A New Method of Trip Distribution in an Urban Area

ANTHONY R. TOMAZINIS, Chief, Transportation Flow Simulation Section, Penn-Jersey Transportation Study, and Assistant Research Professor of City Planning, University of Pennsylvania

This paper introduces a new method of simulating and projecting trip interchanges within an urban area. * The method is based on probability theory but it also utilizes certain aspects of the gravity models. It has been developed with 1947 O-D data of the Philadelphia metropolitan area and it is currently used in analyzing the 1960 data gathered by the Penn-Jersey Transportation Study.

• THE PURPOSE of this paper is to introduce a new method of trip-end distribution within urban areas. Such a method should be able to reproduce substantially the trip interchange between areal units within a metropolitan region. It should be based on sound and rigorous mathematical derivation and founded on theoretical concepts that satisfactorily explain the manner in which individuals and activities form their trip interchange patterns in urban conglomerations. Beyond these rather basic conditions any method dealing with a problem of such wide dimensions should be conducive to wide application under "laboratory conditions" in any part of the country, with various data, and by any qualified traffic engineer or tranportation planner.

At present, there are three accepted methods of trip-end distribution which have been used in various transportation studies (1, 2, 3). (Besides these three methods one should mention the important variance to gravity models introduced by Wynn (4). There is another method developed by the staff of the Detroit Area Traffic Study but it has not, as yet, been utilized by any other study group since then.) In all three methods, however, there are certain shortcomings referring to their basic theoretical framework or their mathematical formulations which oblige researchers either to impose extensive personal subjective judgment or to expand trends and relationships that they know will not necessarily hold true for the future. (It would suffice perhaps to mention here the controversey over the unstable and highly subjective exponent, X, of the distance in the gravity formulas. The same can be said about the value of L of the CATS formula, although this parameter is defined in more sophisticated terms and is within mathematical limits. Similarly, reference to the arguments concerning the stability of projected trends by the Fratar method should also suffice the need here.)

The following method, developed as a research project of the Penn-Jersey Transportation Study is an attempt to meet the stated requirements. The method is primarily based on the probability theory already introduced to the problem by other researchers, and its final step utilizes one of the known methods of successive approximations (e.g., 1). (Credit should be given to the staff of the Detroit Area Traffic Study for developing the first concept of "probability interchange" and to the staff of the Chicago Area Transportation Study, especially Mr. Schneider, for pioneering introduction of the probability theory concepts to the problem of interzonal movement.) The application of the method has produced extremely encouraging results with data from the 1947 Philadelphia-Camden origin-and-destination survey.

---

*Since presentation of this paper, continuing research under the direction of the author has resulted in substantial modification of the mathematical form of the model. The later model is programed for the IBM 7090 computer for the simulation of the entire 1960 pattern of travel in the region.

DERIVATION OF THE RELATIONSHIPS

If a total population or "universe" N is assumed, then within this universe there are two subpopulations H and S. Part of subpopulation H is also part of subpopulation S and vice versa. If one member of the universe N is randomly selected given that this member is also a member of the subpopulation H, what is the probability that this selection is also a member of the subpopulation S? In other words, what is the probability to randomly select a member of the common subpopulations H and S? Then, the probability that the choice is a member of subpopulation H is

$$P_{(H)} = \frac{H}{N} \tag{1}$$

The probability that the choice is a member of the subpopulation S is

$$P_{(S)} = \frac{S}{N} \tag{2}$$

The probability that the choice is a member of both subpopulations H and S is

$$P_{(HS)} = \frac{GHS}{N} \tag{3}$$

in which GHS = the members participating in both subpopulations H and S. If the probability that $P_{(H)}$ will take place is desired, given that $P_{(S)}$ has already taken place, it would be necessary to try to define the conditional probability of $P_{(H)}$ which in terms of formal probability theory is given as

$$P_{(H/S)} = \frac{P_{(HS)}}{P_{(S)}} = \frac{\frac{GHS}{N}}{\frac{S}{N}} = \frac{GHS}{S} \tag{4}$$

in which $P_{(H/S)}$ = the conditional probability of $P_{(H)}$ given $P_{(S)}$. (Almost every advanced text of probability theory treats the concept of conditional probability sufficiently; for instance, (5).)

The concept can easily be presented in graphic terms as in Figure 1 where N is the universe (total population), H is one subpopulation, S is the other subpopulation, and GHS is the common subpopulations.

This concept of conditional probability can now be adapted to the problem. As a starting point there should be an estimate, or projection, of the total trip opportunities (expressed as trip ends) in the urban area, and the trip opportunities within each subarea (district, zone, etc.). This estimation, or projection, should be based on the findings of a trip generation analysis utilizing observed relationships between land use, transportation facilities, and traffic. Then, the total trip opportunities in the urban area is the statistical total population, "universe" N expressed as total trip destinations. Subpopulation H is the number of trip opportunities (destinations) contained in the area bounded by the areal equivalent of the travel time between the district of origin and the district of destination. Thus, if a district of destination is 26 min travel time from the district of origin (centroid to centroid), then the subpopulation H is all the trip opportunities in the area within 26 min from the district of origin. This is defined by the proper "time code." Each time code can include the area between 2, 3, 6, or any agreed-on time interval in minutes. It should be clear that for the purposes here the subpopulation H is the total cumulative trip opportunities in all time codes until the time code of the district of destination is reached.

Subpopulation S is the total trip opportunities (destinations) within the district of destination only. Moreover, subpopulation S is entirely included in subpopulation H. Thus, in this case, the common subpopulation GHS is the same as S, as shown in Figure 2.

Then, utilizing the equation of conditional probability and having population GHS = S,

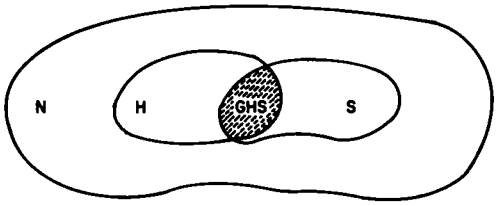

Figure 1.

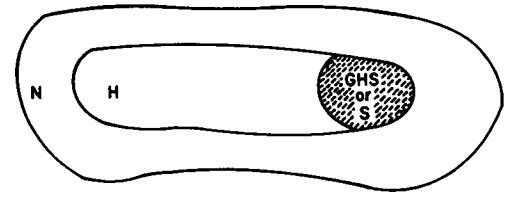

Figure 2.

$$P_{(H)} = \frac{H}{N}\,;\; P_{(S)} = \frac{S}{N}\,;\; P_{(SH)} = \frac{S}{N} = P_{(S)}$$

which gives

$$P_{(S/H)} = \frac{P_{(SH)}}{P_H} = \frac{\frac{S}{N}}{\frac{H}{N}} = \frac{S}{H} \tag{5}$$

Eq. 5 implies that the probability that a trip is going into a district is the ratio between the trip destinations in the district divided by the total trip destinations inside the area delineated by the time code of the district of destination (and having as basis the district of origin). In simpler terms the probability that a trip will go to a district depends on the ratio between the trip opportunites in the district and its competing opportunities. The new concept of competing opportunities is of special importance. According to this concept, the attracting power of a district is conditioned by the number of trip opportunities in the district and then only by the trip opportunities within the same time-distance from the district of origin. The total number of trip opportunities in the metropolitan region enters only indirectly in defining the pertinent probability. This takes place as follows: The summation of the probability of each district within the region should be unity because the total trips distributed should be equal with the trips available in the district of origin, hence:

$$\sum P'_i = 1 \tag{6}$$

in which i represents districts 1, 2 ... n = all the districts into which trips are distributed.

To obtain $\sum P'_i = 1$, the $P_{(S/H)}$ of each district is divided by the summation of all the conditional probabilities, hence:

$$\frac{P_{i\,(S/H)}}{\sum P_{i\,(S/H)}} = P'_i \tag{7}$$

The new, adjusted, probability of each district is then multiplied by the trip origins of the district of origin to obtain the trip interchange. Thus the one-way trips from district X to district Y are given by

$$P'_y \times T_x = T_{xy} \tag{8}$$

in which

P'_y = the adjusted probability of district Y;

T_X = the trip origins in district X; and

T_{XY} = one-way trips from district X to district Y.

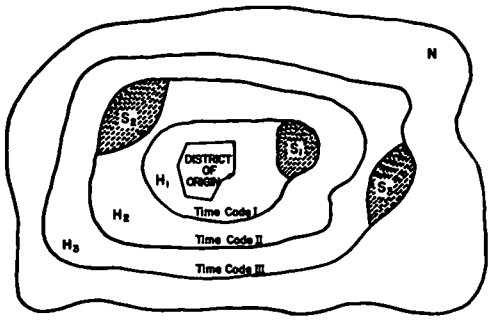

Figure 3

Eq. 8 is practically the final formula of the "theory of competing opportunities." In its general form it expresses the manner in which trips from an area are distributed to any number of other areas. It produces one-way trips given an accurate measurement of trip opportunities in the entire region and every subdivision, and accurate measurements of the time distance between every district of origin and district of destination. (The "time-distance" between districts of course, should be, adjusted for any specific additional impedance which a given link might require; e.g., tolls.) The time-distance measurements are based on the "minimum time path" as it has recently been developed (6, 2:106-107; 7).

When this is repeated for all districts in the metropolitan area, then the total trips going into each district from all the other districts would be estimated. Although the method exactly distributes trips going out of a district, the total trips going into a district might be more or less than the number actually measured or projected through the trip generation techniques. The method has a built-in "balancing out" effect, but it may prove desirable to utilize one of the techniques of successive approximations (e.g., Fratar approximation method) in order to bring estimated and actual trips into a district into perfect balance.

The concept of competing opportunities implies an increase of competing opportunities in a wave-type movement, always increasing the amount of trip opportunities by the amount of opportunities included in each successive time code. Thus the attracting district (potential destination) would have to compete with all the other districts in the same or less time-distance from the district of origin; e.g., with all the potential destinations of equal probability.

The manner according to which this wave-type increase of trip opportunities takes place and the successive derivation of the conditional probabilities is achieved can be shown with the help of Figure 3 and the mathematical formulations which follow, assuming

H_1 = total trip destinations in time code I, including trip destinations of district S_1

H_2 = total trip destinations within time codes I and II, including trip destinations of district S_2

H_3 = total trip destinations within time codes I, II, and III, including trip destinations in district S_3, etc.

then for districts in time code I:

$$P_{(S_1/H_1)} = \frac{S_1}{H_1} = P_1 \text{ and the summation required for the adjusted conditional probability: } \sum P_1 = 1 \tag{9}$$

for districts in time code II:

$$P_{(S_2/H_2)} = \frac{S_2}{H_2} = P_2 \text{ and the summation required for the adjusted conditional probability: } \sum P_2 = 1 - \frac{H_1}{H_2} \tag{10}$$

for districts in time code III:

$$P_{(S_3/H_3)} = \frac{S_3}{H_3} = P_3 \text{ and the summation required for the adjusted probability: } \sum P_3 = 1 - \frac{H_2}{H_3} \tag{11}$$

Thus if there are only three time codes in the urban area, then, the required summation of the conditional probability (Eq. 7) would be

$$\sum P_{i\,(S/H)} = \sum P_1 + \sum P_2 + \sum P_3 = 3 - \left(\frac{H_1}{H_2} + \frac{H_2}{H_3} \right) \tag{12}$$

To generalize this formula and have the summation for any number of time codes, it is assumed that

K = number of time codes

H_1, H_2, H_3 --- H_K = the total trip opportunities in time codes I, II, III --- K

then,

$$\sum P_{i\,(S/H)} = K - \left(\frac{H_1}{H_2} + \frac{H_2}{H_3} \text{ --- } \frac{H_{K-1}}{H_K} \right) = K - \sum_{n=1}^{n=K} \frac{H_{K-1}}{H_K} \tag{13}$$

There are several particular aspects of this new method which should be stressed. First, the method is based on a theoretical concept developed by Bernoulli. This concept, usually called the Bernoulli theorem states that "if the probability of one event is P, the number of attempts is m, and the number of successes n, then the probability, P_K, that the difference between P and the ratio n/m is smaller than any preassigned number (however small) is unity, if the number of attempts is sufficiently large" (for mathematical proof, see 8, ch. 4, pp. 82-116). The essence of this theorem is the basis of the new method when it postulates that the probability of attracting one trip multiplied by the number of origins in the district of origin will turn up the one-way interchange between the two districts (Eq. 8). Such basis, however, sets several conditions, the most important of which is in relation to the minimum number of trips (attempts) required in each case before any distribution be undertaken with reliability. To increase the number of "trips to be distributed," the first step is to aggregate zones into districts. This implies a spatial aggregation that might or might not be of importance, depending on the degree of spatial details required by each study.

A second aspect of aggregation is of increased importance. This involves the aggregation of districts into time codes. Here is a critical point of the new method. The basic objective of the "theory of competing opportunities," as outlined earlier, is the configuration of a function expressing the diminishing probability for interchange at increasing distances from a given district of origin. This probability is defined by the exact number of the competing opportunities in each case. In theory, then, an avoidance of any kind of aggregation would provide a theoretically better estimate of the probability of a trip destination to attract a trip.

Although the new method provides such a potentiality, the reality of the situation makes it impossible. This is so because there is no practical means to determine the actual competing opportunities exactly in each case. Besides, all trip destinations of a particular strata are competing among themselves for the attraction of trip of this strata. The difference is that some of these trip destinations have an edge over the others (higher probability for attraction) because of their closer location. An additional point is that even if the theoretical probability and the computed probability were the

same, there would still be the problem of the empirical discrepancy that should be expected between estimates derived by any theoretical probability and the actual observations, or in other words there are the implications of the Bernoulli theorem at the destination points. A reasonable objective then would be an achievement of a balance (at the point of minimum total error) between the error due to the distortion in the rate of diminution of the probability of attraction of each trip destination and the error due to the expected variation between the theoretical probability and the empirical estimates.

Before the width of time codes can be decided, however, another thought should be explored. The objective here is to simulate human behavior in choosing the destination of a trip within a complex set of trip destinations of a region. This takes place on the basis of two elements of experience; e.g., experience with the transportation system considered, and the ability and sensitivity to count and utilize time in small increments. What is implied is that one cannot have a first time code of 5 min for a mass transit system where "waiting time" alone may be more than 5 min. It also implies that for auto travel or mass transit travel one should not have time codes of odd increments (of say, 6.25 min) but of blocks of time that are easily conceived and frequently used by people in their everyday activities; e.g., time codes of 5 or 10 min.

This consideration of simulating human behavior has proven of substantial importance where best results were achieved with time codes of 5 min driving time in distributing auto trips. In the case of mass transit trips the first time code was 20 min in most of the area and 30 min for the districts at the outskirts of the metropolitan region. The rest of the time codes were of 5 min riding time. This variation of the width of the first time code was in response to the variation of "waiting time" in the system, which ran up to 12 and 18 min, respectively.

(The distribution of mass transit trips presents some additional difficulties which no researcher cared to present until now. For instance, all the known methods of trip distribution have the form of a diminishing function; that is, the interchange decreases when the distance from the origin increases. This is not true, however, in the case of mass transit trips where the frequency of trips has its peak around 25 to 35 min of the origin. Such an "amfitonic" function, however, presents more problems than any auto trip distribution function with its "monotonic" form. The problem was faced in Penn-Jersey, but there was not enough data to incorporate any systematic solution with the 1947 O-D data. For 1960 data, the analysis proceeds in a more systematic manner, and the aggregation of the first time codes vary according to a verified relationship with the "mean length" of transit trips of each district. Of course, a need for a sound projection of the "mean length" of transit trips is implied here, a project that is currently under way in the Penn-Jersey Transportation Study.

The preceding prargraphs imply a need to introduce a dichotomy between auto trips and mass transit trips. For a small city without major mass transit facilities this is not an important dichotomy and need not be carried out. However, when such facilities exist in a large scale, then there is a substantial distortion of reality if the various districts would be arranged from a given point of origin on the basis of a single minimum path based on highway facilities. This is true because "combined" (all modes) minimum paths are difficult to estimate objectively. Thus, it became evident that suitable time codes by auto and mass transit should be established in each case on the basis of the pertinent minimum time paths.

Here, two points should be noted. First, a minimum time path is in essence an accumulation of the minimum penalties (expressed in time) per unit of distance overcome in traveling the links of the path. Thus, tolls or any other particular impedance accruing to the traveler on the link should be incorporated in calculating the minimum time path. Second, there is the evident need "to generate" mass transit trips as part of the original trip generation study. This is a basic requirement if a modal split is going to be undertaken. A method of achieving this mass transit trip generation in terms of "percent of total trips" is currently under study.

A further ramification is derived if purpose of trips is taken into account. This is important for many reasons, one of which is the fact that work trips are mainly undertaken during the peak travel hours. Therefore, a peak-hour travel minimum time path

is more appropriate. The opposite seems more appropriate for all the other purpose trips.

All the preceding particular explanatory points make evident the extensive potentialities of the new method in studying and analyzing conditions of interaction among activities with complex relationships. It also indicates the degree of refinement the method can reach. Its limits and potentialities are stemming from its dual nature, because it combines the rigorous structure of the theory of probability and the ingenious device of gravity models in arranging districts with the same probability of interaction in terms of a spatial permutation on the basis of time-distance from a given point of origin. (Credit should be given for this device to H. Wynn and A. Voorhees for their extensive and pioneering work on gravity models under various names and forms.) This new method combines features of both methods and lies somewhere between the Chicago probability method and "gravity" models. As it stands today the new method succeeds to state the hypothesis objectively and test it without the help of any subjective "correction" of the results through either the utilization of an exponent (as, for example, in the gravity model) derived as an empirical average of widely varying values, or of a parameter (as, for example, in the Chicago method) derived on the basis of an unknown future trip-length distribution and number of intrazonal trips in a manner not clearly demonstrating mathematically significant relationships. Further, the method can take fully into account future variations in all three important components of the problem; namely, the land use at the origin, the land use at the destination, as well as at the land use between and about the point of origin and potential destination.

Changes in the tranportation system are also readily and explicitly incorporated. In fact, the basic concepts of this method are currently utilized in the method of modal split developed by PJ. The same basic concepts are also under scrutiny for application to a land use or regional growth model.

REPORT ON RESULTS

Before presenting a definite example indicating the manner in which the new method is applied, significant points of the results of the empirical investigations carried out with 1947 O-D data can be reported. Figures 4 to 17 indicate the actual and the probability frequencies of seven test districts. For each district the person-trip interchange by auto and mass transit is separately reported. The 1947 O-D survey included 124 districts. From these, seventeen districts were originally selected. Of these, twelve districts were completely analyzed. Three of these districts are located on the New Jersey side of the metropolitan area and the other nine at various locations within the City of Philadelphia. Application of the new method has produced highly satisfactory results for all twelve districts tested when distributing mass transit trips. The results of auto trip distribution for 8 of these districts was also highly satisfactory. Applying the method to simulate auto interchanges for the remaining 4 districts produced less satisfactory results. In these districts the actual interchange is reproduced with less accuracy although the calculated frequency follows the same general form of the actual frequency of interchange.

The four districts with these discrepancies are the three districts located in New Jersey and the district representing the core of the Philadelphia central business district. The calculated auto interchanges of the three New Jersey districts show a more evenly distributed interchange than actually reported. In fact, districts in New Jersey consistently receive more trips from other New Jersey districts than the method predicts. The reverse is true for districts on the other side of the Delaware River which receive less trips than trips predicted with the method. The existence of the Delaware River with a limited number of toll crossings is affecting the trip distribution pattern within the region.

To take the river into account, an additional time penalty of 10 min was established between New Jersey and Philadelphia. This substantially improved the results but did not bring them as close as the results of the Pennsylvania districts. There was, however, no means to improve the results further for many reasons. First, the cor-

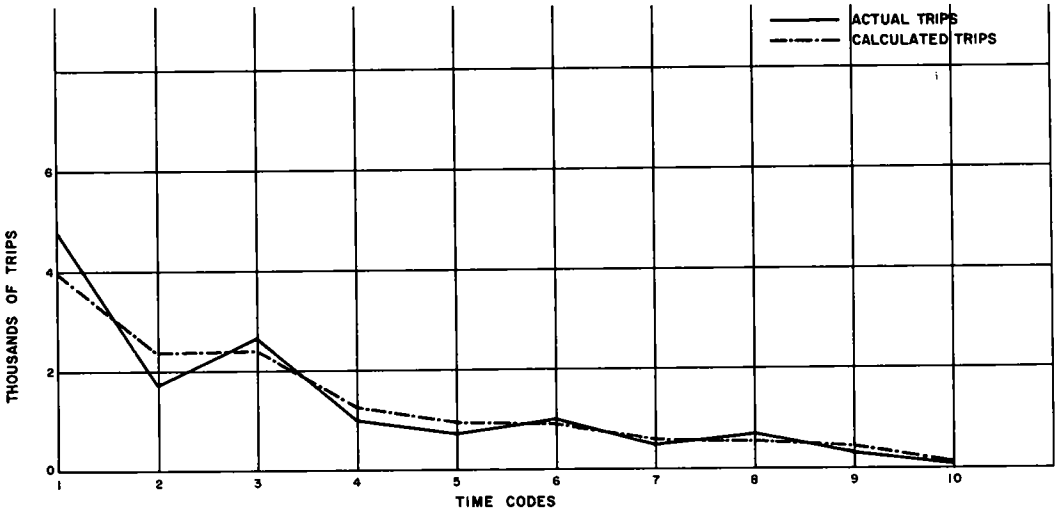

Figure 4. Actual and calculated auto and taxi trip ends distributed by time code, District 012.

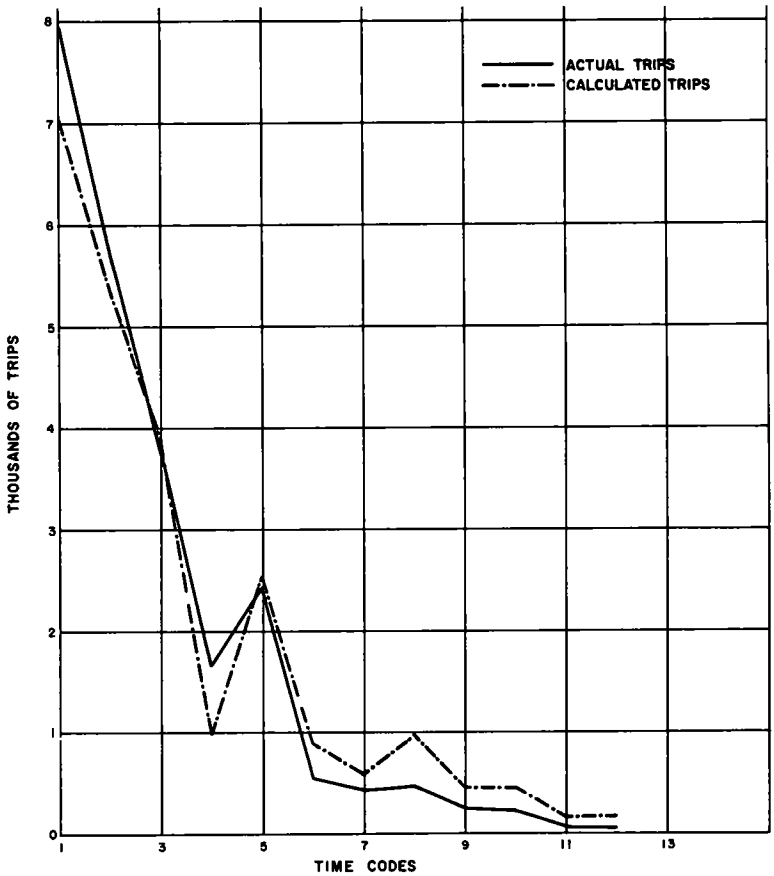

Figure 5. Actual and calculated mass transit trip ends distributed by time code, District 012.

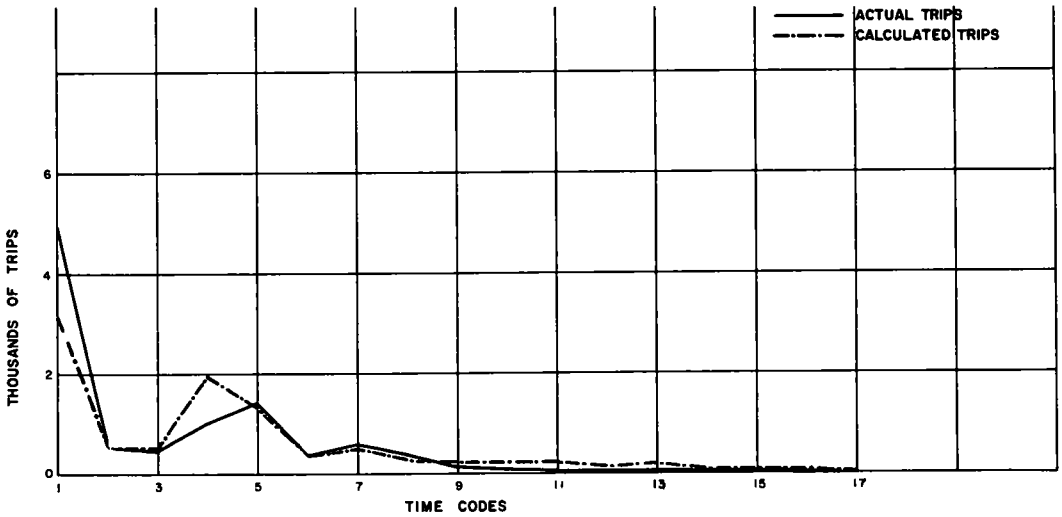

Figure 6. Actual and calculated mass transit trip ends distributed by time code, District 030.

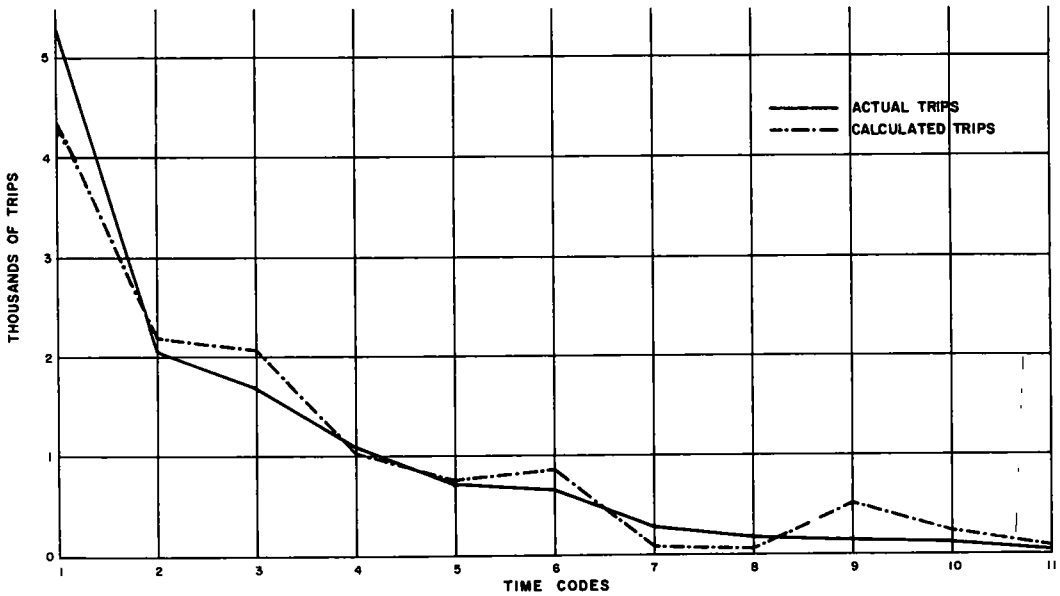

Figure 7. Actual and calculated auto and taxi trip ends distributed by time code, District 030.

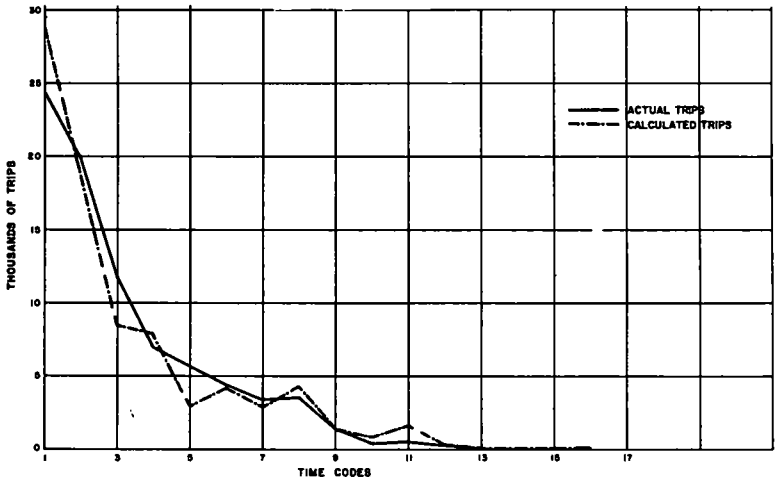

Figure 8. Actual and calculated mass transit trip ends distributed by time code, District 041.

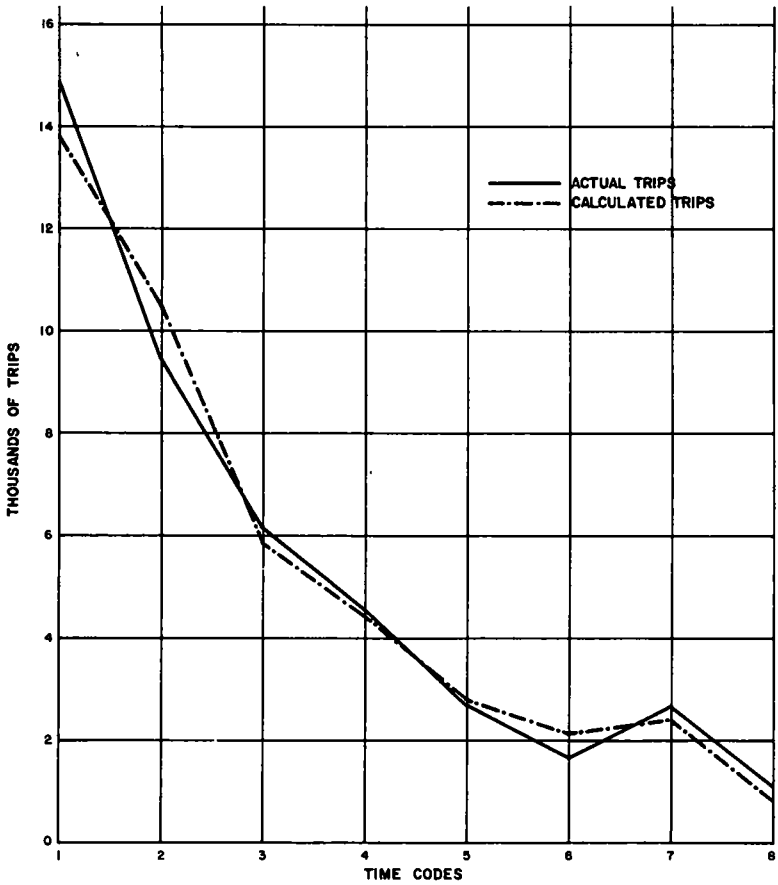

Figure 9. Actual and calculated auto and taxi trip ends distributed by time code, District 041.

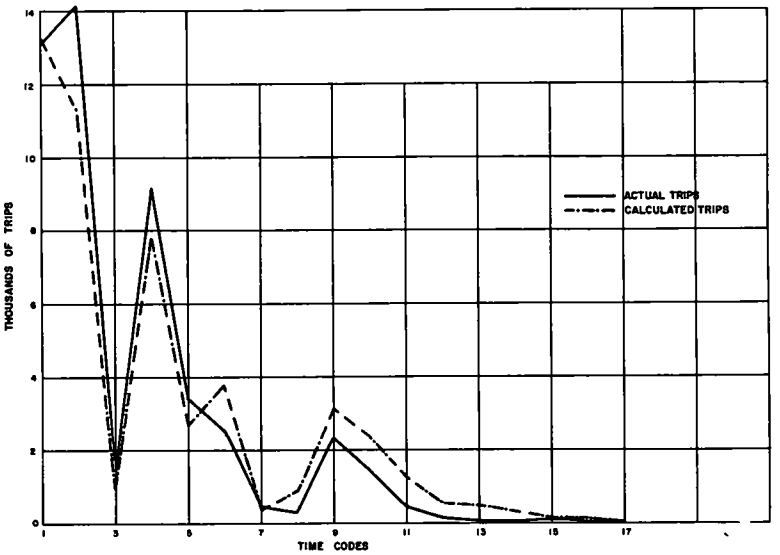

Figure 10. Actual and calculated mass transit trip ends distributed by time code, District 054

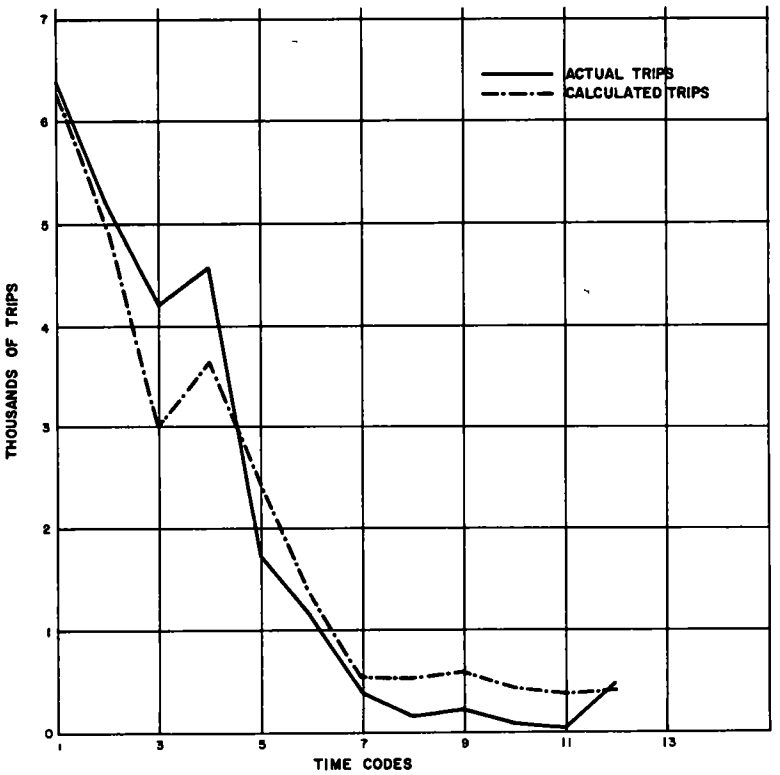

Figure 11. Actual and calculated auto and taxi trip ends distributed by time code, District 054.

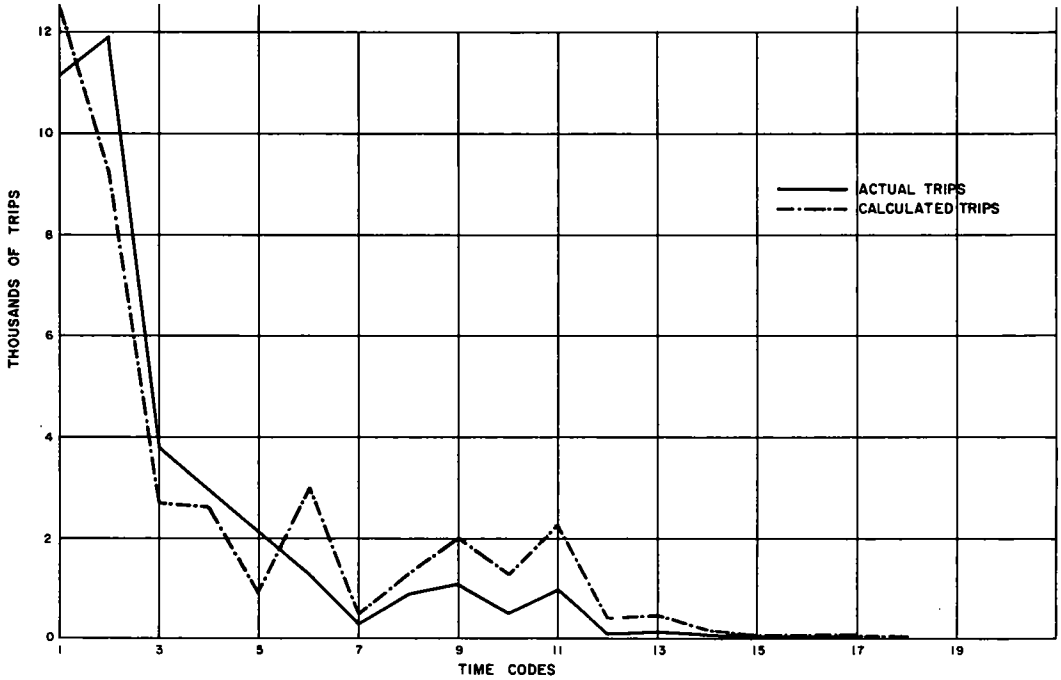

Figure 12. Actual and calculated mass transit trip ends distributed by time code, District 063.

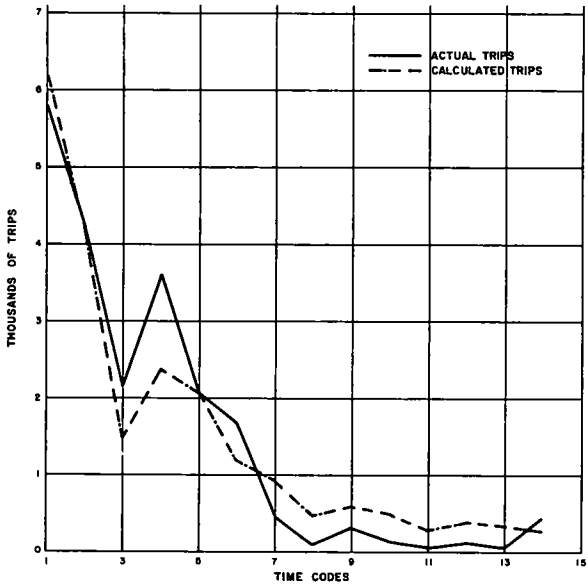

Figure 13. Actual and calculated auto and taxi trip ends distributed by time code, District 063.

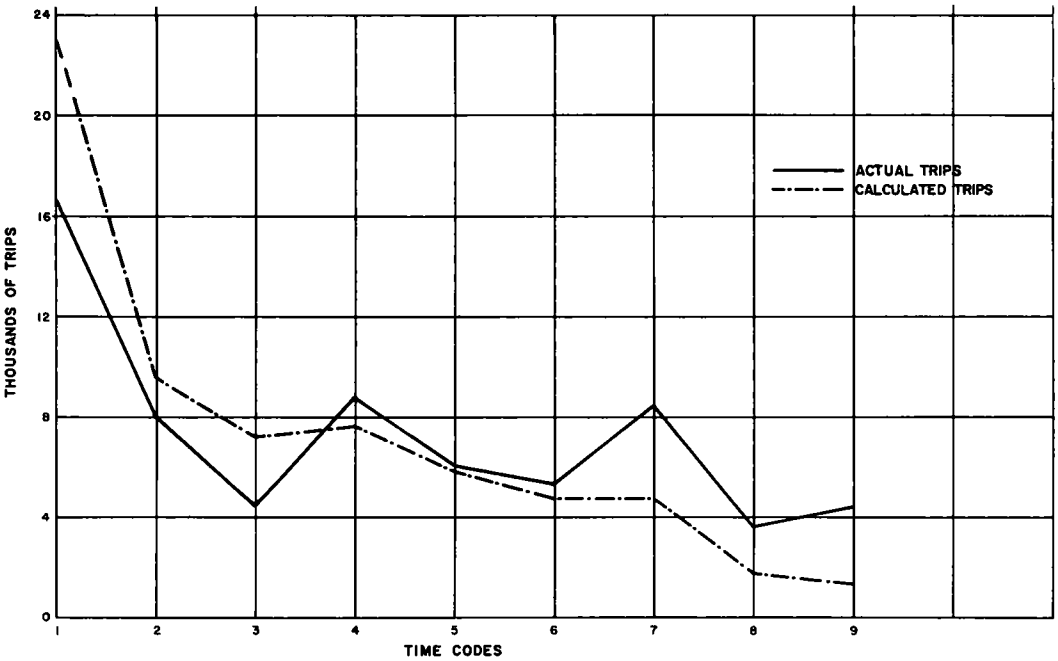

Figure 14. Actual and calculated auto and taxi trip ends distributed by time code, District 000.

don line of the 1947 survey was closer in at the New Jersey side, so even the introduction of a 10-min penalty on the bridges did not substantially alter the relative position of districts for New Jersey origins because within 15 or 20 min all districts in New Jersey could be reached from a bridge. Thus, further time penalties on the bridge could further remove the Pennsylvania districts but it could not produce a difference in the sequence of districts, which, as can be recalled, is crucial in importance.

Second, the 1947 survey reports only unlinked trips. Thus all the trips which consist of two legs—one driving or riding on local bus to Camden and another on the rapid transit to the City of Philadelphia—are shown as two trips, the first of which is shown as having destination the City of Camden, a reporting which is obviously not correct. Third, the results could not be further improved due to an inability to introduce a strata of subclassification of trip interchanges or trip destinations. The 1947 O-D data do not provide for such refinements. Thus, the results are reported as they were reached but the three improvements outlined—all within the model for 1960 trip distribution analysis—would turn up highly satisfactory results when they would be incorporated, as it is done currently with the 1960 data. (Preliminary results of the 1960 data analysis incorporating only the two first improvements and distributing trips only by mode (all purposes aggregated) have already verified this belief.)

The case of district 000, representing the core of the CBD of the City of Philadelphia, is similar in nature. The method reproduces the general form of the distribution but distributes more trips to districts close by and less trips to districts further out than the actual interchange does. No improvement of the original frequency was possible, however, with the 1947 data. This was so because of the reasons of this discrepancy. First, the home interview data do systematically understate non-home-based, intra-CBD trips because trips are reported in the original home interveiw questionnaire by the housewife who usually does not know all the non-home-based trips made by her husband. Second, the parking situation in and around the CBD intervenes to

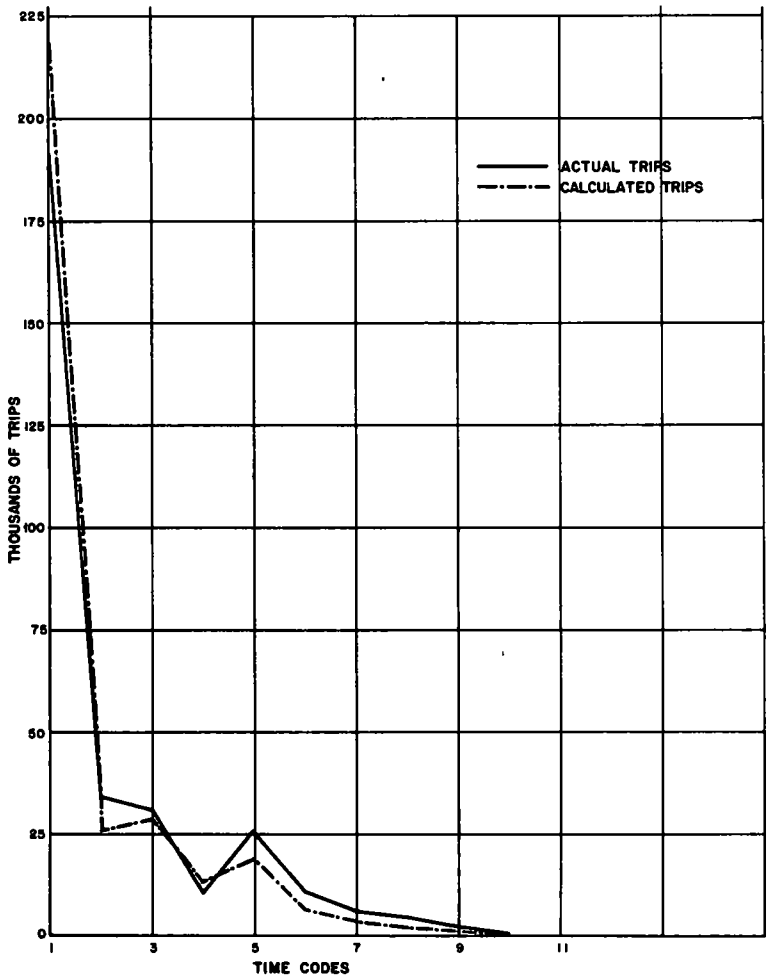

Figure 15. Actual and calculated mass transit trip ends distributed by time code, District 000.

alter circumstances further. One can speculate that some people do not use cars to travel within the CBD and in its vicinity (e.g., within 5 to 15 min driving time) in the City of Philadelphia because it is excessively costly to travel on these links and parking a car may be inconvenient and costly. Thus, in the case of the downtown area, potential auto trips result in either a combined multipurpose trip or in a mass transit trip, or most frequently in a trip on foot.

There is, however, a third important reason. Auto trips made to downtown are not made in the same proportion by the entire population of a region. These trips are primarily made by a certain stratification of the population that can afford the expenses of taking a car into the CBD. The spatial distribution of these people, however, is highly localized in every metropolitan region with a concentration toward the outer areas of the region.

The 1947 O-D data used for this analysis do not permit any of the above reasons to be taken into account. This, of course, is not the case concerning the 1960 data where subdivisions of trip interchanges and trip destinations by population strata is possible. In addition, parking limitations in the CBD can be taken into account in reproducing the inter-downtown auto trips. (Again, preliminary results of the 1960 data analysis

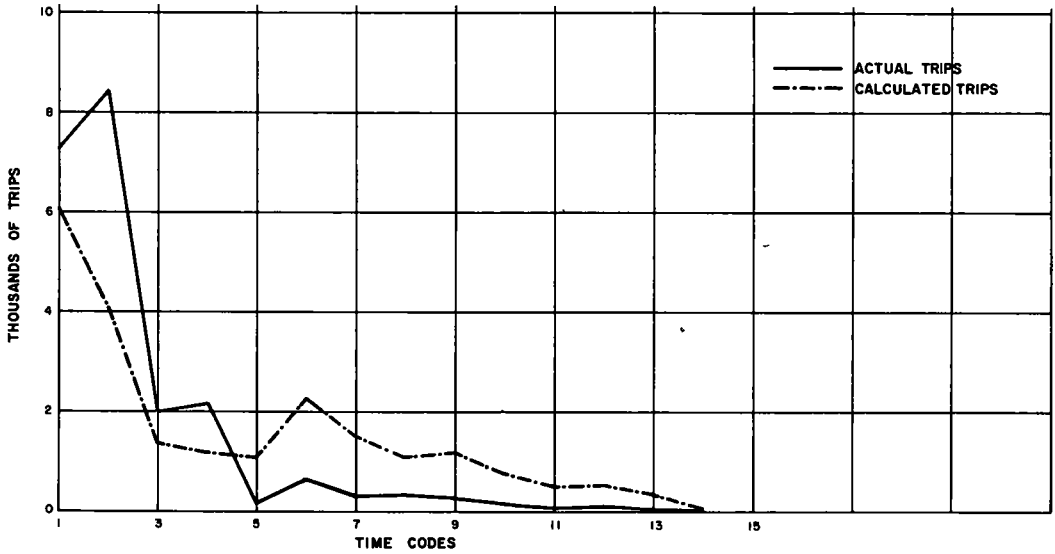

Figure 16. Actual and calculated auto and taxi trip ends distributed by time code, District 421.

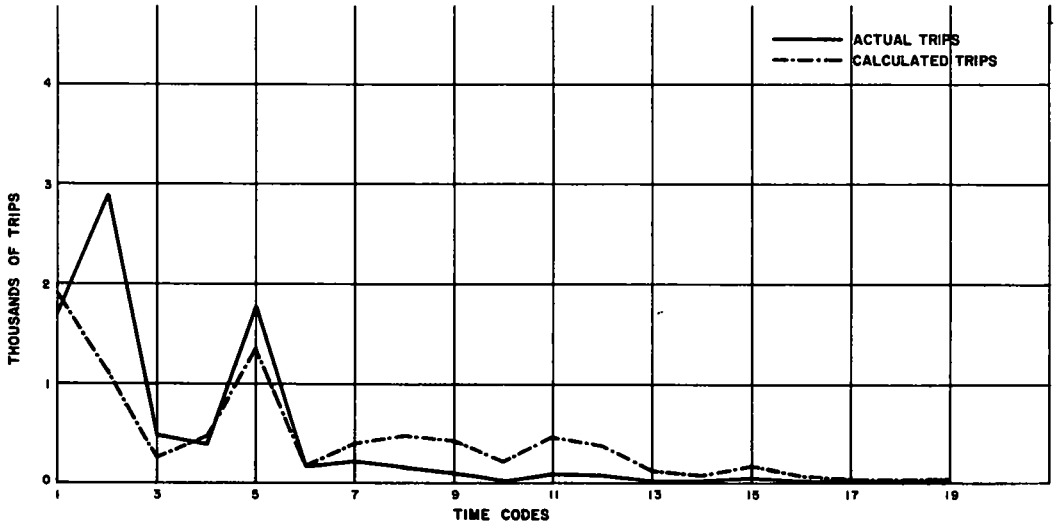

Figure 17. Actual and calculated mass transit trip ends distributed by time code, District 421.

have turned up extremely satisfactory simulation of actual interchanges of auto trips when only a 5-min parking penalty was incorporated for trips going into a congested district. Mass transit trips of the CBD did not present a problem in either case.)

These discrepancies as well as the degree of simulation achieved by the new method are shown by the seven districts reported here. Five of these districts are taken from ones with good results and two (district 000 and New Jersey district 421) are taken from the four districts with less satisfactory results.

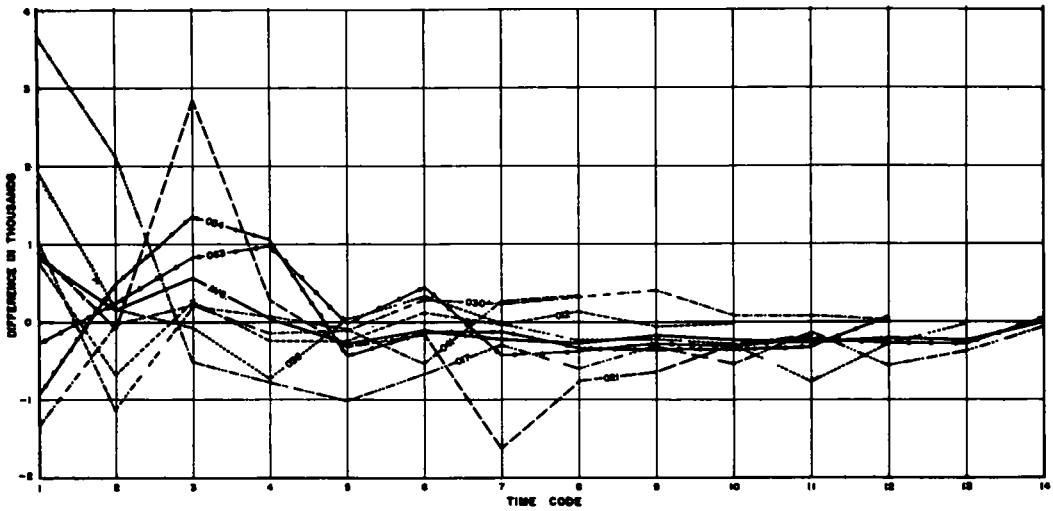

Figure 18. Absolute differences between actual and calculated auto and taxi trip interchanges, by time code.

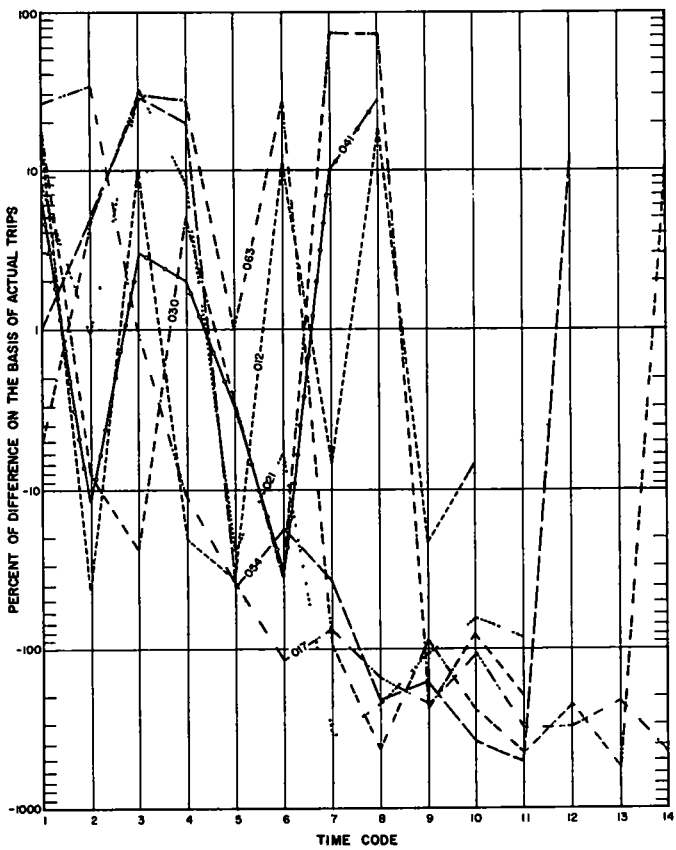

Figure 19.

COMPARISON AND FURTHER TESTS

When the trip-end distribution analysis of these twelve districts was completed, then a number of tests and comparisons were undertaken. First, graphs were compiled as in Figures 18 and 19 where the differences in the interchange of auto and mass transit trips of all districts, grouped by time code, were depicted. The differences were graphed on absolute as well as percent basis. These figures show that the absolute differences are larger at the early time codes, where the total interchange is large, too. This relationship is shown in the second graph where for the early time codes the percent of differences are much smaller than they are at the further out time codes. To study the interchanges in more detail, the graph of Figure 20 was formed. In this graph each interchange of auto trips of the seven reported districts is plotted. In this figure the relation between actual and calculated trips becomes very clear. The reader would notice that as long as the actual interchange is below 1,000 trips the relationship is much farther out from the 45° line which represents a perfect reproduction. This observation is entirely what should be expected according to the theoretical concepts explained earlier because they are variations which the theory of probability expects normally in a phenomenon of stochastic nature as the phenomenon of trip interchange within an urban region. Also, however, most of this variation falls below the 45° line. This means that the calculated trip interchange of longer trips is frequently higher than the reported actual trip interchange. Two reasons stand behind this systematic varia-

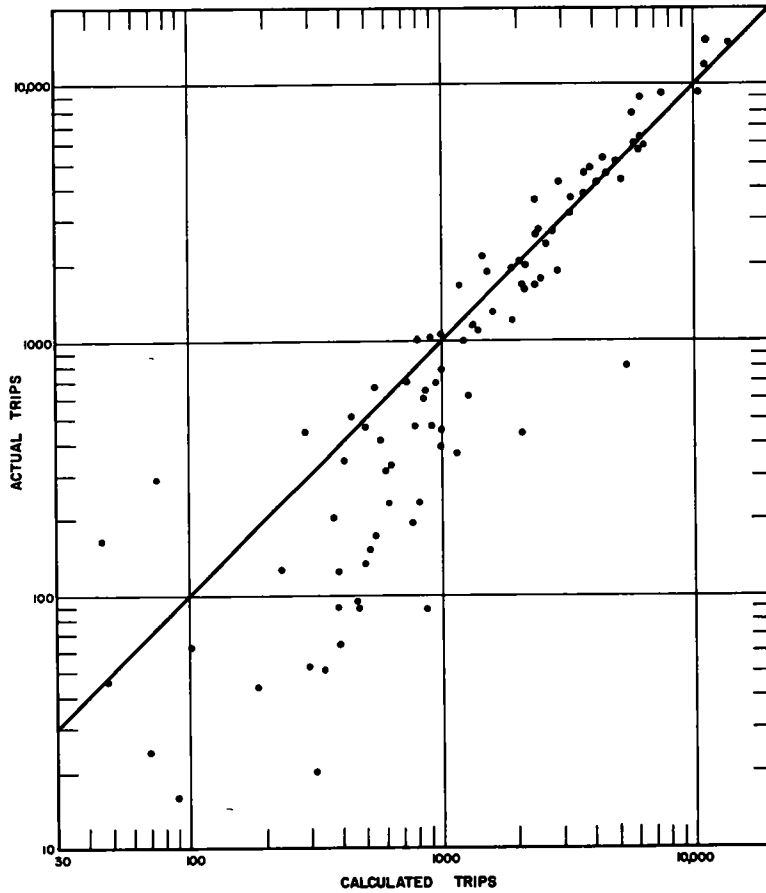

Figure 20.

tion: (a) the fact that the 1947 O-D data were not factored after screenline checks were made, and (b) the fact that in the analysis the internal-external trips (trips going to the stations) were not subtracted.

(Since the basic theory of this method, as well as some other methods, is probabilistic in nature, it would be a mistake to distribute trips on the basis of a "universe N," which misrepresents the situation. For instance, the number of trips going from the city to Washington, D. C., should not be estimated on the basis of the number of trips counted at the entering station. Thus, when trips of a district are distributed, the trip destinations at the stations should be subtracted from the "universe," and the origins that are going out of the metropolitan area subtracted from the district. When this process is completed, then the distribution of the trips from the stations to all dis-

TABLE 1

ACTUAL INTERCHANGES AND PROPORTION OF ERROR OF PREDICTED INTERCHANGES OF VEHICLE TRIPS FOR SEVEN ANALYSIS DISTRICTS BY TIME CODES OUTWARD FROM EACH DISTRICT CENTER (PJTS 1947 O-D data)[a,b]

Time[c] Code	District 012		District 017		District 021		District 030		District 041		District 054		District 063	
	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var.	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var
1	4,734	-16	14,901	- 27	12,129	- 08	5,290	- 18	14,837	- 07	6,367	- 01	5,745	+ 08
2	1,670	+41	7,765	- 35	5,075	+ 01	2,005	- 08	9,427	+ 12	5,161	- 05	4,212	- 03
3	2,652	-10	5,090	+ 00	9,002	- 32	1,675	+ 24	6,082	- 03	4,212	- 29	2,132	- 32
4	1,006	+21	4,325	+ 12	3,655	- 08	1,071	- 05	4,528	- 02	4,571	- 20	3,606	- 35
5	690	+37	2,524	+ 38	1,282	+ 25	701	+ 03	2,682	+ 0[illegible]	1,748	+ 41	2,077	- 01
6	1,032	-12	1,215	+ 127	2,436	+ 06	644	+ 35	1,629	+ 33	1,156	+ 17	1,667	- 29
7	465	+07	1,129	+ 75	438	+ 369	288	+ 74	2,698	- 10	411	+ 38	467	+ 97
8	660	-18	490	+ 165	365	+ 215	168	- 73	1,113	- 28	173	+ 213	88	+ 434
9	339	+22	178	+ 224	614	+ 106	151	+ 241			232	+ 166	310	+ 92
10	45	+07	450	+ 112	473	+ 64	127	+ 82			94	+ 390	133	+ 269
11			43	+ 314	202	+ 84	24	+ 192			64	+ 517	52	+ 471
12			193	+ 302							508	- 13	124	+ 214
13			169	+ 212									50	+ 576
14			16	+ 450									443	- 35

[a] Dotted line separates interchanges below 1,000 trips Such interchanges present more limited stability due to stochastic nature and limited number of events

[b] Var = variation of calculated trips based on actual (e g , calculated trips - actual trips + actual trips)

[c] Each time code equal to 5-min driving time from the district of origin

TABLE 2

ACTUAL INTERCHANGES AND PROPORTION OF ERROR OF PREDICTED INTERCHANGES OF WEIGHTED VEHICLE TRIPS FOR SEVEN ANALYSIS ZONES BY TIME RINGS OUTWARD FROM EACH ZONE CENTER[a,b,c]

Trip Time (min)	Zone 001		Zone 003		Zone 012		Zone 067		Zone 118		Zone 487		Zone 553	
	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var	Actual Trips	% Var
0- 1 9	6,181	+ 57	5,347	+ 59	11,689	+ 33	3,032	- 03	3,522	+ 11	2,079	- 19	15,288	- 20
2- 3 9	7,410	- 26	4,562	+ 18	7,158	+ 96	4,776	+ 33	3,687	+ 52	280	- 02	--	--
4- 5 9	3,459	- 50	2,144	- 02	14,087	- 06	2,439	+ 46	1,403	+ 28	4,357	- 03	9,425	- 36
6- 7 9	3,014	- 04	1,405	+ 21	7,797	- 24	2,823	- 32	3,341	+ 05	971	+ 13	2,460	+ 15
8- 9 9	1,585	+ 21	1,727	- 52	5,622	- 26	1,642	- 08	1,018	- 32	642	+ 111	3,224	+ 10
10-14 9	3,782	+ 37	3,302	- 32	8,769	- 16	3,701	- 27	2,433	- 62	953	- 78	3,620	- 06
15-19 9	3,770	- 16	2,044	- 41	7,110	- 39	1,787	- 01	1,334	- 22	437	+ 112	2,169	+ 31
20-24 9	2,150	- 23	1,359	- 50	5,901	- 47	1,315	- 35	1,555	- 30	291	+ 23	397	+ 210
25-29 9	960	- 13	906	- 57	2,835	- 33	937	- 45	1,170	- 47	425	- 22	535	+ 100
30-34 9	601	+ 20	735	- 56	1,565	- 27	301	- 04	323	- 29	74	+ 405	348	+ 574
35-39 9	942	- 50	272	+ 37	844	- 12	261	- 20	27	+ 441	137	- 01	363	+ 232
40-44 9	314	+ 33	123	+ 00	588	+ 111	107	+ 86	116	+ 40	286	- 64	325	+ 242
45-49 9	410	- 14	88	+ 152	308	+ 18	29	+ 693	85	+ 92	131	- 60	301	+ 00
50-54 9	203	- 24	53	+ 26	246	+ 223	172	+ 35	40	+ 193	42	- 38	206	+ 167
55-59 9	193	- 69	112	+ 09	171	+ 120	36	+ 266	81	- 63	18	- 67	225	+ 21
60 and over	474	+ 45	401	- 11	929	+ 29	183	+ 02	12	+ 675	117	- 75	282	- 03
Total	35,448		24,580		75,619		23,541		20,147		11,240		39,168	

[a] Source Table 54, 6, 2 126

[b] Dotted line separates interchanges below 1,000 trips Such interchanges present more limited stability due to the stochastic nature and the limited number of events

[c] Var = variation of calculated trips, based on actual (e g , calculated trips - actual trips + actual trips)

tricts is taking place. For each district, the trips coming from the stations are duplicated, and thus the total number of trip origins and destinations in each district is achieved. "Though trips" (e.g., station to station) are estimated separately on the basis of a predetermined and commonly agreed-on "area growth factor.")

For the analysis of 1960 data the first phenomenon (of wider discrepancies of smaller interchanges) cannot be improved very much, as long as the investigation proceeds on the zone or small district level. Over-all check distributions on the basis of large districts can achieve, of course, a substantially closer simulation of interchanges at all levels. As for the under-reporting of certain classifications of trips, it is not expected to be especially troublesome because the 1960 household trip file has been factored on the basis of the results of the screenline checks and the internal-external trips were subtracted and analyzed separately.

After these tests were completed, a comparison of the results reached in PJ with the detailed results reported by CATS was undertaken. Tables 1 and 2 give these results. It is evident from these tables that the results reached by the new method are generally at least as good as the results reported by the CATS method.

Finally another test was carried out along the lines suggested by Brokke and Mertz (9) and followed in the Wilbur Smith studies (e.g., 10, 11). The results are given in Table 3 and in a graphic form in Figure 21. This figure should be compared with the results reported in the Wilbur Smith studies (10, pp. 188-191; 11, pp. 137-140) and by Brokke and Mertz (9, pp. 63-68) in their testing of the Detroit and Fratar methods in the Washington, D.C., area. By comparing these results, and observing Figure 21, it is once more verified that (a) the accuracy of simulation (and henceforth projection) increases substantially when the magnitude of interchange increases, and (b) this accuracy does not level off at a point of 10 percent but it continues to increase when the magnitude of trip interchange increases.

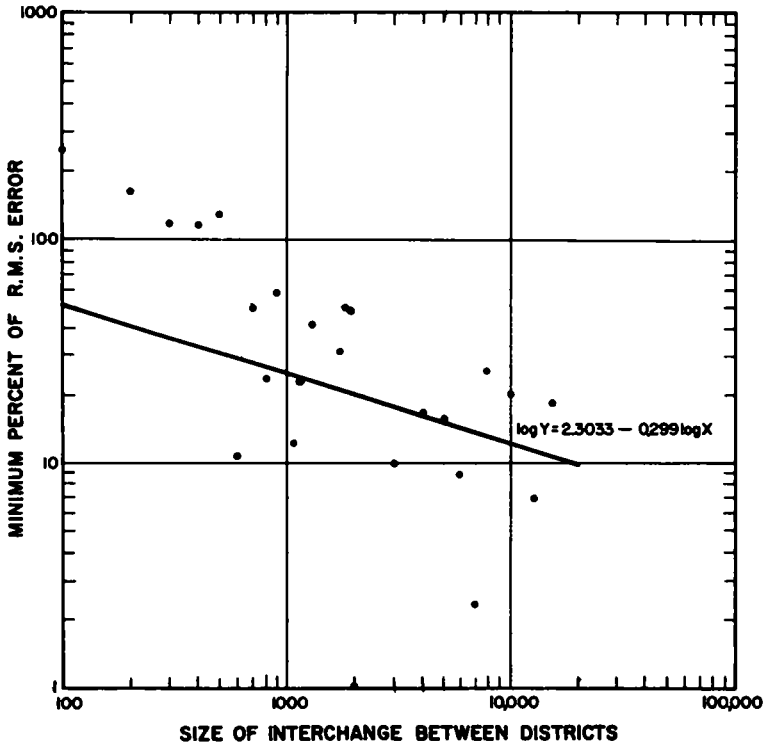

Figure 21.

TABLE 3

PERCENT OF ROOT MEAN SQUARE OF EACH GROUP OF INTERCHANGE,[a] AUTO TRIPS

Interchange	RMS	RMS Percent of Interchange	No. of Pairs in Group
0-100	248.24	248.24	12
100-200	347.26	173.63	8
200-300	375.10	125.03	4
300-400	480.58	120.14	5
400-500	682.44	136.49	7
500-600	66.00	11.00	1
600-700	350.34	50.05	5
700-800	199.87	24.28	2
800-900	519.73	57.75	2
1,000-1,100	147.61	13.42	3
1,100-1,200	273.67	22.80	3
1,200-1,300	541.72	41.67	2
1,600-1,700	539.63	31.74	4
1,700-1,800	723.00	40.17	1
1,800-1,900	753.80	39.67	2
1,900-2,000	37.00	0.02	1
2,000-3,000	311.59	10.39	6
3,000-4,000	657.43	16.44	4
4,000-5,000	774.81	15.50	6
5,000-6,000	528.88	8.81	5
6,000-7,000	158.12	2.26	2
7,000-8,000	2,136.00	26.72	1
9,000-10,000	2,089.19	20.89	3
10,000-12,500	911.00	7.01	1
12,500-15,000	2,684.25	17.90	2

[a] RMS can also be called "standard error of estimate" because it is same estimate (12, pp. 259-260).

(There is a slight difference, however, between the two sets of data. Brokke and Mertz compare projected trips from 1948 to 1955 to actual trips counted by the 1955 O-D survey. Instead, the author compares actual trips with simulated trips at the same date, having the first set of data as the basis for deriving the second set of data with no projection problems entering into account at all.)

CONCLUSIONS

Concluding this phase of the analysis the author reviewed the steps taken. The original dissatisfaction with existing methods of trip distribution led to a new approach with emphasis on structural simplicity and objective projection capabilities. As a result it was decided to test this method further with the 1960 data, a job which is currently under way.

EXAMPLE

Assume that there is a small city of about 150,000 to 200,000 people. Figure 22 shows the assumed over-all outline of the urban development, with two major highways crossing at the central business district. The entire area is divided into districts for the purpose of a transportation study. Assume that the city has no mass transit system and all trips consist of auto trips. A total of 24 districts are outlined on the basis of particular features and characteristics. A trip generation study is then carried out,

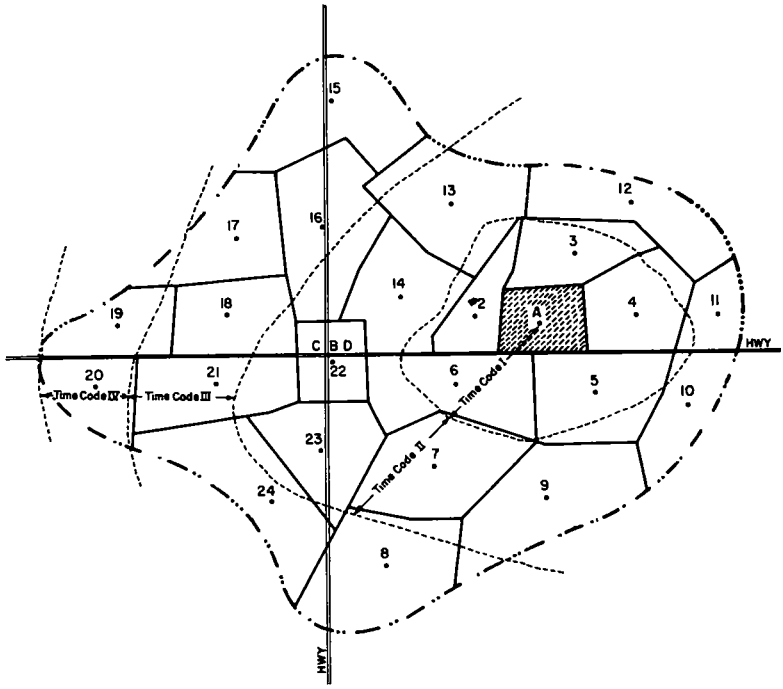

Figure 22.

TABLE 4

TIME CODES AND TRIP ENDS
(on the basis of District A)[a]

District A Origins	Time Code I		Time Code II		Time Code III		Time Code IV	
	Dist.	Destinations	Dist.	Destinations	Dist.	Destinations	Dist.	Destinations
17,000	A	16,500	7	19,000	15	17,000	19	6,000
	2	14,000	9	4,000	16	36,000	20	9,500
	3	12,000	10	3,000	17	18,000		
	4	4,000	11	2,500	18	14,000		
	5	11,000	12	6,500	21	28,000		
	6	17,000	13	14,500	24	11,000		
			14	26,000	8	7,500		
			22	65,000				
			23	46,000				
Time Code Total		74,500		186,500		131,500		15,500
Cum Total		74,500		261,000		392,500		408,000

[a] Thus, on the basis of these data H_1 = 74,500; H_2 = 261,000, H_3 = 392,500; H_4 = 408,000.

and on the basis of such a study, auto trip-ends generated in each district are estimated.

Suppose that it is desired to find out how the trip origins of district A are distributed to the other 23 districts. On the basis of the average travel time between district A and each other district, a set of time codes can be established expressing distance in minutes between district A and all other districts. Suppose the time codes are defined as an interval of 5 min travel time and the area boundaries are reached in 4 time codes as shown in Figure 22. The data are ready to enter in Table 4. The trip destinations in each district take the place of S in Eq. 5. The total trip destinations in each time code takes the place of H in the equation.

Table 5 gives the districts, the trip destinations in each district, and the time code of each district. Then the original conditional probability of each district ($P_{(S/H)}$ = S/H) is established. This column is summed up and the denominator of Eq. 7 $\left[\sum P_{i\,(S/H)}\right]$ is established. The summation is then checked with the help of Eq. 13.

TABLE 5

ESTIMATED INTERCHANGE BETWEEN DISTRICT A AND ALL OTHER DISTRICTS

District	Destination	Time Codes	$P_{(S/H)} = \frac{S}{H}$	$\frac{P_{i\,(S/H)}}{P_{(S/H)}} = P'_{i\,(S/H)}$	Est. Trip Ends[a]	Actual Interchange Between A and j[b]
A	16,500	1	0.2215	0.1060	1,802	
2	14,000	1	0.1879	0.0899	1,528	
3	12,000	1	0.1611	0.0771	1,311	
4	4,000	1	0.0537	0.0257	437	
5	11,000	1	0.1477	0.0707	1,202	
6	17,000	1	0.2282	0.1092	1,856	
7	19,000	2	0.0728	0.0348	592	
8	7,500	3	0.0219	0.0105	179	
9	4,000	2	0.0153	0.0073	124	
10	3,000	2	0.0115	0.0055	94	
11	2,500	2	0.0096	0.0046	78	
12	6,500	2	0.0249	0.0119	202	
13	14,500	2	0.0556	0.0266	452	
14	26,000	2	0.0996	0.0476	809	
15	17,000	3	0.0433	0.0207	352	
16	36,000	3	0.0917	0.0439	746	
17	18,000	3	0.0459	0.0220	374	
18	14,000	3	0.0357	0.0171	291	
19	6,000	4	0.0147	0.0070	119	
20	9,500	4	0.0233	0.0111	189	
21	28,000	3	0.0713	0.0341	580	
22	65,000	2	0.2490	0.1191	2,025	
23	46,000	2	0.1762	0.0843	1,433	
24	11,000	3	0.0280	0.0134	228	
Total	408,000		2.0904[c]	1.000	17,000	17,000

[a] Origins in A $\times P'_{1(S/H)}$.

[b] If available.

[c] $\sum P_{(S/H)}$.

The next column of Table 5 presents the adjusted probability P'_i of each district which is reached by applying Eq. 7. The next step is to multiply the adjusted conditional probability of each district with the trip origins of the originating district (district A in this case). This is the application of Eq. 8. The result is given in the next column and is compared with the actual interchange, if known. Accuracy of results depends on accuracy of measurements of time-distance between districts and trip opportunities within each district. Limits of accuracy can vary depending on the degree of significance each study accepts. If desired, the entire process can be programed for an electronic computer.

REFERENCES

1. Fratar, T., Jr., "Vehicular Trip Distribution by Successive Approximations." Traffic Quart. (Jan. 1954).
2. Voorhees, A.M., and Morris, R., "Estimating and Forecasting Travel for Baltimore by Use of a Mathematical Model." HRB Bull. 224, 105-114 (1959).
3. Lynch, J.T., Brokke, G.E., Voorhees, A.M., and Schneider, M., "Panel Discussion on Inter-Area Travel Formulas." HRB Bull. 253, 128-135 (1960).
4. Wynn, F.H., and Linder, C.E., "Tests of Interactance Formulas Derived from O-D Data." HRB Bull. 253, 62-85 (1960).
5. Feller, W., "An Introduction to Probability Theory and Its Applications." Vol. 1, Ch. 5, Wiley (1960).
6. "Final Report." Chicago Area Transportation Study.
7. Moore, E.F., "The Shortest Path Through a Maze." Internat. Symposium on the Theory of Switching, Harvard Univ. (1957).
8. Fry, T., "Probability and Its Engineering Uses." Von Nostrand, 12th ed.
9. Brokke, G.E., and Mertz, W.L., "Evaluating Trip Forecasting Methods with an Electronic Computer." HRB Bull. 203, 52-75 (1958).
10. "A Highway Planning Study for the St. Louis Metropolitan Area." Wilbur Smith and Assoc. (1959).
11. "Kansas City Metropolitan Area Origin and Destination Survey." Wilbur Smith and Assoc. (1959).
12. Mills, F.C., "Statistical Methods." Holt (1955).