for 2 days before starting actual field work. Excellent publicity was obtained through newspapers and radio and television stations. Letters explaining the purpose of the study were sent to occupants of all dwelling units selected as samples, as well as to truck owners and taxicab operators.

TABLE 3

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

INTERNAL STUDY
UNIT COST DATA

INTERNAL SURVEY	Date Started	Date Completed	Units of Work Accomplished	Total Man Hours	Total Cost	UNIT COST	
						Field	Total
MAP PREPARATION	10-9-50	2-3-51	-	1081	$ 1425 01 2/	$ -	$ -
SAMPLING							
Dwelling Units 1/	10-9-50	1-12-51	8621	3786	4509 80	-	523
Trucks	12-7-50	1-9-51	2519	286	309 45	-	123
Taxis	2-3-51	2-10-51	82	10	11 00	-	.134
Total Sampling			11222	4081	4830 25	-	430
INTERVIEWING							
Dwelling Units - Office	11-6-50	2-3-51	-	1406	2440 37	-	-
Dwelling Units - Field	11-6-50	1-15-51	8086	7949	8742 95	1 081	1 383
Trucks - Office	1-7-51	3-1-51	-	198	401 43	-	-
Trucks - Field	1-7-51	2-24-51	1945	2075	2569 10	1 321	1 527
Taxis - Office	2-3-51	2-10-51	-	4	5 12	-	-
Taxis - Field	2-3-51	2-10-51	82	247	312 39	3 810	3 872
Total Interviewing			10113	11879	14471 36	1 149	1 431
CODING (Interviews)							
Coding Index 3/	11-16-50	1-3-51	-	376	430 84	-	-
Coding - Dwelling Unit	11-15-50	1-20-51	8086	3476	3476 00	-	430
Coding - Truck	1-10-51	3-1-51	1945	1114	1165 86	-	599
Coding - Taxi	2-24-51	2-26-51	82	83	83.50	-	1 018
Total Coding			10113	5049	5156 20	-	51
SCREEN LINE							
Office	1-22-51	3-3-51	-	169	209 08	-	-
Manual Counts	1-22-51	2-8-51	8	281	280 63	35 079	-
Machine Counts	1-22-51	2-17-51	8	113	113 00	14 125	-
Total Screen Line			8	563	602 71	-	75.338
PUNCHING & VERIFYING (Cards)							
Dwelling Unit Interviews	1-17-51	2-20-51	60992	706	929 82	-	015
Truck Interviews	4-1-51	4-25-51	13115	139	178 60	-	014
Taxi Interviews	4-1-51	4-25-51	4009	27	36 45	-	009
Machine Rental			-	-	83 77	-	-
Total Punching & Verifying			78116	872	1228.64	-	016
TOTAL INTERNAL SURVEY FIELD OPERATIONS			10113	23525	$27714 17	-	2 740

1/ Field inventory method Includes procurement of Land Use information

2/ Supervision included in Total Costs

3/ Coding Index Costs prorated to each phase of survey.

Truck interviews were assigned by geographical areas for the convenience of personnel and resulted in considerable saving of time.

Taxi interviews were taken directly from manifest sheets turned in by the drivers. This made for both an easy and accurate method of obtaining the information.

6 a.m. to 10 p.m., with the exception of four of the higher-volume stations which were operated for a full 24-hr. period. A total of 70 stations were scheduled, 26 around the six suburban cities and 44 around the metropolitan area. Interviews were obtained from drivers of vehicles moving in both directions and were obtained from

TABLE 4

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

EXTERNAL STUDY
UNIT COST DATA

EXTERNAL SURVEY	Date Started	Date Completed	Units of Work Accomplished	Total Man Hours	Total Cost	Unit Cost	
						Field	Total
OFFICE	10-25-50	3-1-51	-	898	$ 1212 84 1/	$ -	$ -
INTERVIEWING	10-25-50	1-24-51	107128	5206	6735 01	-	063
MANUAL COUNTS	10-25-50	1-24-51	46	561	619.67	-	13 471
MACHINE COUNTS	10-25-50	1-24-51	46	381	439 66	-	9 558
CODING (Interviews)							
Coding Index 2/			-	501	574.45	-	-
Coding	10-31-50	3-1-51	107128	6521	6543 05	-	.061
Total Coding			107128	7022	7117 50	-	066
PUNCHING & VERIFYING (Cards)							
Labor Costs	2-20-51	4-25-51	107128	756	992 60	-	.009
Machine Rental			-	-	66 54	-	-
Total Punching & Verifying			107128	756	1059 14	-	.010
TOTAL EXTERNAL SURVEY FIELD OPERATIONS			-	14824	$ 17183 82		

1/ Supervision included in Total Costs

2/ Coding Index costs prorated to each phase of survey

Unit costs of interviews in the internal survey, including both office and field costs, were: dwelling unit interviews, $1.38; truck interviews, $1.53; and taxi interviews, $3.87.

EXTERNAL SURVEY

The external origin-and-destination survey consisted of two parts: (1) around the metropolitan area, and (2) around six isolated communities within the county. Interview stations were located on all highways and county roads carrying more than 200 vehicles per day. The lower volume stations were operated first with a small crew sufficient in number to man 3 stations per day. These men later formed the nucleus of larger crews to operate the higher-volume stations. On roads carrying less than 500 vehicles per day the survey was operated from 10 a.m. to 6 p.m. On the remaining stations the hours were from

95.5 percent of all vehicles which passed the stations. A total of 107,128 interviews were obtained in the external survey. This operation required 5,206 man-hours and cost 6 cents per interview.

Hourly automatic recorders were installed at all interview stations for a period of 24 to 48 hrs. at the time the interviews were being obtained. Five control stations were selected where automatic recorders were installed on the external cordon line for the entire period of the survey. These control counts were included with other machine counts on the external cordon line in the cost analysis and amounted to $9.56 per station.

ACCURACY CHECK

The reliability and completeness of the internal and external surveys were checked by a comparison of expanded trip data at crossings of a screen line previously es-

tablished for the purpose with the actual ground counts made during the survey. The viaducts across the Trinity River made an ideal screen line. Comparison of data indicated satisfactory survey coverage of approximately 90 percent of trips crossing the screen line. Screen-line costs were included in the cost data for the internal survey, since this phase was handled by internal personnel. Total screen line cost was $602.71, which gives a unit cost of $75.34 for each of the eight points checked.

PARKING SURVEY

The area of the parking survey comprised 140 blocks, on which were located 110 public lots, 73 private lots, 30 garages and 3,003 legal curb spaces. Inventory of parking facilities was completed with a crew of four men, plus the active participation of the parking supervisor.

Two maps were prepared, one showing the location of all curb faces and their code numbers and the other showing the location and code number of all off-street facilities. A sufficient number of 18-by 18-in. reproductions were made to provide each interviewer with a map to insure his location at the proper facility.

Information as to origin, destination, purpose of trip, and place and time of parking was obtained from 59,210 drivers of vehicles parking in the central business sector between 10 a.m. and 6 p.m. Manual classification counts were made of all vehicles crossing the internal cordon line which delimited the central business sector from 7 a.m. to 6 p.m. Four automatic traffic recorders on major streets crossing the internal cordon line furnished control information for the period of the survey. The final unit cost of the parking survey averaged 56 cents per interview. This cost included all phases of the parking survey. Inventory costs were $1.40 per block of curb face and $3.10 for each off-street facility. Curb interviews cost 22 cents each while off-street interviews cost 20 cents each. Forty-four 16-to-24 hr. cordon-line manual counts were made at a cost of $16.66 each and four control stations were operated with automatic traffic recorders for the entire period of the survey at a cost of $120.75 each.

TABLE 5

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

INTERNAL & EXTERNAL STUDY

ANALYSIS & REPORT
UNIT COST DATA

ORIGIN & DESTINATION REPORT INTERNAL & EXTERNAL SURVEYS	Date Started	Date Completed	Units of Work Accomplished	Total Man Hours	Total Cost	Unit Cost	
						Field	Total
SORTING & TABULATING (Cards)							
Labor Costs	3-10-51	9-30-51	201399	452	$ 684 80	$ -	$ .0034
Machine Rental	3-10-51	9-30-51	-	-	498 97	-	-
Total Sorting & Tabulating			201399	452	1183 77	-	006
ANALYSIS	3-10-51	11-1-51	-	1885	2518 75	-	-
PREPARATION OF REPORT							
Typing			-	104	127.03	-	-
Drafting			-	978	1270 11	-	-
Reproduction (Books)			800	-	1820 96	-	2 28
Total Preparation			-	1082	3218 10	-	-
TOTAL - METROPOLITAN AREA REPORT			-	3419	6920 62	-	-
TOTAL - INTERNAL & EXTERNAL FIELD OPERATIONS			-	38349	44897 99	-	-
TOTAL - METROPOLITAN AREA TRAFFIC SURVEY			-	41768	51818 61	-	-
ADMINISTRATION 1/					$ 7523.24		
GRAND TOTAL					$59341 85		

1/ 65% Total Administrative Costs

ZONING

For the purpose of analysis and conformance with tabulating machine usage, the area was divided into sectors, districts, zones, and subzones, designated numerically with a four-digit code. The metropolitan area was divided into sectors radiating from the business district of the city and designated by Numerals 0 through 7. Each sector was divided into districts, zones, and subzones. In the county area between the metropolitan area and the county limits, the sector lines were extended as district boundaries and this intermediate area was coded as Sector 8, with further divisions into districts, zones and subzones. The second digit indicated the sector of the city to which this county district was adjacent. Adjoining counties and major cities in Texas were coded 9000 to 9600 and the United States was coded by states from 9700 to 9900. This system of zoning greatly simplified both coding, sorting and tabulating.

TABLE 6

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

COUNTY TRAFFIC SURVEY
UNIT COST DATA

COUNTY TRAFFIC SURVEY	Date Started	Date Completed	Units of Work Accomplished	Total Man Hours	Total Cost	Unit Cost	
						Field	Total
OFFICE	10-25-50	2-1-51	-	149	$ 192 35	$ -	$ -
INTERVIEWING	10-25-50	11-13-50	16155	1265	1376 03	085	-
MANUAL COUNTS	10-25-50	11-13-50	26	214	218 63	-	8.41
MACHINE COUNTS	10-25-50	11-13-50	26	156	160 63	-	6 18
CODING INDEX 1/ CODING (Interviews)	10-31-50	12-15-50	- 16155	62 976	71 80 979 93	- -	- 061
Total Coding			16155	1038	1051 73	-	065
PUNCHING & VERIFYING Labor Costs Machine Rental	2-20-51	4-25-51	16155 -	111 -	145 40 10 03	- -	009 -
Total Punching & Verifying			16155	111	155 43	-	010
TOTAL FIELD OPERATIONS			-	2933	3154.80	-	-
REPORT							
SORTING & TABULATING (Cards) Labor Costs Machine Rental	12-1-51	2-1-52	16155 -	23 -	44 95 16 30	- -	003 -
Total Sorting & Tabulating			16155	23	61 25	-	004
ANALYSIS	12-3-51	4-1-52	-	110	161 49	-	-
PREPARATION Typing Drafting Reproduction (Books)	1-5-52 12-1-51 2-29-52	3-31-52 4-1-52 4-15-52	- - 250	34 354 -	43 38 495 59 431 38	- - -	- - 1 73
Total Preparation			-	388	970 35	-	-
TOTAL COUNTY REPORT			-	521	1193 09	-	-
TOTAL COUNTY TRAFFIC SURVEY			-	3454	$ 4347.89	-	-
ADMINISTRATION 2/					$ 578 71		
GRAND TOTAL COUNTY TRAFFIC SURVEY					$ 4926 60		

1/ Coding Index costs prorated to each phase of survey

2/ 5% Total Administration Cost

CODING, PUNCHING AND VERIFYING

A coding index was prepared in which all streets were listed alphabetically.

TABLE 7

DALLAS METROPOLITAN AREA TRAFFIC SURVEY

PARKING STUDY
UNIT COST DATA

PARKING SURVEY	Date Started	Date Completed	Units of Work Accomplished	Total Man Hours	Total Cost	Unit Cost Field	Unit Cost Total
INVENTORY							
Curb - Office	10-10-50	10-21-50	-	23	$ 53.61	$ - -	$ -
Curb - Field	10-10-50	10-21-50	140	139	142.07	1 015	1 398
Off Street - Office	10-21-50	11-18-50	-	88	115 26	-	-
Off Street - Field	10-21-50	11-18-50	213	479	545 84	2.563	3.104
Total Inventory			353	729	856 78	1 949	2 427
INTERVIEWING							
Curb - Office	11-16-50	1-27-51	-	363	717 94	-	-
Curb - Field	11-16-50	1-23-51	32736	6465	6621 86	202	224
Off Street - Office	1-7-51	3-1-51	-	1108	1536 73	-	-
Off Street - Field	12-16-50	1-25-51	26474	3557	3657 50	138	.196
Total Interviewing			59210	11493	12534 03	.174	.212
CODING (Interviews)							
Coding Index 1/				313	359 03	-	-
Coding - Curb	12-21-50	2-13-51	32736	2694	2695 25	-	.082
Coding - Off Street	2-9-51	3-1-51	26474	1766	1766 00	-	067
Total Coding			59210	4773	4820 28	-	.081
CORDON COUNTS							
Manual Counts	11-23-50	1-24-51	44	733	733 00	-	16 66
Machine Counts	11-17-50	1-20-51	4	467	483 00	-	120 75
Total Cordon Counts			48	1200	1216 00	-	25.33
PUNCHING & VERIFYING (Cards)							
Curb Interviews	2-16-51	3-24-51	32736	300	385 84	-	012
Off Street Interviews	3-4-51	3-24-51	26474	184	241 74	-	009
Machine Rental			-	-	45.09	-	-
Total Punching & Verifying			59210	484	672.67	-	011
SORTING & TABULATING (Cards)							
Machine Operators	3-24-51	9-1-51	59210	337	503 72	-	008
Machine Rental			-	-	285.58	-	-
Total Sorting & Tabulating			59210	337	789 30	-	013
PARKING REPORT							
Preparation	3-24-51	11-30-51	-	729	1020 28	-	-
Analysis	3-24-51	10-1-51	-	1282	2261 82	-	-
Reproductinn	12-1-51	1-15-52	800	-	1965 50	-	2 46
Total Parking Report			-	2011	5247 60	-	-
CHIEF SUPERVISOR	10-10-50	2-1-51	-	758	1835 43	-	-
OTHER SUPERVISION	10-10-50	2-1-51	-	1427	1613 36	-	-
Total			59210	23212	$29585 45	-	499
ADMINISTRATION 2/					$ 3472.27		
GRAND TOTAL			59210		$33057.72		$ 558

1/ 1/4 of total cost of Coding Index.

2/ 30% of total Metropolitan Area Survey Administrative Costs

All principal cities and towns within the state, major buildings within the city, outlying shopping areas and neighborhoods, schools, theatres, industrial establishments, and housing projects were also included. When completed, the coding index comprised approximately 175 pages and was bound with a spiral binding and heavy cover to enable easy access to all pages.

The cost of the coding index was $1,486.12, but it is believed that its preparation produced a saving in coding costs sufficient to justify the expenditure, as well as to promote a greater degree of accuracy. Coding index costs were prorated to each phase of the survey and included in unit coding costs.

Punching and verifying were started as soon as coding and checking were completed for a tract or station. A total of 260,609 cards were used to record data from all phases of the survey.

All work through coding was done in the Dallas field office. Some punching was done by the City of Dallas, but the majority was done in the Planning Survey Accounting Section in Austin. Unit cost of coding for the Internal Survey was 51 cents per interview; dwelling unit coding cost 43 cents per interview; truck interviews cost 60 cents each, and taxi interviews cost $1.02.

The external survey coding cost 7 cents per interview, and coding for the parking survey cost 8 cents per interview.

Average cost of punching and verifying cards for all phases of the survey was 1.2 cents per card.

TABULATING, ANALYSIS AND PUBLICATION

Tabulating, analysis, drafting, and reproduction were done in Austin. Three reports were printed covering (1) Metropolitan Area Traffic Survey, (2) Dallas County Traffic Survey, and (3) Parking Survey.

COST ANALYSIS

Unit costs for some of the more-important phases of the survey have been mentioned previously. The final cost of the internal and external traffic study was $59,341.85; the parking study, $33,057.72; and the county traffic study, $4,926.60, or a total cost of $97,326.17. A detailed man-hour requirement and cost analysis record of all phases of the project is outlined in Tables 1 to 7, inclusive.

Table 1 shows the distribution of the total cost to each phase of the survey; Table 2 the individual costs, time of beginning and ending and man-hours for each item of the administrative costs; and Tables 3 through 7 unit costs, man-hours, units of work accomplished and dates of starting and completion for all phases of the internal, external, county, and parking surveys.

APPLICATION OF DATA

The final report was not complete until 18 mo. after the initiation of the study, but tabular data were placed in use as it became available. Preliminary data were furnished the city and state within 2 mo. after completion of field work by reproduction of tracing paper work sheets for origin-destination data. Photostatic copies of the parking tables were also prepared for early use of these data. This preliminary information proved most timely and useful to both the city and state.

The City of Dallas, immediately upon completion of the study, set up an office to use these data as a basis for the (1) revision of its major thoroughfare plan, (2) development of a network of expressways for the metropolitan area, (3) determination of adequate standards of right of way and paving, and (4) development of data to guide future street programs. Population trends have been established and extended to a 20-to-30-yr. period. The increment of population increase from 1950 to 1970 or 1980 is being distributed to survey zones on the basis of type of zoning, plans for the extension of utilities, building trends and other similar factors. This expansion of the origin-destination data was made possible by the complete and current land-use information obtained in the sampling procedure.

The Texas Highway Department has used the survey data as a basis for the location of expressway routes, the design of intersections and the assignment of traffic to future routes in the area.

Postcard Method of Obtaining Origin and Destination of Traffic and Comparison with Roadside-Interview Method

ROBLEY WINFREY, Formerly Research Professor of Civil Engineering, Iowa State College

● THROUGH cooperative financing, money became available to conduct origin-and-destination traffic surveys in the urban area of Ames, Iowa, during the fall of 1949. Contributors to the project were the Engineering Experiment Station of Iowa State College, the Iowa Highway Commission, and the City Plan Commission of Ames. The project was administered by the Engineering Experiment Station.[1]

The objectives of the project were three-fold: (1) To assemble origin-and-destination information useful to the highway commission, the college, and the City of Ames in connection with their ordinary traffic handling and transportation planning, (2) to collect field data by which research could be conducted on the improvement of methods of making origin-and-destination studies of traffic, and (3) to compare the field postcard method of getting origins and destinations with the roadside-interview method.

Field stations were established on the basis that no attempt would be made to cover the interior traffic in Ames, other than that between the Fourth Ward area to the west and the Wards 1, 2, and 3 to the east (see Fig. 1 for the location of the following 13 stations):

Station 1, US 69 at the north city limits
Station 2, US 30 at the east city limits
Station 3, US 69 at the south city limits
Station 4, US 30 at the west city limits
Station 11, Lincoln Way (US 30) east of Beech Avenue
Station 12, Sixth Street east of Squaw Creek Bridge
Station 13, Thirteenth Street, east of Stange Road
Station 21, Wallace Road entrance to I. S. C. campus
Station 22, Knoll Road entrance to I. S. C. campus
Station 23, Morrill Road entrance to I. S. C. campus
Station 24, Welch Road entrance to I. S. C. campus
Station 25, Union Drive entrance to I. S. C. campus
Station 26, Pammel Drive (Cemetery) entrance to I. S. C. campus

Postcard studies for the full 24-hr. day were planned for only the primary road stations at the city limits, stations 1 to 4, and the three east-west connections within Ames, Stations 11, 12, and 13. However, because of the unfinished paving on Thirteenth Street, the postcards were not passed out at Station 13. About 2 weeks following the handing out of postcards, roadside interviews were conducted at all stations for a 16-hr. period from 7 a. m. to 11 p. m.

FIELD PROCEDURES

Traffic-volume counts were taken with traffic recorders by the highway commission daily by hours from September 27 through November 13. These counts provided the basis for determining the traffic pattern for each day of the week and trend of the volume of traffic for the fall period. Further, these counts afforded a basis for assignment of personnel to the field stations so that the traffic could be handled with the minimum of delay.

Because of the paving of the Sixth Street extension from Riverside Drive west to the I. S. C. campus at Osborn Drive, the paving of Thirteenth Street from east of Squaw Creek west to Stange Road,

[1]The over-all planning of the field work was handled by Mark Morris and Carl Schach of the Iowa Highway Commission, with the latter in active charge of the field work. Robley Winfrey handled the recruitment of personnel and the office details in the preliminary work and all of the analysis of the results. All of the IBM punching, sorting, and tabulating was handled by the Iowa Highway Commission through Schach.

TABLE 1

SUMMARY DATA ON POSTCARDS PASSED OUT

Full 24-hour day, October 17 to 25, 1949

Station	Time Started	Time Ended	Number Cards Handed Out			Remarks
			Inbound	Outbound	Total	
1	3 p. m Thurs 20th	11 p. m. Thurs 20th	370	331	701	Rained out, rerun
1	7 a. m. Mon. 24th	7 a. m Tues. 25th	1,459	1,617	3,076	
2	7 a. m. Mon. 17th	7 a. m Tues 18th	2,007	1,882	3,889	
3	7 a. m. Tues. 18th	7 a. m. Wed. 19th	2,185	2,092	4,277	
4	3 p m. Thurs 20th	11 p m. Thurs 20th	487	533	1,020	Rained out, rerun
4	7 a. m. Mon. 24th	7 a. m Tues 25th	1,767	1,841	3,608	
11	3 p. m Wed. 19th	3 p m. Thurs 20th	5,485	5,610	11,095	
12	3 p. m Wed. 19th	3 p. m Thurs 20th	2,802	2,633	5,435	
TOTAL			16,562	16,539	33,101	
Less reruns			857	864	1,721	
NET TOTAL			15,705	15,675	31,380	

and the paving of Stange Road, Thirteenth Street, south to Pammel Drive, the starting of the field work and the timing of the work was somewhat delayed over what would have been desired from the standpoint of weather and the normal October traffic.

Publicity for the field work was given by a lead story in the Ames Daily Tribune just prior to the starting of each phase. Spot announcements were made over the local Radio Stations WOI and KASI. The staff of Iowa State College was sent a special letter through the college campus mail stating the objectives of the survey and asking for cooperation in returning the postcard, either through campus mail or through the postoffice mail.

The ordinary field party of six persons for passing out the postcards consisted of a party chief, a postcard passer for each direction, a traffic recorder, a timer, and a traffic director. The number of personnel was varied according to the traffic volume; the night shift was reduced to about three to five men.

The roadside-interview phase was conducted by the same field parties, plus many additional wives and students. For the one day, Wednesday, November 2, when interviews were taken at Stations 11 and 12, the field parties were supplemented

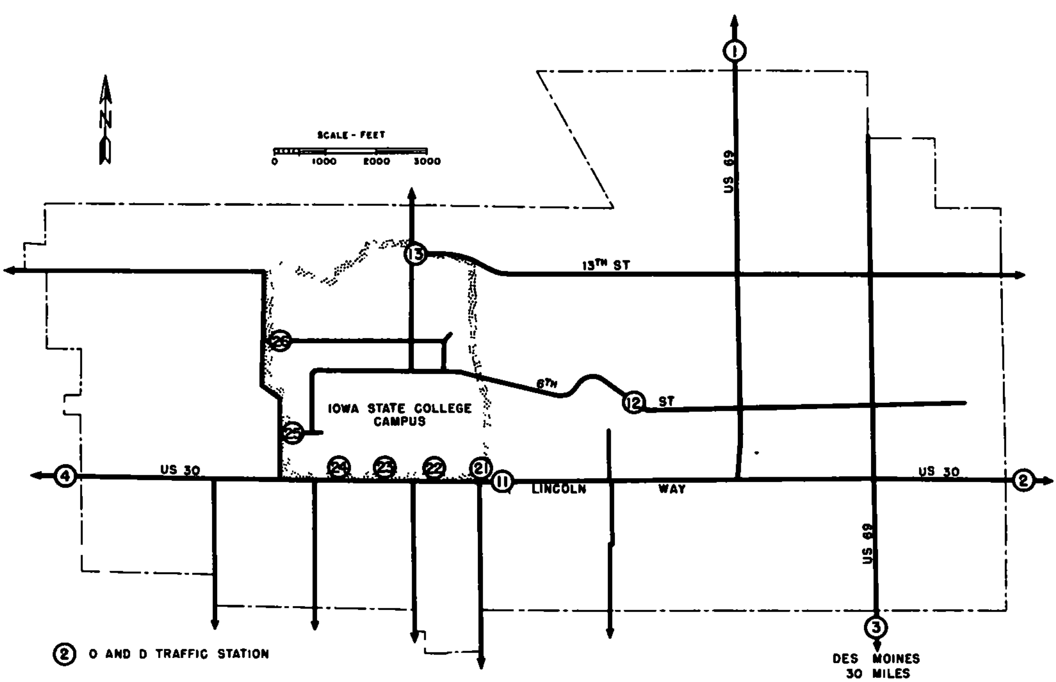

Figure 1. Outline of the corporate limits of Ames, Iowa, showing the location of Iowa State College campus and the traffic stations. Postcards were handed out at only Stations 1, 2, 3, 4, 11, and 12.

by senior girls from the Ames High School.

All interviewers, male and female, were given about 1 hr. of assembled instruction on how to interview and a field practice consisting of the taking of 10 to 30 interviews.

interviewers handled traffic in both directions. Station 11 was operated as two separate stations, one for each direction.

A timer called out the time each 2 min. over a public-address system. The interviewers recorded the time called on the

FIRST CLASS
PERMIT NO 156
(Sec 34 9 P L.&R.)
AMES, IOWA

BUSINESS REPLY CARD
No Postage Stamp Necessary if Mailed in the United States

2c—POSTAGE WILL BE PAID BY—

Engineering Experiment Station
Marston Hall, Campus
Iowa State College
Ames, Iowa

PLEASE FILL OUT AND MAIL EACH CARD GIVEN YOU DURING THIS TRAFFIC SURVEY OF AMES, IOWA

Your filling out and mailing this card is appreciated The information will be used in improving traffic in the Ames urban area

In questions 1, 2, and 3, indicate the place by street address, building name, or firm name if the place is within Ames or the ISC campus, for other places give only the city and state.

1 From what place did you last leave just before receiving this card? ________________
____

2 To what place did you first go just after receiving this card? ________________
____

3 Where was your ultimate destination if it was not the place given in question 2?
____

4 How many people were there in the vehicle, including the driver? ____________

5 In what state is the vehicle registered? ________________, if Iowa, what is the county number? ________________

6 In what city and state is the vehicle kept overnight when it is not on a trip?
____

7 Type of vehicle (check one)
☐ Car ☐ Panel or pickup truck
☐ Sta Wagon ☐ Medium weight truck
☐ Taxicab ☐ Heavy weight truck
☐ Bus ☐ Tractor-semi-trailer
☐ Motor car transport

8 For trucks and trailer combinations, give total number of axles ______________

9 Purpose of the trip (check one)
☐ Going to work ☐ Student going to or from school
☐ Coming from work
☐ To or from shopping ☐ Delivery of persons
☐ Business trip ☐ Delivery of things
☐ Social & recreational ☐ Attending meetings

Thank you—
Iowa State Highway Commission
Iowa State College, Engr Exp Sta.
City of Ames, City Plan Commission

Lincoln Way and
Beech Avenue

Eastbound

№ 5926

Figure 2. Specimen postcard passed out to drivers. Each card is identified as to station, direction, and time of day through its serial number.

In operation at the station, the traffic was controlled by the traffic directors, who endeavored to stop the vehicle at the position of the interviewer who was to make the interview. On the interview line 2 to 14 interviewers worked at a time, the number present being adjusted by the party chief to fit the traffic volume. With the exception of Station 11, individual

lines of their data sheet as it was called. New data sheets were started at the beginning of each clock hour.

The questions asked were: Origin of trip? Destination of trip? Purpose of trip? Commodity carried (for trucks)? Place where vehicle was owned?

Information recorded by observation was: type of vehicle, number of axles

(trucks only), state of registration, county of registration (Iowa only), and number of persons in passenger cars.

The personnel consisted of the regular staff of the Traffic and Planning Section of the Iowa Highway Commission, which furnished party chiefs and traffic directors, wives of students at Iowa State College, and some male students.

TABLE 2

SUMMARY OF NUMBER OF POSTCARDS RETURNED AMES URBAN AREA O AND D TRAFFIC SURVEY

October - November 1949

1949 Date		At Station	U S. Mail	Campus Mail	Daily Total	Cumulative Total
Oct.	17	34	12	0	46	46
	18	29	563	11	603	649
	19	112	900	31	1,043	1,692
	20	234	2,629	255	3,118	4,810
	21	4	3,278	478	3,760	8,570
	22	16	1,500	6	1,522	10,092
	23 - Sunday					
	24	62	1,068	54	1,184	11,276
	25	37	1,022	116	1,175	12,451
	26	0	1,059	53	1,112	13,563
	27	6	541	5	552	14,115
	28	0	344	13	357	14,474
	29	0	196	7	203	14,675
	30 - Sunday					
	31	0	225	0	225	14,902
Nov.	1	0	108	10	118	15,018
	2	5	49	6	60	15,078
	3	0	45	4	49	15,127
	4	0	31	0	31	15,158
	5	0	24	0	24	15,182
	6 - Sunday					
	7	3	28	0	31	15,213
	8	0	43	0	43	15,256
	9	0	27	0	27	15,283
	10	0	20	0	20	15,303
	11 - P. O. Closed					
	12	0	15	0	15	15,318
	13 - Sunday					
	14	0	11	0	11	15,329
	15	0	11	0	11	15,340
	16	0	0	0	10	15,350
	17	0	4	0	4	15,354
	18	0	1	1	2	15,356
	19	0	0	0	0	15,356
	20 - Sunday					
	21	0	0	5	5	15,361
	22	0	0	9	9	15,370
	23	0	0	9	9	15,379
	24	0	0	0	0	15,379
	25	0	0	0	0	15,379
	26	0	0	1	1	15,380
	27 - Sunday					
	28	0	0	1	1	15,381

THE POSTCARD

Figure 2 is a specimen postcard given out to the drivers of the vehicles. The design and wording of the card was checked with a dozen or so individuals before it was printed. Even so, several items could have been made clearer to the drivers, judged on the basis of the returns. The cards were identified by station and direction, so that identification of those returned would be positive. Further, the cards were numbered serially from 1 up

TABLE 3

COMPARISON OF POSTCARDS HANDED OUT AND RETURNED BY HOURS

STATIONS 1, 2, 3, AND 4 ALL CLASSES OF VEHICLES

Clock Hour	Direction 1, Inbound			Direction 2, Outbound			Both Directions
	Cards Handed Out	Usable Cards Re-turned	Percent-age Re-turned	Cards Handed Out	Usable Cards Re-turned	Percent-age Re-turned	Percent-age Re-turned
7-8 A	640	262	40 9	361	172	47 6	43 4
8-9	460	208	45 2	419	195	46.5	45.8
9-10	470	211	44.9	464	201	43 3	44.1
10-11	472	208	44.1	439	189	43 1	43 6
11-12 N	478	220	46 0	427	185	43 3	44.8
12-1 P	420	186	44.3	424	192	45 3	44.8
1-2	500	195	39 0	486	187	38.5	38.7
2-3	488	178	36 5	521	208	39 9	38.3
3-4	485	188	38 8	532	234	44.0	41.5
4-5	600	238	39 7	705	266	37 7	38.6
5-6	535	229	42.8	719	285	39 6	41.0
6-7	370	154	41 6	392	134	34.2	37 8
7-8	363	147	40 5	316	119	37 7	39 2
8-9	192	67	34 9	219	82	37 6	36.3
9-10	199	64	32.2	209	74	35 4	33 8
10-11 P	192	61	31 8	226	78	34.5	33.3
16 hr Total	6,864	2,816	41 0	6,859	2,801	40 8	40 9
11-12 M	136	53	39 0	160	55	34.4	36.5
12-1	95	26	27 4	82	23	28.0	27 7
1-2	45	12	26 7	52	19	36 5	32.0
2-3	28	14	50 0	25	5	20 0	35.8
3-4	24	10	41 7	25	8	32.0	36.7
4-5	45	18	40.0	40	6	13 3	28.2
5-6	43	12	27 9	70	21	30.0	29.2
6-7 A	138	42	30.4	119	40	33 6	31 9
8 hr. Total	554	187	33.8	573	177	30 9	32.3
24 hr Total	7,418	3,003	40 5	7,432	2,978	40 1	40 3

TABLE 4

COMPARISON OF POSTCARDS HANDED OUT AND RETURNED BY HOURS

STATION 11 ALL CLASSES OF VEHICLES

Clock Hour	Direction 1, Westbound			Direction 2, Eastbound			Both Directions
	Cards Handed Out	Usable Cards Re-turned	Percent-age Re-turned	Cards Handed Out	Usable Cards Re-turned	Percent-age Re-turned	Percent-age Re-turned
7-8 A	430	214	49 8	279	132	47 3	48.8
8-9	269	123	45 7	297	149	50 2	48.1
9-10	263	118	44.9	266	120	45 1	45 0
10-11	274	129	47 1	286	128	44.8	45.9
11-12 N	356	159	44.7	364	173	47 5	46.1
12-1 P	366	173	47 3	382	151	39 5	43 3
1-2	342	138	40.4	348	145	41 7	41.0
2-3	285	109	38.2	281	107	38.1	38.2
3-4	295	159	53 9	342	190	55 6	54.8
4-5	383	201	52.5	469	235	50 1	51.2
5-6	498	267	53.6	508	284	55.9	54.8
6-7	401	196	48.9	312	137	43 9	46 7
7-8	318	154	48 4	368	193	52.4	50 6
8-9	241	110	45 6	251	118	47 0	46.3
9-10	224	112	50 0	289	130	45 0	47 2
10-11 P	225	105	46 7	235	106	45.1	45.9
16 hr Total	5,170	2,467	47 7	5,277	2,498	47 3	47 5
11-12 M	124	50	40.3	153	64	41 8	41.2
12-1	68	24	35 3	85	30	35.2	35.3
1-2	0^r	0	--	0^r	0	--	--
2-3	10^r	1	--	3^r	0	--	--
3-4	15	6	--	12_r	8	--	51.6
4-5	11	5	--	18^r	5	--	34.5
5-6	7^r	1	--	8^r	1	--	--
6-7 A	80	30	37 5	64	24	44.4	40.3
8 hr. Total	315	117	37 1	333	132	39 6	38.4
24 hr Total	5,485	2,584	47 1	5,610	2,630	46.9	47 0

r = rain, some vehicles not given cards.

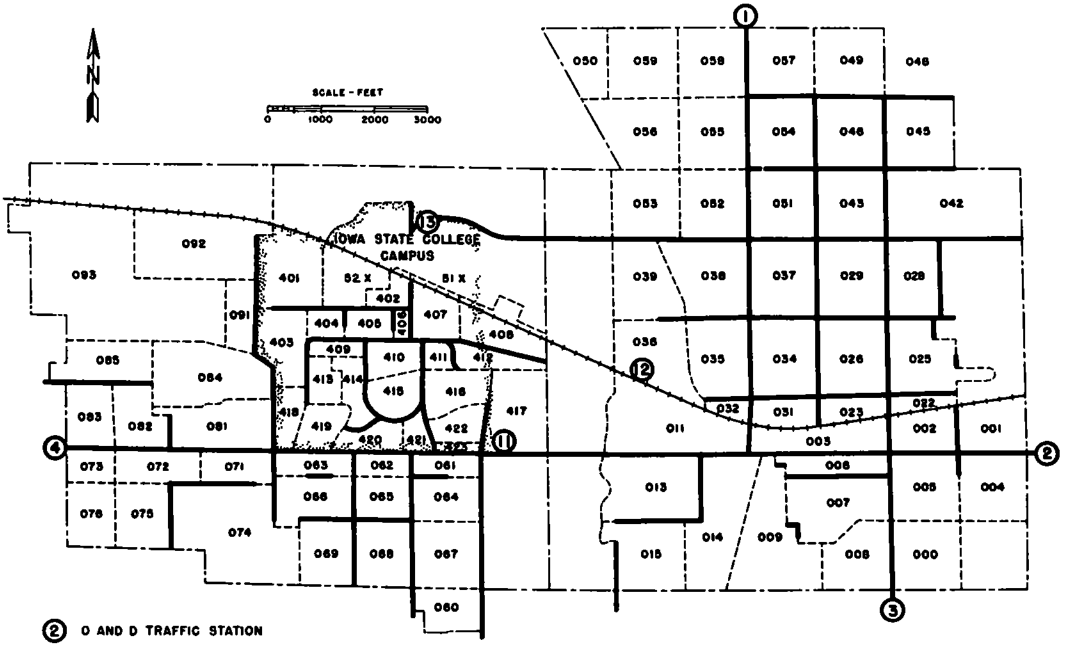

Figure 3. Zones and tracts for coding the origins and destinations within the City of Ames.

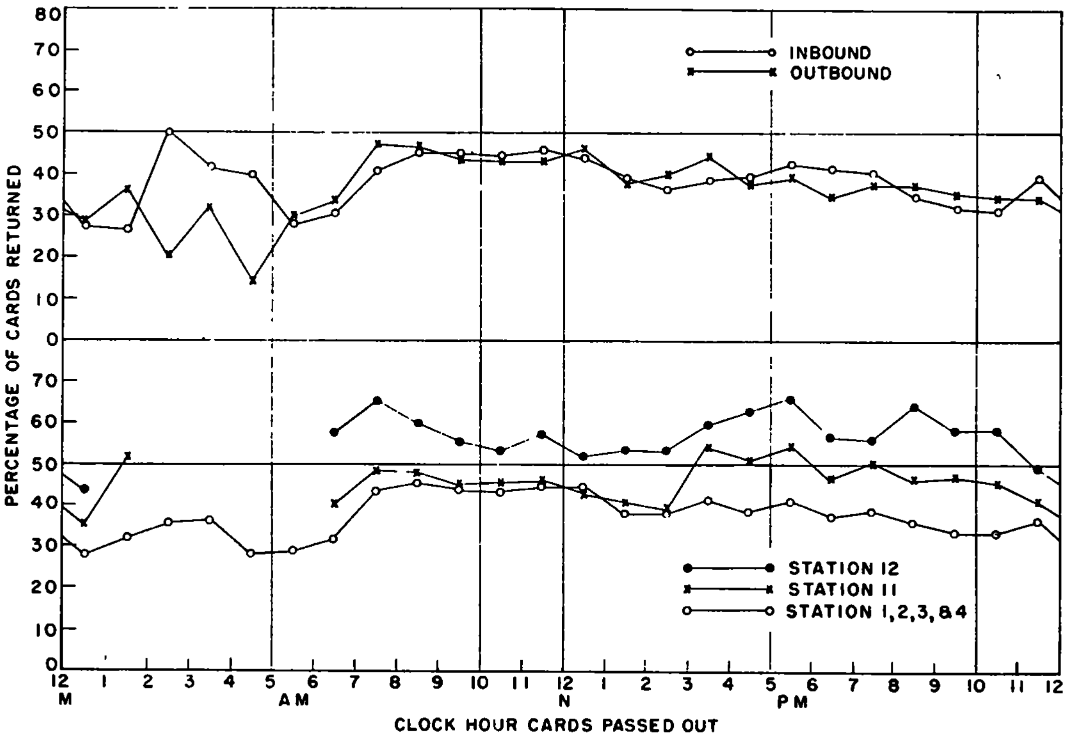

Figure 4. Percentage return of postcards by clock hours for combined Stations 1, 2, 3, and 4 and for Stations 11 and 12 for all classes of vehicles.

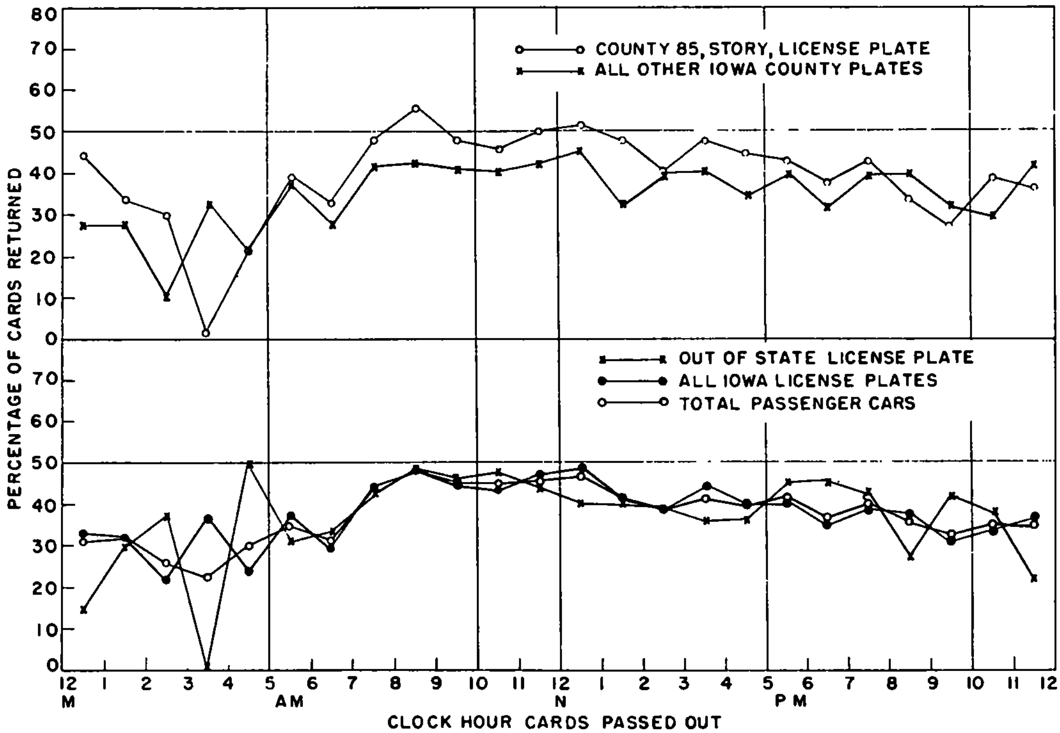

Figure 5. Percentage of postcards returned by passenger cars for Stations 1, 2, 3, and 4 combined for both directions by license plate.

TABLE 5

COMPARISON OF POSTCARDS HANDED OUT AND RETURNED BY HOURS

STATION 12 ALL CLASSES OF VEHICLES

Clock Hour	Direction 1, Westbound			Direction 2, Eastbound			Both Directions
	Cards Handed Out	Usable Cards Returned	Percentage Returned	Cards Handed Out	Usable Cards Returned	Percentage Returned	Percentage Returned
7-8 A	352	230	65 3	99	64	64. 6	65. 2
8-9	173	104	60 1	98	59	60 2	60. 1
9-10	126	73	57 9	107	57	53 3	55. 8
10-11	112	59	52 7	129	69	53. 5	53 1
11-12 N	153	81	52. 4	296	178	60 1	57 7
12 - 1 P	277	139	50 2	198	110	55. 6	52. 4
1-2	207	119	57 5	163	79	48. 5	53. 5
2-3	111	66	59 5	113	54	47 8	53 6
3-4	127	66	52 0	169	111	65 7	59 8
4-5	207	135	65 2	300	185	61. 7	63 1
5-6	240	149	62. 1	324	223	68. 8	66. 0
6-7	157	85	54. 1	138	83	60 1	56. 9
7-8	157	85	54 1	118	70	59. 3	56. 4
8-9	102	65	63 7	99	64	64. 6	64. 2
9-10	99	59	59 6	87	49	56. 3	58. 1
10-11 P	81	48	59 3	102	59	57 8	58. 5
16 hr Total	2, 681	1, 563	58. 3	2, 540	1, 514	59. 6	58. 9
11-12 M	39	18	46 2	44	23	52. 3	49 4
12-1	19	9	--	20	8	40 0	43. 6
1-2	5	0	--	8	0	--	--
2-3	4	2	--	2	1	--	--
3-4	3	2	--	3	2	--	--
4-5	5	2	--	0	0	--	--
5-6	4	4	--	4	0	--	--
6-7 A	42	24	57. 1	12	7	--	57 4
8 hr Total	121	61	50. 4	93	41	44. 1	47. 7
24 hr Total	2, 802	1, 624	58. 0	2, 633	1, 555	59 1	58. 5

for each station and direction. These serial numbers were recorded by the timer at the end of each 2-min. interval throughout the 24-hr. day so that each card returned could be identified within intervals of 2 min. as to the time of day it was handed out.

Table 1 gives the schedule of passing out cards and the number given out at each station. Table 2 gives the number of cards returned by days.

There are minor discrepancies in the count of total vehicles by hours as tallied in the manual count and classification and the number of postcards passed out. The field parties were instructed to record the serial number of the top card in the deck each 2-min. interval through the day upon call from the timer, and the manual classifier and enumerator was instructed to start a new tally line with each 2-min. interval. Neverthelesr there are variations of one to three vehicles between the two counts. Some of these variations result from mistakes in the numbering of the cards by the printer; some represent skipping and duplicating on the part of the field personnel. With the exception of

those differences in the two counts (which are attributed to the fact that during rain, cards were not passed out at certain time periods, although the manual count of traffic volume continued), the differences in counts is not of importance to the analyses reported.

Examples of typical codes for origins or destinations are:

043 813 On Kellogg Avenue in the 1300 block (Ames, zone 4, tract 3, Kellogg Street, 13 hundred block)

414009 Beardshear Hall, I.S.C. campus (I.S.C. campus, zone 1, tract 4, bldg. 009)

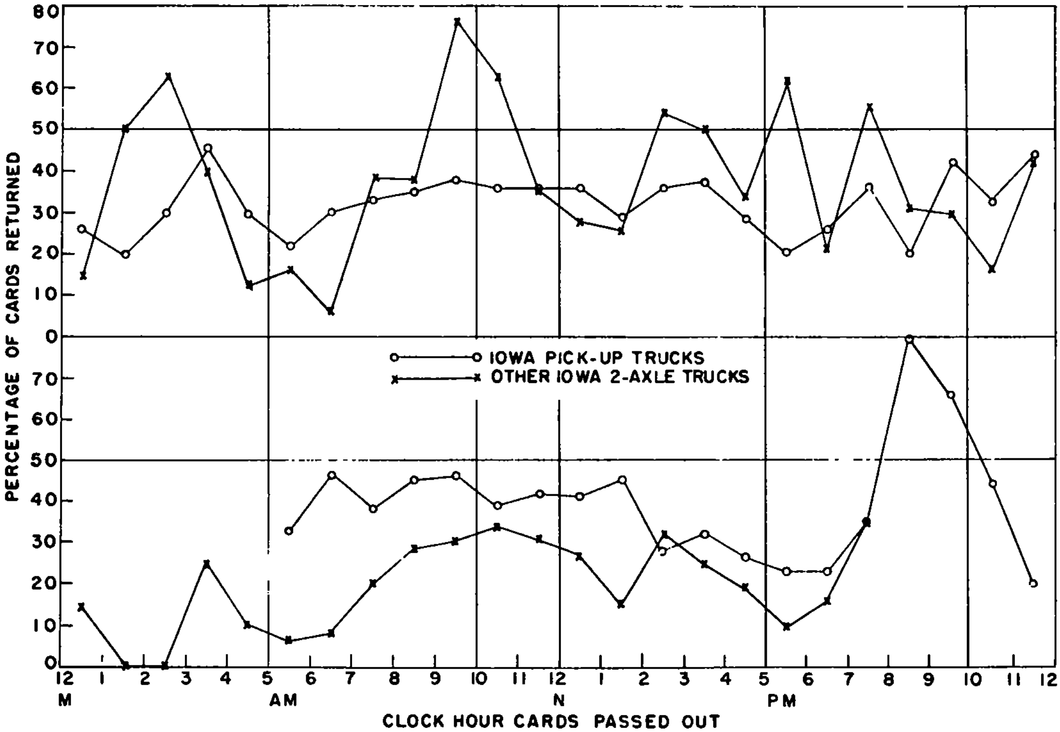

Figure 6. Percentage of postcards returned by trucks for out-of-state and Iowa registration plates for Stations 1, 2, 3, and 4 combined for both directions.

OFFICE WORK

A system of numerical codes was developed so that the analysis of the returns could be handled by IBM business machines. The city was divided into 10 geographical zones, with each zone divided into 10 or fewer tracts of about three blocks square. Within the tract the origins and destinations were coded to the street and block. Thus, the six-number code permits of determining the location in Ames of the origin and destination to the specific block length of street. The campus of Iowa State College was divided into 23 areas covered in the first three numbers of the code; the last three numbers designate the specific building. Figure 3 gives the zones and tracts for Ames.

514 762 Dwelling 762 in North Pammel Court (Pammel Court, north section, area 4, House No. 762)

185,056 Nevada, Iowa, (a town of more than 1,000 population; 85 is Story County, and 056 is Nevada)

177,027 Des Moines, Iowa

285,676 Ontario, Iowa, (a town of less than 1,000 population; 85 is Story County, and 676 is Ontario)

327,000 State of Minnesota (out of state, 27 is Minnesota, and 000 is any place in Minnesota)

The postcards returned were sorted by hand into stations and directions, and then arranged in serial number. Coding sheets were prepared by numbering vertically the 24 lines on the 11-in. dimension of a 17- by 11-in. ruled sheet. Consecutive serial

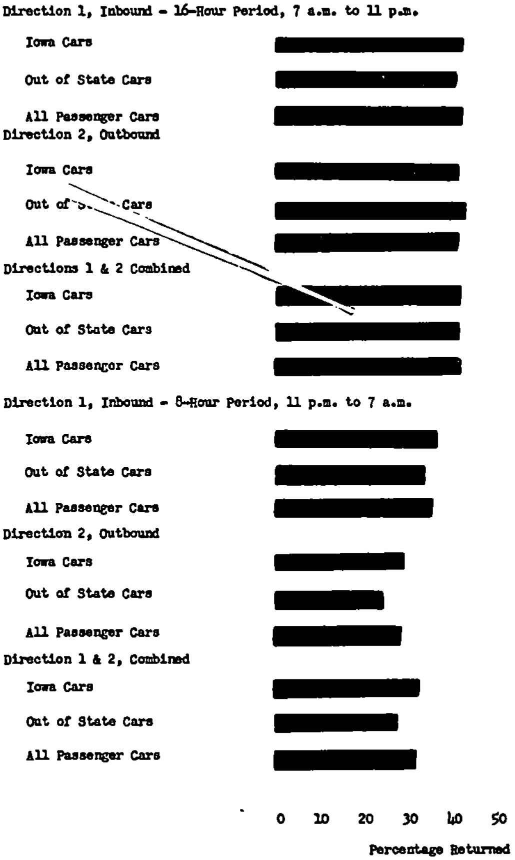

Figure 7. Percentage of postcards returned for passenger cars by registration plate, Stations 1, 2, 3, and 4, combined.

numbers were used from 1 up to cover the entire series of cards passed out at each station. Coding was then done directly on these sheets on lines corresponding to the serial number of the cards. Thus, the 24 lines on a coding sheet show the serial numbers of all cards passed out; the coded lines thereon represent the usable cards returned. The hour and 2-min. interval during which the cards were passed as recorded on the coding sheet were obtained by reference to the field record of the serial numbers and times.

The coding was not checked other than by inspection of the completed sheets to see if the origin and destination looked reasonable for the station and direction.

Neither was the punching of the cards verified. Once the cards were punched they were sorted by origin and then listed. Again, by inspection of these listings for reasonableness the further main errors were eliminated by re-examination of the coding sheets and original postcards.

DIFFICULTIES WITH THE RETURNED INFORMATION

The difficulties had with the information contained on the cards were mainly what would be expected. When dealing with run-of-the-mill vehicle drivers, it is expected that information gathered in a traffic survey by the use of postcard questionnaires would contain discrepancies, and that sometimes wanted information would be missing. On the whole, however, the information was satisfactorily submitted. The record (Table 2) of cards returned daily totals 15,381 cards. The 15,381 includes 619 cards from Stations 1 and 4 which were not used, since the passing out was repeated later when it became necessary to suspend operations on account of rain. Available net cards totaled 14,762, of which 14,374 were coded. Not every one

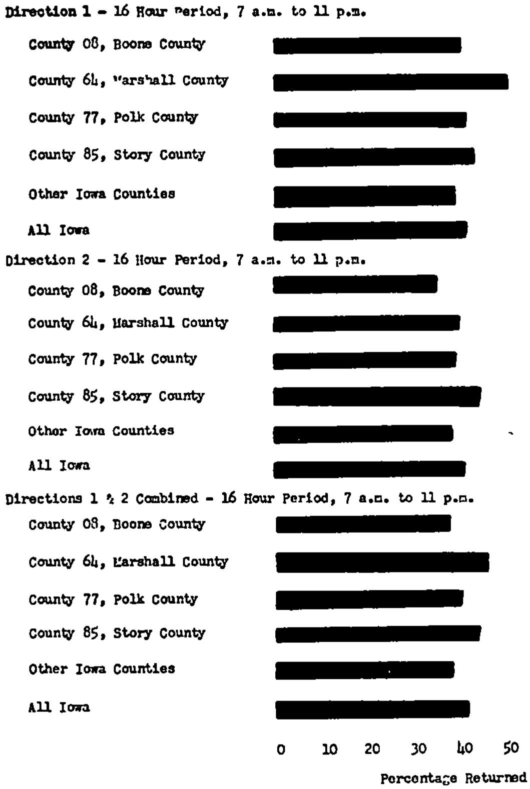

Figure 8. Percentage of postcards returned for passenger cars by Iowa registration plate, Stations 1, 2, 3, and 4, combined.

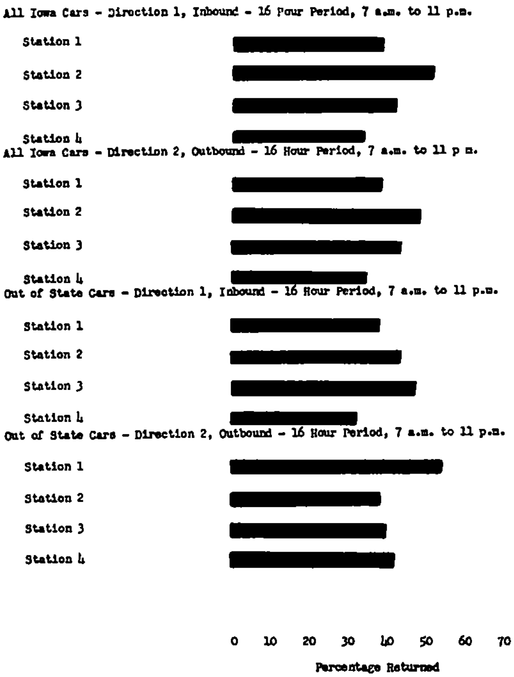

Figure 9. Percentage of postcards returned for passenger cars by registration plate, for each exterior station.

of the cards punched was 100 percent usable, though at least either the origin or destination was determined. These figures indicate that 97.37 percent of the cards returned were coded and punched.

From the standpoint of an origin-and-destination traffic survey, the main item that gave trouble was the description of the origin and the destination. Many cards had these items either incomplete, stated in general terms, confused, or a definitely wrong geographical location. The following answers are types of answers frequently encountered: a street number without giving the city; Ames, Iowa, without giving the street number; home to work; to Bill Smith's place; to a location north of Ames for a vehicle given the card when headed south at the south entrance to Ames (Station 3).

Considerable of the difficulty with the origin and destination is attributed to the fact that many drivers received more than two cards during a given day and possibly at more than one station. When filling out these cards, the driver did not realize that each card was for a particular station, direction of travel, and minute of the day. As a result, the correct trip was recorded on the wrong card. The coders salvaged many of these cards by a process of matching; that is, all questionable cards were compared in terms of the indicated trip, type of vehicle, time of day, and handwriting.

RESULTS OF THE ANALYSIS

Inasmuch as this study of the postcard method was directed toward those aspects of the method which would help to evaluate its reliability, there is no attempt to analyze the results strictly to determine the origin and destination of the traffic passing the six stations. Rather, the results are analyzed to determine any bias in the return of the cards and a comparison of the origin and destinations with those obtained by the roadside-interview method.

RETURN OF CARDS BY HOURS OF THE DAY

Tables 3, 4, and 5 give by hours the

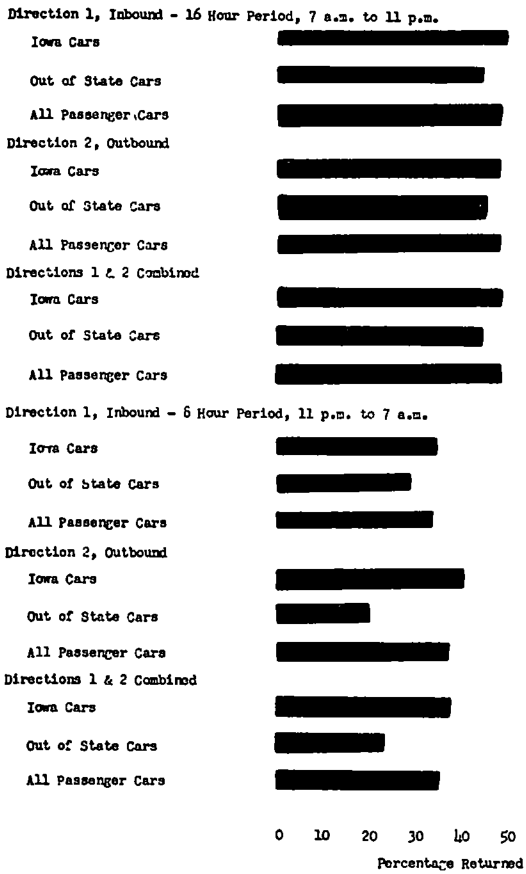

Figure 10. Percentage of postcards returned for passenger cars by registration plate, Station 11.

cards handed out and the usable cards returned by all classes of vehicles for each hour of the 24-hr. day. Summary curves of the percentage of cards returned each hour is given in Figure 4.

There appears to be no significant difference in the four exterior stations in the percentage of the cards returned by outgoing vehicles and incoming vehicles. There is, however, some trend downward in the percentage returned from the hour

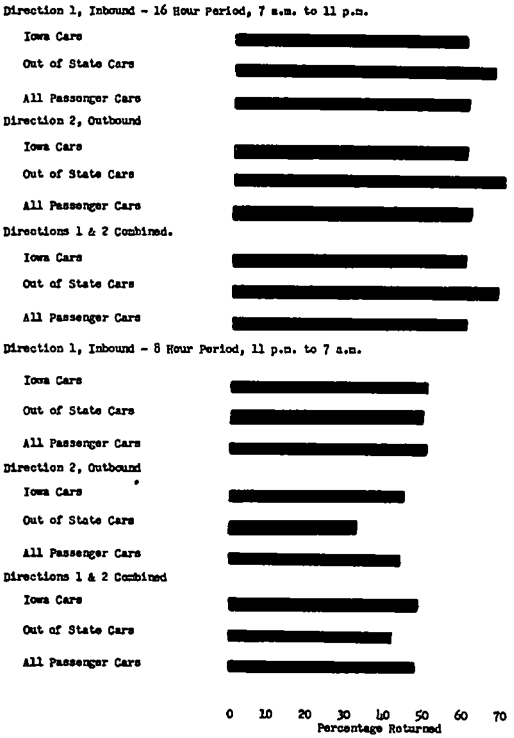

Figure 11. Percentage of postcards returned for passenger cars by registration plate, Station 12.

from 7 to 8 a.m. to the same hour the next day. For the four stations, the 8-hr. period from 11 p.m. to 7 a.m. averaged 32.3 percent returned for both directions and all vehicles as compared to 40.9 percent for the 16-hr. period from 7 a.m. to 11 p.m. The number of cards returned in the individual hours of the 8-hr. period was so small as to render unreliable the percentage returned for any given hour.

Figure 4 gives the highest percentage return from Station 12, next from Station 11, and the lowest from Stations 1, 2, 3, and 4 combined. Station 12, the Sixth Street connection between the eastern area

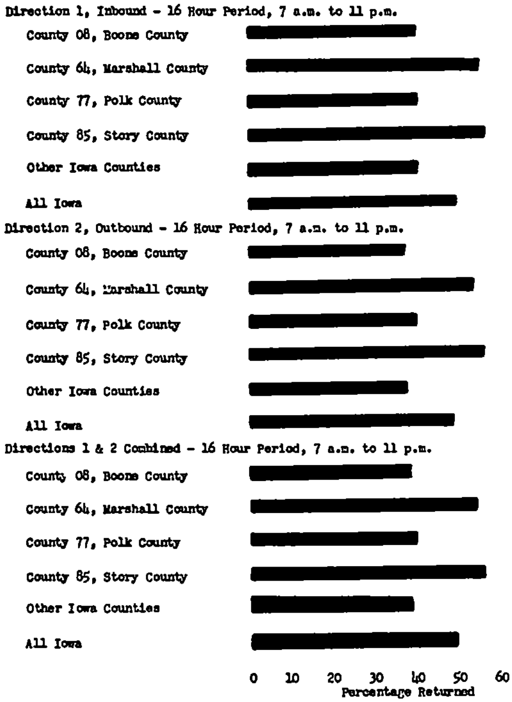

Figure 12. Percentage of postcards returned for passenger cars by Iowa registration plate, Station 11.

of Ames and the college campus, is used largely by college employees. These employees received a letter stating the purposes of the survey and asking for their cooperation. Further, perhaps the college

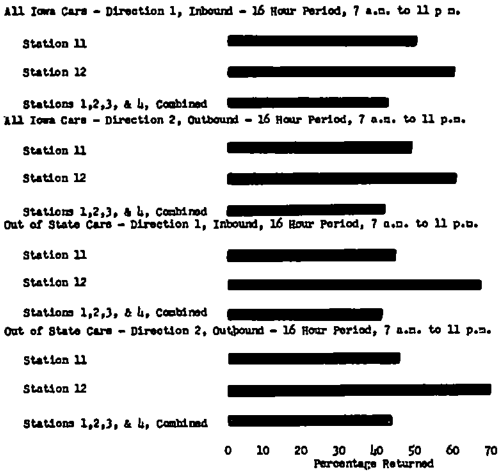

Figure 13. Percentage of postcards returned for passenger cars by registration plate, for interior Stations 11 and 12 and for exterior Stations 1, 2, 3, and 4, combined.

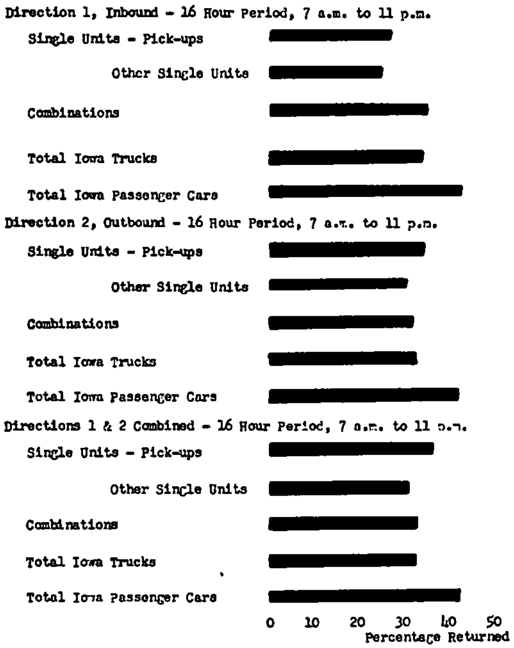

Figure 14. Percentage of postcards returned for Iowa trucks, and Iowa passenger cars by registration plates, Stations 1, 2, 3, and 4, combined.

people were somewhat more inclined to cooperate in a study of this type than was the average run of drivers, since the project was in charge of one well known to them. The returns from Station 11 also included many cards from college employees. Stations 1, 2, 3, and 4 passed ordinary traffic, such as moves into and out of typical Midwestern cities of 15,000 to 25,000 population.

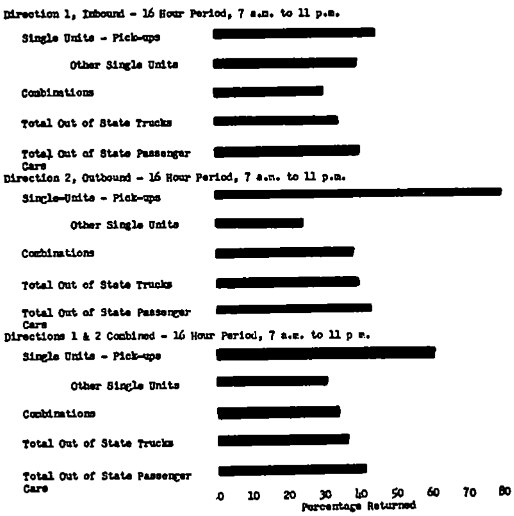

Figure 15. Percentage of postcards returned for out of state trucks and out of state passenger cars by registration plates, Stations 1, 2, 3, and 4, combined.

Attention is directed to the rather high position of the curve for Station 11 at 3 to 4 p.m. as compared to the other hours. From Table 1 it is seen that this station

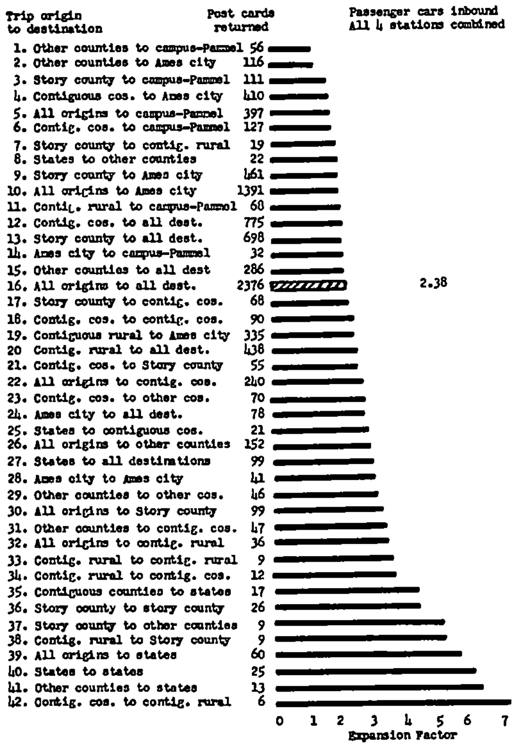

Figure 16. Postcard expansion factors for trips between origin-and-destination groups. The roadside-interview trips between the same origin-and-destination groups are used as the base (100 percent trips) to which the postcard trips were expanded. Expansion factors are omitted for trip groups of less than 25 roadside interviews and less than 6 postcards returned.

was started at 3 p.m. There may be some indication that the percentage of cards returned is somewhat higher the first hour or hours of operation of the station than throughout the remainder of the day. This high point is also found for the exterior stations at 7 to 8 a.m., the beginning hour. Station 12, on the other hand, shows its high hours at 7 to 8 a.m. and 5 to 6 p.m., though the passing out of cards was started at 3 p.m. Station 12 is greatly different in

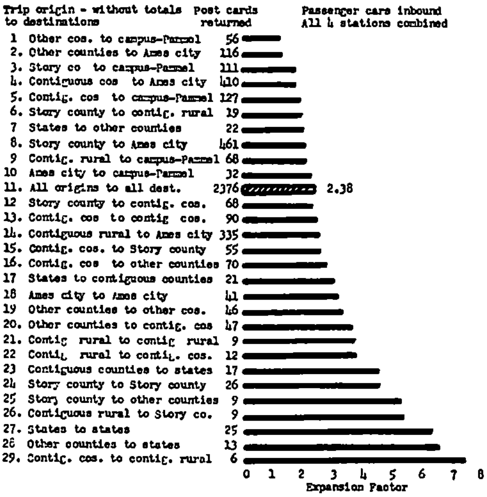

Figure 17.

character of traffic than Station 11, for the reason that it carries only local passenger cars and light commercial trucks between the college campus and the eastern area of Ames, including the main business district.

Figure 5 shows that the returns from the passenger cars bearing an 85 (the local Story County) license plate returned a higher percentage of cards than did the passenger cars bearing license plates from all other Iowa counties combined. The percentages for the hours of 11 p.m. to 7 a.m. are erratic because of low number of cards returned. The downward trend of the curve from 7 a.m. to midnight is evident on both the Story County and other county returns.

The lower graphs of Figure 5 indicate that there is little difference in the percentage returns hour for hour between the returns from out-of-state registered passenger cars and those registered in Iowa. An important fact to keep in mind, however, is that in Ames, Iowa, because of the Iowa State College, there are many out-of-state

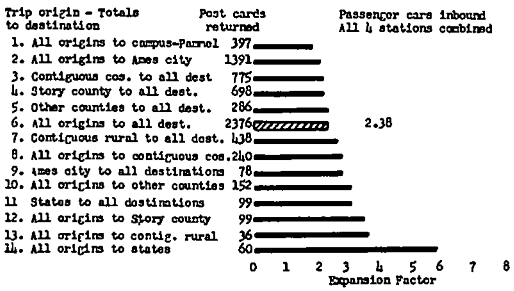

Figure 18.

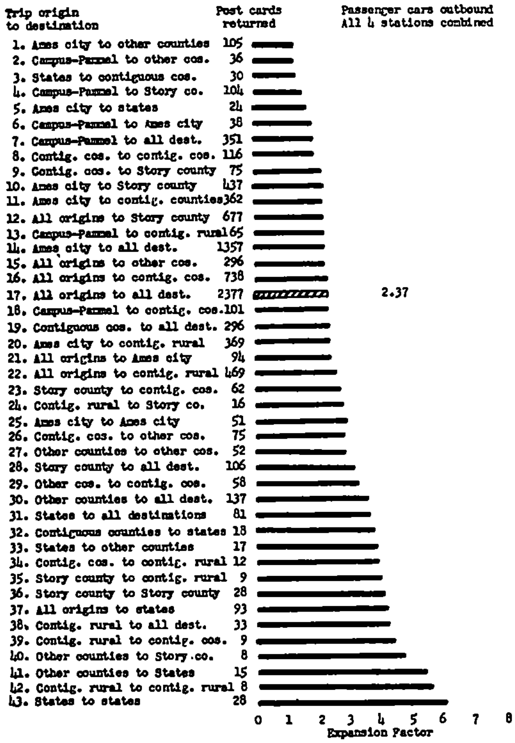

Figure 19.

licensed passenger cars in operation daily which are permanently kept in Ames, at least from September to June. In the fall of 1951 there were 510 out-of-state cars registered with the college. About 1,650 Iowa cars from other than Story County were registered and 2,955 Story County cars. These cars, however, would not

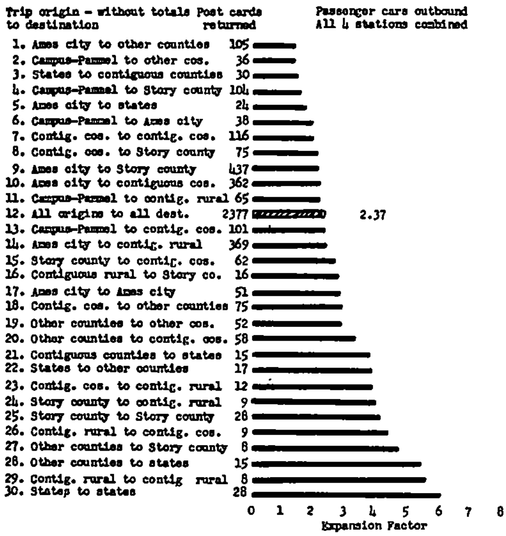

Figure 20.

pass Stations 1, 2, 3, or 4 frequently during week days.

The curves of Figure 6 for trucks disclose but little significant difference in postcards returned by hours because of the small number of cards passed out and returned. Perhaps the curve for the pick-up trucks in the lower section is higher than the curve for other two-axle trucks because the pick-up vehicles are mostly of local ownership and use.

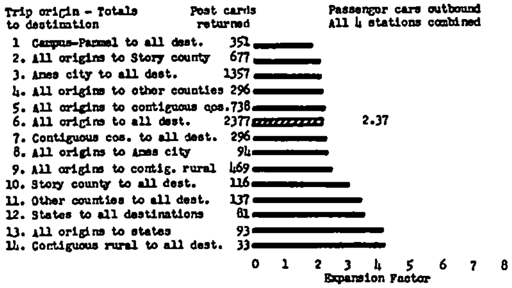

Figure 21.

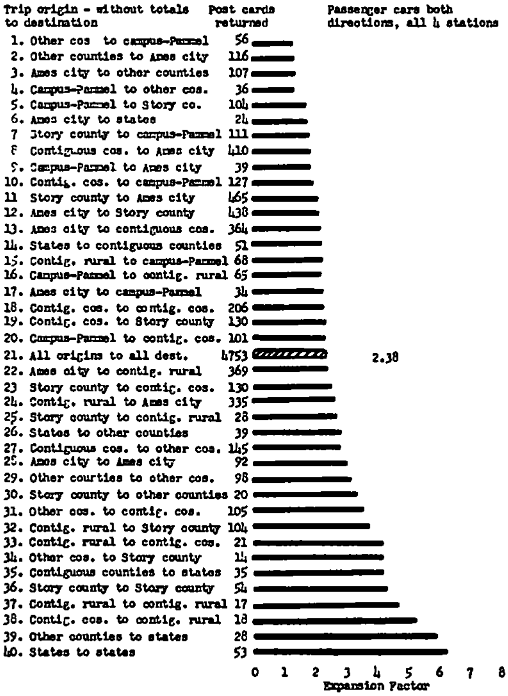

Figure 22.

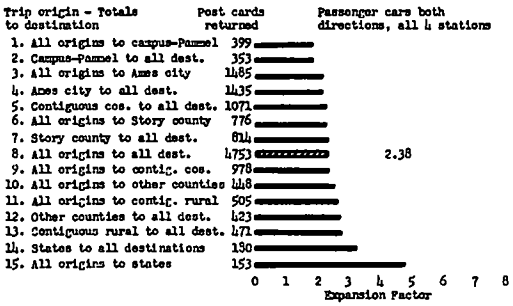

Figure 23.

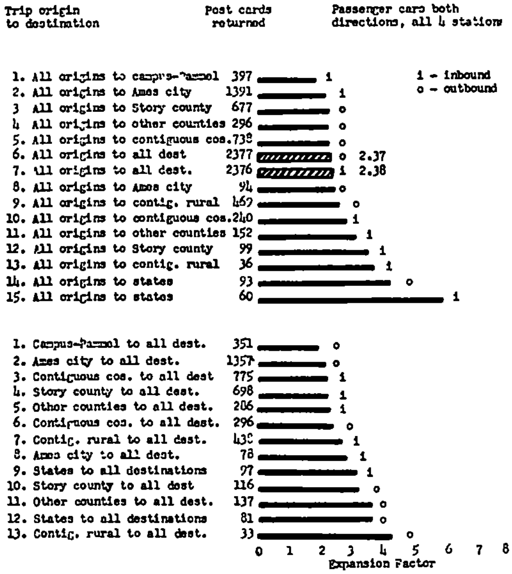

Figure 24.

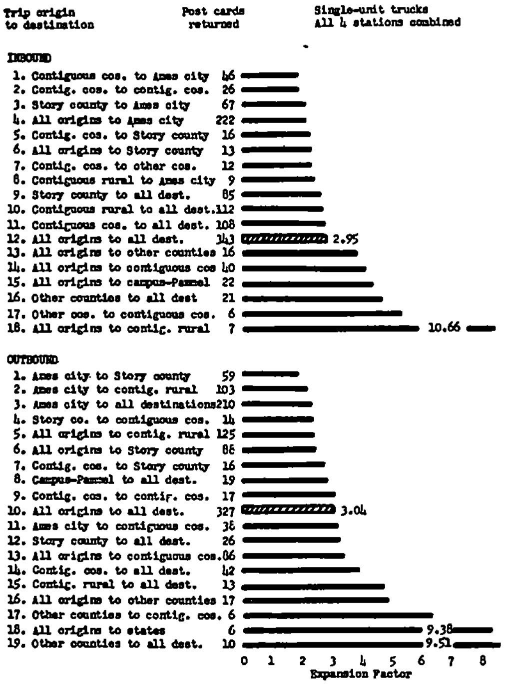

Figure 25.

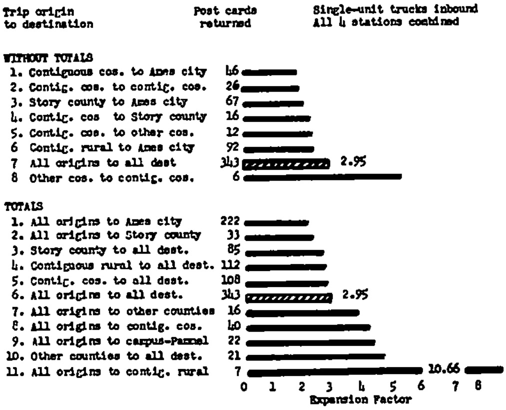

Figure 26.

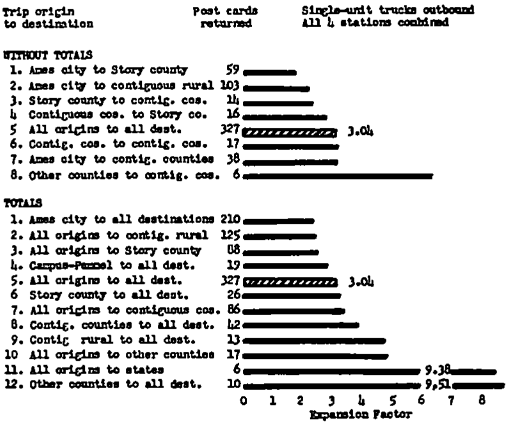

Figure 27.

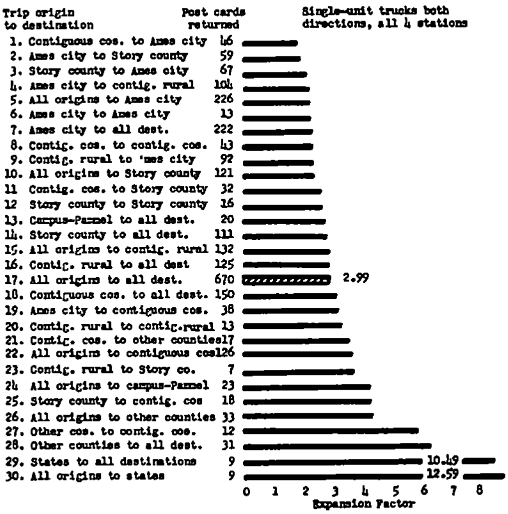

Figure 28.

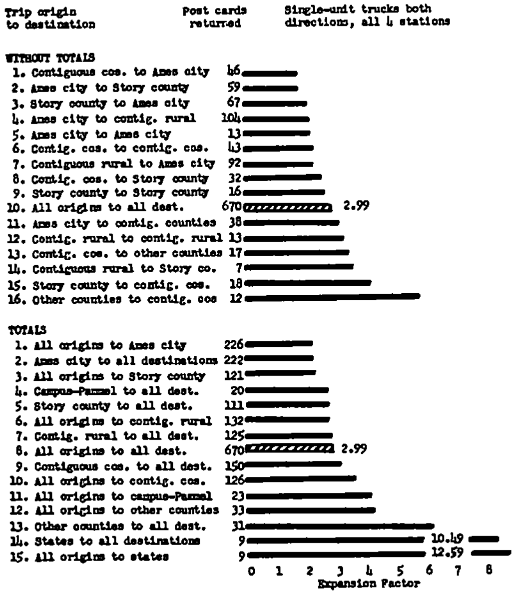

Figure 29.

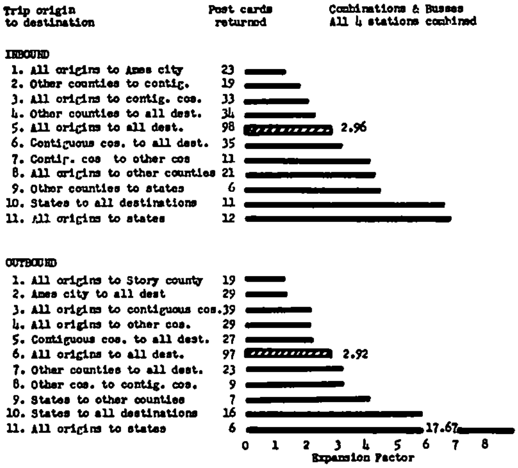

Figure 30.

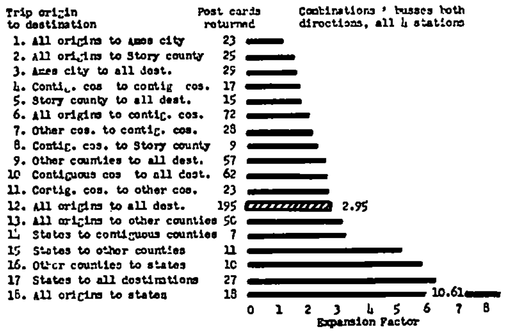

Figure 31.

Return By Station, Class of Vehicle and Registration Plate

Figures 7 through 15 set forth the percentage of cards returned by classes of vehicles by registration plate, direction, and station for the 16-hr. and the 8-hr. periods for all six stations.

Figure 7 indicates that the percentage of passenger cars returning cards was about 42 percent for the 16-hr. period and only 32 percent for the 8-hr. period, all four exterior stations combined and directions combined.

The comparison in Figure 8 is by counties, three counties contiguous to Story,

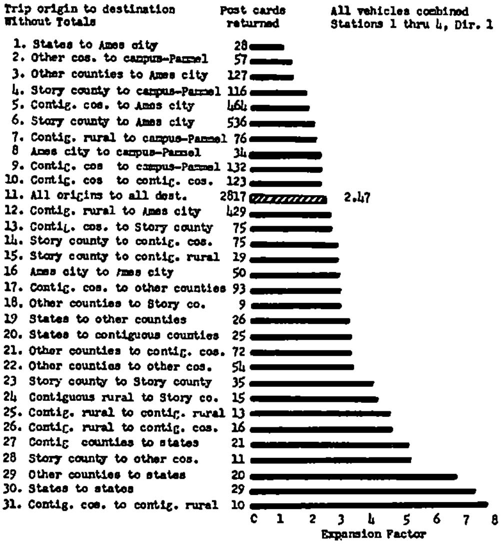

Figure 32.

the local Story (County 85), and other Iowa counties. The variations in the percentage returned from passenger cars is lowest for Polk County (to the south from Station 3) and other counties, with Marshall County, which is east of Story County, having a percentage about equal to that of Story County.

Figure 9 shows considerable variation in the percentage of cards returned by exterior stations, though no particular difference in the return by direction, nor between Iowa registered cars and out-of-state registered cars. There is no apparent reason seen for the variation in percentage by stations.

For interior Station 11, the US 30 east-west highway through Ames, Figure 10 shows about a 49-percent return from Iowa

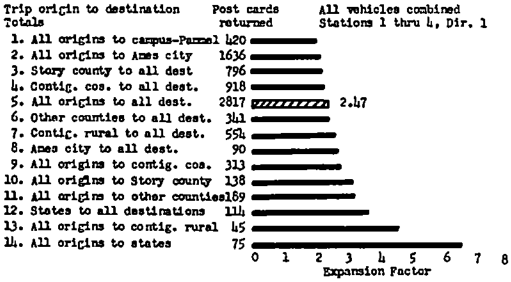

Figure 33.

registered passenger cars for both directions for the 16-hr. period as compared to 45 percent for the out-of-state cars. Similar percentages for the 8-hr. period are 38 and 24 percent, respectively.

The college-Ames east-west connection, Station 12, indicates the reverse percentages between Iowa and out-of-state registered cars as shown for Station 11, with only 61 percent returns from the Iowa cars and 69 percent from the out-of-state. Again, the 8-hr. period produced a small percentage return falling to 49 percent for the Iowa cars and 42 percent for the out-of-state. For Station 12, however, it should be kept in mind that the out-of-state registered cars as well as the out-of-county registrations are quite largely local cars that are based at Ames for the school year.

In Figure 12, Marshall County and Story County registered passenger cars again show up with high percentages as compared with other counties.

Figure 13 affords opportunity to com-

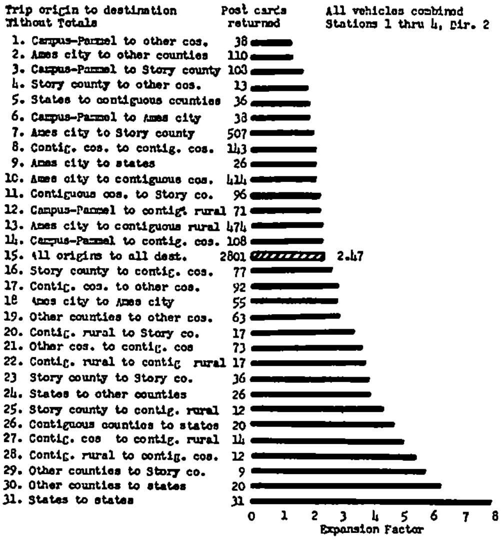

Figure 34.

pare the returns from passenger cars for the 16-hr. period by interior and exterior stations. For both the Iowa and out-of-state registrations, the highest returns were received from Station 12 with Station 11 being in the middle, and the combined exterior stations producing the lowest percentage returns.

Figure 14 shows that the percentage returns for Iowa registered vehicles from the exterior stations are about 10 percentage points greater from passenger cars than from commercial vehicles. The pick-up truck had a higher percentage than did the other classes of trucks.

Figure 15 illustrates about the same trend for out-of-state trucks as is shown in Figure 14 for Iowa registrations.

TABLE 6

SUMMARY OF ADJUSTMENT OF ROADSIDE INTERVIEWS FOR STATIONS 11 AND 12 TO STANDARD VOLUME 16-HR. PERIOD

Vehicle Class	Actual Number of Interviews: Interior-Interior Trips	Actual Number of Interviews: Ext-Ext and Int-Ext Trips	Actual Number of Interviews: Total Trips	Adjusted to Standard Volume: Interior-Interior Trips	Adjusted to Standard Volume: Ext-Ext and Ext-Int Trips	Adjusted to Standard Volume: Total Trips
Station 11, Direction 1						
Pass Cars	2,491	1,432	3,923	3,153	1,428	4,581
Single Unit	287	221	508	327	213	540
Combinations	8	77	85	7	72	79
Total	2,786	1,730	4,516	3,487	1,713	5,200
Station 11, Direction 2						
Pass. Cars	2,522	1,305	3,827	3,298	1,283	4,581
Single Unit	298	239	537	327	213	540
Combinations	6	72	78	6	73	79
Total	2,826	1,616	4,442	3,631	1,569	5,200
Station 12, Direction 1						
Pass. Cars	1,728	249	1,977	1,918	242	2,160
Single Unit	99	19	118	108	22	130
Combinations	0	0	0	0	0	0
Total	1,827	268	2,095	2,026	264	2,290
Station 12, Direction 2						
Pass Cars	1,748	280	2,028	1,987	293	2,280
Single Unit	103	27	130	104	26	130
Combinations	0	0	0	0	0	0
Total	1,851	307	2,158	2,091	319	2,410

COMPARISON OF ORIGINS AND DESTINATIONS OBTAINED FROM THE 100-PERCENT ROADSIDE INTERVIEWS WITH THE POSTCARD RETURNS

As stated in the introduction, roadside interviews for the 16-hr. period of 7 a.m. to 11 p.m. were conducted at each of the stations at which postcards were distributed. The roadside interviews were conducted about 2 weeks after the postcard field work had been completed and most of the cards were received. As near as can be judged, the traffic conditions were comparable during the days of the two separate surveys, except that on Wednesday, November 2, 1949, the day that the

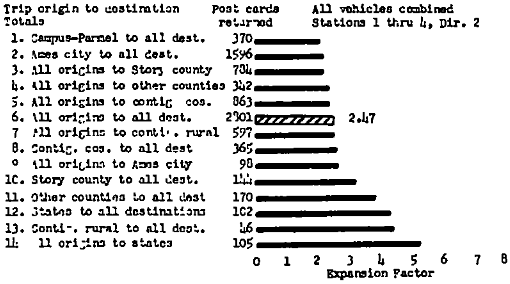

Figure 35.

two interior stations, 11 and 12, were operated for interviews, the day was unseasonably cold, windy, and generally disagreeable to the field parties as well as to the traffic. As a result of this bad weather, the traffic on Station 11 was 1,442 trips short of a normal week day, and Station 12 was 447 trips short. A detailed study of the traffic at these stations as observed in the manual count and classification the day postcards were handed out, the hourly volumes from the automatic recorders, and the roadside interviews indicates that there was no lessening of the through traffic, of the interior-exterior traffic, nor of the exterior-interior traffic on this Wednesday. Practically the entire shortage came from local passenger trips interchanging be-

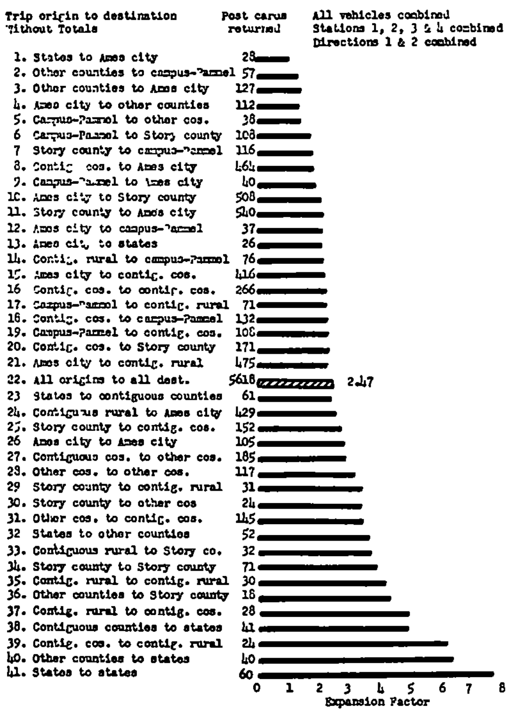

Figure 36.

tween the eastern and western areas of Ames. Consequently, the roadside interviews were adjusted accordingly (see Table 6 for a summary of these adjustments).

The origin-and-destination distributions for Stations 1, 2, 3, and 4 as determined from the roadside interviews were used without adjustment in comparing with the similar results obtained from the postcard returns.

expansion factor, by which the postcard returns could be expanded to equal the number of identical trips obtained in the 100-percent interviews for the 16-hr. period. See Table 7 for a specimen table of the data and calculations. This method of comparison assumes that the traffic and origin-and-destination distribution on the days of the two separate surveys was comparable. So far as could be determined, such assumption is valid.

TABLE 7

SPECIMEN TABLE SHOWING CALCULATION OF EXPANSION FACTORS FOR COMBINED STATIONS 1, 2, 3, AND 4, DIRECTION 1, INBOUND, PASSENGER CARS

	Destinations								
Origins	Ames City (1)	Campus Pammel (2)	701-704 (3)	County X85 (4)	Contiguous Counties (5)	Other Counties (6)	States (7)	Total (8)	Error (9)
0. Errors	19.4	4 4	1 0	2 4	1.9	1 2	1 9	34 2	2.0
1. Ames City 00X-099	131.5	74.5	11.2	5 0	0	1 0	0	223.2	
2. Campus Pammel 4XX-5XX	0	0	0	0	0	0	0	0	
3 701-702-703-704	865 2	144 9	33 9	49 1	45 8	14.1	8 8	1,163 0	1.2
4. County X85	959.7	201.0	38 3	118 1	163 6	48.4	21 3	1,555 1	4 7
5 Contiguous Co.	762 7	252 5	44 8	148 1	227 1	199.0	77 0	1,715.6	4 4
6. Other Counties	162 6	75 8	1.0	19 8	170 0	155.2	84 8	669.2	
7. States	24 0	7.2	0	4 1	63 9	45.6	159.7	304.5	
8 Total	2,925 1	760 3	130 2	346.6	672.3	464 5	353.5	5,664.8	12 3
1. Ames City 00X-099	41	32	0	1	2	2	0	78	
2 Campus Pammel 4XX-5XX	1	1	0	0	0	0	0	2	
3 701-702-703-704	335	68	9	9	12	3	2	438	
4 County X85	461	111	19	26	68	9	3	698	
5 Contiguous Co.	410	127	6	55	90	70	17	775	
6. Other Counties	116	56	2	6	47	46	13	286	
7 States	27	2	0	2	21	22	25	99	
8 Total	1,391	397	36	99	240	152	60	2,376	1
1. Ames City 00X-099	3 21	2 33	–	–	–	–	0	2.86	
2 Campus Pammel 4XX-5XX	–	–	0	0	0	0	0	–	
3. 701-702-703-704	2.58	2.13	3 77	5 46	3.82	–	–	2 66	
4. County X85	2.08	1.81	2.02	4 54	2.41	5.38	–	2 23	
5. Contiguous Co.	1.86	1 99	7 47	2 69	2.52	2.84	4.53	2 21	
6 Other Counties	1.40	1 35	–	–	3.62	3.37	6.52	2 34	
7. States	–	–	0	–	3.04	2.07	6 39	3.08	
8 Total	2.10	1.92	3.62	3 50	2.80	3 06	5.89	2.38	

NOTE: Upper section of table gives the roadside interview trips between each origin and destination multiplied by the adjusting factor to raise the number of interview trips to the number of post cards handed out.

The middle section gives the number of cards returned for each origin and destination pair.

The lower section gives the expansion factor for each origin and destination pair, upper section divided by the middle section

The object of the comparison of the origins and destinations of the traffic as obtained by the two field methods was to determine whether there was any unequal returns of the postcards which could be attributed to the origin and destination of the trips. In order to accomplish this comparison, all postcard returns by origin and destination pairs were summarized for identical pairs of origins and destinations as obtained by the roadside interviews. The ratio of the postcard trips to the roadside-interview trips for each origin-destination group thus produced a conversion or

In making the comparison of results obtained by the two methods, attention was focused on larger classes of trips, that is, trips between larger areas or zones, rather than the smaller tracts or individual counties or cities. In general the classifications used as origins and as destinations are: (1) states other than Iowa, (2) counties of Iowa other than Story and other than those contiguous to Story, (3) counties contiguous to Story, (4) Story County, (5) rural area adjacent to Ames city, (6) Ames city area,

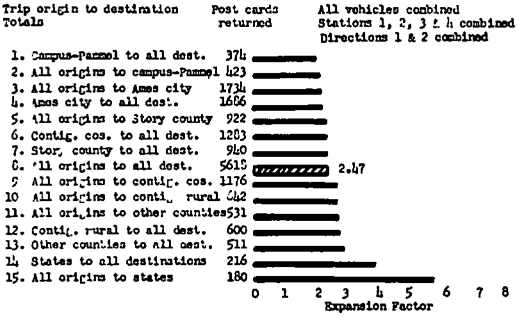

Figure 37.

and (7) campus and Pammel Court area of Iowa State College.

This combination of origins and of destinations massed the data sufficiently so that the number of trips in most trip groups was large enough to give a reliable indication of the expansion factor. Further, in order to eliminate trip groups in which the total number of trips was small, the charts presented in the series of figures do not include trips between origin-and-destination pairs which had less than 25 trips in the roadside interviews or less than 6 postcards returned.

Exterior Stations

Figures 16 to 24 present in detail the comparison of the number of passenger car trips as obtained in the roadside interviews and in the postcard survey.

Figure 17, passenger cars inbound from all four exterior stations, is a typical spread of the expansion factor, in this case, from about 1.38 to 7.46 for the 29 listings of origin-and-destination pairs, with 2.38 as the factor for all trips combined. For these four stations combined, an expansion on the basis of 2.38 for all trips would state the number of trips from other counties to the campus-Pammel area as 172.5 percent of the true number and state the number of trips from the contiguous counties to the contiguous rural area as only 31.9 percent of the correct total. The other listings of origins to destinations would lie in between these two extremes.

Of the total of 2,376 cards returned (Fig. 17) only 593 (Items 9 to 14) would be expanded to within a 10-percent-plus-or-minus range from the roadside-interview total, should the expansion be based on the ratio of cards returned to total cards passed to 100 percent of the traffic.

A closer inspection of Figures 16 through 37 indicates that, in general, the pairs of origins and destinations having one terminal interior to the exterior stations have a lower expansion factor (a greater percentage return) than do these trips having both termini exterior to the stations; this result is particularly evident for those trips from one state through Iowa to another state.

Just how serious is this wide variation in expansion factors as related to particular classes of trips depends, of course, on each origin-and-destination survey and the specific applications of the results. That the number of cards returned by the passenger-car drivers was controlled to an appreciable extent by the particular trip, origin to destination, is evident from this analysis.

Figures 25 through 29 present the expansion factors for single-unit trucks. For these trucks, the expansion factor runs as high as 12.59, or 4.2 times the average factor. The general trends and sequence of items on the graphs is similar to that for the passenger cars.

Figures 30 and 31 give the expansion factors for the combination vehicles and busses. The number of busses is so small

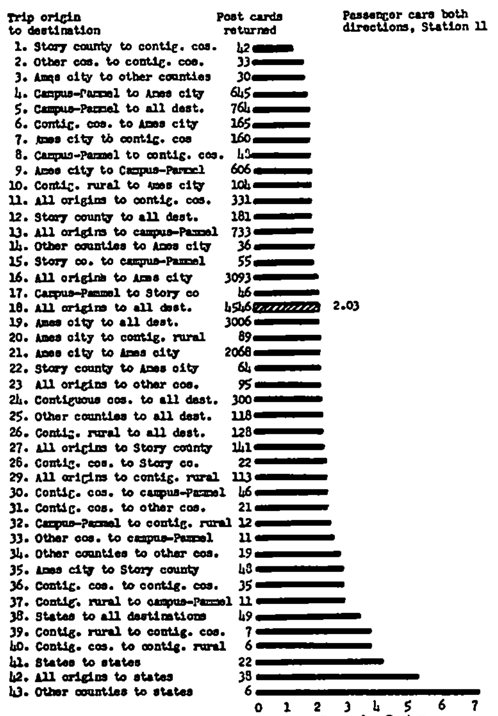

Figure 38.

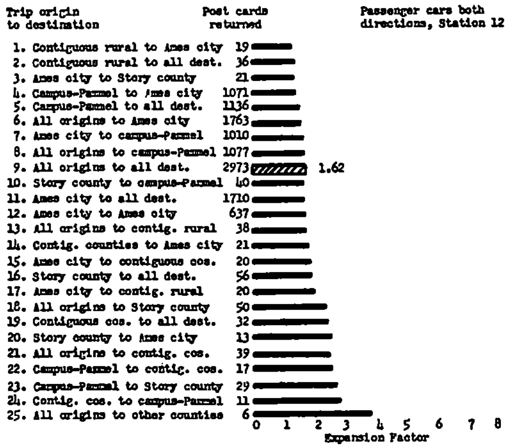

Figure 39.

(less than 5 percent) that these data represent the tractor-semitrailer class almost exclusively. The over-all expansion factor of 2.95 compares to 2.99 for the single-unit trucks, and 2.38 for passenger cars. The range of the expansion factor for inbound combinations varies from 1.35 to 6.85 and for outbound the range is 1.30 to 17.67. The generally small numbers of trips makes the expansion of the postcard returns for combination vehicles somewhat more uncertain than for the two other classes of vehicles.

The results for the combined four exterior stations, both directions, and all three classes of vehicles is given in Figures 36 and 37. With an over-all factor of 2.47, the range of 41 items of origin-and-destination pairs is from 1.12 to 7.65. A range from 10 percent below and above the average factor of 2.47 would include Items 15 through 24 (Fig. 36) totaling 2,129 trips or 37.9 percent of the total of 5,618.

Stations 11 and 12

Figures 38 through 42 show the expansion factors for Stations 11 and 12. There is a greater uniformity of rate of return from the several origin-and-destination pairs, as shown by Figure 38 for passenger cars, both directions, Station 11, than was achieved from the four exterior stations. The range of the expansion factor of the 43 items (including the lines of totals) is from 1.28 to 7.26. A range of 10 percent, plus and minus, from the average factor of 2.03 includes Items 6 through 26 for a total of 3,441 trips from a total of 4,546 postcards returned.

Station 12, on Sixth Street between the campus and eastern Ames, has an over-all expansion factor of 1.62 for both directions and a favorable range of from 1.30 to 3.82 (Fig. 39). A 10-percent variation covers Items 5 through 13 for a total of 1,687 trips from a total of 2,973 cards returned from passenger cars. The more-favorable range and distribution of the expansion factor for Station 12 no doubt is partially accounted for by the high percentage return of 58.5.

For single-unit trucks (Fig. 41) the range of the expansion factor is greater than for passenger cars and there is a less clustering of the items about the average expansion factor of 2.74 for Station 11.

Of the many charts presenting the expansion factors, Figure 42 for single-unit trucks on Station 12, presents the most-favorable distribution of factors. True, not many of the origin-and-destination pairs qualified (25 or more trips in the roadside interviews and at least 6 postcards returned), but such as are presented are within a range of 10 percent plus or minus of the average factor. The expansion factor for trips inbound to the campus is 2.80 and for the outbound trips the factor is 2.63. The reverse order of magnitude could be expected because of the convenience of returning the cards from the college offices through campus mail.

DISCUSSION OF RESULTS

The following items are presented as being significant results brought out in the

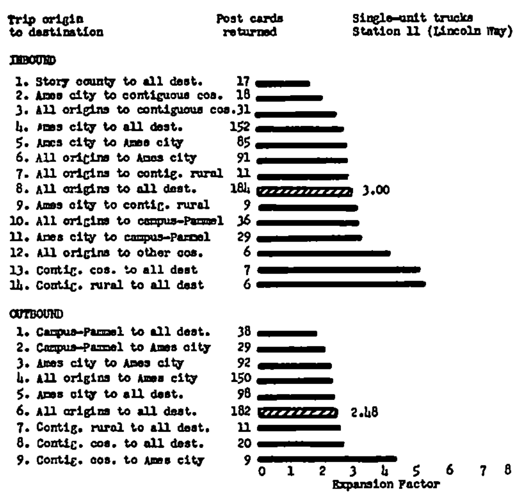

Figure 40.

preceding presentations of the information assembled from the postcard origin-and-destination study and a comparison of the results therefrom with similar results obtained by roadside interviews:

1. There is some tendency to a reducing percentage of postcards returned starting with the morning hour of 7 to 8, and continuing to midnight.

2. The 8-hr. period from 11 p.m. to 7 a.m. produced a significantly less percentage of returns than did the 16-hr. period from 7 a.m. to 11 p.m.

3. About equal percentages of returns were received from the inbound and outbound traffic streams. At least between the various hours and various stations there was no consistent pattern of differences.

4. A smaller percentage return of cards was received from the drivers of trucks than from drivers of passenger cars.

5. For the two interior stations, the locally owned and operated vehicles produced a higher percentage return than did the vehicles bearing a non-Story County license plate.

6. There was considerable variation in percentage return of cards, station to station, county to county of registration, but no explanations are evident of the cause of the variations. The one exception is that the high percentage return of 58.5 percent from Station 12 is attributed to special letters sent to the college employees and that many drivers through Station 12 were most willing to cooperate in the project because of personal relations, friendly and professional, with those in charge of the survey.

7. For the traffic volumes encountered and for the wide dispersion of origins and destinations, many pairs of origins and destinations had so few cards returned that

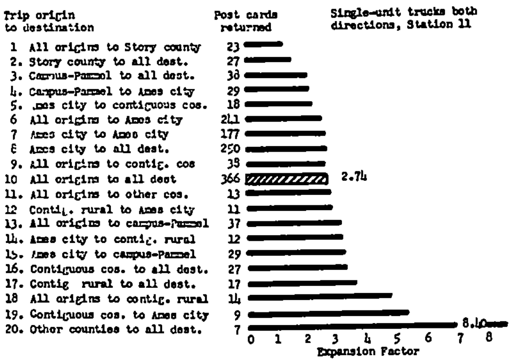

Figure 41.

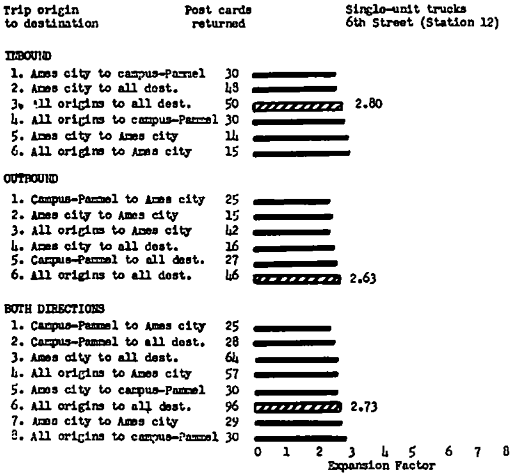

Figure 42.

any expansion of the small return would produce unreliable information. This situation was particularly true of the single unit trucks and of the combination vehicles.

8. There is a smooth range of expansion factors from roughly 1.3 to 8.0 for the postcards returned for particular pairs of origins and destinations to bring these returns up to the number of same trips found in the roadside interviews at the same stations for the 16-hr. period of 7 a.m. to 11 p.m.

9. There seems to be no specific rela-

Conclusions and Final Discussion

From the results of this study of the field postcard method of gaining origin-and-destination information of traffic, there is evidence that the expansion factors should be carefully determined, preferably by classes of vehicles, registration plate, and by hours of the day. Even so, there is no assurance that the expanded results would be in an acceptable agreement with the actual distribution of trips by origins and destinations. This study (Figs. 16 to 42) developed a wide variation of expansion factors for the various origin-and-destination-trip groups and without sufficient consistency that the tion between the expansion factor and the geographical location of the origin-and-destination pairs, other than, in general, the factor is lower when one termini is interior to the station than when both termini are exterior. Generally, trips from one state across Iowa to another state resulted in higher expansion factors than did other types of trips.

correct factors could be anticipated and

assumed ahead of knowing the true distribution of the trips.

Unfortunately, with the postcard method there is no means available of checking the correctness of the expanded returns once they are expanded to trips between specific origins and destinations.

There may be applications of the postcard method for which the approximations of the traffic distribution by origins and destinations as gained by the field postcard method would be sufficient, or surveys in which such high return of cards is obtained as to render errors of expansion within acceptable tolerances. Even so, it would appear that some roadside interviews should be conducted in order to get a basis of evaluation of the reliability of the postcard returns.

The returns for the 8-hr. night period were so small, that expansion of a station to a 24-hr. day on the basis of the 16-hr. survey would be as accurate as expanding the questionable small return for the light night and early morning hours.

Although the serial numbering of the cards and the manual enumeration of traffic in 2-min. periods was adopted for purposes of research and comparison with the roadside interviews, similar provisions would be advantageous for other surveys. Perhaps a 5- or 10-min. interval would be desirable rather than the 2-min. period used in this survey.

Sampling Methods for Roadside Interviewing

IRWIN MILLER, Graduate Research Assistant, Purdue University;
P.E. IRICK, Assistant Professor of Mathematics, Purdue University;
H.L. MICHAEL, Research Engineer, Joint Highway Research Project, Purdue University;
R.M. BROWN, In Charge Metropolitan Area Traffic Survey,
State Highway Commission of Indiana

PART I. SAMPLING THEORY

PART I of the paper is concerned with an appraisal of the sampling errors in the estimates of the trip frequencies for the various cells of an origin-destination traffic-survey tabulation.

A mathematical exposition is given of the expected errors when the sampling is done at random, by time clusters, and by volume clusters. The paper also discloses the results of an empirical investigation into the sampling errors that actually arose when various sampling methods were applied to the origin-destination tabulations of the Lebanon and Kokomo, Indiana, surveys. These experimental results were in general accord with the theory.

It is concluded that on the average, for the large number of estimates involved in any one survey, the theory of random sampling will satisfactorily explain the errors which arise from the various sampling methods proposed. The results given make it possible to predict the average errors and probability limits for these errors when a particular amount and type of sampling has been done. Conversely, one can determine the amount of sampling necessary to keep the sampling errors within specified probability limits.

PART II. PRACTICAL APPLICATION

IN this second part of the paper, the practical application of sampling to actual field conditions is presented. Station arrangement and sampling procedure for obtaining samples of 50 percent and 25 percent as developed during the conduct of an origin-destination survey at Richmond, Indiana, are discussed. Operational procedure for two sampling methods, volume cluster and time cluster, is given for two sample sizes and for locations having various roadway conditions. Sampling on two-lane, three-lane, and four-lane highways carrying traffic volumes ranging from 1,000 to 12,000 vehicles per day is shown to be both practical and economical. The advantages of each of the two sampling methods for various locations are enumerated.

It is concluded that the use of systematic sampling in the taking of roadside interviews is not only practical and economical but that it has definite operational advantages. The conclusions indicate that a predetermined sampling procedure will place interviewing on a business-like basis that (1) is recognized by the vehicle operator and the community; (2) promotes efficient operation of the interviewing stations; and (3) produces statistically sound results.

● IN recent years, origin-and-destination studies have been of prime importance in the planning, design, and operation of highway facilities. The data from such surveys provide a basis for determining the demand, location, type, and magnitude of new facilities, as well as important facts concerning their operation.

The methods of obtaining the desired data have necessarily been of recent development. Data collection within the survey area has been by means of home interviews, postcards, observations of license plates, route interviews, or combinations of two or more of these. All have been used with varying degrees of

success, but only a few actual studies of the relative accuracy and the advantages and disadvantages of each of these methods have been reported. Only the home interview method has received important study toward placing it on a scientific, statistically sound basis (5). One of the purposes of such studies has been directed toward the proper selection of an adequate sample that would be just large enough to produce accurate results in the most economical manner.

Origin-and-destination data of persons entering or leaving the survey area are generally collected by means of roadside interviews at an external cordon. The methods ordinarily used to obtain the necessary information at the cordon result in the selection of a sample which, it is assumed, has been selected in a random manner. In a large majority of the past surveys it is probable that an unbiased sample was selected. A high percentage of the total traffic, however, was often interviewed, so as to minimize the possibilities of obtaining a bias. It is clearly recognized that such almost total sampling is unwarranted if the sample selection is truly random.

The "Manual of Procedures for the Metropolitan Area Traffic Studies" of the Bureau of Public Roads states that "Representative samples of traffic passing each station in both directions should be interviewed." The same manual suggests that in selecting the representative sample:

> No attempt should be made to select for interview a definite predetermined percentage of vehicles passing each station. The number of interviews obtained will depend entirely on the number of interviewers assigned to the station and the number of vehicles passing. The recommended policy is to assign sufficient personnel to each station to interview approximately 50 percent of the traffic passing during peak periods. During the remaining hours of operation the crew will, of course, be able to interview a greater percentage of vehicles passing the station.

The question, however, arises if the selection of such a sample is truly unbiased. Do the interviewers, as good as they may be, have the ability to select an unbiased sample in the field? Do some interviewers select a preponderance of obviously easy interviewees, local vehicles, vehicles containing friends or relatives, or vehicles containing attractive girls? Do some interviewers select too-few foreign vehicles, trucks, or vehicles whose drivers were previously interviewed? Are interviewers prone to pass some vehicles through the station in the interest of less work for themselves? Under such a method, are the vehicles that are interviewed selected randomly over the hour or over that period for which factors are later to be computed?

These questions, undoubtedly, cannot always be truthfully answered in a manner which assures an unbiased sample. Furthermore, the question arises as to what size of sample will give adequate results. Certainly there can be no warrant to take a 75-percent sample if a smaller sample will give the necessary information as accurately as required.

It is generally recognized that a method which would assure the selection of a random sample would be of great value to roadside interviewing. The purpose of this paper is to examine the errors which could be expected from the use of various sampling rates and sampling methods and the practicality of these methods under field conditions.

PART I. SAMPLING THEORY

Irwin Miller and P. E. Irick

● AT a particular external-survey station, incoming traffic can be considered to originate at the station and have d, say, different destinations. Outgoing traffic destined for this station can have d different origins if d+1 is the total number of internal tracts and external stations. Vehicles passing through a single station can have one of second origin-destination (O. D.) combinations. For any one of these O. D's we define the universe to consist of the N vehicles which could, by virtue of the direction of traffic flow, have had the specified O. D. during the time that the survey station is open. One universe serves for d O. D. cells in the (rectangular) O. D. tabulation. If this universe is sampled in some predetermined manner, the sample estimates of the universe O. D. frequencies are subject to chance variation. It is our

purpose to appraise these errors of estimation. We shall first consider the sampling theory that evolves when one makes certain assumptions about the universe and the nature of the sampling. Secondly, we shall show the results of an empirical investigation into these sampling errors.

RANDOM SAMPLING

Let T be the number of vehicles in a universe of N vehicles which have a particular O.D. If a sample of n vehicles is chosen at random from the N, and t of these have the O.D., then

$$\hat{T} = \frac{N}{n} t \qquad (1)$$

is an unbiased estimate of T. The sampling variance of $\hat{T}$ is known (1, p115) to be

$$\text{var}(\hat{T}) = \left(\frac{N - n}{N - 1}\right) \frac{N^2 pq}{n} \qquad (2)$$

where p = T/N, the universe proportion having the O.D., and where q = 1 - p. We shall consider the coefficient of variation of $\hat{T}$, CV($\hat{T}$), to be a measure of the percentage error in $\hat{T}$ which arises because of sampling variation. This measure is defined by

$$CV(\hat{T}) = \frac{[\text{var}(\hat{T})]^{1/2}}{T} \qquad (3)$$

where Equation 3 gives the decimal equivalent of the percentage error. Substitution of Equation 2 into 3 gives

$$CV(\hat{T}) = \left[\left(\frac{N - n}{N - 1}\right) \frac{N}{n} \left(\frac{1}{T} - \frac{1}{N}\right)\right]^{1/2} . \qquad (4)$$

If the sampling rate is designated by r=n/N, and if N is large relative to T, we have the approximate formula

$$CV(\hat{T}) = \left[\left(\frac{1 - r}{r}\right) \frac{1}{T}\right]^{1/2} . \qquad (5)$$

Since the universe is not defined until after the period of the survey, there is no opportunity to construct a chance mechanism for the selection of a random sample of n vehicles. One might regard any subset of the universe as being a random sample, but in such a case there would be no assurance that the foregoing formulas were applicable. As soon as one turns to some objective method for selecting the subset, then it becomes necessary to study the sampling variances associated with this method. We shall discuss various cluster sampling methods.

CLUSTER SAMPLING

General Theory

Let the universe be divided into G clusters of vehicles, and let the sample consist of g of these clusters. Each cluster is to be completely interviewed, and so there is no sampling variation within clusters. It could be that the g clusters are drawn at random, or that they are chosen systematically, taking every k^{th} universe cluster into the sample. Let t_i be the number of vehicles in the i^{th} cluster (i = 1, 2, . . . , G) which have the specified O.D. and let $t = \sum_{i=1}^{g} t_i$ be the number of vehicles with the O.D. which are found in the sample. If the expected or average value of each t_i, with respect to all possible samples of the type selected, is T/G, where $T = \sum_{i=1} t_i$, then

$$\hat{T} = \frac{G}{g} t \qquad (6)$$

is an unbiased estimate of T. The sampling variance of $\hat{T}$ is given (1, p. 193) by

$$\text{var}(\hat{T}) = \left(\frac{G - g}{G - 1}\right) \frac{G^2}{g} \text{var}(t_i) \qquad (7)$$

where

$$\text{var}(t_i) = \frac{1}{G} \sum_{i=1}^{G} \left(t_i - \frac{T}{G}\right)^2 . \qquad (8)$$

The cluster to cluster variance, var (t_i), is a universe parameter peculiar to the universe and the clustering device which has been used. Formula 7 is valid when the g clusters in the sample have been chosen at random from the G available clusters. If the sample clusters are taken systematically (2, p. 157),

$$\text{var}(T) = \left[1 + \frac{2}{g} \sum_{j=1}^{g-1} (g-j)\, \rho_{jk}\right] \frac{G^2}{g} \text{var}(t_i) \qquad (9)$$

where ρ_{jk} is the serial correlation coefficient defined by

$$\rho_{jk} = \qquad (10)$$

$$\sum_{i=1}^{k(g-j)} \left(t_i - \frac{T}{G}\right)\left(t_{i + jk} - \frac{T}{G}\right) \Big/ k(g - j)\, \text{var}(t_i).$$

If the ρ_{jk} values are negative, the implication is that sample clusters with unusually high t_i values will be compensated for by other clusters in the sample with lower than expected t_i. In fact, if the ρ_{jk} are sufficiently negative, $var(\hat{T})$ from systematically chosen clusters will be less than $var(\hat{T})$ from randomly chosen clusters. The equality of these two variances occurs when

$$\sum_{j=1}^{g-1} (g-j)\, \rho_{jk} = \frac{-g(g-1)}{2(G-1)} \quad . \qquad (11)$$

There is empirical evidence, to be mentioned later, that Equation 11 may be satisfied on the average for the various O.D. cells in a given survey when particular types of clusters are used. More investigation of this sort should be done, but in the present discussion we shall use Equation 7 to describe the sampling variation in $\hat{T}$. Substituting Equation 7 into 3 we have the expected error for cluster sampling.

$$CV(\hat{T}) = \left[\left(\frac{G-g}{G-1}\right)\frac{G^2}{g}\ \frac{var(t_i)}{T^2}\right]^{1/2} . \qquad (12)$$

TABLE 1

CHARACTERISTICS OF SAMPLING VARIATION IN TIME CLUSTER MODELS

Time Cluster	G	$E[var(t_i)]$	$[CV(\hat{T})]^2$ g=1	g=2	g=3	95% limits for $var(t_i)$	$[CV(\hat{T})]^2$ g = 1	g = 2	g = 3
30 min	2	25T	$\frac{1}{T}$			96T	$\frac{3.89}{T}$		
15 min.	4	19T	$\frac{3}{T}$	$\frac{1}{T}$		49T	$\frac{7.81}{T}$	$\frac{2.60}{T}$	
10 min	6	14T	$\frac{5}{T}$	$\frac{2}{T}$	$\frac{1}{T}$	.31T	$\frac{11.07}{T}$	$\frac{4.43}{T}$	$\frac{2.21}{T}$

If a sample of g clusters has been taken, $var(t_i)$ can be estimated by

$$var(t_i) = \frac{(G-1)}{G(g-1)} \sum_{i=1}^{g} \left(t_i - \frac{t}{g}\right)^2 . \qquad (13)$$

The question arises as to whether one can make any appraisal of $var(t_i)$ before a sample is taken. This can be done only if one assumes some mathematical model or theoretical mechanism for any particular clustering procedure. One mechanism described in the literature (1, p. 198) assumes that the clusters obtained in any one sample represent a random selection from all possible clusterings of the N universe

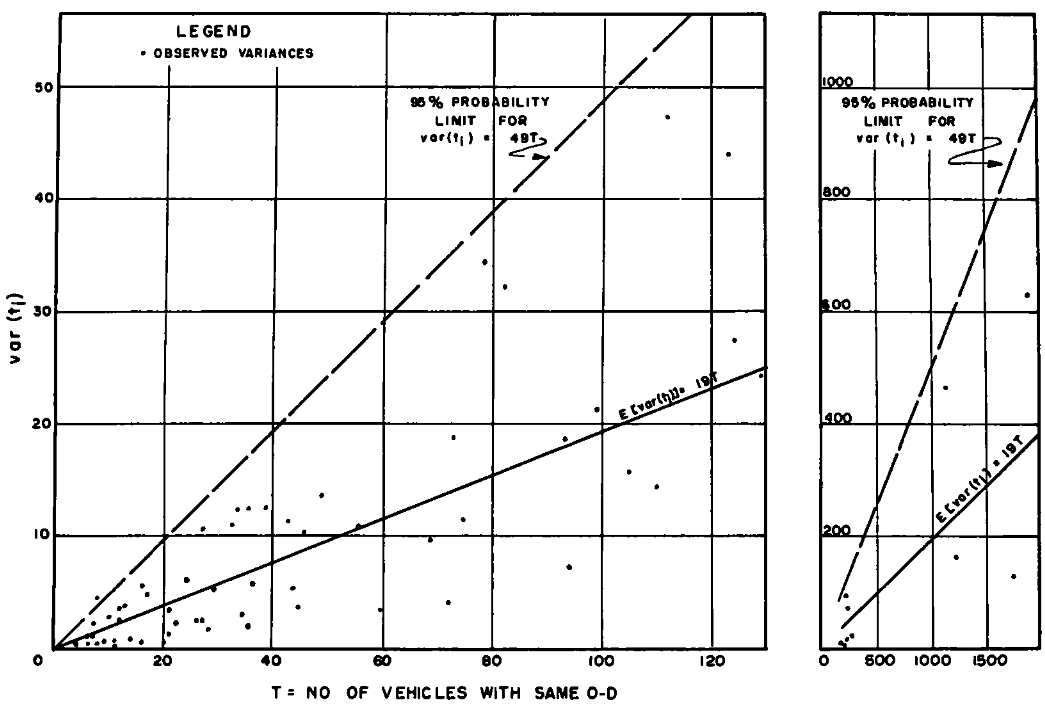

Figure 1. Observed quarter-hour-cluster variances (Lebanon universes).

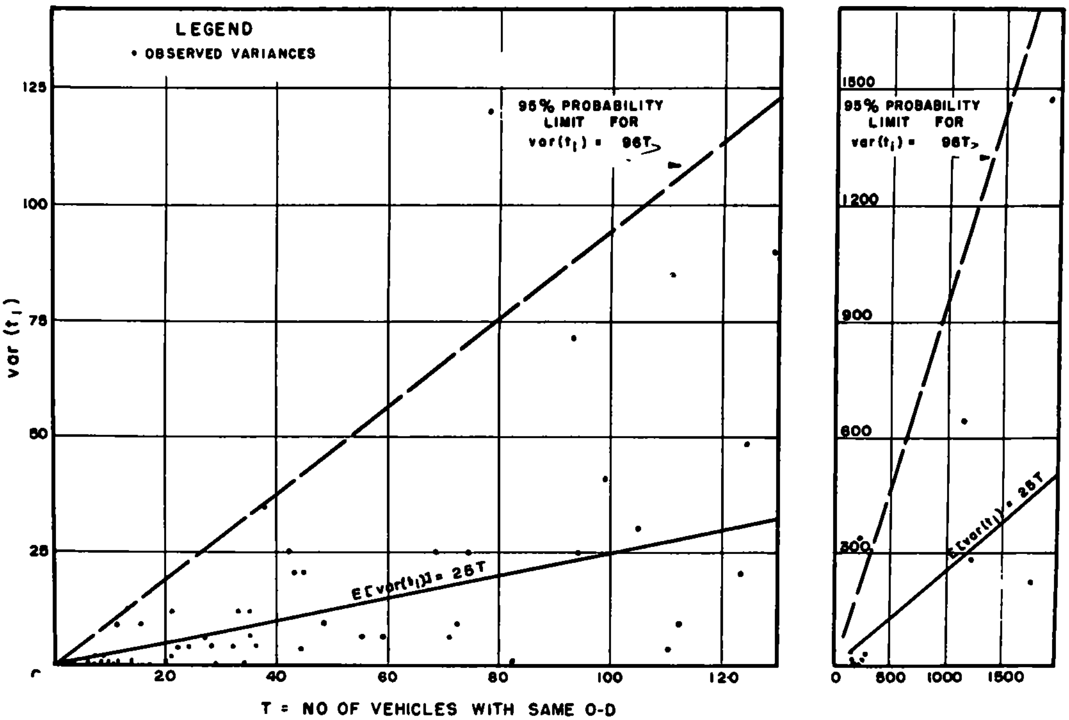

Figure 2. Observed half-hour-cluster variances (Lebanon universes).

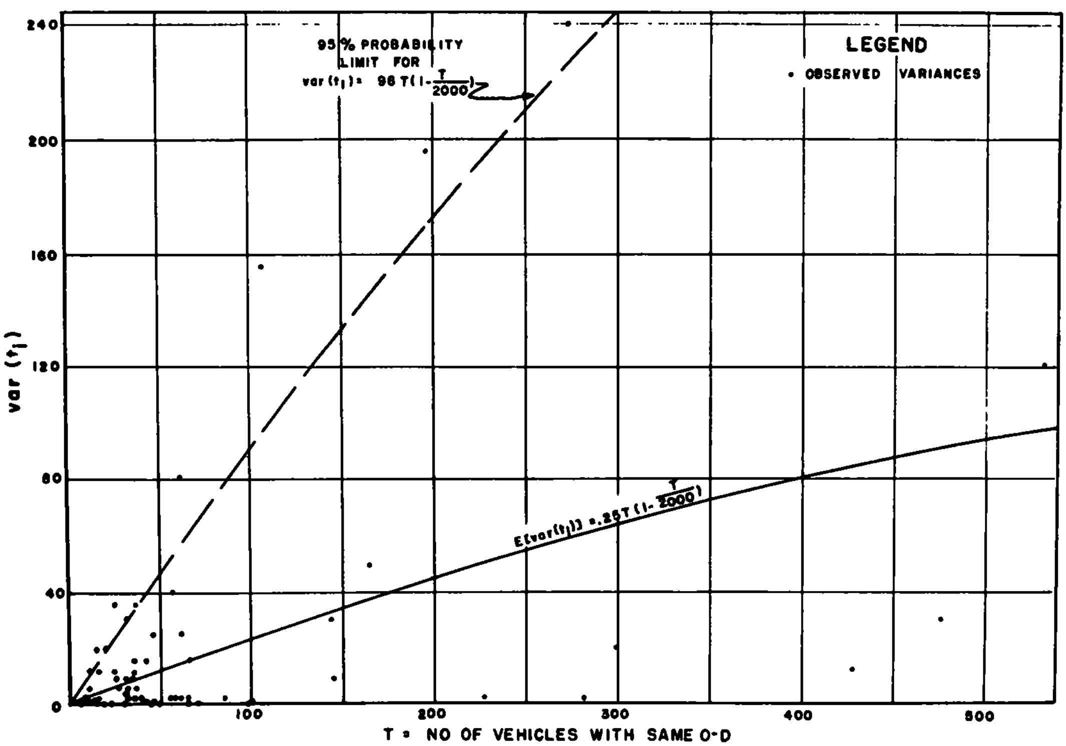

Figure 3. Observed volume-cluster variances (Kokomo universes).

cars into the G clusters. This supposition leads to an average or expected value

$$E\left[var(t_i)\right]\left(\frac{G-1}{G^2}\right)T(1-\frac{T}{N}) . \qquad (14)$$

We shall discuss two additional models which seem to be intimately related to practical sampling procedures.

Time Clusters

In this method of sampling the survey period is divided into hours, then each hour into the same number of equal time intervals, say 15 min. Let G be the number of intervals in each hour, and t_i be the total number, over all hours, of vehicles which have the O.D. in question and which are found in the i^{th} time interval. Then the extent of sampling variation, where the sample is to consist of g from the G clusters, is dependent upon $var(t_i)$ as in Equation 7 or 9 depending upon how the g clusters are selected. To estimate $var(t_i)$ we shall assume that the T cars having the O.D. represent independent "trials" such that there is probability 1/G for each car to fall into any one of the G intervals. Under this assumption it can be shown that the expected value of each t_i is T/G, and from this that Equation 6 gives an unbiased estimate of T. Furthermore, from Equation 8,

$$var\,(t_i) = \frac{T}{G^2}\chi^2, \text{ where } \chi^2 = \sum_{i=1}^{G}(t_i - \frac{T}{G})^2/\frac{T}{G}. \qquad (15)$$

χ^2 follows the chi-squared probability law with G - 1 degrees of freedom. Since the expected value of χ^2 is G - 1, we have

$$E\left[var(t_i)\right] = \left(\frac{G-1}{G^2}\right)T, \qquad (16)$$

and the corresponding expected error is, from Equations 16 and 12,

$$CV(\hat{T}) = \left[\left(\frac{G-g}{g}\right)\frac{1}{T}\right]^{1/2} . \qquad (17)$$

Probability limits for $var(t_i)$ can be set using tables of chi-square. Table 1 shows several values for Equation 16, upper 95-percent probability limits for $var(t_i)$, and corresponding values of the squared coefficient of variation for several sample sizes.

Among the various O.D. cells in any one survey, it is almost certain that one will encounter varying degrees of lack of independence of the type assumed. It would seem that any theoretical model can only be true "in the large" and that its appropriateness can best be judged in the light of empirical evidence. Figure 1 shows quarter-hour variances for 68 cells in an O.D. tabulation to be described later (Lebanon). Figure 2 shows the half-hour variances for the same cells. The solid and dotted lines are plotted from expected values and probability limits given in Table 1. The serial correlations for the quarter hour clusters in these cells varied considerably, but averaged to be -0.33, which happens to be the value specified by Equation 11.

Tables 2 and 3 give the data for Figures 1 and 2 respectively. The t_i values are vehicle counts over a 12-hr. period of time.

Volume Clusters

Suppose that the universe of N cars is divided into consecutive sets of c cars each so that there are N/c (approximately) such sets. We may take c to be the number of interviewers on duty at the survey station. Let g such sets be interviewed, then G - g sets waved through the station without being interviewed, etc. Then the universe has been clustered into G clusters, g of which appear in the sample. The i^{th} such cluster has t_i vehicles which have the O-D in question and $(N/G)-t_i$ which do not. We assume that, for every vehicle, the probability of having the O-D, p=T/N, is independent of the probability of being in any particular cluster, and that the latter probability is 1/G. This model then describes a 2 x G contingency table (2, p. 227) and, under the assumptions,

$$\chi^2 = \frac{G^2}{qT}\,var(t_i) \qquad (18)$$

has nearly the probability distribution of chi-squared with G - 1 degrees of freedom. The expected value of $var(t_i)$ is as in Equation 14, and using this average value in Equation 12 we obtain

$$CV(\hat{T}) = \left[\left(\frac{G-g}{g}\right)\left(\frac{1}{T}-\frac{1}{N}\right)\right]^{1/2} \qquad (19)$$

where $\hat{T}$ is calculated from Equation 6. Again we have considered the systematic

TABLE 2

QUARTER HOUR VARIANCES - LEBANON DATA

T	t_1	t_2	t_3	t_4	var(t_1)	ρ 12	T	t_1	t_2	t_3	t_4	var(t_1)	ρ 12
4	1	0	1	2	5	-1 00	36	11	5	10	10	5 5	- 18
6	1	1	3	1	5	00	38	6	7	10	15	12 2	- 63
6	1	2	0	3	1 2	60	42	5	11	14	12	11 2	- 82
7	2	0	3	2	1 2	- 05	43	14	12	10	7	5 1	- 58
7	1	2	4	0	2 2	- 49	44	13	11	8	12	3 5	- 85
8	2	2	3	1	5	00	45	11	16	7	11	10 2	- 06
8	0	3	0	5	4 5	78	48	15	12	15	6	13 5	33
9	3	3	2	1	7	- 82	55	19	11	14	11	10 7	41
10	1	3	5	1	2 8	- 82	59	15	12	15	17	3 2	- 87
11	3	3	3	2	2	- 33	68	19	20	12	17	9 5	- 53
11	2	3	2	4	7	63	71	17	21	17	16	3 7	- 69
12	2	1	6	3	3 5	- 42	72	22	11	18	21	18 5	- 46
12	5	2	4	1	2 5	80	74	15	17	24	18	11 2	- 82
13	4	6	1	2	3 7	- 69	78	16	12	23	27	34 2	-1 00
14	4	4	2	4	8	- 33	82	20	20	13	29	32 2	- 01
16	8	3	2	3	5 5	- 64	93	20	18	27	28	18 7	- 99
16	4	4	3	5	5	00	94	27	25	22	20	7 2	- 72
17	4	5	7	1	4 7	- 33	99	17	26	29	27	21 2	- 71
20	5	4	5	6	5	-1 00	105	31	27	20	27	15 7	- 93
21	5	4	5	7	1 2	- 89	110	31	22	26	31	14 2	- 86
21	4	3	7	7	3 2	- 96	112	30	29	36	17	47 5	05
22	7	6	3	6	2 2	- 78	123	35	22	39	27	44 2	77
22	8	5	4	5	2 2	- 78	124	28	27	40	29	27 5	- 35
24	10	4	4	6	6 0	- 67	129	26	29	38	36	24 2	- 99
26	8	5	8	5	2 2	1 00	195	47	47	47	54	9 2	- 67
27	9	7	5	6	2 2	- 94	198	46	52	48	52	6 8	85
27	7	4	12	4	10 7	41	203	54	47	56	46	18 7	- 02
28	8	8	5	7	1 5	- 67	205	64	57	38	46	99 7	-1 00
29	5	10	5	9	5 2	95	226	62	65	66	43	76 2	21
32	11	3	7	11	11 0	- 82	236	53	59	58	66	21 5	30
33	14	6	8	5	12 2	24	1111	303	278	286	244	461 2	22
34	7	9	7	11	2 8	64	1218	303	323	305	287	162 8	-1 00
35	12	9	11	3	12 2	24	1758	425	429	457	437	130 8	- 97
35	10	10	7	8	1 7	- 92	1895	478	508	472	437	636 2	-1 00

TABLE 3

HALF HOUR VARIANCES - LEBANON DATA

T	t_1	t_2	var(t_1)	T	t_1	t_2	var(t1)
4	1	3	1.0	36	16	20	4 0
6	2	4	1 0	38	13	25	36 0
6	3	3	0	42	16	26	25.0
7	2	5	2 2	43	26	17	20 2
7	3	4	.2	44	24	20	4 0
8	4	4	0	45	27	18	20 2
8	3	5	1.0	48	27	21	9 0
9	6	3	2 2	55	30	25	6 2
10	4	6	1.0	59	27	32	6 2
11	6	5	.2	68	39	29	25 0
11	5	6	2	71	38	33	6 2
12	3	9	9 0	72	33	39	9 0
12	7	5	1.0	74	32	42	25 0
13	10	3	12 2	78	28	50	121.0
14	8	6	1 0	82	40	42	1 0
16	11	5	9.0	93	38	55	72 2
16	8	8	.0	94	52	42	25.0
17	9	8	.2	99	43	56	42 2
20	9	11	1 0	105	58	47	30.2
21	9	12	2 2	110	53	57	4.0
21	7	14	12 2	112	59	53	9 0
22	13	9	4.0	123	57	66	20 2
22	13	9	4 0	124	55	69	49.0
24	14	10	4 0	129	55	74	90.2
26	13	13	0	195	94	101	12 2
27	16	11	6 2	198	98	100	1 0
27	11	16	6.2	203	101	102	.2
28	16	12	4 0	205	121	84	342 2
29	15	14	.2	226	117	109	16 0
32	14	18	4.0	236	112	124	36 0
33	20	13	12.2	1111	581	530	650 2
34	16	18	.2	1218	626	592	289.0
35	21	14	12 2	1758	864	894	225.0
35	20'	15	6.2	1895	986	909	1482 2

selection of the g clusters to be equivalent to the corresponding random selection. Comparison of Equation 19 with 17 and Equation 18 with 15 shows that Table 1 can be used for this volume cluster model provided the factor $q = 1 - \frac{T}{N}$ is inserted as a multiplier for each entry there. Figure 3 and the accompanying data in Table 4 shows 68 observed values of var(t_1) from Kokomo universes where G = 2, g = 1, N = 2,000 (approx.). These variances, however, are not from true volume clusters as will be pointed out later. Table 1, modified as just stated, gives the solid and dotted curves in Figure 3.

In summary it should be pointed out that Formula 14 substituted into Equations 12, 16, and 19 are very similar, and in fact, that all three are practically equivalent to Formulas 4 and 5. This is so because the assumptions made for cluster sampling have led to the conclusion that we shall expect cluster sampling errors to average to be the same as random sampling errors even though the two types of errors may be different for a single universe. Having discussed the theoretical sampling variation to be expected through the use of these methods, we now turn our attention to the observed sampling errors which arose when these methods were applied.

DESCRIPTION OF EMPIRICAL INVESTIGATIONS

The results of census surveys taken at Lebanon and Kokomo, Indiana, comprise the universes from which samples were drawn in an empirical study of sampling variances. It was impracticable to interview every vehicle when traffic was unusually heavy, but over 90 percent of the traffic was interviewed in each survey, making it reasonable to regard these surveys as complete. Table 5 shows the O. D. frequencies for the Lebanon survey.

Traffic at Lebanon was interviewed at eight different external stations, and the area was divided into 24 internal tracts. Vehicles passing through Station 8, for example, could have had 24 internal-external O. D. combinations, and 14 through-trip combinations (7 inbound and 7 outbound)

or 62 in all. For the O.D. Cell 8-010, the universe consists of N = 4703.4 vehicles inbound through Station 8. For the O-D Cell 8-3 the universe has been taken to consist of the 4,703.4 vehicles which went through Station 8 in the inbound direction, plus the 947.7 vehicles passing through Station 3 in the outbound direction, or 5,651.1 vehicles in all. As can be seen in Table 5, the respective T values for these cells are 91.6, and 184.8. These values are not actual counts, since the counts were adjusted for day to day and other variations by means of several factors (3).

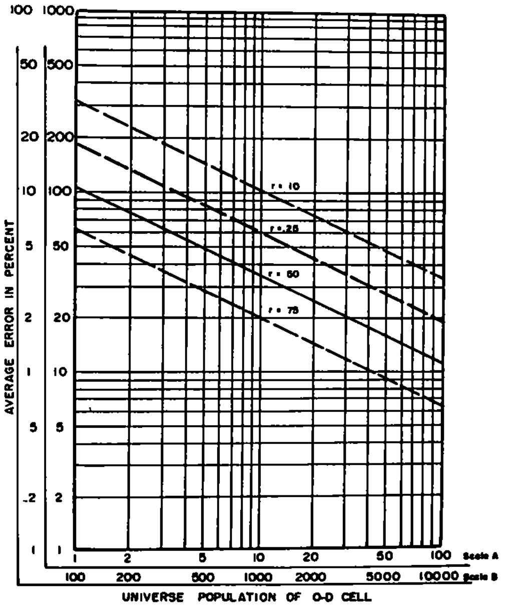

Figure 4. Expected errors due to uniform sampling (for various sampling rates).

Types of Samples

The data for each universe interview was recorded on a punch card. The card showed the quarter hour of the hour in which the interview was taken and serial numbers were assigned to the cards within the quarter hours. However, these serial numbers were not necessarily in the order of occurrence of the interview, due to the fact that several interviews were being conducted simultaneously. Consequently, time-cluster samples were limited to rates of 25, 50, and 75 percent, and it was felt reasonable only to consider the volume cluster sample which compares even and odd serial numbers. Such rates are de-

TABLE 4

VOLUME CLUSTER VARIANCES - KOKOMO DATA

T	t_1	t_2	$var(t_i)$	T	t_1	t_2	$var(t_i)$
6	2	4	1.0	39	21	18	2.2
7	3	4	.2	40	19	21	1.0
7	3	4	2	42	20	22	1.0
8	5	3	1 0	42	20	22	1.0
9	6	3	2.2	42	17	25	16.0
10	5	5	0	45	23	22	.2
10	4	6	1 0	46	28	18	25.0
11	8	3	6 2	46	22	24	1.0
11	2	9	12 2	49	24	25	.2
14	8	6	1 0	55	29	26	2.2
15	12	3	20.2	57	19	38	40.2
15	8	7	.2	57	30	27	2 2
17	10	7	2.2	61	29	32	2 2
17	12	5	12.2	62	40	22	81.0
18	9	9	.0	62	36	26	25.0
19	10	9	2	65	33	32	.2
21	6	15	20 2	65	31	34	2.2
23	12	11	.2	66	29	37	16.0
25	16	9	12.2	71	35	36	.2
26	10	16	9.0	85	44	41	2.2
26	7	19	36.0	97	48	49	2
27	11	16	6.2	100	49	51	1 0
29	14	15	2	107	41	66	156.2
30	18	12	9.0	143	77	66	30 2
30	13	17	4 0	144	69	75	9.0
31	14	17	2 2	164	89	75	49 0
31	10	21	30.2	196	112	84	196.0
31	13	18	6 2	227	112	115	2.2
34	20	14	9 0	273	152	121	240 2
35	14	21	12.2	281	139	142	2 2
35	16	19	2.2	299	145	154	20.2
36	14	22	16 0	427	210	217	12.2
37	24	13	36.0	475	243	232	30.2
37	21	16	6.2	532	255	277	121.0

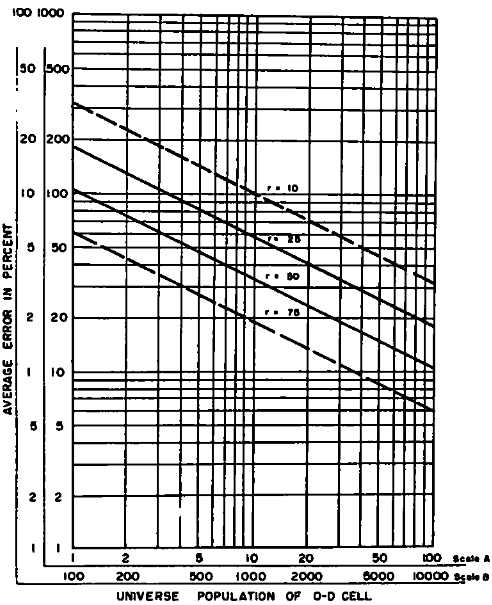

Figure 5. Expected errors due to quarter-hour sampling.

TABLE 5

ORIGIN AND DESTINATION UNIVERSE TABULATION OF NUMBER OF VEHICULAR TRIPS FOR LEBANON BY-PASS "BEFORE" STUDY

Internal Tracts - Destinations × External Stations - Origins

Internal Tracts - Destinations	1	3	5	8	10	11	13	14
001	3 7	1 1	1 1	2 2	1 5	1 3	0 0	0 0
002	16 7	2 3	1 3	15 3	1 2	0 0	2.2	6 3
003	6 7	4 7	11.1	24 9	5 3	4.0	1.3	10.9
004	16 5	17 7	29 2	40 4	23 3	10 1	4 9	35 0
005	5 5	2 4	10 2	24 8	5 9	10 0	9.2	15 1
006	2 4	12 9	11 3	30 9	14 5	7 8	11 7	18 9
007	1 0	21.1	15 6	9 0	5.6	0.0	0 0	1 0
008	39 4	40 9	62 3	51 6	13 2	22 5	14.3	59 8
009	8 1	17 1	19 4	42 4	6 2	7 4	12 8	24 7
010	41 1	50 9	57 0	91 6	33 3	42.7	26 6	67 1
011	2 4	3 9	0 9	14 1	7.1	3 0	1 8	12.2
012	8.5	21.4	17 0	32 3	6 1	3 2	3 3	18 8
013	4 1	19.0	29 7	11 9	5 6	7 3	7 2	18 1
014	20 8	49.4	48 5	71 1	11 7	13 9	7 9	30 5
015	163 2	193.9	203 0	278 7	97.3	100 3	74 5	167 6
016	12 9	26 5	76.3	39 9	18 2	29 7	29 8	33.3
017	16 8	16 0	28 4	84 9	18.5	36 9	15 6	27 3
018	2 2	2.8	11 0	7 2	1.1	0 0	0 0	0.0
019	15 1	17 9	27 0	56 5	18 4	19 0	6 0	23 8
020	6 5	3 7	4.5	29 2	5 3	1 1	4 7	16 5
021	9 1	20 8	33 0	46 3	18 5	16 7	10 9	35 8
022	1 3	1 0	16 5	6 2	1 3	0 0	0 0	2 9
023	2 3	2 9	6.3	19 9	6 7	5 4	18 4	6 5
024	1 1	1 3	2 6	20 5	9 7	2 7	0 0	3 8

Internal Tracts - Origins × External Stations - Destinations

Internal Tracts - Origins	1	3	5	8	10	11	13	14
001	5.4	0 0	1 2	4 4	0 0	0.0	0.0	0 0
002	7 3	4 5	1 5	16 6	4 5	2 4	1.1	12.0
003	10.2	9.6	11 7	30 4	3.0	4.5	3 7	9 8
004	15 9	19 9	32 3	34 5	16 9	6 5	4.3	18.4
005	6 1	1.1	5 5	32 5	4.4	2.0	11.9	9.2
006	5 5	10.4	9.8	32.2	8.0	9.3	9.3	29 0
007	4 3	7.8	10 0	2.4	3.1	0.0	0 0	1.1
008	49 9	55 9	59 2	54 8	15 9	15.6	17.8	54 3
009	14.5	21.5	20 0	48 9	10 5	3 4	11 7	24.8
010	36 1	48.9	76 1	83 7	30 6	21 0	20 9	68 9
011	5.4	4.3	0.0	13 9	5.5	5.1	2.7	13 0
012	11 1	27 7	13 9	25 8	7 1	1.7	3 6	21 8
013	4 9	23 5	25 4	10 4	7 3	7 4	5 0	22.2
014	25.2	37.8	40.9	63 9	14 9	11 5	12 2	31 7
015	134 1	129 9	199 5	242 5	78.1	104 8	72 1	47 9
016	14 3	26 8	61.4	41 6	31 8	38 1	25.0	29 1
017	17 4	26 9	29 0	77 4	25 4	23 3	19 7	36 0
018	4 1	1 0	5 3	7 7	3 8	0.0	0.0	2.3
019	15 5	22 4	34 2	52 3	19 3	17 8	11 4	32 2
020	7 4	1 3	6 7	23 7	9.3	1.2	2 2	11 8
021	16 2	14 6	30 3	41 9	18 8	17.9	12 7	40 3
022	3 0	0 0	8 7	9.2	1.3	0 0	0 0	1 2
023	1 0	2.3	6 9	19 0	5 5	5 2	13 9	4 6
024	4 0	0 0	3 0	25.5	2.5	0 0	0 0	3 2

External Stations - Origins × External Stations - Destinations

External Stations - Origins	1	3	5	8	10	11	13	14
1		12 8	67 8	4175 6	102 8	10.6	3 6	0 0
3	4.4		26 4	206 8	220 2	7 6	6 6	31 6
5	62.0	27 2		38 4	65 6	14 4	15 6	111 0
8	3949 6	184 8	25 2		29 2	8 4	8 0	498 2
10	132 4	168.8	62 6	30 0		3 2	1 2	198 4
11	13 2	4 6	9 4	11 4	2 6		0.0	12 0
13	0.8	7.2	18 0	9 0	1 2	1 2		1 0
14	0 0	44 2	131 6	523 8	218 0	8 0	2.4	

sirable, however, in that they permit reasonable frequencies to appear in cells with small T values. The following samples were selected for study.

Sample 1 is a volume cluster sample, each cluster consisting of $\frac{N}{2}$ vehicles. The sample selected represented the even serial numbers, and thus is a 50-percent sample with G = 2, and g = 1.

Sample 2 is a quarter-hour time-cluster sample. A 50-percent sample was obtained by choosing vehicles which were interviewed in alternate directions by quarter-hour time intervals. Thus, G = 4 and g = 2 in this case.

Sample 3 differs from Sample 2 only in the use of a half-hour time interval, or in that G = 2 and g = 1. Both Samples 2 and 3 would be expected to provide 50-percent samples, since the inbound universe was sampled in the periods during which the outbound universe was not sampled, and vice versa.

Sample 4 combines both the volume cluster technique of sample 1 and the time cluster sampling of sample 3. In this sample the even numbered vehicles were interviewed every other half hour, yielding an expected 25 percent sample.

Sampling Errors

After a drawing of any of these samples, $\hat{T}$ was computed from Equation 6 for each cell, and compared with the corresponding universe frequency, T. The relative error of estimate in this cell is given by the formula

$$e = \frac{|\hat{T} - T|}{T}. \tag{20}$$

Table 6 shows the relative errors in percent arising from Sample 2 drawn from the Lebanon universes. Similar tables have been constructed for the remaining samples from both surveys (4).

The variance of t_i has been computed directly from the universe data in the case of cluster sampling. Graphs of these values versus T are typified by Figures 1, 2, and 3, and show that var(t_i) is approximately proportional to T. Then it follows from the theory that

$$CV(\hat{T}) = \sqrt{\frac{K}{T}} \tag{21}$$

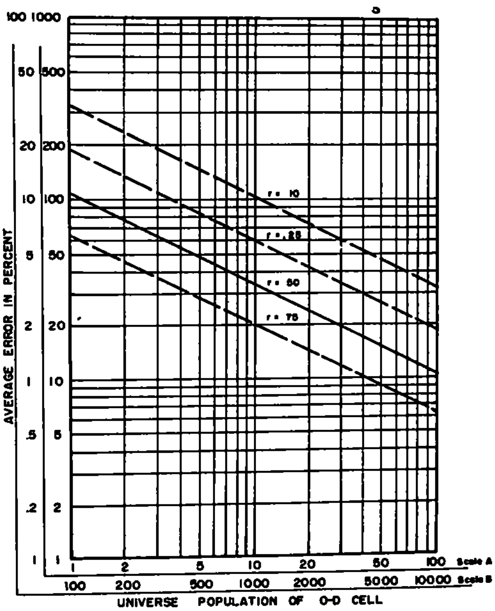

Figure 6. Expected errors due to half-hour sampling.

where K is a constant which will be left for the data itself to determine. In order to estimate K, let us first make the following change of variables. Let

$$Y = \left[CV(\hat{T})\right]^2 ; \quad X = \frac{1}{T} . \tag{22}$$

When this is done, Equation 21 becomes

$$Y = KX. \tag{23}$$

Now each complete sample, as illustrated in Table 6, gives rise to s numbers e_j, one for each O. D. cell, and the universe tabulation (e.g., Table 5) provides s corresponding values T_j. (The subscripts j are used to indicate that these values came from the j^{th} O.D. cell). Since e_j is the only available estimate of $CV(\hat{T})$ in the j^{th} cell, we shall compute $Y_j' = e_j^2$ our estimate of Y_j, from this number. K is estimated by the method of least squares, using the formula

$$\hat{K} = \sum_{j=1}^{s} X_j Y_j' \Big/ \sum_{j=1}^{s} X_j^2 . \tag{24}$$

Under the assumption that the errors arising in the j^{th} cell from all possible samples of the type selected are normally distributed with mean zero and standard deviation $CV(\hat{T}_j)$, it is found (4) that the variance of $\hat{K}$ can be estimated with the aid of the statistic

$$Var\ \hat{K} = 2\hat{K}^2 \sum_{j=1}^{s} X_j^4 \Big/ \left(\sum_{j=1}^{s} X_j^2\right)^2 . \tag{25}$$

This normality assumption has been validated (4) from the observed data for cells having more than 25 vehicles in the universe. We should not expect normal distribution of errors for the smaller cells.

Computation of $\hat{K}$ together with its variance affords a way to evaluate the results of the plans, as well as a means for estimating $var(t_1)$. Table 7 exhibits these results for the four samples under investigation with respect to the Lebanon and Kokomo universes. The computations have been carried out for Sample 4 in the manner outlined, although no theory is presented here for this type of sample. The general theory related to this method appears in the literature (1, p. 144).

From Table 1 it can be noted that $\hat{K}$ values for samples 2 and 3 were expected to have been 1.00 with upper probability limits of 2.60 and 3.89 respectively. Table 7 indicates that the corresponding average empirical values for $\hat{K}$ were 1.18 and 1.34. The magnitude of the standard errors of these $\hat{K}$ values implies that the empirical estimates of K might vary considerably with more experimentation of the type done here, but also makes it credible that the observed sampling errors are fairly well in accordance with the theory which resulted in Table 1. On the other hand, since all 4 $\hat{K}$ values from Samples 2 and 3 were greater than 1,

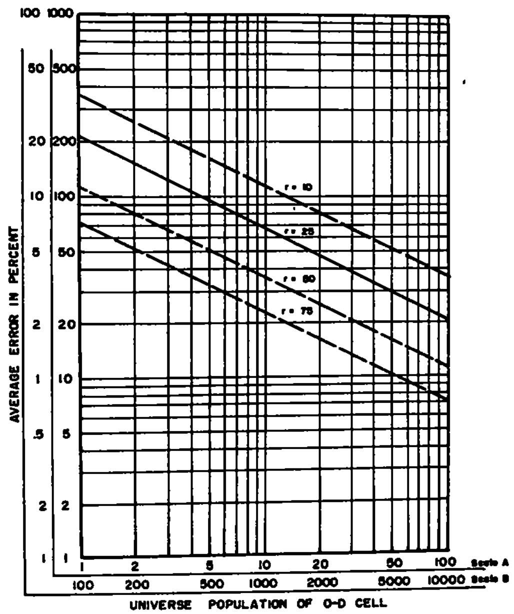

Figure 7. Expected errors due to two-stage sampling.

TABLE 6

OBSERVED ERRORS (IN PERCENT) OF ORIGIN AND DESTINATION TABULATION
SAMPLE 2 FOR LEBANON BY-PASS "BEFORE" STUDY

External Stations - Origins (rows: Internal Tracts - Destinations)

Internal Tracts - Destinations	1	3	5	8	10	11	13	14
001	100 0	100 0	120 0	9 1	100 0	100 0	0 0	0 0
002	9 8	100 0	116 7	45 3	100 0	0 0	100 0	71 0
0Q3	6 1	47 8	9 1	1 6	5 0	35 0	116 7	64 8
004	28 4	16 1	11 1	20 1	17 4	48 0	12 5	4 0
005	63 0	8 3	72 0	13 1	17 2	10 2	44 4	34 7
006	8 3	26 6	51 8	6 6	16 7	21 1	3 4	29 0
007	100 0	36 5	61 0	77 3	28 6	0 0	0.0	100 0
008	47 9	24 8	12 1	12 2	10 8	46 8	22 5	6 1
009	12 5	47 6	10 4	2 4	58 1	70.3	27 0	16 4
010	12 3	8 0	11 0	2 9	29 3	8 1	9 2	10 9
011	8 3	26.3	100 0	8.6	2 9	100 0	11 1	5 0
012	21 4	26 4	47 6	42 1	16 7	100 0	18 8	24 7
013	10 0	22.3	22 4	27 1	17 9	47 2	38.9	20 2
014	30 1	20 9	13 8	4 8	39 7	15 9	12 8	32 7
015	19.1	6 0	1 2	5 4	7 5	17 0	16 8	5 4
016	21 9	29 0	18 4	35 5	7 8	2 7	6 8	0 6
017	7 2	31.6	12 9	12 4	19 8	29 7	10 4	10 4
018	9 1	100 0	48.1	61.1	100 0	0 0	0 0	0 0
019	20 0	2 3	9 0	19 7	22 0	5.3	63 3	11.1
020	71 9	100 0	40 9	16 0	69 2	100 0	8 7	19 8
021	2 2	30.1	12 3	13 2	15 4	73 2	20 4	25 4
022	100 0	100 0	42 0	16 1	116 7	0 0	0 0	35.7
023	18 2	107 1	22 6	10 2	39 4	51 9	5 5	28 1
024	100.0	116 7	100 0	34 7	27 1	23 1	0 0	36 8

External Stations - Destinations (rows: Internal Tracts - Origins)

Internal Tracts - Origins	1	3	5	8	10	11	13	14
001	107 7	0 0	0.0	109 5	0 0	0.0	0.0	0.0
002	5 1	104 5	100 0	26 1	54.5	0.0	0 0	10 3
003	10 9	6 4	6 5	5 8	100 0	9 1	100.0	10 1
004	11 5	0 9	11 2	11 9	9 2	25 0	42 9	16 5
005	36 1	120 0	44 4	2 9	66 7	100 0	15 0	37 3
006	45 2	12 5	35 4	11 9	29.8	4 6	5 1	12.9
007	100 0	3.3	33 3	33 3	40.0	0 0	0 0	120 0
008	22 3	10 4	36 7	4 9	20 5	18 4	21.5	16 4
009	8 0	5 5	20 4	11 5	14 0	29 4	34 9	25 8
010	0 5	2 3	3 7	11 4	20 8	7 3	31 4	7 3
011	30 3	42 9	0 0	40 7	18 4	54 5	44 4	28 0
012	9 3	46 2	98 7	7 6	51.4	0 0	56 5	7.9
013	29 2	26 4	33.3	4 8	36 1	69 4	8 3	0 8
014	3 4	4 1	23 5	6 8	43.8	58 9	27 7	3.9
015	6 4	1 1	8 9	0.9	20 1	14 0	3 6	3.5
016	3.9	11.8	6 6	8.0	47.1	8 1	0 8	40 8
017	17 6	14 8	9 9	2 3	27.9	26 7	9 4	20 3
018	50 0	0 0	53.8	51 8	31 6	0.0	0 0	9 1
019	9 6	16 0	17 6	6 4	25 7	23 5	2 5	12.0
020	51 0	116 7	37 8	0 7	55 6	100 0	37.5	10 8
021	47 1	11 3	37.8	6 6	27 7	57.4	3 5	20 5
022	40 0	0 0	64 3	61 4	116 7	0 0	0.0	100 0
023	100.0	0 0	32 7	16 4	36 4	8.5	3 0	21 4
024	45 0	0 0	6 7	2 2	31 6	0.0	0 0	31 3

External Stations - Destinations (rows: External Stations - Origins)

External Stations - Origins	1	3	5	8	10	11	13	14
1		37 5	7 4	0 4	0 4	20 8	38 9	0 0
3	95 5		10 6	2 2	2 2	10 5	33 3	13 9
5	7 4	18 4		33 3	0.9	59.7	15 4	21 1
8	1 1	6 8	15 9		10 3	2 4	10 0	2 9
10	16 9	0 5	8 0	40 0		43.8	100 0	6 3
11	27 3	8 7	100 0	63 2	0 0		0 0	13 3
13	100 0	69 4	42.2	100 0	100.0	83 3		100.0
14	0.0	0 5	9 3	11 5	4 4	72.5	8 3	

there is indication that the average sampling errors may be somewhat greater than shown in Table 1, but are well below the probability limits for average error shown there. It would follow from the theory that $\hat{K}$ for Sample 1 should turn out to be somewhat smaller than 1. It is not clear whether its failure to do so can be simply attributed to the chance fluctuation expected in $\hat{K}$, or to the uncertainty as to the true nature of sample 1 relative to volume clusters. Sample 4 produced $\hat{K}$ values of the order expected in 25 percent cluster sampling as indicated in Table 1.

Empirical Expected Error Curves

In Equation 21 we may take the logarithm of each side to obtain

$$\log[CV(\hat{T})] = \tfrac{1}{2} \log K - \tfrac{1}{2} \log T. \quad (26)$$

When plotted on double logarithmic coordinates, this function is a straight line with slope $-\frac{1}{2}$ and ordinate intercept $\sqrt{K}$. In accordance with the theory it is clear that K is a function of $\frac{1-r}{r}$, where r is the sampling rate. We write

$$\hat{K} = K^*\left(\frac{1-r}{r}\right) \quad (27)$$

and substitute for K in Equation 26 to have

$$\log[CV(\hat{T})] = \tfrac{1}{2} \log K^*\left(\frac{1-r}{r}\right) - \tfrac{1}{2} \log T. \quad (28)$$

Figures 4 to 7 illustrate Equation 28 for the four samples studied. The solid lines on these figures are obtained from Equation 28 when the observed average $\hat{K}$ values are used. Dotted lines on these figures represent extrapolations to other sampling rates via Equation 27 and 28. The K^* values are also shown in Table 7.

TABLE 7

VALUES OF K AND STANDARD ERRORS

Sample No	Universe	Actual Rate of Sampling	$\hat{K}$	Avg $\hat{K}$	K^*	Standard Error of $\hat{K}$
1	Lebanon Kokomo	.499 500	1.24 1 21	1 22	1 22	0 22
2	Lebanon Kokomo	496 501	1 21 1 16	1.18	1.18	0 21
3	Lebanon Kokomo	483 491	1.28 1 41	1 34	1 28	0 22
4	Lebanon Kokomo	242 247	4 59 3.93	4.26	1.38	0 92

Probability Limits for Error

All of the discussion thus far has been with respect to the average sampling error to be expected. In order to give probability limits for these errors we must assume some probability distribution for the errors which can arise when a particular universe is sampled repeatedly in the same way. It is more or less conventional to suppose that 95 percent of such errors would not exceed $2CV(\hat{T})$. If the errors are normally distributed, this supposition is very close to being correct. These $2CV(\hat{T})$ probability limits are easily obtained from Figures 4 to 7 since one can simply double the $CV(\hat{T})$ corresponding to a particular sampling rate curve for a given value of T. Other multiples of $CV(\hat{T})$ can be used for different probability limits. For example, we should expect 90 percent of all possible sampling errors in a given cell to be less than 1.65 $CV(\hat{T})$, or 99 percent to be less than 2.58 $CV(\hat{T})$. The assumption of normal distribution of errors is not valid for small, and consequently unimportant, T values, say for T less than 20 trips.

When one plots the observed errors of Table 6 (relative to Table 5) on Figure 5, it turns out that about 3 percent of the 632 errors are beyond 2 $CV(\hat{T})$ limits.

PART II PRACTICAL APPLICATION

H. L. Michael and R. M. Brown

● THE theory study resulted in a group of curves (Figs. 4, 5, 6, and 7) which are of great interest. These curves indicate, for the various methods, the average error that can be expected for any origin and destination table cell volume (that number of trips in one direction between an origin and a destination area) for various sampling rates. The data for any other sampling rate can also be obtained by interpolation. In order to use these curves for a proposed survey, two decisions must be made: (1) What minimum O-D cell volume is going to be important in this survey? (2) What average error is allowable for this minimum cell volume?

In order to answer the first of these questions, one must consider the purpose of the survey and that minimum volume of traffic which will receive important consideration because it would affect any recommendations that might be made. It is generally recognized that the low-volume traffic movements contribute very little to the overall analysis of the data or the recommendations made from that data. This does not mean that these volumes of 1, 10, 20, 40, or 50 movements are disregarded entirely, but only that they do not materially affect the survey recommendations. These small volumes will also be obtained, but they will, on the average, contain larger percentage errors.

After the minimum, important O. D. cell volume has been selected, the average error due to sampling that will be allowable must be selected. In the taking of any sample, other than a 100-percent sample, some error due to not interviewing every person will probably occur. Some O. D. cells will have less volumes than the expanded sample data indicate; other cells will have larger volumes than the sample results show. If these errors are considered as a percentage of the actual O. D. cell volume with regard only to their magnitude and not to their sign (minus percent or a plus percent) and are averaged, the average error as used in this paper will result. If the signs of the errors are considered, the average error will, of course, tend to be zero as the negative errors will tend to cancel out the positive errors.

The selection of this average allowable error will also only be for the minimum important cell volume selected. Higher volumes, according to the curves, for the same rate of sampling will have lower average errors when expressed as percentages. The percentage error allowed will again depend on the use of the survey and on the effect that such error will have on the final results. Since many O. D. cell volumes are often used together in the analysis of data, it should be noted that the errors, since some are negative errors and some are positive errors, often tend to cancel each other. Furthermore, an average error of 10 percent in a volume of 100, for example, would be only ten, a variation that might be less than the variation found if the interviewing were conducted at a different time.

It should also be noted that the total

traffic volume at an interview station has little to do with the selection of a sampling rate. The effect of the total volume is negligible if the cell frequency is small relative to the total volume. Furthermore, the sampling error in a cell which is large compared to volume is overestimated by not taking volume into account. A sampling rate, therefore, is on the safe side when it is based only on an important minimum O.D. cell volume and on an allowable average error for that volume. It is evident, therefore, that the selection of larger tracts with correspondingly larger traffic volumes would be advantageous for accuracy reasons. Here again the uses of the survey data and the accuracy needed must be considered.

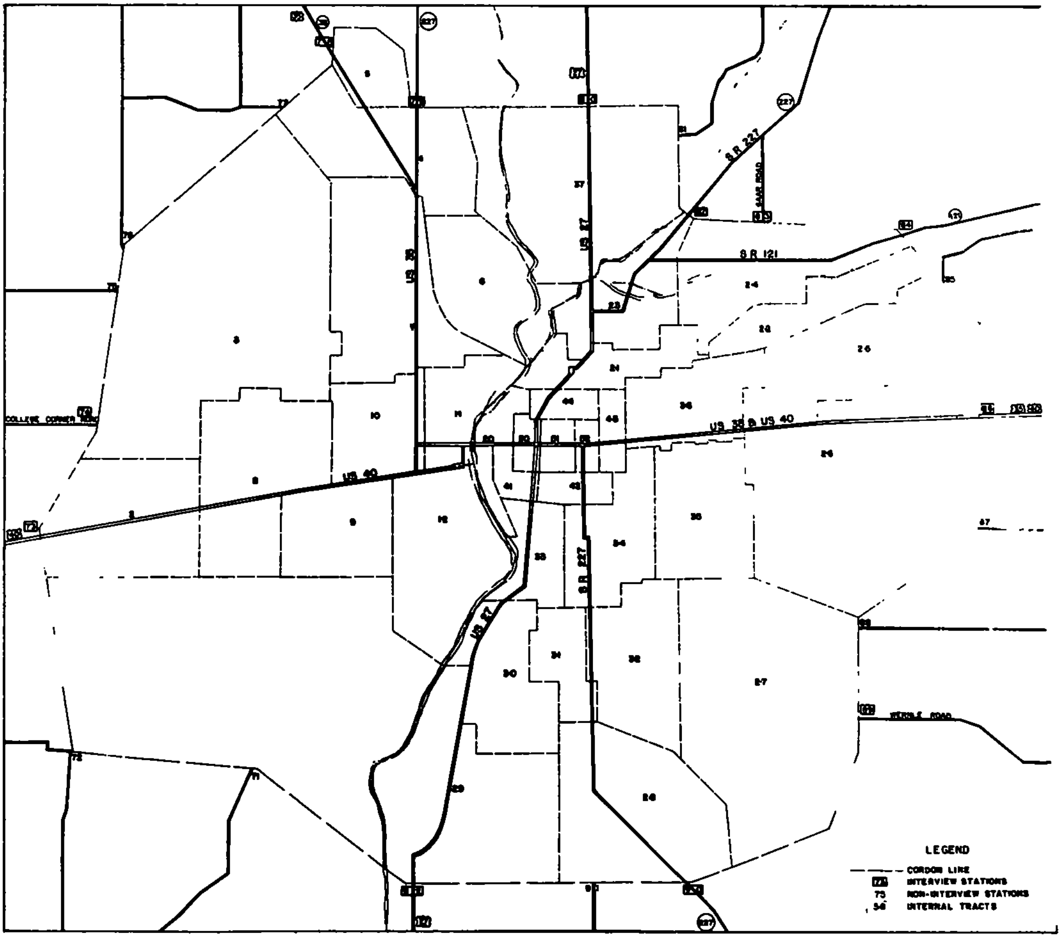

Figure 8. Richmond, Indiana, major routes and survey area.

The curves developed from the Lebanon and Kokomo data and shown in Figures 4, 5, 6, and 7 are, of course, only truly applicable for those two cities. There is no evidence, however, to indicate that these two cities are distinctly different in the manner of traffic characteristics from hundreds of other cities in the United States. This fact, together with the weight of mathematical theory indicated in Part I, seems to indicate that the curves as developed are applicable to surveys in other cities.

Use of Sampling Curves, Example

The State Highway Commission of Indiana authorized a complete origin-and-destination study for the City of Richmond, for the summer of 1952. This offered an excellent opportunity to apply the various sampling methods. The representatives of the cooperating agencies, the Metropolitan Area Traffic Survey Unit of the State High-

TABLE 8

ROAD AND TRAFFIC CONDITIONS AND SAMPLING METHOD USED
RICHMOND TRAFFIC SURVEY

Sta No	Road Conditions Surface	Shoulders	Land use	Average 24-Hour Traffic Volume	Inter-viewing Hours	Sampling Plan Used
71	Gravel	Narrow	Open Fields	262	None	None
72	Gravel	Narrow	Open Fields	137	None	None
73	3-Lane Concrete	Wide	Open Fields	11,907	6am - 6am	Vol Cluster
74	Asphalt	Narrow	Open Fields	369	6am -10pm	100% Sample
75	Gravel	Narrow	Open Fields	61	None	None
76	Gravel	Narrow	Open Fields	353	None	None
77	Gravel	Narrow	Open Fields	69	None	None
78	Concrete	Medium	Open Fields	5,454	6am - 6am	Time Cluster*
79	Asphalt	Narrow	Open Fields	1,114	6am -10pm	100% Sample
80	Asphalt	Narrow	Residential	3,898	6am - 6am	Vol Cluster*
81	Gravel	Narrow	Open Fields	181	None	None
82	Asphalt	Narrow	Open Fields	1,684	6am - 1am	Time Cluster*
83	Asphalt	Narrow	Open Fields	460	6am -10pm	100% Sample
84	Asphalt	Narrow	Open Fields	2,162	6am - 1am	Time Cluster*
85	Gravel	Narrow	Open Fields	288	None	None
86	Divided 4-Lane Concrete	Wide	Open Fields	13,773	6am - 6am	Time Cluster
87	Gravel	Narrow	Open Fields	93	None	None
88	Gravel	Narrow	Open Fields	134	None	None
89	Asphalt	Narrow	Open Fields	552	6am -10pm	100% Sample
90	Asphalt	Narrow	Open Fields	2,671	6am - 1am	Vol Cluster*
91	Gravel	Narrow	Open Fields	250	None	None
92	Asphalt	Narrow	Open Fields	2,586	6am - 1am	Vol Cluster*

*100% sample taken during night hours of low volume count

Richmond, Indiana, June - July, 1952

way Commission of Indiana, the Bureau of Public Roads, and the Joint Highway Research Project of Purdue University, agreed to such sampling. It was further agreed that the minimum important cell volume would be considered to be 100 vehicles and that a 10-percent average error would be allowable for this cell volume. A total of 39 tracts for Richmond, a city of 38,000 people, was selected on the basis of land use and intensity of development.

A study of the curves for quarter-hour sampling (Fig. 5) and uniform sampling (Fig. 4) shows that for a cell volume of 100 and an error of 10 percent a sampling rate of slightly greater than 50 percent is necessary. A sampling rate of 50 percent was therefore chosen as one which would be practical in application. An average error of 10 percent in cell volumes of about 110 could be expected from this sampling rate.

Richmond, is a typical midwestern city. A principal east-west highway (US 40) passes through the city and four other state and federal routes pass through or terminate in the city (see Fig. 8). A cordon line around the developed area intercepted 22 roads and highways. The traffic volumes on these roads ranged from less than 100 to more than 13,000 vehicles on an average weekday (see Table 8). The roads varied from narrow, two-lane, gravel roads to modern, four-lane divided highways (see Table 8). These great differences in traffic volumes and roadway conditions permitted a study of the various sampling methods under many conditions.

The selection of the stations which were operated was made according to the usual criteria of traffic volume. A few stations on county roads were operated that had total volumes of less than 500 in order to obtain the traffic pattern for county roads. Four other stations with volumes around 2,000 vehicles per day were operated for 19 hr., instead of the usual 16 hr. because of rather large volumes of traffic between 10 pm. and 1 am. The exact location of all stations was made on the basis of the established criteria for safety.

The personnel required to operate the largest station was the determining factor for the number of interviewers hired. Each interviewer was considered capable of taking 40 interviews per hour. The number of interviewers required for each station for each hour for a 50-percent sample was thus determined (see Table 9). The number of men provided for each shift was then determined on the basis of the men needed during most hours of each shift. This method provided a rest period for usually one man at a time during the low-volume hours of the shift. Since the peak period of 3 to 6 pm. and the period 10 to 1 am. usually required several more men than other hours of these shifts, additional men were hired and assigned as necessary to the various stations during these hours (see Table 9). The necessary manual counters and crew chiefs required for the various stations were also determined. One counter and one crew chief were normally assigned to each station, except that on those

TABLE 9

PERSONNEL REQUIRED PLAN FOR A TYPICAL STATION
RICHMOND TRAFFIC SURVEY

Station No 78 — Hours of Operation 6A M to 6 A M
Sampling Rate 50% — Sampling Method Time Cluster

Time	Pre Survey Volume Count	Men Required Inter-viewers	Counters	Crew Chiefs	Remarks
6-7 am	447	6	1	1	Use Following Crew
7-8 am	305	4	1	1	1 - Supervisor*
8-9 am	262	4	1	1	1 - Crew Chief
9-10 am	304	4	1	1	1 - Counter
10-11 am	408	5	1	1	4 - Interviewers
11-12 N	203	3	1	1	
12-1 pm	290	4	1	1	
1-2 pm	301	4	1	1	
2-3 pm	327	4	1	1	Use Following Crew
3-4 pm	549	7	1	1	1 - Supervisor*
4-5 pm	527	7	1	1	1 - Crew Chief
5-6 pm	403	5	1	1	1 - Counter
6-7 pm	290	4	1	1	4 - Interviewers
7-8 pm	237	3	1	1	Note 2 additional
8-9 pm	186	3	1	1	interviewers from
9-10pm	177	3	1	1	3pm - 6pm Release 1 interviewer at 8pm
10-11pm	160	2	1	---	Use Following Crew
11-12M	148	2	1	---	1 - Supervisor*
12-1 am	124	2	1	---	1 - Counter
1-2 am	33	2	1	---	2 - Interviewers
2-3 am	51	2	1	---	Note From 1 am - 5 am
3-4 am	43	2	1	---	take 100 % Sample
4-5 am	40	2	1	---	Supervisor to direct
5-6 am	120	2	1	---	traffic as necessary

Operate also Station 102 on this day

*Also Supervisor of Above Station
Supervisor to assist in interviewing during under-manned hours

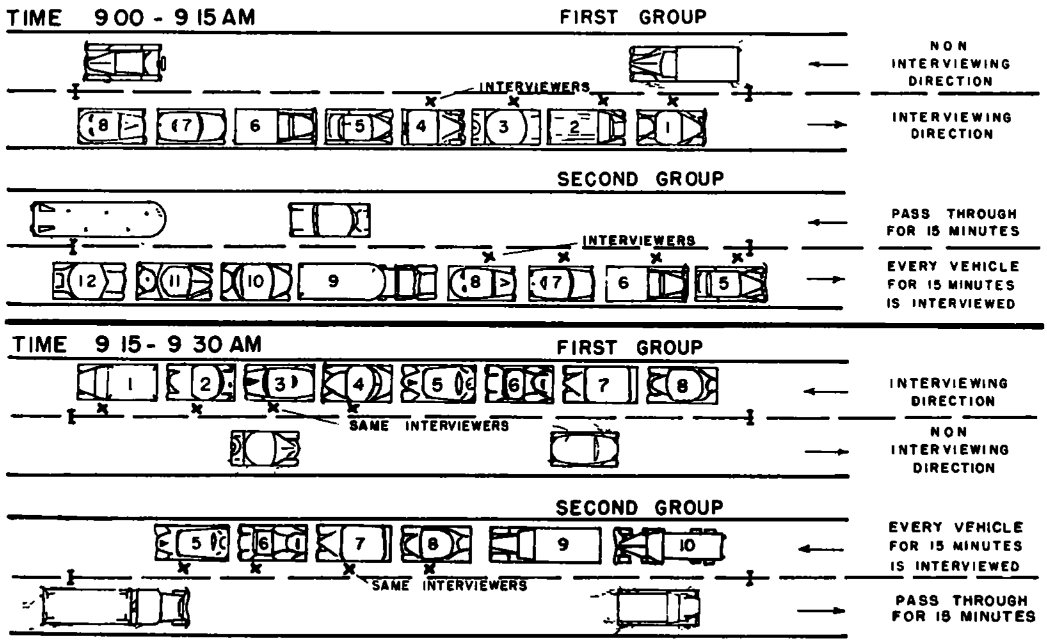

Figure 9. Typical time-cluster method of interviewing.

stations with more than 6,000 vehicles per day two counters and two crew chiefs were assigned. The crew chiefs also acted as traffic directors at heavy stations, especially during hours of peak volumes. Each shift's daily operation for all stations was also directed by a regularly employed supervisor. Typical station interview parties are also indicated in Table 9.

The preparation of such a personnel plan permitted the hiring and assignment of personnel in a most efficient and economical manner. The entire survey was conducted in seven interviewing days by a total crew for all shifts of 30 men (includes interviewers, manual counter, crew chiefs, and supervisors).

SAMPLING PROCEDURE AT THE ROADSIDE

Three sampling methods were used during the Richmond survey: (1) quarter-hour samples (time clusters), (2) uniform samples (volume clusters), and (3) combination of (1) and (2).

The method used at each station was predetermined. Only one method of sampling was used at a station so as to minimize the confusion that might occur if sampling methods were changed hourly. Sampling was conducted, at all stations that had daily traffic volumes greater than 1,200 vehicles, during all hours that traffic volumes were expected to exceed 30 vehicles per hour in either direction. Sampling was not done at other stations or during other hours because one interviewer for each direction had to be present anyway and he could easily interview all traffic.

Time-Cluster Sampling

The time-cluster sample as used here refers to the selection of a sample of the total traffic volume on the basis of a selected time interval. Any time interval could, of course, be chosen for this method. Every vehicle that arrives at the station during a 1-min., 5-min., 15-min., or other chosen time period is interviewed and then no vehicles are interviewed during a corresponding time period, the length of which is dependent on the selected sampling rate. If, for instance, a sampling rate of 50 percent is chosen, then interviewing is conducted for all those vehicles that arrive at the station during 50 percent of the total time. The time interval chosen for the Richmond survey was 15 min. During one 15-min. period all vehicles that arrived at the station from one direction were interviewed and no vehicles were interviewed in that direction during the

next 15-min. period. This process was then repeated for all periods of the day. The interviewers, however, were not idle every other quarter hour, but interviewed traffic first for 15 min. in one direction and then for 15 min. in the other direction. A typical example of this procedure is shown in Figure 9.

Each station was arranged so that interviewing for both directions could be carried on from the same location. The station arrangement was exactly the same as if both directions would have been interviewed at the same time, except that the stop sign for the direction not being interviewed was covered during the periods when the traffic could proceed without stopping. On four-lane roads the interviewing was conducted in two lanes in each direction while on other roads, including three-lane roads, interviewing was done in one lane only for each direction.

Adequate personnel were assigned each shift, with added personnel during peak hours, to interview all traffic in the one direction. At the end of one time period in one direction and the simultaneous beginning of a time period in the other direction, interviews were of necessity taken in both directions at the same time. It is necessary in the time cluster sample to interview all traffic that arrives during the time interval. This means that at the end of some time periods there will be some vehicles that arrived before the end of the period that are waiting to be interviewed. They must be interviewed. In the other direction, there are vehicles arriving after the beginning of the next period which also must be interviewed. This problem was solved at Richmond by taking a portion of the interviewers from the direction being interviewed about 1 or 2 min. before the end of the period and placing them in position to interview the traffic from the other direction at the beginning of the next period. The other interviewers continued in the first direction until all vehicles that arrived before the end of the period had been interviewed. They then began interviewing in the other direction. This system operated smoothly at Richmond, and in every case the changeover was begun and completed in less than 5 min. (2 min. prior and 3 min. after). Usually the total delay was only a matter of 2 or 3 min.

The crew chiefs, of which there was always one and, on heavy stations, two, in addition to their normal duties of interview collection and review and general supervision of the station were given the following responsibilities: (1) direct traffic as necessary (primarily traffic not being interviewed but passed through the station and, on four-lane roads, that traffic being interviewed in more than one lane); (2) announce the end of one time period and

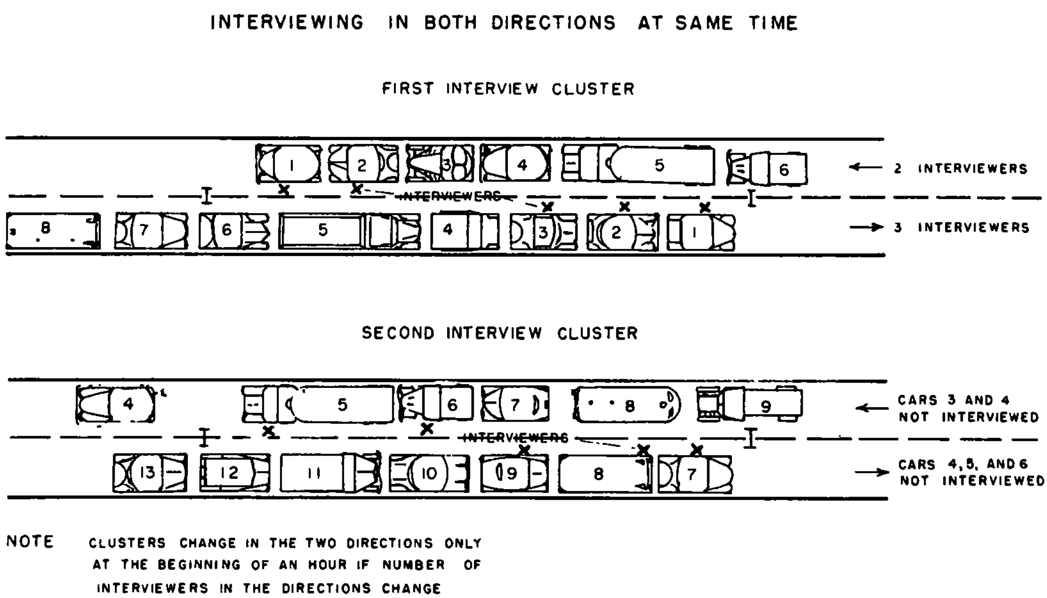

Figure 10. Typical volume-cluster method of interviewing.

the beginning of the next; (3) indicate to the interviewers the first and last car to be interviewed in each time period; and (4) select the number of specific interviewers to be taken from interviewing in the one direction before the end of the time period to get in position to begin interviewing in the other direction (number dependent on traffic volumes in both directions).

The crew chief or chiefs also had the assistance of the supervisor during peak periods and at heavy stations. All personnel were instructed to adhere to exactly 15-min. intervals.

Volume-Cluster Sampling

The volume-cluster sampling consists of interviewing consecutive vehicles in clusters or groups of a predetermined number and then not interviewing a cluster or group of vehicles of a size that is dependent on the sampling rate. The size of the cluster to be interviewed could be one or more. A single cluster size was not used in the Richmond survey, but a cluster equal to the number of interviewers working in each direction was used. Cluster sizes or number of interviewers, however, were not changed during any one hour of operation but only at the beginning of an hour, if necessary. Cluster sizes depended on the volume of traffic expected (from previous observations and counts) in any one direction and were not the same during many hours for the two directions of one station. The actual clusters used ranged from one (every other vehicle for a 50-percent sampling rate) to six. Since all interviewing by this method was at a 50-percent rate, a cluster of vehicles equal in number to the cluster interviewed was passed through the station without being interviewed. Figure 10 illustrates this type of sampling.

The station arrangement and personnel to operate the station were the same as for time cluster sampling. Interviewers were placed in the two directions by the crew chief or chiefs as volumes in each direction warranted, except that any changes were made only at the beginning of an hour. The first interviewer in each direction and the crew chief during heavy hours determined the first vehicle in each cluster that was interviewed and directed those vehicles in the cluster that were not interviewed through the station. Each interviewer made one and only one interview for each cluster that was interviewed. This helped to maintain an accurate check of those vehicles interviewed and, at the same time, distributed the responsibility of getting accurate cluster sizes. All personnel were instructed to adhere strictly to the predetermined cluster size with no exception, except in the case of emergency vehicles.

Combination Time-and-Volume-Cluster Sampling

In order to test a method of obtaining a 25-percent sample, a combination of the time-and-volume-cluster methods was used. This method simply obtained a sample by interviewing volume clusters within a time period. In short, interviews were taken during 15-min. intervals in one direction only and then not taken in that direction for the next 15 min. Instead of interviewing every vehicle, however, within the 15-min. period, only one half of the traffic was interviewed. This was accomplished by interviewing a cluster of traffic equal in number to the number of interviewers and then not interviewing an equal-sized cluster.

The procedure of station operation was the same as that for the time-cluster operation, with the modification that it was necessary for the first interviewer or the crew chief to pass a cluster through without interviewing and then to select the first car in each cluster to be interviewed. Only a cluster size of four was used with this method. A separate crew was used for operation in the other direction during alternate 15-min. periods. This procedure eliminated a quick change of direction for interviewers at the end of one period and the beginning of another, and gave each group of interviewers a short period of rest during each half hour.

APPRAISAL OF SAMPLING METHODS

The success of these various methods was established after the first day for each method. The taking of a predetermined sample not only proved to be practicable but, it appeared, actually improved the interviewing process. The general public understands that there is such a thing as sampling and that it certainly should be more economical to in-

TABLE 10

ACTUAL SAMPLING RATES BY STATION
FOR PERIOD THAT SAMPLING WAS CONDUCTED
RICHMOND TRAFFIC SURVEY

50% Sampling Rate Planned

Station	Sampling Method	Number Hours Sampled	Manual Count	Number Inter-viewed	Actual Sampling Rate
73	Vol. Cluster	24	10669	5304	49.7
74	100% Sample	16	290	290	100.0
78	Time Cluster	20	5091	2637	51.8
79	100% Sample	16	891	888	99.7
80	Vol. Cluster	18	3535	1811	51.2
82	Time Cluster	14	1237	611	49.4
83	100% Sample	16	413	410	99.3
84	Time Cluster	14	1439	725	50.4
86	Time Cluster	21	9954	5077	51.0
		3	1486	365	24.6*
89	100% Sample	16	476	475	99.8
90	Vol. Cluster	14	1944	967	49.7
92	Vol. Cluster	14	1856	924	49.8

*25% Sampling rate planned

vestigate only a fraction of the total than to interview the whole. The public generally does not recognize the details necessary to take a sample; but from responses and discussions with motorists at Richmond, they view with approval a sampling method they can understand. The time-cluster and volume-cluster methods of sampling offered them that evidence. They could "see" the methodical, business-like selection of vehicles in 15-min. cluster or in one to six vehicle clusters and then the release of an equal-sized cluster. Time and time again the vehicle operators expressed approval of, in their words, "the business-like manner" in which the interviewing was conducted.

The selection of a fraction of the total also seemed to salve the ire of many of those residents who made many trips through the station in one day. They knew they had a gambler's chance to get through without being interviewed. They were gratified whenever they "won" and accepted the interview in a sportsmanlike manner when they "lost."

The actual selection of the sample did not prove difficult by any of the methods. Table 10 indicates the actual sample selected during those hours when sampling was actually used. Table 11 shows the hour-by-hour sampling rates for two stations. Variances in the time-cluster rates were expected because of the actual variances that occur for traffic volumes in 15-min. periods. The volume cluster, if all procedure had been exactly followed, would be almost exactly 50 percent for every station. Miscounts of the number of vehicles interviewed, or of the cluster passed through, or of the total at the station, or the failure of vehicles to stop are causes for variations. These causes, as indicated by Tables 10 and 11, were evidently controlled quite well. The use of competent crew chiefs and head interviewers did much to minimize the possibility of miscounts.

Although no consideration was given to vehicle type in the selection of the sample, approximately 50 percent of each type of vehicle was obtained. Table 12 lists the interviewing rate by vehicle type for two interview stations.

The use of the crew chief as a traffic director resulted in a speedier movement of vehicles through the station. He directed the vehicles to be interviewed to position speedily and efficiently; this left the interviewer more time to concentrate on taking the interview and properly recording it. He also kept the traffic flowing smoothly when not being interviewed. It appeared that the methods operated best when, especially during peak hours, he was the one who counted the vehicles that were not interviewed or selected the first and last vehicles that were interviewed. A crew chief as a traffic di-

TABLE 11

ACTUAL SAMPLING RATES BY HOUR AT TYPICAL STATIONS
FOR RICHMOND TRAFFIC SURVEY

Station No	73			86		
Sampling Method	50% Volume Cluster			50% Time Cluster		
Hour	Manual Count	Number Inter-viewed	Actual Sampling Rate	Manual Count	Number Inter-viewed	Actual Sampling Rate
6-7 am	522	264	50.6	511	284	55.6
7-8 am	507	254	50.1	454	229	50.4
8-9 am	529	265	50.1	507	255	50.3
9-10 am	558	279	50.0	492	269	54.7
10-11 am	554	280	50.5	540	273	50.6
11-12 N	545	266	48.8	600	300	50.0
12-1 pm	561	283	50.4	582	287	49.3
1-2 pm	548	264	48.2	616	306	49.7
2-3 pm	565	275	48.7	598	321	53.7
3-4 pm	663	331	49.9	684	366	53.5
4-5 pm	776	392	50.5	767	398	51.9
5-6 pm	706	351	49.7	717	372	51.9
6-7 pm	559	270	48.3	632	290	45.9
7-8 pm	605	299	49.4	594	294	49.5
8-9 pm	550	283	51.5	668	322	48.2
9-10 pm	446	216	48.4	574	134	23.3*
10-11 pm	365	181	49.6	492	124	25.2*
11-12 M	290	142	49.0	420	107	25.5*
12-1 am	200	99	49.5	242	128	52.9
1-2 am	131	65	49.6	189	101	53.4
2-3 am	106	50	47.2	114	54	47.4
3-4 am	121	61	50.4	105	61	58.1
4-5 am	110	58	52.7	122	68	55.7
5-6 am	152	76	50.0	220	99	45.0

*25% Sampling rate planned

rector appeared to be a must. Most of his other duties, such as collection and review of the interviews, were handled by the shift supervisor on heavy volume stations.

The use of a predetermined sample also minimized the confusion at interview stations. Each man knew what he was to do, who he was to interview, and who he was not to interview. During peak hours there was no doubt or confusion as to how long the line of waiting traffic should get before releasing some or how many to pass through without interviewing, since the interviewers were instructed to take the predetermined sample according to the predetermined method at all times. During peak hours some back-up of traffic did occur. It appeared, however, that a line of traffic formed almost instantly (such as would occur from a large group arriving simultaneously), then stayed about the same length for a short period, and gradually disappeared within 5 to 10 min. This cycle, during peak periods, then often repeated itself. The maximum delay for any vehicle under either method was slightly less than 5 min. This maximum delay occurred only during very short periods of time during peak hours on high volume stations (12,000 vehicles per day). Many times no delay except for interviewing occurred, and during most hours of the day the maximum delay was two to 3 min.

TABLE 12

SAMPLING RATES FOR TYPICAL STATIONS BY TYPE OF VEHICLE FOR RICHMOND TRAFFIC SURVEY

Station No	73			86		
Sampling Method	50% Volume Cluster			50 % Time Cluster		
No. of Hours Sampled	24			21		
Vehicle Type	Manual Count	Number Interviewed	Sampling Rate	Manual Count	Number Interviewed	Sampling Rate
Passenger Cars	8455	4199	49 7	7808	3975	50 9
Light Trucks	700	340	48 6	465	247	53 1
Medium Trucks	480	258	53 8	514	263	51 2
Heavy Trucks	970	479	49 4	1104	560	50 7
Buses	64	28	43 8	63	32	50 8

The volume-cluster method appeared to work most efficiently and smoothly on two-lane roads where four interviewers or less in each lane handled the volume of traffic. More interviewers than this resulted in a decrease in the number of interviews an interviewer took per hour because of the longer lapse between interviews, due to passing a large cluster through the station. The interviewing efficiency did not become serious, however, unless more than six interviewers per direction were used, since the traffic director speeded the movement through the station. This method worked better than the time cluster on narrow roads with poor shoulders, because all interviewers were permanently positioned and required less necessary space in which to operate. On the three-lane road this method, with interviewing in one lane per direction, worked quite well with as many as six interviewers on line in one direction.

The time-cluster method also worked best where the number of interviewers was four or less per interviewing lane. With this method, since all vehicles were taken for a selected time period, speed of interviewing was most important in order to prevent long lines of waiting traffic. This method required a little more working space for the interviewers, since they were required to change position every time period and since one direction of traffic was moving freely through the station. A little spare space in the center of the highway was advisable for obvious safety reasons. On the four-lane highway, with two interviewing lanes for each direction, this method worked excellently. A short time before the end of the time period, the inside lane was closed to interviewing and the interviewers on that lane got into position to interview in the outside lane of the other direction. After all required vehicles were interviewed in the first direction, the remainder of the interviewers there transferred to the inside lane of the second direction. Traffic in the direction that was not interviewed passed through the station in the outside lane. This made it unnecessary for interviewers to cross lanes of moving traffic. A careful plan of action for the short period at the end of one time period and the beginning of the next was extremely important for the time-cluster method in order to minimize confusion and to select all the proper vehicles in the sample.

A time period of any length could, of course, be chosen. The 15-min. period was selected for this survey for at least two reasons:

1. The original study of the Lebanon and Kokomo data was made for 15-min.

time clusters, since the field data were separated only for each quarter hour. This time period, incidentally, proved to be the most accurate of all the methods studied with this data.

2. A shorter time period would have necessitated more frequent directional changes of interviewing. This has definite disadvantages as it increases chance of confusion and also has the statistical objection of increasing the number of times that the selection of the first and last vehicle is left up to an individual who may err as to its proper selection.

The selection of a shorter time period, however, might reduce the length of any waiting traffic lines. It appears that several studies of a shorter time cluster should be made.

It must be admitted that time- (and possibly volume-) cluster samples would have biases if a succession of cars would all have the same origin-destination. If such movements were more likely to happen at a certain time than at another, a bias might occur if interviewing always began at a certain hour on the hour. One way to try to minimize such biases would be to leave to chance (such as spin a needle around the clock) when the interviewing should begin. However, biases eliminated in some origin-destinations might be put into others in this way. It appears, therefore, that on the average starting at the beginning of an hour would be as good as could be done. From a practical point of view, it certainly would be best.

The combination time-and-volume-cluster sample, although only used for 3 hr., appeared to operate smoothly and efficiently. A 25-percent sample was taken without delaying traffic for long periods and with a small group of interviewers who obtained an average number of interviews per hour. Under the volume-cluster plan, the number of interviews per interviewer for this rate would probably be less because of the necessarily long waits between interviews while large clusters of traffic not being interviewed passed through the station. On the other hand, the use of the time-cluster method only would necessitate periods of time when the interviewers would not work at all. The use of the two methods together minimizes these disadvantages, but at the same time makes a changeover to another direction at the end of a time period difficult. A solution to this problem might be to allow a short time lapse between periods when no interviews for either direction would be taken. This, of course, would reduce the sampling rate and would have to be considered in obtaining the desired size sample.

It must be emphasized that thorough planning is necessary to insure proper operation of any of these sampling plans. The required number of interviewers must be determined from presurvey counts and then that number selected and thoroughly instructed. Competent crew chiefs and head interviewers are also a necessity. Above all, every one connected with the operation must realize the importance and understand the selection of the sample.

CONCLUSIONS

1. If the traffic passing through a roadside interview survey station is sampled in clusters, on either a time or volume basis, then, for the numerous estimates of O. D. trip frequencies associated with this traffic flow, the overall sampling variation is similar to that given by the theory of random sampling. When a delineation is made among the different types of cluster samples, experimental results from two independent surveys are in general accord with the corresponding sampling theory derived from basic assumptions on the nature of the sampled universes.

2. The expected percentage sampling error, $CV(\hat{T})$, for an estimate, $\hat{T}$, of a true O. D. frequency, T, can be expressed by the relation

$$CV(\hat{T}) = (K/T)^{1/2}$$

where K depends primarily upon the sampling rate, and, to a lesser degree, upon the method of sampling. Empirical values of K were obtained in this study and may be used to estimate the magnitude of sampling errors and their limits in probability. Conversely, the results can be used to select a sampling plan when a specified amount of sampling variation is to be allowed for a particular origin-destination survey.

3. It is evident that the use of a predetermined sampling process is both practicable and beneficial. A 50-percent sample was obtained with a minimum of confusion in a business-like manner that was recognized and appreciated by the partici-

pating vehicle operators and the community.

4. The sampling procedure promoted efficient operation of the interviewing station by minimizing confusion, and assured the selection of a sample whose chance variation could be appraised.

5. The time-cluster method performed more smoothly than the volume method on roads where traffic volumes could be handled by three or four interviewers per interviewing lane and where adequate space for safe movement of interviewers between interview lanes was available.

6. The volume-cluster method operated more efficiently than the time cluster method on roads where traffic volumes required five or six interviewers or where roads and shoulders were narrow.

7. Either method operated well where only one or two interviewers were required.

8. The combination time-and-volume-cluster method appears to have possibilities where small samples (25 percent or less) are adequate.

9. A great deal of thought and study is necessary to determine the sampling accuracy which is necessary for origin-destination studies. Specifically, the question that requires study is: What is the allowable average error (or a certain probability error) for various origin-and-destination, rectangular-table, cell volumes? It may be that an average error of 10 percent for cell volumes of 100 may be more accurate than generally necessary.

10. Thorough planning of station operation and the assignment of adequate and competent personnel are required. This should include presurvey hourly traffic counts and excellent instruction of personnel.

11. The use of some sampling procedure not only will reduce the number of personnel required to conduct the field work but will also reduce the number of cards to be coded, key punched, and tabulated. Such reduction probably will be in direct proportion to the number of interviews taken. This not only reduces the cost of O. D. surveys but reduces the time necessary to complete such a study.

12. A number of other studies on similar or other sampling methods as well as studies on necessary size of sample would prove of benefit to the conduct of future origin-and-destination surveys.

ACKNOWLEDGEMENTS

The material presented in this paper was made possible by the cooperation of many individuals. The authors are especially indebted to K. B. Woods and A. K. Branham of the Joint Highway Research Project, Purdue University, for their constant encouragement and helpful guidance: to A. D. May, Jr., formerly of the Joint Highway Research Project, for his interest and assistance in the initial phases of the study; and to S. F. Coffman, Bureau of Public Roads, for his invaluable advice and suggestions during the practical application study. Special acknowledgement is also given to many departments of the State Highway Commission of Indiana for their assistance and cooperation. To the many individuals who assisted with the preparation of the data and the report, the authors also wish to express their personal gratitude.

REFERENCES

1. Deming, W. E., Some Theory of Sampling, J. Wiley and Sons Inc., New York (1950).

2. Kennery, J. F. and Keeping, E. S., Mathematics of Statistics, Part Two, D. Van Nostrand Co., Inc. (1951).

3. Michael, H. L., "Traffic and Engineering Report of the Before and After Surveys of the Kokomo By-Pass", Thesis submitted to Purdue University in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering (1951).

4. Miller, I., "The Application of Sampling Techniques to Highway Surveys", Thesis submitted to Purdue University in partial fulfillment of the requirements for the degree of Master of Science (1952).

5. Manual of Procedures for the Metropolitan Area Traffic Studies, Public Roads Administration (1946).

6. Manual of Traffic Engineering Studies, Association of Casualty and Surety Co. (1945).

Sampling Procedures for Roadside Interviews in Origin-and-Destination Traffic Surveys

ROBLEY WINFREY, Formerly Research Professor of Civil Engineering, Iowa State College, and ROBERT J. HANSEN, Assistant City Traffic Engineer, Spokane, Washington

IN the origin-and-destination traffic study of the main routes at Ames, Iowa, in the fall of 1949, the roadside interviews, 7 a. m. to 11 p. m., were recorded in time intervals of 2 min. It is possible, therefore, to reconstruct the traffic flow through any of the interview stations in road sequence of vehicles, except within the 2-min. intervals.

Five samples were drawn from the IBM cards. The resulting trips between pairs of origin-and-destination zones were compared with the trips found in the 100-percent interviews.

Systematic samples of 10 and 20 percent were drawn by selecting for each direction of travel every 10th and every 5th trip in sequence throughout the 16-hr. period for the four external stations on US 30 and US 69 at the city limits. Statistically, these two samples adhered to the characteristics of a normal distribution and are of value in predicting the probable maximum percent errors to be expected from comparable samples.

The third sample was drawn from Station 3 only on a time-controlled basis by taking all vehicles reaching the interview station in the first 2 min. of each 10-min. period. This sample had about the same characteristics as the systematic 20-percent sample.

Since these three samples were selected by a procedure difficult to maintain at a roadside station, the fourth sample was taken on a time-and-size-controlled procedure that could be easily performed at the roadside. This sample was satisfactory and of about the same quality as the time-controlled sample. These third and fourth samples were expanded to the 10-min., 1-hr., and 16-hr. 100-percent traffic volumes. The 10-min. and the 16-hr. basis gave better results than the 1-hr. expansion.

The fifth test was applied to traffic in one direction only. The results indicate that for these stations at Ames, Iowa, acceptable results could be obtained by sampling one direction only.

These brief results are insufficient, even for Ames, to determine the best basis of taking an origin-and-destination sample at roadside-interview stations. The results do show, however, that the time-and-size-controlled sample of about 20 percent of the total traffic should yeild acceptable results. When the traffic by direction balances each other in interchanges between pairs of origin-and-destination zones, the unidirectional method of sampling could be used.

● IT is the purpose of this paper[1] to present the results of a study of the sampling of traffic at four external origin-and-destination interview survey stations at Ames, Iowa. In November 1949, 100 percent of the entering and leaving traffic was interviewed for a 16-hr. period, all information being recorded in 2-min. intervals. These roadside interviews are now used in testing sampling procedures. The specific objectives are to arrive at a field procedure of interviewing and a size of sample that will satisfactorily produce acceptable trip data between origins and destinations at a cost of time and money much less than required by the 100-percent interviews.

This paper deals solely with the Ames origin-and-destination survey. The IBM punch cards in which the trips were assigned to 2-min. intervals in actual time sequence for the 16-hr. day, 7 a. m. to

[1]This paper is based upon a thesis submitted by Robert J. Hansen as partial fulfillment of the requirements for the degree, master of science, major in highway engineering, at Iowa State College, June 1952. Acknowledgments are extended to Mark Morris and Carl Schach of the Iowa State Highway Commission for their valuable assistance in both the field work and office analysis.

11 p. m., were used in the different sampling methods.

To establish a reliable sampling method, the many variations in the composition of traffic and in the origins and destinations should be taken into consideration, as well as the proper placement in the field of personnel to handle the traffic and to interview the drivers. The size of the sample should be determined from the total number of trips between pairs of origin-and-destination zones or tracts within which the major interests lie.

A satisfactory method of sampling traffic at roadside-interview stations should eliminate the human error as much as possible as well as provide for a reliable sample. The desirable number of interviews, stated as a percentage of the total traffic, at periods of minimum traffic flow, may be greater than at peak periods of flow. Statistically speaking, when the size of population is increased, a sample size in percentage of the population can be decreased without decreasing the degree of accuracy. Thus, in practice, three interviewers may be used to take a 30-percent sample at periods of minimum flow, but as traffic is increased, they may take only a 15-percent sample without causing a decrease in the reliability of the sample. On the other hand, a greater probable error in the sample may be tolerated when the traffic volume is relatively low because of the lesser importance of the low traffic volume to applications of highway design and traffic control.

Devising a method of sampling traffic for origin and destination interviews should consider the following criteria: (1) a reliable minimum size of sample to simplify the survey and reduce the number of interviews, coding personnel, and tabulating machine time to a minimum, consistent with the desired results; (2) a method of selecting the vehicles to be stopped such that the minimum of vehicular delay is involved; (3) a flexible method of sampling the traffic stream workable on all forms of daily traffic patterns; and (4) the proper placement at the field station of interview personnel for maximum efficiency of sampling, accuracy of data, and safety.

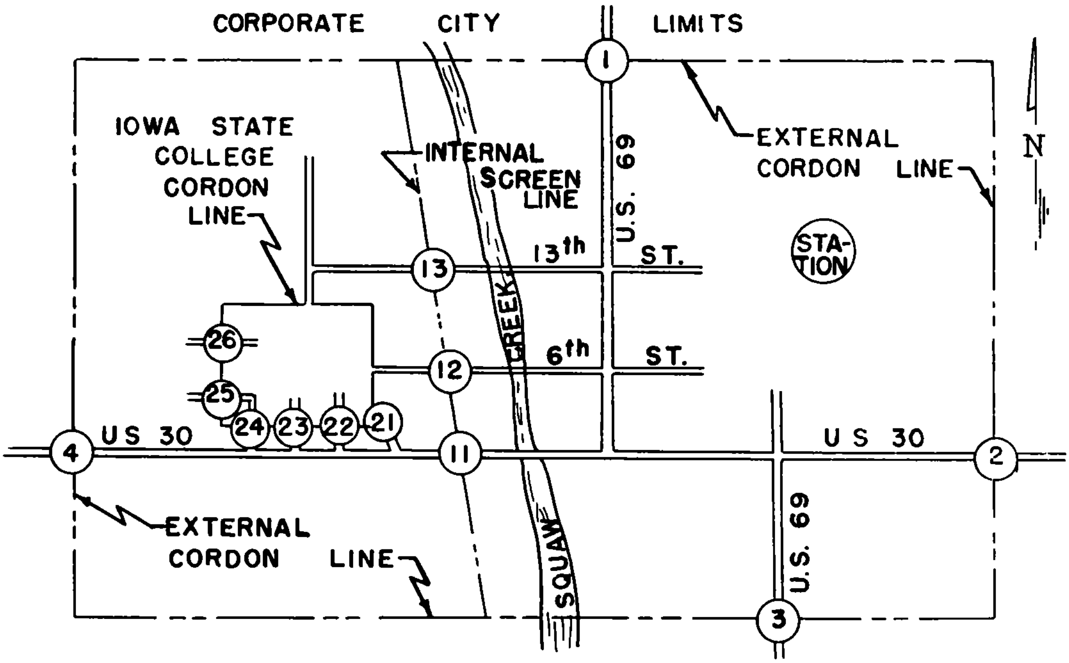

Figure 1. Roadside-interview stations of the Ames origin-and-destination traffic survey.

THE AMES ORIGIN-AND-DESTINATION SURVEY

The following is a description of the

city of Ames, Iowa, the origin-and-destination survey, and overall objectives of the survey.

Ames, population 23,000, has the normal traffic generators of the business area, a small, light industrial area, and residential areas, but in addition, it has the Iowa State College and the Iowa State Highway Commission. The city is divided into two main sections by a ½-mi.-wide strip of low land adjacent to the Squaw Creek, running predominantly from the north to the south through the city (Fig. 1). The west section of the city consists of Iowa State College, a residential area, and a small business area. The east section consists of the main business area, industrial area, residential areas, and the Iowa State Highway Commission. The two sections of the city are joined by three connecting east-west routes, US 30 and two major street extensions. Iowa cities are predominantly light industrial areas and shopping communities serving farm areas. Ames is different in character, however, since over a third of the 23,000 population consists of employees and students of Iowa State College.

The Ames origin-and-destination survey was conducted in late October and early November, 1949: (1) as an aid to planning a transportation system with which to serve vehicular traffic, (2) to make improvements in traffic regulations, and (3) to aid in research on methods of conducting origin-and-destination studies.

The survey was conducted under the auspices of the Iowa State Highway Commission, the Iowa State College Engineering Experiment Station, and the Ames City Plan Commission. Availability of funds limited the origin-and-destination survey to inbound and outbound traffic on the four primary highway entrances into the city, the intracity traffic on the three routes connecting the east and west sections of the city, and the entrances to the college campus.

The roadside postcard method and the roadside-interview method were used two weeks apart to furnish traffic data of the same population for research study of methods of gaining origins and destinations. An external cordon was established about the city on the four entering primary highways, and an internal screen line was established across the three connecting routes of the two sections of the city. For an additional check on the external and internal stations, a cordon was established about the college campus, including six entrances in addition to the two stations on the screen line. The designation of roadside stations is shown in Figure 1. A home interview was conducted for origins and destinations of the trips made by residents of Pammel Court, a married veterans' housing unit within the college campus. Neither the roadside postcard method[2], the Pammel Court home interview, or the roadside interviews at the college entrances are discussed in this paper.

The roadside interviews were conducted on a 16-hr. basis, 7 a.m. to 11 p.m., on Monday through Thursday, excluding Friday and weekends because of college social activities. These interviews were taken at stations 1, 2, 3, 4, 11, 12, 13, 21, 22, 23, 24, 25, and 26. All vehicles were stopped and drivers were interviewed. The information obtained included origin, destination, number of occupants in the vehicle, Iowa county (or state) of vehicle registration, type of vehicle, purpose of trip, and the time of day in which the driver was interviewed recorded for each 2-min. interval throughout the day.

The postcard field work was conducted about two weeks prior to the roadside interviews at Stations 1, 2, 3, 4, 11, and 12, and operated for a 24-hr. day at each station. As the postcards were handed out, observers counted and classified each vehicle by its type and place of registration, recording the information by 2-min. intervals during the 24-hr. period. Information asked for on the postcards was approximately the same as that asked for in the roadside interviews.

Ames was divided into 10 zones and each zone divided into 10 or fewer tracts for coding of origins and destinations. Each tract was approximately ¼ mi. square. The zones were composed of homogeneous land use areas with their borders along major streets, railroad tracks, streams, or land boundaries.

The office work for the surveys was conducted in the usual manner by checking, correcting, and coding each trip. The coded data were punched on IBM cards for analysis.

[2]For a report on the postcard survey, see "Postcard Method of Obtaining Origin and Destination of Traffic and Comparison with Roadside-Interview Method" on page 10 of this bulletin.

An example of the diversity of trips is shown in Figure 2 for all entering traffic at Station 3 for the 16-hr. period, 7 a. m. to 11 p. m. In the overall origin-and-destination study 38,416 vehicles were stopped for driver interviews at 13 stations. A total of 31,598 postcards were handed out at six stations to drivers, of which 14,375 were returned.

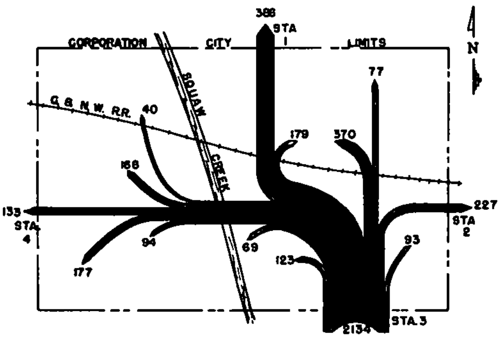

Figure 2. Zone destination of trips of all traffic entering through Station 3 from 7 a.m. to 11 p.m.

METHODS OF SAMPLING TRAFFIC

Interviewing 100 percent of the traffic for origins and destinations requires a great amount of time, cost and personnel. One scheme to reduce the size of these items would be a desirable sampling method in which accuracy and reliability are retained. In conducting an origin-and-destination survey by sampling traffic, one should know the resulting probable error in the number of trips between given pairs of origin-and-destination zones or tracts which would result from the chosen size of sample. Statistical analysis of a known universe can be used as a means of estimating the probable error of a proposed sampling scheme if applied to a traffic stream of similar character.

If correct conclusions are to be reached, the sample must be representative of the total trips. Statisticians have developed satisfactory methods of sampling to gather maximum information. But with samples, however, one faces the possibility that the result obtained may depart from the total population sampled. Fortunately, the uses or origin-and-destination traffic data are such that high degrees of exactness are not essential.

Traffic between pairs of origin-and-destination zones varies in volume, time distribution, composition, trip purposes, and directional peak hours. One pair of zones may have its peak periods at 7 a. m., 11 a. m., and 4 p. m. while another pair may lag an hour. It is apparent that a sample must be drawn throughout the day in order that it will be representative and reasonably accurate. Peak-hour traffic is composed of a more homogeneous form of trip, while the off-peak-hour trips are somewhat more variable in origin and destination. The taking of a larger percentage size of sample during off-peak hours than during peak hours may be desirable for two reasons: (1) the non-homogeneity of the traffic may require a larger sample, and (2) the interview personnel may be kept working at a constant rate to eliminate idle time.

Probable errors can be estimated through the process of statistical evaluation. The standard deviation $\underline{s}$ is an important tool in computing the coefficient of variation $\underline{C}$, standard deviation of the mean $s_{\bar{x}}$, and the confidence limits.

The confidence limits, or fiducial limits, are the extreme points in which the mean of a population may vary for a given probability of correctness. Theoretical conditions must be met, however, if one's inferences are to be correct when using a sampling method.

The terms "total trips" and "whole population" include all types of vehicles. No differentiation is made between trucks or passenger cars. The sampling methods recommended do not include methods of sampling for truck origin and destination separate from other vehicles.

STATISTICAL RESULTS AND EVALUATION

Five methods of sampling were investigated and are discussed. From two of these methods, estimates of probable errors for various numbers of trips between pairs of origin-and-destination zones were established. The five methods are as follows: (1) systematic 10-percent samples of traffic, (2) systematic 20-percent samples of traffic, (3) time-controlled samples of traffic, (4) size-and-time controlled samples of traffic, and (5) samples of unidirectional traffic. In the first four of the above named sampling methods, the

results were calculated separately for each direction of traffic. That is, inbound trips were considered independently from outbound trips, even though the origin and destination of the inbound trip was the destination and origin, respectively, of the outbound trip.

Systematic 10- and 20-Percent Samples of Traffic

The 10-percent and 20-percent sampling methods are systematic schemes of stopping every 10th and every 5th vehicle traveling in the same direction to be interviewed regardless of class, make, or position in the traffic stream. It is assumed that these systematic samples are random samples, because vehicles do not pass an interview station in any schematic manner controlled by origins and destinations. The trips within a 2-min. interval as arrayed for this sampling study are in random order. Within a given 2-min. interval, the interviews can be placed in actual road sequence only when all interviews were taken by the same interviewer, and further, the IBM cards were deliberately randomized within the 2-min. intervals.

Interviewing every 10th or 5th vehicle is not feasible in the field, especially when the specified vehicle to be stopped appears in the middle of a platoon of vehicles. The value of this method for the purpose of this report is its adaptability for use in statistical procedures in estimating probable errors.

Statistical analysis was made possible by sorting the 100-percent interviews for each direction into time sequence based on the 2-min. intervals. For the 10-percent sample, ten groups were then selected, each group being a 10-percent sample of the traffic by direction. Group 1 was composed of the 1st, 11th, 21st, 31st, etc. trip; Group 2 was composed of the 2nd, 12th, 22nd, 32nd, etc. trip. Group 3 was composed of the 3rd, 13th, 23rd, 33rd, etc. trip and so on for each of the 10 groups. Each sample group was then sorted by origins and destinations. The result was ten estimates of the number of trips between each pair of origin and destination zones. The ten estimates were then used in computing the standard deviation, coefficient of variation, and confidence limits by the standard root-mean-square method, where the number of estimates $\underline{n}$ is 10 and the degrees of freedom $\underline{df}$ is 9.

The 20-percent systematic sample was drawn from the IBM listing in a manner similar to that used in drawing the 10-percent sample, but choosing every fifth trip to make up a 20-percent sample. The 1st, 6th, 11th, 16th, etc. trip was selected from the listing of the whole population to make up group 1; group 2 was made up of the 2nd, 7th, 12th, 17th, etc. trip, and similarly to Groups 3, 4, and 5. In the 20-percent sample, the number of trips between each pair of zones had five estimates of the true number. The standard root-mean-square analysis was performed using $\underline{n} = 5$ and $\underline{df} = 4$.

Figure 3 shows the relationship of the computed standard deviations to the number of trips, origin to destination, for the several pairs of zones[3]. Separate curves were drawn with respect to inbound and outbound directions, but since the curves were identical only one curve is shown in Figure 3.

As would be expected, the range and standard deviation are related. If several samples are drawn from a normally distributed population, each sample having nine or ten items, the average of their ranges is about three times the standard deviation of the whole population. As a check of the accuracy of the computed standard deviations, Figure 4, for the 10-percent sample, graphically compares the actual range versus the theoretical range computed from standard deviations. Since there is no large difference between the two curves, it is assumed that the samples were drawn from a normal population and the standard deviations are correct.

The graphical presentation of the standard deviation and coefficient of variation, Figures 3 and 5, respectively, indicate from observation that as the number of trips between pairs of origin-and-destination zones increase the accuracy and reliability increase. These curves are approaching a horizontal line where a smaller percentage sample would produce results of the same accuracy. As expected, the curves for the 20-percent sample reach a flatter stage than the 10-percent sample

[3] The curves of Figure 3 were drawn by judgment with the expectation of fitting them statistically at a later date. As of this writing, however, this work remains to be accomplished.

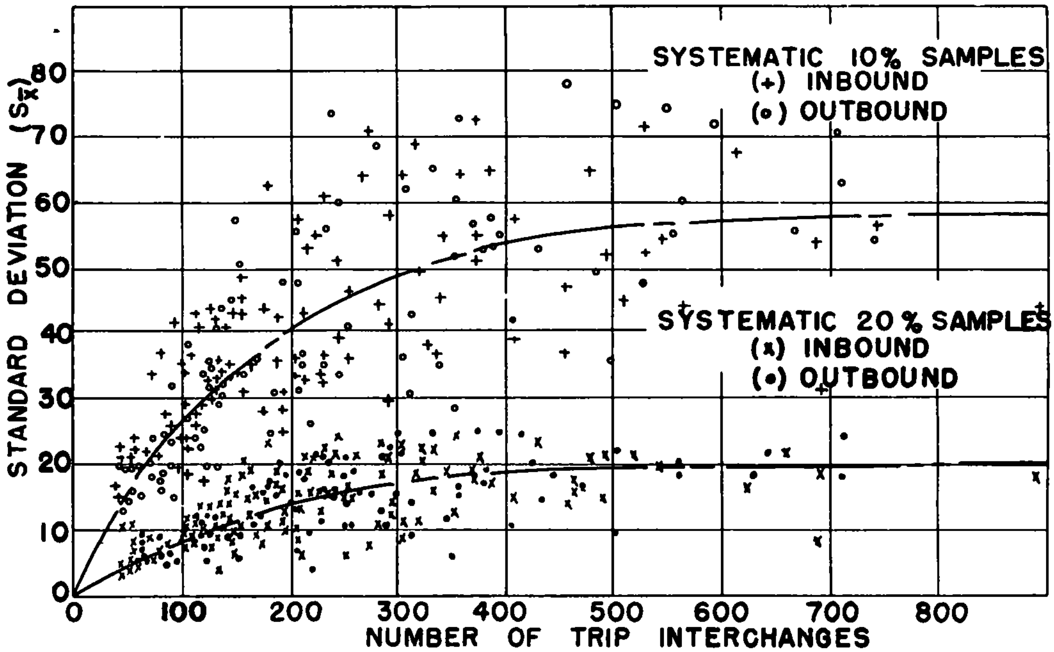

Figure 3. Standard deviations of the systematic 10- and 20-percent samples computed from the number of trip interchanges between pairs of origin-and-destination zones; 50 pairs of zones having less than 100 trip interchanges are omitted from the plot. Stations 1, 2, 3, and 4 are included.

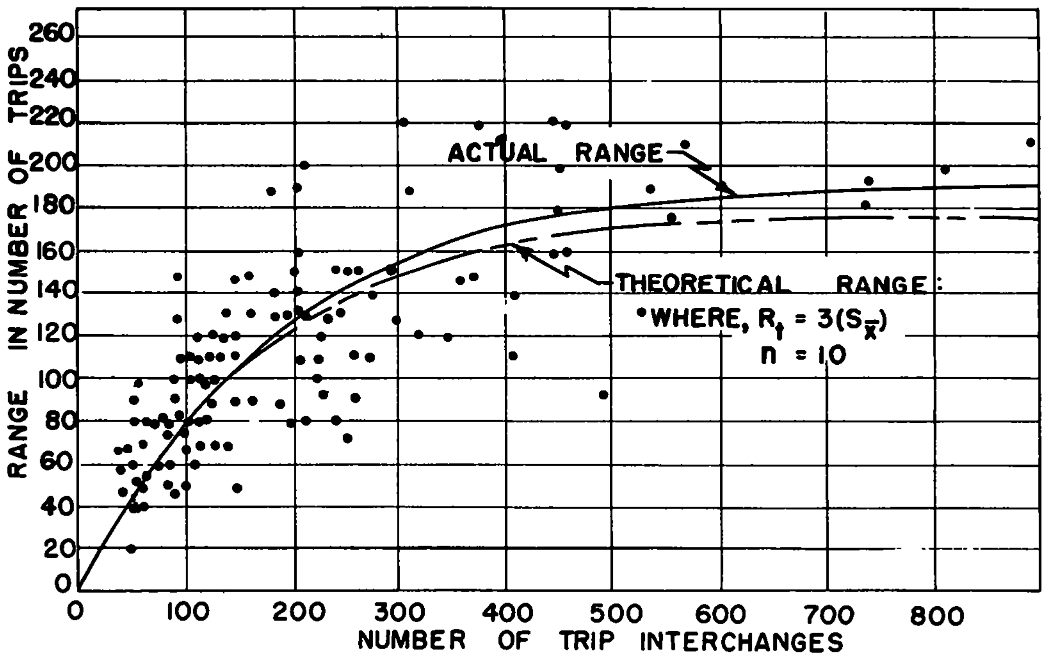

Figure 4. Comparison of actual range of number of trips to the theoretical range as a check of the standard deviation and normalcy of the population of the systematic 10-percent-sampling method.

curves. Thus, the larger sample produces results of lesser probable error.

The coefficients of variation vary according to the subject being sampled, and samples of vehicular traffic appear to have related coefficients of variation. The coefficients of variation computed for 10- and 20-percent sampling of traffic are comparable to those found by Morris for automatic recorder count of traffic volume[4]. It is apparent, therefore, that some relation must exist between samples of traffic for origin and destination and short counts of volume used to estimate average daily traffic.

Figure 6 represents graphically the probable maximum error that one may expect using either a 10- or 20-percent random sampling method. The curves were computed from the standard error of the mean and converted to percent error. A 99.9-percent-confidence limit was used to include all errors resulting from a standard deviation higher than the average curve shows.

Questions have arisen as to the maximum allowable error in the number of trip interchanges between specific pairs of origin-and-destination zones. When a 10-percent or a 20-percent error is specified as tolerable, is it acceptable to allow the same probable error in the number of trips between a pair of zones having a large number of trips as allowed for a less important pair of zones? Because of the relative less importance of the pairs of zones having but few trip interchanges, it should be acceptable to allow the maximum percent error to increase as the number of interchange trips decreases. The curve of Figure 6 is a likely criteria to follow for an allowable or tolerable maximum percent error deviation from the actual.

The systematic 10- and 20-percent samples furnish useful tools by which estimates of probable errors can be made. They show the maximum probable error introduced by sampling a specified number of trip interchanges between pairs of zones. However, this systematic sampling method is not a practical field method because of the difficulties in arranging to interview precisely every tenth or fifth vehicle. To operate an interview station with desirable efficiency and maximum convenience to the drivers, the vehicles should be stopped in groups or chosen singly at times opportune to the traffic flow.

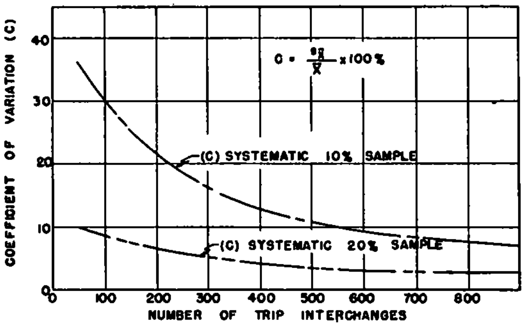

Figure 5. Curves of the coefficient of variation computed from the smooth standard deviation curves of Figure 3.

Time-Controlled Samples of Traffic

Although it is not feasible in the field to interview every fifth or tenth vehicle in the traffic stream, time-controlled samples of traffic, another form of systematic sampling, may be used. The method consists of interviewing all vehicles which reach the interview station during, say, a chosen 2-min. period in each 10-min. period throughout the entire day. Although, under this plan, interviews would be taken during 20 percent of the clock time during the day, the number of vehicles interviewed might vary plus or minus from 20 percent of the total number of vehicles passing the station.

A time-controlled sample was taken for external Station 3 by choosing the vehicles reaching the station during the first 2 min. of every 10-min. period from 7 a.m. to 11 p.m. The original survey provided for listing the trips in their respective 2-min. interval to facilitate studies by time-controlled and other methods. The selections of trips were taken from the inbound and outbound travel separately. Three separate expansions of the sample to the known number of trips in the 16-hr. period were made. The expansions were based on the 10-min., 1-hr., and 16-hr. ratios of the actual 100-percent vehicle interviews to the number of interviews chosen for the sample in each of the three time intervals. Thus, for the 10-min. expansion, 96 expansion

[4]Morris, Mark. Standard Deviation and Coefficient of Variation of Automatic Recorder Counts. Highway Research Board Proceedings, Vol. 30: 336-337. 1951.

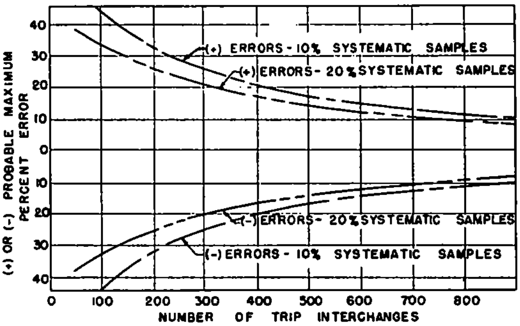

Figure 6. Theoretical, probable maximum-percent error for the 10- and 20-percent systematic samples as computed from the curves of Figure 3 using the standard error of the mean.

factors were used; 16 were used for the 1-hr. expansion; and only one expansion factor was used for the 16-hr. expansion.

The 16-hr. totals for trips between specific pairs of zones as obtained by expansion are shown in Table 1. The 10-min. and 16-hr. expansion factors indicate about the same reliability, and each is better than the 1-hr. expansion. The fact that the rate of vehicles passing the interview station increased and decreased materially from one 10-min. period to another, would indicate a greater accuracy for the 10-min. expansion factor, but such advantage is not significantly apparent.

Probably the main disadvantage of the time-controlled method of sampling traffic would be that of keeping the interview personnel busy at something useful during the short-time intervals that traffic is bypassed. Further, the 2-min. periods are not always proportional in traffic volume to the 10-min. period volume. This method was investigated not that it might be a possible field method but to determine whether or not acceptable results could be obtained from samples taken throughout the day on a time interval basis.

Time - and - Size - Controlled Samples of Traffic

The methods of sampling previously explained were used to establish curves from which to predict maximum probable errors. These curves may be so used because they possess the characteristics of a random sample. The next step then is to develop a method of field sampling of the traffic stream which is workable from the standpoints of personnel organization and of handling the traffic, yet a method producing a representative sample.

One workable method of sampling traffic from the standpoint of field operation is that of selecting time-and-size-controlled samples. A time-and-size-controlled sample would be made up of a predetermined number of vehicles chosen from the traffic stream for a given time interval. As used herein, the number of vehicles to interview in each time interval is predetermined as a stated percentage of the vehicles passing the interview station in the preceding time interval. For example, assuming that a 20-percent overall sample is desired, if 30 vehicles passed the interview station in a given 10-min. period, 20 percent of 30, or six vehicles, would be interviewed in the succeeding 10-min. period. But no specified vehicle would be predetermined to be interviewed; the party chief would have the six vehicles chosen on the basis of the least inconvenience to the traffic and maximum efficiency of his party. The six vehicles would be time spaced so that interview personnel would not be rushed; the choice of vehicles should vary location to location within the platoons and from single vehicles.

For comparison of the time-and-size-controlled samples of traffic with the maximum percent errors established by the systematic 20-percent samples, a test sample controlled both on time and size and again by direction was drawn from the 16-hr., 100-percent interviews of external Station 3. A 20-percent sample was established as the size sample required from each 10-min. interval in each direction. The sample was so chosen from the IBM listing that the number of vehicles from each 10-min. interval throughout the 16-hr. day was 20 percent of the vehicles passing the interview station in each preceding 10-min. interval.

The sample from the 16-hr. listing by 2-min. intervals of the trips passing Station 3 was drawn by selecting the first trip listed in each of the five 2-min. groups comprising the 10-min. period, plus additional selections from the heaviest intervals when more than five trips per 10-min. interval were required. For instance, if seven vehicles were required, five would be drawn as the first vehicle of every 2-min. group with an extra vehicle from each of the two heaviest 2-min. groups. If

TABLE 1

RESULTS OF 10-MIN., 1-HR., AND 16-HR. EXPANSION FACTORS FOR THE TIME-CONTROLLED SAMPLING METHOD

This 20-percent sample consists of the vehicles reaching station 3 during the first 2 min. out of each 10-min. consecutive period for the 16-hr. day.

Zone*		100 percent inter-	Expanded Sample								
			10-minute expansion factors			1-hr expansion factors			16-hr expansion factors		
Origin	Destination	view trips	Trips	Error, percent Plus	Error, percent Minus	Trips	Error, percent Plus	Error, percent Minus	Trips	Error, percent Plus	Error, percent Minus
Inbound Toward City											
A	D	483	478		1.0	462		4.4	460		4.8
A	E	145	188	29 7		198	36.6		180	24.1	
A	F	80	104	30.0		99	23 8		100	25.0	
A	G	223	190		14.8	196		12.1	194		13 0
A	H	153	126		17 6	143		6.5	144		5.9
A	I	141	181	28.4		177	25 5		183	29.8	
B	F	60	38		36.7	40		33 3	50		16.7
B	G	133	138	3 8		135	1.5		127		4.5
C	F	145	145	0.0	0.0	139		4.1	133		8.3
C	G	270	251		7 0	248		8.1	260		3.7
C	H	79	77		2.5	84	6.3		78		1.3
X	X	222	218		1 8	213		4.1	225	1.4	
Total		2134	2134	91.9	81 4	2134	93 7	72.6	2134	80 3	58.2
Total percent				173 3			166.3			138.5	
Outbound From City											
D	A	500	482		3.6	465		7 0	484		3 2
E	A	153	165	7 8		152		0 7	152		0.7
F	A	108	91		15.7	102		5.6	96		11.1
G	A	237	241	1.7		251	5.9		253	6.8	
H	A	153	134		12.4	122		20 2	124		18.9
I	A	138	205	48.6		212	53.6		208	50.7	
F	B	50	39		22.0	33		34.0	34		32.0
G	B	100	104	4.0		119	19 0		118	18.0	
F	C	158	137		13.3	139		12.0	146		7.6
G	C	246	240		2.4	234		4.9	248	0.8	
H	C	91	72		20.9	55		39 6	56		38.5
X	X	230	254	10.4		280	21 7		245	6.5	
Total		2164	2164	72.5	90 3	2164	100 2	124.0	2164	82.8	112.0
Total percent				162.8			224.2			194.8	
Both Directions Combined											
AD	DA	983	960		2.3	927		5 7	944		4.0
AE	EA	298	353	18.5		350	17 4		332	11 4	
AF	FA	188	195	3.7		201	6.9		196	4.3	
AG	GA	460	431		6.3	447		2.8	447		2.8
AH	HA	306	260		15 0	265		13 4	268		12.4
AI	IA	279	386	38.4		389	39 4		391	40.1	
BF	FB	110	77		30.0	73		33 6	84		23.6
BG	GB	233	242	3 9		254	9 0		245	5.2	
CF	FC	303	282		6.9	278		8.3	279		7 9
CG	GC	516	491		4.8	482		6 6	508		1.6
CH	HC	170	149		12.4	139		18.2	134		21.2
X	X	452	472	4.4		493	9.1		470	4.0	
Total		4298	4298	68.9	77.7	4298	81.8	88.6	4298	65.0	73.5
Total percent				146.6			170.4			138.5	

*The letter designations refer to geographical areas as follows:

A and D — States and Counties other than Iowa and Story

B and E — Story County (Ames is located near the western edge of Story County).

C — Rural area contiguous to Ames city

F — City of Ames south of the C and N W. Ry. tracks and east of Squaw Creek plus the rural contiguous area to the east and south.

G — City of Ames north of C and N W Ry tracks and east of Squaw Creek plus the north contiguous rural area.

H — All of city of Ames west of Squaw Creek plus the west contiguous rural area.

I — Iowa State College campus including Pammel Court.

X — Miscellaneous trips including the interchange trips of less than 50 trips.

only four vehicles were wanted, none was drawn from the lightest 2-min. group.

It must be remembered here that there is but little similarity in the manner by which these test samples were chosen and the manner in which actual field samples would be taken except the number of vehicles would be the same. The 100-percent field interviews were recorded by 2-min. periods during the entire 16-hr. day. Because more than one interviewer was working at a time, the interviews cannot be placed in exact time sequence within a 2-min. period, but the 2-min. periods are in actual road time sequence. Within each 2-min. period the IBM listing of the trips is strictly random. To make certain of this, the cards were sorted in such a way that they were in random order. Therefore, the test sample drawn on the basis of the first trip listed in each 2-min. period constitutes a random sample for that 2-min. period.

The test sample as drawn was 24.3 percent of the inbound trips and 23.3 percent of the outbound trips. The samples exceed 20 percent of the total traffic because a fraction of a vehicle was added to some 20-percent calculations to round off the number of vehicles required to whole vehicles.

The highest 10-min. volume of trips in one direction was 53, resulting in a test sample of eleven vehicles for the succeeding 10-min. interval. One interviewer for each direction could probably have handled such a field sampling method because the peak hour called for no more than eleven interviews in any 10-min. interval.

The samples obtained by this method were expanded to the known 16-hr. volume in a similar manner as were the time-controlled samples, that is, by 10-min., 1-hr., and 16-hr. expansion factors. As shown in Table 2, there are no important differences in the separate results obtained by the three expansion factors. Since the selections by 10-min. periods were taken proportionately according to the increase and decrease in vehicular flow, the three methods of expansion would be expected to result in about the same number of trips for each pair of zones. Table 2 shows some minor variation. However, the expansion should be based upon the 10-min. time periods whenever the sample for a time period varied during the day from the chosen fixed percentage of the traffic.

Although both Tables 1 and 2 for specific pairs of zones show a few percentage errors larger than desirable limits, the general range is considered to be acceptable. Both samples gave better results with the inbound direction than with the outbound. The combined directions effects an improvement over the separate directions. The time-and-size-controlled sample of Table 2 shows slight improvement over the time-controlled sample of

Table 1. However, before either of these methods can be adequately evaluated, additional test samples should be drawn. Particularly they need to be applied to interview stations having larger traffic volumes than exist at Ames.

It has been shown that a constant percentage sample based on time and traffic volume will produce reliable results for an analysis of trip origins and destinations. In addition, the method allows enough flexibility so that the percent sample at minimum flow periods may be increased over the constant percent taken at peak flows. This increase in sample during periods of light volume and less homogeneity is desirable as a means of increasing the reliability.

The time-and-size-controlled method of sampling appears to possess the criteria of an acceptable method so that this method may be adopted as a satisfactory method of sampling traffic. Reliability will come from the random sample as chosen by the party chief using a control based on a short interval of time, 10 min. or so. Additional reliability may be achieved by expanding to a 16-hr. or 24-hr. total on the basis of hourly or shorter periods, depending upon the percentage size of sample and its selection. An explanation of the field procedure follows in a succeeding section.

Samples of Unidirectional Traffic

The four sampling methods discussed relate to samples of traffic taken from both directions, in which the destinations of inbound trips were not considered to be the same as the origins of outbound trips. In these test samples of 20 percent of the traffic in each direction, the results produce a sample of 20 percent of the total traffic. Assuming equal volume in each direction for the day, a 20-percent sample from only one direction would be a sample of 10 percent of the total traffic. When the traffic flow is relatively of the same volume in each direction and the origin and destination of the inbound trip are the destination and origin of the outbound trip, a sample of a specific number of interviews from one direction only should be as adequate as a sample of the same total number of interviews, but drawn half from each direction. Sampling of traffic in one direction only would simplify both the field and office procedures and perhaps increase the accuracy of the office work by eliminating certain chances for confusion.

Barkley[5] mentioned surveys where only one direction of traffic was sampled. The outbound traffic was used in most instances so that the route traveled through the urban area being studied cound be obtained from the driver.

Through study of the preceding methods of sampling, i.e., 10-percent systematic sampling, 20-percent systematic sampling, time-controlled sampling, and time-and-size-controlled sampling, data have been made available for comparing the inbound and outbound trips with their respective interchange groups.

Statistical computations used in comparing observed data with expected data are made possible by the use of the statistic, chi-squared. Anderson[6] states that chi-squared is a measure of the degree to which a series of observed frequencies of occurrence deviate from corresponding theoretical frequencies. Values of chi-squared giving a probability of less than 0.05 are generally accepted as an indication that the discrepancies are too great to be attributed to chance.

Each origin-and-destination group of outbound trips was tested with its related group of inbound trips for each of the four external stations. The expected values used in the statistical evaluation was the mean of the two directions for each group. The computed results of Station 1 are shown in Table 3. The total of 14 chi-squares of Station 1 was 15.60 which indicates a probability value greater than 0.30. Stations 2, 3, and 4 were computed in a similar manner and their probability values were greater than 0.70, 0.20, and 0.90, respectively. No significant differences in trip distribution by direction exist in these sets of data, because the probability values are all greater than the 0.05 value. These results are contrary to Anderson's paper, where the one station he tested resulted in a difference too great to be considered reliable.

The results of the preceding four methods of sampling found in Figure 3 and Tables 1 and 2, were analyzed with re-

[5]Barkley, R. E. Origin and Destination Surveys and Traffic Volume Studies, Highway Research Board Bibliography No. 11, Wash., D. C., U. S. Govt. Printing Office 1951.

[6]Anderson, O. K. Statistical Evaluation of Origin-Destination Data. Traffic Engineering. Vol. 22, No. 5: 183-187. February 1952.

TABLE 2

THE 10-MIN., 1-HR., AND 16-HR. EXPANDED TRIPS FOR THE TIME-AND-SIZE-CONTROLLED SAMPLE FROM STATION 3

Zone*		100 percent inter-view trips	Expanded Sample								
			10-minute expansion factors			1-hour expansion factors			16-hour expansion factors		
			Trips	Error, percent		Trips	Error, percent		Trips	Error, percent	
Origin	Desti-nation			Plus	Minus		Plus	Minus		Plus	Minus
Inbound Toward City (24.3 percent sample)											
A	D	463	430		11.0	440		8.9	450		6.8
A	E	145	141		2.8	146	0.7		148	2.1	
A	F	80	73		8.8	76		5.0	74		7.5
A	G	223	258	15.7		252	13.0		259	16.1	
A	H	153	145		5.2	144		5.9	152		0.7
A	I	141	152	7.8		154	9.2		152	7.8	
B	F	60	54		10.0	57		5.0	54		10.0
B	G	133	147	10.5		152	14.3		144	8.3	
C	F	145	104		28.3	110		24.1	103		29.0
C	G	270	277	2.6		278	3.0		284	5.2	
C	H	79	75		5.1	73		7.6	70		11.4
X	X	222	278	25.2		252	13.5		244	9.9	
Total		2134	2134	61.8	71.2	2134	53.7	56.6	2134	49.4	65.4
Total percent				133.0			110.3			114.8	
Outbound From City (23.3 percent sample)											
D	A	600	473		5.4	470		6.0	474		5.2
E	A	153	90		41.2	96		37.3	91		40.5
F	A	108	145	34.3		145	34.3		138	27.8	
G	A	237	279	17.7		287	21.1		280	18.1	
H	A	153	178	16.3		182	19.0		181	18.3	
I	A	138	110		20.3	111		19.6	116		15.9
F	B	50	62	24.0		58	16.0		56	12.0	
G	B	100	102	2.0		101	1.0		108	8.0	
F	C	158	176	11.4		169	7.0		168	6.3	
G	C	246	271	10.2		268	8.9		284	15.4	
H	C	91	117	28.6		122	34.1		116	27.5	
X	X	230	161		30.0	155		32.6	152		33.9
Total		2164	2164	144.5	96.9	2164	141.4	95.5	2164	126.2	95.5
Total percent				241.4			236.9			221.7	
Both Directions Combined											
AD	DA	983	903		8.1	910		7.4	924		6.0
AE	EA	298	231		22.5	242		18.8	239		19.8
AF	FA	188	218	16.0		221	17.6		212	12.8	
AG	GA	460	537	16.7		539	17.2		539	17.2	
AH	HA	306	323	5.6		326	6.5		333	8.8	
AI	IA	279	262		6.1	265		5.0	268		3.9
BF	FB	110	116	5.4		115	4.5		110	0.0	0.0
BG	GB	233	249	6.9		253	8.6		252	8.2	
CF	FC	303	280		7.6	279		7.9	271		10.6
CG	GC	516	548	6.2		546	5.8		568	10.1	
CH	HC	170	192	12.9		195	14.7		186	9.4	
X	X	452	439		2.9	407		10.0	396		12.4
Total		4298	4298	69.7	47.2	4298	74.9	49.1	4298	66.5	52.7
Total percent				116.9			124.0			119.2	

* The letter designations refer to geographical areas as follows:

A and D — States and Counties other than Iowa and Story

B and E — Story County (Ames is located near the western edge of Story County)

C — Rural area contiguous to Ames city

F — City of Ames south of the C. and N.W. Ry. tracks and east of Squaw Creek plus the rural contiguous area to the east and south.

G — City of Ames north of the C. and N.W. Ry. tracks and east of Squaw Creek plus the north contiguous rural areas.

H — All of city of Ames west of Squaw Creek plus the west contiguous rural area.

I — Iowa State College campus including Pammel Court.

X — Miscellaneous trips including the interchange trips of less than 50 trips.

spect to direction. The same maximum probable error for each direction was found for each of the four methods. Since the distribution of trips by pairs of origin-and-destination zones was materially the same for each direction, one can conclude that samples of one direction should produce the same reliability of results as would samples from both directions. The 100-percent interview trips given in Tables 1 and 2 and the expanded samples show no great differences between the inbound and outbound totals for each pair of zones.

Many routes carry greater volumes in one direction than in the other. But when the directional volumes are approximately equal, surely most trips in one direction necessitate return trips in the opposite direction, unless Sunday trip or tourist travel is involved. A 100-percent unidirectional sample of traffic would be approximately a 50-percent sample of the whole population at most traffic stations. Where there is little influence by direction of travel on the total interchange of trips between pairs of origin-and-destination zones, unidirectional sampling would be appropriate.

The advantages of sampling traffic in only one direction outnumber the advantages of two-directional sampling. The interview personnel can work at a more constant rate taking a set percentage of interviews at peak hours with increasing percentage samples at minimum flow hours by a function of the number of personnel available. It is agreed, however, that for 100-percent samples of traffic, the interview personnel may be shifted from one side of the road to the other for efficient handling of peak-hour flows. The bypassing of vehicles that are not to be interviewed presents a problem in location of the station and safety of operation. It is apparent that two-directional sampling of traffic would require two vehicular lanes for each direction, or four lanes. One-direction interview stations could be located where only three lanes were possible and where two-direction interviews could not be maintained simultaneously.

When a two-direction sample of traffic is taken at external stations, all through trips are interviewed twice so that half of such interviews, being duplicates, would not be used in the final analysis. However, all interviews of a one-direction sample can be used, and, assuming that the origin and destination are not influenced by direction, the same reliability would prevail as if both directions were sampled.

RECOMMENDED FIELD SAMPLING PROCEDURE

The basic preliminary investigations for locating cordon lines and roadside stations, and the preparation for coding and analysis required to complete an origin-and-destination study using samples of traffic are the same as when every vehicle is to be interviewed. Interviewing to obtain a time-and-size-controlled sample of traffic differs from 100-percent interviewing only in the number of vehicles stopped for interviewing. Samples taken from only one direction were shown to be reliable for the Ames origin-and-destination study.

One-direction sampling is recommended in conjunction with the time-and-size-controlled sampling method where existing conditions have the same characteristics as Ames.

Size Samples of Traffic

The routes under study may have somewhat different compositions of traffic. Peak-hour volumes and their occurrences throughout the day may differ location to location the same as do origins and destinations. The number of pairs of origin-and-destination zones of interest and the number of interchange trips for each pair also may vary station to station. The size of the sample of traffic and the procedure for taking the sample, therefore, must be flexible enough to meet with these variable conditions. The degree of accuracy of a sample is related both to the absolute size of the sample and its size relative to the population from which it is taken. The maximum size of the sample that can be taken, in turn, is dependent upon the number of interview personnel assigned to the roadside station.

A 20-percent sample of traffic in one direction would be an adequate size of sample for the external cordon located about Ames or any city of the same size with about the same characteristics. However, in other localities a thorough study of the existing facilities and conditions must be made before the size of sample can be determined. The city should be divided into homogeneous zone designations and the zones divided into tracts. The size of the sample should be determined from the number of trips between pairs of origin-and destination zones, but if the number of trips were known a survey would not be required. The desirable size of sample is a function of the number of tract designations within the area to be studied, the total number of trips to and from each tract of interest, and the vehicular volume passing the proposed roadside station.

A traffic engineer familiar with the area and the driving habits of the population, probably could estimate reasonably well the number of trips generated by a given tract. The percentage size of sample would be controlled by the tract of interest having the smallest number of trips and the degree of accuracy desired. For instance, if tract 037 was estimated to generate 100 trips which pass through Station 3, then a 20-percent sample would produce only 20 such trips. These 20 interviews would produce a minimum reliability of 68 percent (Fig. 6). Should this reliability be acceptable for the tract, a 20-percent sample could be taken. Combining adjacent tracts to produce a larger number of trips, would increase reliability and is an acceptable procedure as long as each land area of specific interest does not lose its identity.

Method of Sampling the Traffic

Field procedures when interviewing traffic by time-and-size-controlled sampling methods would not be the same as when 100 percent of the traffic is stopped for interview. Vehicles would be bypassed throughout the day, thus requiring an interview lane and a bypass lane for each direction of traffic interviewed. For a route normally carrying one lane of traffic in each direction, four lanes would be required at the roadside station to avoid waiting and congestion. Again, where feasible, samples taken from one direction only are recommended so that roadside stations may be located for efficiency and safety, using only three lanes.

No specified vehicle should be predetermined to be interviewed, but the number of interviews within a time interval would be predetermined. The time interval may be 10, 15, 20, or 30 min. in length depending on the sharpness of the peaks in the daily volume curve. The vehicles to be drawn from the traffic flow should be controlled by the party chief on the basis of the least inconvenience to the traffic and maximum efficiency of his party. As platoons of vehicles approach the roadside station, the choice of vehicles to be interviewed should not necessarily be the first or the last vehicles in line. The choice should vary, using vehicles of different locations within the platoons.

At peak periods of the day, samples of traffic could be a constant percentage of the vehicle stream. At minimum-flow periods, the percentage size of sample may be greater than the percentage at peak flows as a means of increasing the reliability. The number of interviews to be made during the minimum flow periods can be handled by not more than the number of interviewers that would be adequate at peak hours because there would not likely be need to

TABLE 3

CHI-SQUARED TEST FOR SIGNIFICANT DIFFERENCES IN TRIP DISTRIBUTION BY DIRECTION AT STATION 1 AS COMPUTED FROM THE 100-PERCENT SAMPLE

$$\text{Chi-square} = \frac{2(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

Interchange Groups		Number of Trips			Chi-square Value
O	D	Inbound	Outbound	Expected	
01	11	505	465	485	1.65
01	14	90	67	78.5	3.37
01	12-13	63	66	64.5	0.07
01	15-16	68	85	76.5	1.89
02	11-12	99	108	103.5	0.39
02	13	61	65	63	0.13
02	14	141	133	137	0.23
02	15-16	77	70	73.5	0.33
03	11-12-13	127	125	126	0.02
03	14	164	152	158	0.46
03	15-16	85	61	73	3.95
Total 01 Group		377	376	376.5	0.00
Total 02 Group		384	346	365	1.98
Total 03 Group		726	686	706	1.13
				Total Chi-Square =	15.60*

*Corresponding value of probability is greater than 0.30, inferring no significant differences of trip distribution between outbound and inbound traffic.

interview more vehicles per hour at low flows than during the peak flows.

At Station 3 of the Ames origin-and-destination study, one interviewer could handle a 20-percent sample at peak hours in one direction since the maximum flow was 53 vehicles in a 10-min. interval. The maximum 16-hr. count was 2,164 vehicles inbound. If one interview per minute was taken during the 16-hr. period by the interviewers, there would be a possible 960 interviews or 44 percent of the vehicles interviewed in one direction. Thus, Station 3 could be handled on a unidirectional basis with a party of five or six persons for each 8-hr. shift, consisting of one party chief, one interviewer, one traffic enumerator and classifier, one or two traffic directors, and one relief. The traffic enumerator and classifier may also act as the timer who records the traffic flows by short-time intervals and informs the interviewers of the predetermined number of interviews to be taken in the next time interval. The party chief probably would choose to rotate the assignments and to assist himself in the work as needed.

Standard procedures for choosing vehicles from the traffic stream and establishing a given number of personnel to interview a set percentage of vehicles is difficult to formulate. Rather, the operation of each interview station should be tailor-made to fit the conditions determined from preliminary counts, observations, and experience.

SUMMARY

A reliable method of sampling traffic for interviews of origins and destinations to decrease the cost, time, and personnel is needed. The Ames, Iowa, origin-and-destination study conducted in 1949, where 100 percent of the traffic at four external stations and three interior screen line stations were interviewed and recorded on a 2-min. basis has furnished the data for testing various sampling schemes.

The systematic 10- and 20-percent sampling methods investigated resulted in useful tools by which estimates of probable errors of other methods can be made. Criteria for which maximum errors of specified number of trips between pairs of origin-and-destination zones have been set.

Reliability, or the maximum allowable percent error, varies with the number of trips between pairs of zones or tracts. Based on the systematic 20-percent sample and a 99.9-percent probability, groups of 100 trips were found to have a maximum probable error of 32 percent, while pairs of zones of 800 trip interchanges were found to have a maximum probable error of only 9 percent. However, in field operations it is not feasible to stop exactly every fifth or tenth vehicle for interview, so a more flexible procedure of selecting the vehicles to be interviewed is desirable.

The time-controlled sampling method was investigated to determine whether or not samples can be based on time periods taken throughout the day. It is concluded that there is significant reliability in samples taken during the first 2 min. of every 10-min. period throughout the day. However, this method, also, is not practical in field operation because the interviewers would not be busy during periods when all vehicles were bypassed.

The time-and-size-controlled method of sampling traffic was found to produce reliable trip origins and destinations. The method allows flexibility in field operations so that the percent sample at minimum flow periods may be increased over that taken at peak flows. This time-and-size-controlled method of sampling traffic simplifies the analysis, shortens coding time, reduces the number of field personnel, produces a minimum of vehicular delays, possesses flexibility, and provides accurate data. The time-and-size-controlled method of sampling traffic for inter-

views is recommended for origin-and-destination studies.

Unidirectional samples of traffic were discussed in conjunction with the use of any of the preceding four methods of sampling. Statistical computations used in comparing observed data with expected data were made possible by the use of the statistic, chi-squared. The preceding tests compared trips between pairs of zones by direction and found no significant differences in results. Probably the same reliability would be had from a sample taken from only one direction as would be had from a sample from both directions.

The advantages of sampling only one direction by the time-and-size-controlled method far outweigh the advantages of two-directional sampling where there is little influence by direction of travel on the origin and destination. A 20-percent sample of traffic in one direction, also being a 10-percent sample of traffic from both directions, gave reliable results when applied to the origin-and-destination study of the Ames external stations.

This paper does not suggest that the problem of sampling traffic for origin-and-destination surveys is solved. The contrary is true. The recommended method of sampling is only in the development stage. Field-sampling procedures cannot be perfected on paper. Field work is needed to further test the proposals herein.

More data are available at the Iowa State College Engineering Experiment Station to continue research on the internal Stations 11, 12, and 13 which concern intracity travel. Further, additional sampling studies are needed on external stations under different conditions.

This brief and incomplete study, however, does show that the sampling procedure can be applied to origin-and-destination studies at a great saving in time and cost without sacrificing reliability of results.

PUBLICATIONS OF THE HIGHWAY RESEARCH BOARD

Sponsored by the Department of Traffic and Operations

	Price per copy
Bulletin 6: Report of Committee on Uses of Highway Planning Survey Data including Special Papers (1947) 40 p.	.45
Bulletin 15: Parking, Committee Report and Three Papers (1948) 31 p.	.60
Bulletin 16: Expressways, Committee Report and Three Papers (1948) 21 p.	.45
Bulletin 17: Highway Planning (1948) 45 p.	.60
Bulletin 19: Parking (1949) 78 p.	.90
Bulletin 25: Controlled Access Expressways in Urban Areas (1950) 45 p.	.60
Bulletin 32: One-Way Streets (1950) 39 p.	.60
Bulletin 33: Use of Parking Meter Revenues (1951) 30 p.	.60
Bulletin 36: Pavement Marking (1951) 28 p.	.45
Bulletin 41: Traffic Surveys by Post Cards (1951) 32 p.	.60
Bulletin 48: Off-Street Parking: Legislative Trends, Administrative Agencies (A Summary Report) (1952) 46 p.	.60
Bulletin 50: Weighing Vehicles in Motion (1952) 29 p.	.45
Bulletin 60: Road-User Characteristics (1952) 71 p.	.90
Bulletin 61: Traffic Assignment (1952) 75 p.	1.05
Bulletin 72: Directional Channelization and Determination of Pavement Widths (1953) 53 p.	.75
Bulletin 73: Driver Characteristics and Accidents (1953) 60 p.	.90
Bulletin 74: Traffic-Accident Studies (1953) 57 p.	.90
Bulletin 76: Origin and Destination Surveys: Methods and Costs (1953) 68 p.	.90
Bibliography 1: The Effect of Limiting Access Expressways on Existing Street Systems (1947) 15 p., mimoegraphed	.15
Bibliography 2: A Selected Bibliography on Highway Safety (Annotated) (1947) 51 p.	.45
Bibliography 2 (Supplement No. 1): A Selected Bibliography on Highway Safety (Annotated) (1949) 12 p.	.15
Bibliography 4: Uses of Highway Planning Survey Data (Annotated) (1948) 23 p.	.30
Bibliography 11: Origin-Destination Surveys and Traffic Volume Studies, by Robert Emmanuel Barkley (containing a review of the literature and an annotated bibliography) (1951) 277 p. (postage outside U. S. A. $0.30)	3.00
Special Report 3: Traffic Research Problem Statements (1952) 37 p.	.45
Special Report 5: Channelization — The Design of Highway Intersections at Grade (1952) (size 11" x 17"), plastic binding, 250 p. (postage outside U. S. A. $0.50).	6.00

The Highway Research Board is organized under the auspices of the Division of Engineering and Industrial Research of the National Research Council to provide a clearinghouse for highway research activities and information. The National Research Council is the operating agency of the National Academy of Sciences, a private organization of eminent American scientists chartered in 1863 (under a special act of Congress) to "investigate, examine, experiment, and report on any subject of science or art."