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 informa-

tion 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

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.

cent 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
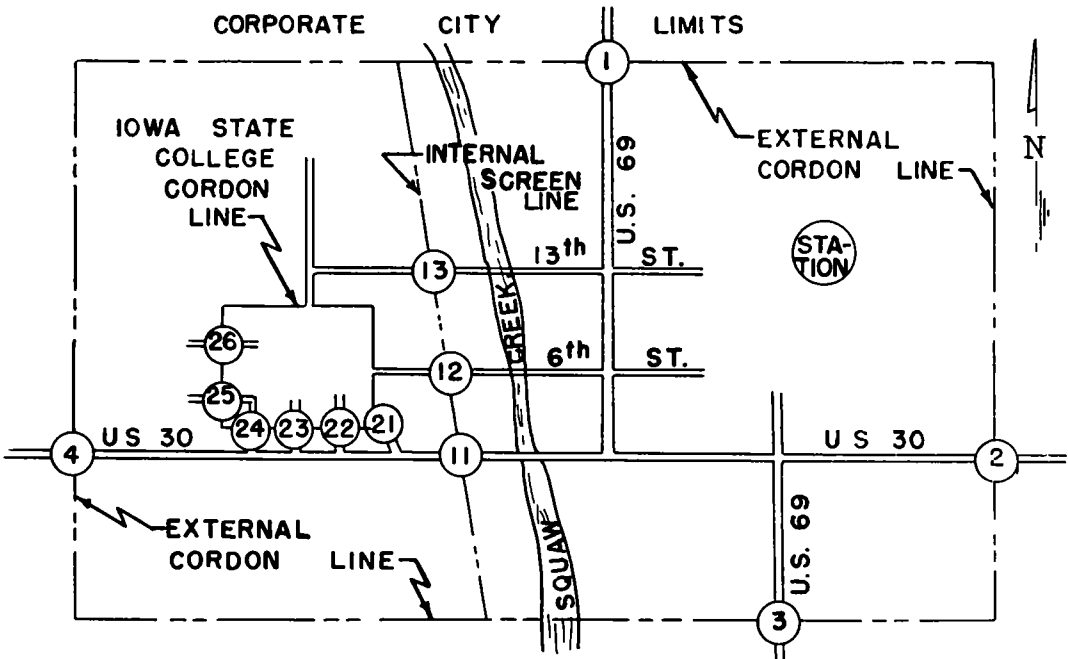

Figure 1. Roadside-interview stations of the Ames origin-and-destination traffic survey.

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-per-

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.

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 $\frac{1}{2}$-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 $\frac{1}{4}$ 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.
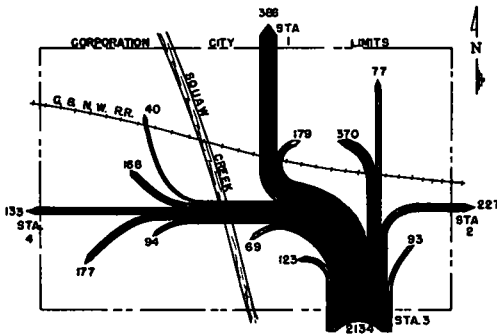

Figure 2. Zone destination of trips of all traffic entering through Station 3 from 7 a.m. to 11 p.m.

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 s is an important tool in computing the coefficient of variation $\underline{C}$, standard deviation of the mean $\underline{s_{\bar{x}}}$, 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 n is 10 and the degrees of freedom 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 $n = 5$ and $df = 4$.

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.
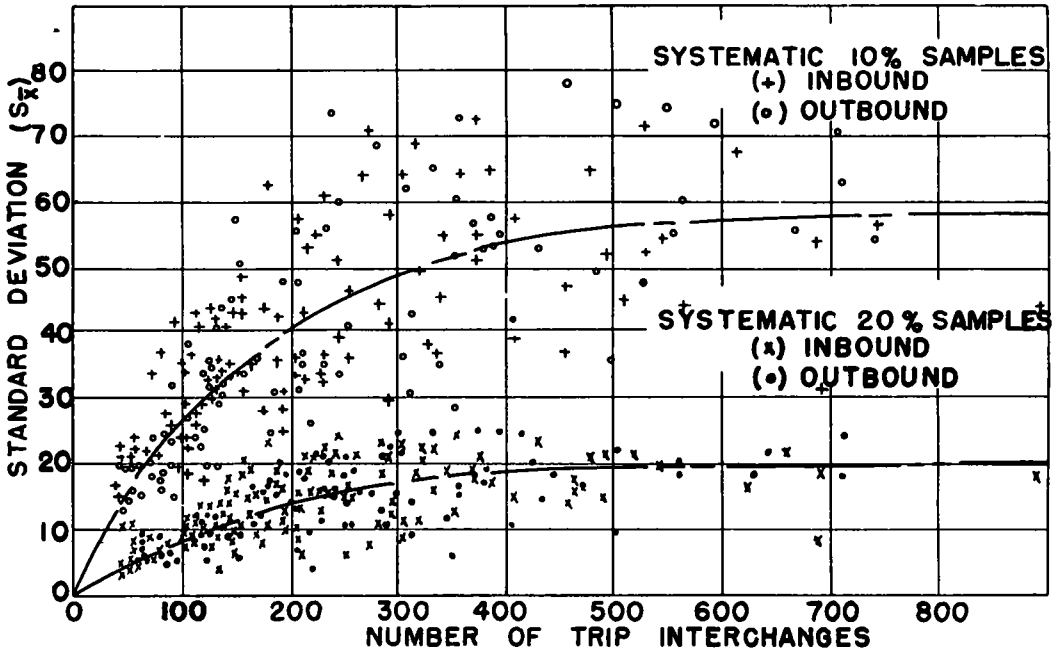
Figure 3. Standard deviations of the systematic 10- and 20-percent samples computed from the number of trip interchanges between pairs of origin-and-destination zones; 50 pairs of zones having less than 100 trip interchanges are omitted from the plot. Stations 1, 2, 3, and 4 are included.
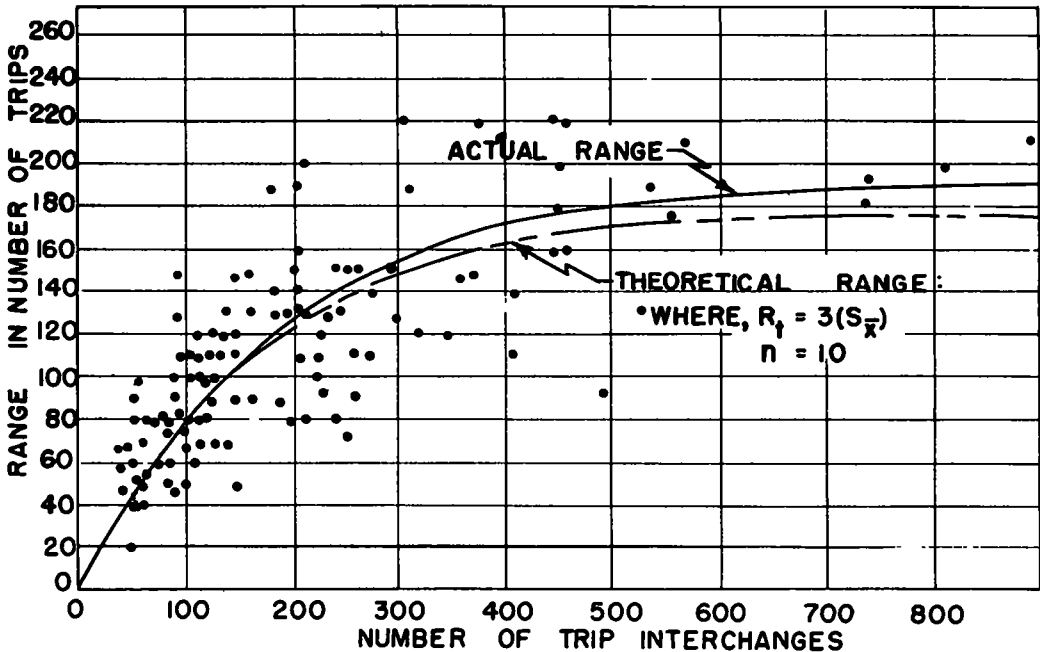
Figure 4. Comparison of actual range of number of trips to the theoretical range as a check of the standard deviation and normalcy of the population of the systematic 10-percent-sampling method.

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

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

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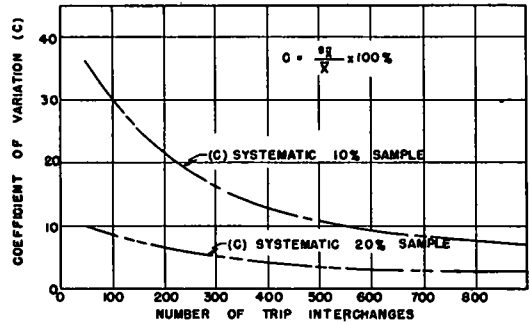
Figure 5. Curves of the coefficient of variation computed from the smooth standard deviation curves of Figure 3.

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

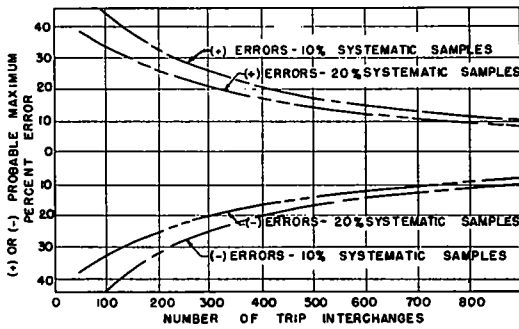Figure 6. Theoretical, probable maximum-percent error for the 10- and 20-percent systematic samples as computed from the curves of Figure 3 using the standard error of the mean.

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- view trips	Expanded Sample								
			10-minute expansion factors			1-hr expansion factors			16-hr expansion factors		
Origin	Desti- nation		Trips	Error, percent Plus	Minus	Trips	Error, percent Plus	Minus	Trips	Error, percent Plus	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		3.3
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	248	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	6.6	
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	68.0	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 could 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		
Origin	Desti-nation		Trips	Error, percent		Trips	Error, percent		Trips	Error, percent	
				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	7 5
A	F	80	73		8. 8	76		5. 0	74		
A	G	223	258	15. 7		252	13 0	5 9	259	16. 1	0. 7
A	H	153	145		5. 2	144	9. 2		152		
A	I	141	153	7 8		154	9. 2	5. 0	152	7 8	10. 0
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		8. 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	500	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 6	
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	5. 2		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	963	903		8. 1	910		7 4	924		6. 0
AE	EA	298	231	22. 6		242		16. 8	239	12.6	19. 8
AF	FA	188	218	16. 0		221	17 6		212	12.6	
AG	GA	460	537	16. 7		539	17 2		539	17 2	
AH	HA	306	323	8. 6		328	6. 6		333	8. 6	
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		188	9. 4	
X	X	452	439		2. 9	407		10. 0	396		12. 4
Total		4298	4298	69 7	47 3	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 uni-

directional 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

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

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, in-
ferring 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 inter-
viewers, there would be a possible 960
interviews or 44 percent of the vehicles
interviewed in one direction. Thus, Sta-
tion 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 enumer-
ator and classifier, one or two traffic di-
rectors, and one relief. The traffic enu-
merator and classifier may also act as the
timer who records the traffic flows by
short-time intervals and informs the inter-
viewers 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 ve-
hicles from the traffic stream and estab-
lishing a given number of personnel to
interview a set percentage of vehicles is
difficult to formulate. Rather, the opera-
tion 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 sta-
tions 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 flexi-
ble 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 sam-
ples taken during the first 2 min. of every
10-min. period throughout the day. How-
ever, 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 re-
liable trip origins and destinations. The
method allows flexibility in field opera-
tions so that the percent sample at mini-
mum 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 person-
nel, produces a minimum of vehicular
delays, possesses flexibility, and provides
accurate data. The time-and-size- con-
trolled 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.