Gravity Model Theory Applied to a Small City Using a Small Sample of Origin-Destination Data

BOB L. SMITH, Associate Professor, Department of Civil Engineering, Kansas State University

The research was concerned with the use of the gravity model in a small city (Hutchinson, Kansas; 37,000 population) and was conducted to study the feasibility of using a small clustered sample of O-D interviews to estimate the gravity model parameters, trip production, trip attraction, and travel time factors. The model was calibrated using the estimations of trip productions and attractions and was compared to the trip distribution obtained from a complete O-D survey. In addition, a gravity model was calibrated using O-D productions and attractions and the resulting distribution was also used in the comparison of results.

The information utilized was obtained from interviews from 402 dwelling units in 14 selected zones; the complete O-D survey was made up of 2,528 interviews obtained from a 20 percent sample of dwelling units from the 83 zones in the survey area. The study was limited to consideration of auto-driver trips which were internal in nature. Three trip purposes—home-work, home-other, and non-home based—were studied in detail.

The study was concerned with present-day traffic rather than with the estimation of future traffic and although the research was not specifically aimed at future estimation, the data used in the development of estimating equations for attractions and productions were those which could be expected to be quickly and economically obtainable and to estimate reasonably well for the future.

The method of using the clustered sample (reduced sample size in the 14 zones) enables adequate estimation of existing zonal trip productions. The estimates of attractions and non-home productions, although not as good as the estimates of home-based trip production, appear to be adequate for use in the planning process. On the basis of experience in Hutchinson, it should be possible to develop for smaller cities excellent estimating equations for both trip productions and attractions using the data from a complete O-D interview study. The travel time factors were satisfactorily estimated using data from the clustered sample study. Analysis of the results of the gravity model distribution, using productions, attractions, and travel time factors based on the clustered sample, indicates that the existing trip distribution as measured by the complete O-D survey data can be adequately reproduced.

•SINCE the end of World War II, the growth of urban areas and the increase in automobile ownership have created new and involved transportation problems. A large number of transportation studies have been made in the last ten years to make available factual data which would describe existing problems and serve as a basis for estimating future problems. These studies have resulted in general agreement that urban traffic patterns are a function of (a) the type and extent of the transportation

Paper sponsored by Committee on Origin and Destination.

facilities available in the area; (b) the pattern of land use in an area, including the location and intensity of use; and (c) the various social and economic characteristics of the people who make trips (1). A significant effort has been made to develop a transportation planning process which uses these interrelationships to provide quantitative information on the travel demands created by alternate land-use patterns and transportation systems in any urban area. This information can be used by various agencies to make decisions concerning improvements in transportation networks to satisfy present and future travel demands and to promote desirable land development patterns.

The planning process must be capable of estimating, within limits of acceptable accuracy, the zonal trip interchanges for the alternate land-use patterns and transportation systems which might reasonably be expected to develop in an area. The information obtained from home interview origin-destination surveys coupled with information on the existing land-use configuration and transportation system gives an adequate picture of the existing travel patterns in an area. However, it is the future travel demands with which we are most interested and the present-day data must in some way be extrapolated to the future.

Studies of travel habits have led to the development of mathematical formulas or "traffic models" which can satisfactorily reproduce zonal trip interchange estimates from comprehensive home interview traffic studies. If present-day zonal interchanges can be estimated within acceptable limits of accuracy and these interchanges are dependent on measurable characteristics of the urban area, it follows, that if it is possible to estimate future urban characteristics (intensity and type of land use, the distribution of job opportunities, and the economic status of the residents), it should be possible to estimate future zonal interchanges. This is subject, of course, to the possibility that, for a given set of identical circumstances for the present and the future, higher or lower trip generation rates may result because of a change in the amount of travel per vehicle. Several formulations of traffic models have been developed for the estimation of future interchanges, particularly in large metropolitan areas, but much additional research is needed to evaluate and verify the various models in cities of all sizes.

The mathematical traffic model offers estimates of likely consequences in terms of traffic patterns for various alternative land-use configurations and transportation systems. There are a number of different traffic models currently being used in transportation studies but the most widely used model to date is the so-called gravity model. This model is based on the adaptation to the movements of human beings of Newton's law of gravity, which states that the gravitational force exerted between two bodies in space is in direct proportion to the masses of the two bodies and inversely proportional to the square of the distance between them. To apply the gravity model theory to a given city, it generally is considered necessary to conduct, as a minimum, a comprehensive origin-destination (O-D) survey and to calibrate or adjust the model to reproduce, at an acceptable level, the trip distributions found in the O-D survey. The model is then used to distribute trips with various configurations of land use and transportation alternatives that would logically be expected to develop in the future.

PURPOSE AND SCOPE

This research was concerned with the use of the gravity model in a small city and was conducted to study the feasibility of using a small sample of home interviews taken in selected O-D zones to estimate the gravity model parameters of trip production, trip attraction and travel time factors. One gravity model was calibrated using the estimations of trip productions and attractions obtained from the small sample and another was calibrated using those obtained from the comprehensive home interview O-D survey. The resulting distributions were compared.

In the sample study, information obtained from interviews from 402 dwelling units in 14 zones was used. The comprehensive O-D study consisted of 2,528 interviews obtained from a 20 percent sample of dwelling units from all 83 zones in the survey area.

The study was limited to the consideration of auto-driver trips which had both origin and destination within the survey area shown in Figure 1. The auto-driver trips

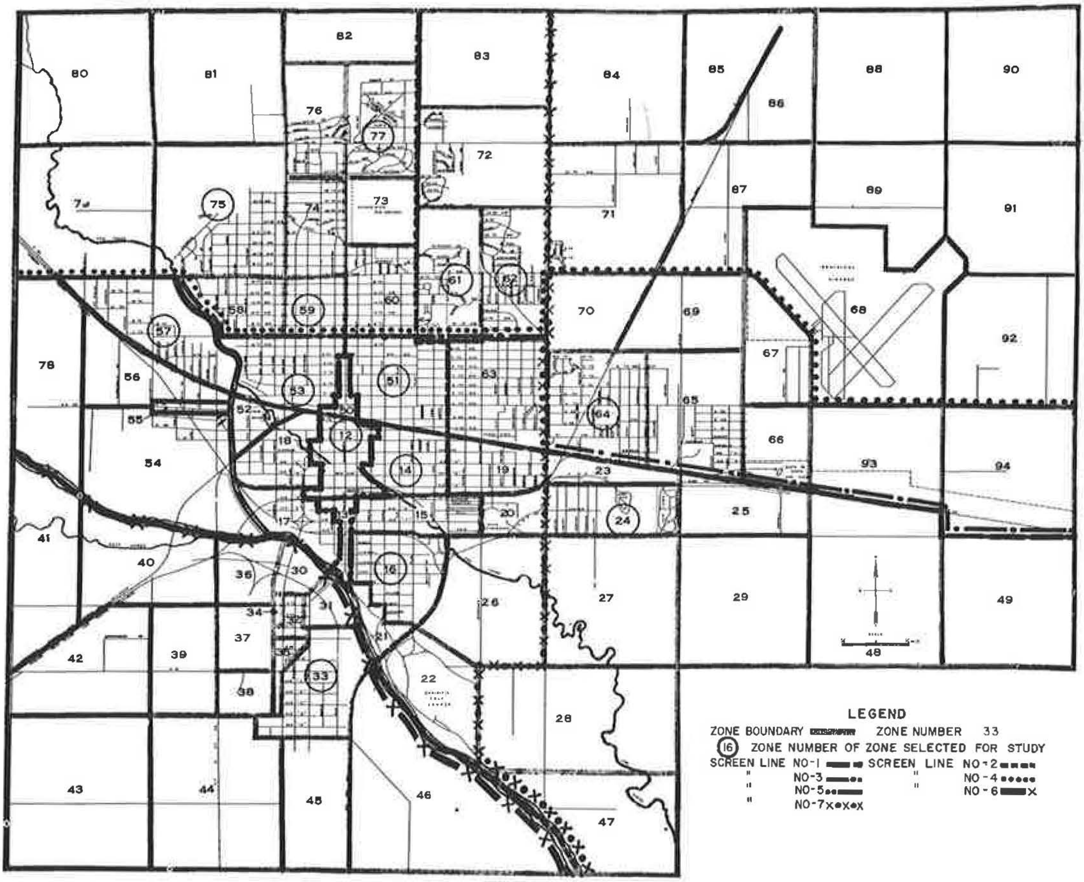


Figure 1. Hutchinson, Kans., metropolitan area zone map showing screenlines and selected zones for study.

were classified according to trip purpose, three of which (home-work, home-other, and non-home-based) were studied in detail.

The study was concerned with present-day traffic rather than with the estimation of future traffic. However, the data used in the development of estimating equations for attractions and productions could be obtained quickly and economically and could be expected to estimate reasonably well for the future.

GRAVITY MODEL THEORY AND USE

The gravity model theory as proposed by Voorhees (2) stated that the trip interchange between zones is directly proportional to the relative attraction for trips of each of the zones and inversely proportional to some function of the spatial separation between zones.

Stated mathematically, the gravity model formulation as used in its earlier applications is as follows:

$$T_{i-j} = P_i \frac{\dfrac{A_j}{\left(d_{i-j}\right)^b}}{\dfrac{A_i}{\left(d_{i-i}\right)^b} + \dfrac{A_j}{\left(d_{i-j}\right)^b} + \dots \dfrac{A_n}{\left(d_{i-n}\right)^b}} \qquad (1)$$

where

T_{i-j} = trips produced by zone i and attracted to zone j;

P_i = trips produced by zone i;

A_j = trips attracted by zone j;

d_{i-j} = spatial separation between zones i and j, generally expressed as total travel time between zones i and j; and

b = an empirically determined exponent expressing average areawide effect of spatial separation between zones on amount of trip interchange.

Early research by Voorhees and others indicated that the exponent, b, varied between 0.6 and 0.8 for work trips in areas of different population size (3).

In response to studies indicating a need for a variable exponent and other refinements, the form of the gravity model formula was changed to the following in which the distribution is generally handled on a basis of various trip purposes:

$$T_{i-j} = \frac{P_i A_j F_{i-j} K_{i-j}}{\sum_{x=1}^{n} A_x F_{i-x} K_{i-x}} \qquad (2)$$

where

T_{i-j} = trips produced in zone i and attracted to zone j;

P_i = trips produced by zone i;

A_j = trips attracted by zone j;

F_{i-j} = an empirically derived travel time factor expressing average area wide effect of spatial separation on trip interchange between zones (The measure of distance or spatial separation between zones is usually the total travel time between the centroids of zones i and j. The use of this factor to express the effect of distance between zones on the zonal trip interchange, rather than the previously used inverse exponential function of time, greatly simplifies the computational requirements of the model and provides for the consideration that the effect of spatial separation generally increases as the separation increases, particularly for some trip purposes);

K_{i-j} = a specific zone-to-zone adjustment factor to allow for incorporation of effect on travel patterns of defined social or economic linkages not otherwise accounted for in gravity model formulation; and

n = total number of zones.

In dealing with the gravity model, confusion often exists among the terms productions, attractions, origins, and destinations. With the exception of trips classified as non-home-based, the number of trips produced refers to the number of trips originating in and returning to a given zone; the number of trips attracted refers to the number of trips arriving at and departing from a given zone. For non-home-based trips, origins and destinations are, respectively, productions and attractions. The gravity model, in the determination of T_{i-j}, deals with trip interchange between zones with no reference to the direction of movement. The trip interchange between zones is often referred to as non-directional or two-way, as opposed to directional trips or trips which start in zone i and end in zone j. Some models use the one-way trip and deal with origins and destinations. Figure 2 is a schematic diagram of the process of determining zonal trip productions and attractions.

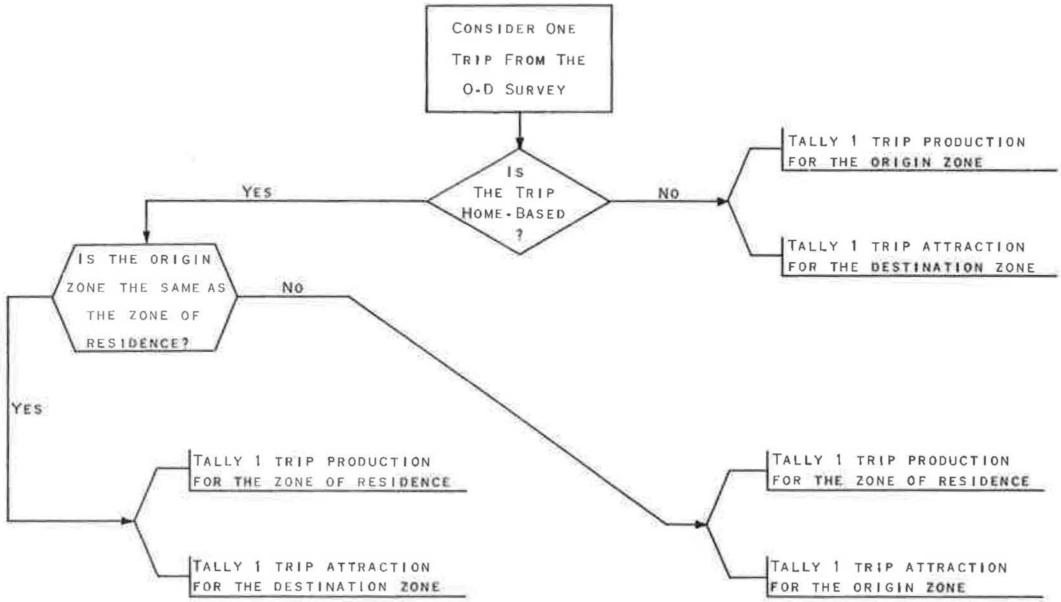


Figure 2. Schematic diagram of process of determining zonal trip productions and attractions.

STUDY PROCEDURE

To carry out the objectives of the research, the metropolitan area of Hutchinson, Kansas, was chosen for the study. In 1961, the city had a population of 37,873 and the metropolitan area a population of approximately 41,000 persons. At the initiation of this project this was the only smaller city in Kansas in which both an internal origin-destination survey as well as a land-use study had been made. Origin-destination and some land-use data were available for Topeka, Wichita and Kansas City, Kansas, but these were the three largest metropolitan areas in Kansas and were not typical in size of smaller Kansas cities. Table 1 indicates all cities in Kansas with a population over 10,000. There were 24 cities outside of metropolitan areas with a population between 10,000 and 50,000 and only 3 cities with a population over 50,000. It was believed that the results of the study, if aimed at these smaller cities, would be of greatest value in Kansas since they so outnumber the larger metropolitan areas.

In 1959, the Kansas State Highway Planning Department, in cooperation with the U. S. Bureau of Public Roads and the City of Hutchinson, conducted a comprehensive home interview O-D survey and a complete land-use study in the Hutchinson metropolitan area. The O-D survey was conducted in accordance with standard procedures prescribed by the Bureau. The internal survey was made by the home interview method in which a 1 in 5 (20 percent) dwelling unit sample was taken. The data gathered in the internal O-D survey and the land-use study were used in this research.

Among the data collected for each surveyed dwelling unit in the internal O-D survey were number of persons, number of employed persons, number of cars owned, age groups, number of vehicular trips, trip purposes at origins and destinations, and mode of travel for each trip. The land-use study recorded the following major groupings of land use by zone in 1,000's of square feet: residential; manufacturing; retail trade; wholesale and warehouse; transportation; construction; personal, business, repair services and office; government and utility; other open space (streets, alleys, rivers, and lakes); and recreation and institution. The land-use categories recorded within each major grouping were as given previously (10, Appendix C).

The IBM 1620 computer and various allied tabulating equipment were extensively used in this study. A list of computer programs written for and used in this research

TABLE 1

CITIES OF KANSAS HAVING POPULATION OVER 10,000[a]

City	Population
Wichita	247,557
Kansas City[b]	126,236
Topeka	120,799
Salina	43,090
Overland Park[b]	40,796
Hutchinson	37,873
Prairie Village[b]	26,873
Lawrence	26,132
Leavenworth	23,707
Manhattan	21,410
Junction City	20,944
Pittsburg	18,737
Great Bend	17,885
Coffeyville	17,030
Emporia	16,763
Liberal	14,806
Newton	14,704
Arkansas City	14,696
Dodge City	13,303
Parsons	13,014
El Dorado	12,614
Garden City	12,575
Hays	12,301
Atchison	12,126
Independence	11,387
Shawnee[b]	11,387
Ottawa	11,237
Olathe	10,776
Chanute	10,666
Winfield	10,522

[a]Populations as of Jan. 1, 1962; reported by county assessors and compiled by Kansas State Board of Agriculture.
[b]Included within Kansas City, Kansas, metropolitan area.

was published previously (10, Appendix B). A write-up of each such program developed can be obtained from the Kansas Highway Commission.

Preparation of O-D Survey Data

The information obtained from the internal O-D survey (hereafter referred to as the O-D data or O-D survey data) was, for the most part, transferred to tabulating machine punch cards and was available to the researcher from the beginning of the project. The cards were of two general types. Card 1, the dwelling unit card (only one card 1 existed per sampled dwelling), contained information on the zone in which the dwelling unit was located, the number of persons living in the dwelling unit, the number of cars owned by these persons, and information on the number of trips made on the day (the trip day) before the interview. Card 2, referred to as the trip card, contained information on the location zone of the home (or dwelling unit), the zone of origin and the zone of destination of the trip, the land-use category at both destination and origin of the trip, mode of travel, the number of persons in the car, and the purpose of travel of each end of the trip. There was a card 2 for each trip recorded at a sampled dwelling unit.

Classification of Trips by Purpose.— In most model studies the trips have been studied by grouping them into a number of trip purposes. After an examination of the O-D survey data, it was initially decided that the five trip purpose groupings (home-work, home-social-recreation, home-shopping, home-miscellaneous, and non-home) would be studied instead of the O-D survey trip purpose categories (work, business, medical-dental, school, social-recreation, change travel mode, eat meal, shopping, home, and serve passenger). Eventually it was found more satsifactory to use only three trip purpose categories, however, because of the small numbers involved.

In an origin-destination survey, one trip ends and another begins every time a person changes his mode of travel, a driver stops to serve a passenger, or a trip maker reaches a destination. In the first two cases, if each of these trips were analyzed separately, the relationships among the actual starting point, the destination and the purpose of the trip would be lost. It would also be difficult to relate the type and intensity of trip making to the type and intensity of land use. Consequently, it is desirable to combine or link those trips which have a change travel mode or serve passenger purpose in order to preserve the relationship between the purpose of the trip and the destination of the trip.

Trip linking may not be necessary in all cases. In many small cities where change travel mode trips may be small in number because of lack of transit facilities and where serve passenger trips may also be small in number because of the absence of car pools, trip linking may be unnecessary (1). However, in studying the Hutchinson data it was found that although change travel mode trips were negligible, the serve passenger trips made up approximately 23 percent of all auto driver trips. Consequently, it was considered necessary to link those trips. In this case, it was judged more expedient to link by hand than to prepare a computer program for the process, although trip linking in a large metropolitan survey area would, no doubt, be most efficiently carried out by the use of a high-speed computer. The U. S. Bureau of Public Roads, in fact, has a trip linking program for use on the IBM 1401.

In the linking process for Hutchinson about 2, 400 trips were lost. That is, there were 2, 400 trips that made up a part of a journey but were not meaningful to the major trip purpose. With the serve passenger trips linked or converted into meaningful purposes, the original ten trip purposes were combined into five categories.

Because of the numbers of trips involved and the differences in treatment of trip productions and attractions, two general classifications of trips are usually made: (a) the trips in which one end is the home (home-based trips), and (b) the trips in which neither end is the home (non-home-based trips). Data in Tables 2 and 3 indicate that the major purposes of home-based trips were work, shopping, and social-recreation.

Using these three categories and combining all other home-based trips into one category and all non-home-based trips into another resulted in the trip purpose groups with percentages of trips in each group as shown in Table 4. During the processes of developing estimating equations for trip productions and attractions and calibrating model by purpose, the relatively small numbers of trips in the home-based social-recreation, shopping and miscellaneous trips appeared to be responsible for much of the variability of results. Therefore, the trip purpose groups were further combined into home-work, home-other, and non-home trips as shown in Table 4. The discussions throughout the remainder of this report relate to these three groups.

Additional Data Obtained.--The number of employed persons per dwelling unit was determined from the home interview sheets and was placed in each dwelling unit card.

TABLE 2

AUTO-DRIVER TRIPS BY RESIDENTS OF INTERNAL AREA, SHOWING O-D SURVEY RECORDED PURPOSE OF TRIP

From	To	(1) Work	(2) Business	(3) Medical-Dental	(4) School	(5) Social Recreation	(6) Change Travel Mode	(7) Eat Meal	(8) Shopping	(0) Home	Total
Work	(1)	2,497	615	51	128	593	21	2,609	802	8,788	16,104
Business	(2)	656	903	31	51	394	5	57	742	1,861	4,700
Medical-Dental	(3)	31	31	5	5	52	-	10	127	295	556
School	(4)	432	77	20	167	233	-	259	148	1,879	3,215
Social Recreation	(5)	422	252	26	66	1,505	16	103	736	5,678	8,804
Change Travel Mode	(6)	31	10	5	-	21	-	5	31	97	200
Eat Meal	(7)	2,446	73	15	217	140	-	5	107	352	3,355
Shopping	(8)	328	468	31	98	640	10	114	1,685	5,813	9,187
Home	(0)	9,395	2,226	387	2,523	5,332	128	339	4,786	5	25,121
Total		16,238	4,655	571	3,255	8,910	180	3,501	9,164	24,768	71,242

TABLE 3

PERCENT AUTO-DRIVER TRIPS BY RESIDENTS OF INTERNAL AREA, SHOWING O-D SURVEY RECORDED PURPOSE OF TRIP

From \ To	(1) Work	(2) Business	(3) Medical-Dental	(4) School	(5) Social Recreation	(6) Change Travel Mode	(7) Eat Meal	(8) Shopping	(0) Home	Total
Work (1)	3.51	0.86	0.07	0.18	0.83	0.03	3.66	1.13	12.34	22.61
Business (2)	0.92	1.27	0.04	0.07	0.55	0.01	0.08	1.04	2.61	6.59
Medical-Dental (3)	0.04	0.04	0.01	0.01	0.07	0	0.01	0.18	0.41	0.77
School (4)	0.61	0.11	0.03	0.23	0.33	0	0.36	0.21	2.64	4.52
Social Recreation (5)	0.59	0.35	0.04	0.10	2.11	0.02	0.14	1.03	7.97	12.35
Change Travel Mode (6)	0.04	0.01	0.01	0	0.03	0	0.01	0.04	0.14	0.28
Eat Meal (7)	3.43	0.10	0.02	0.30	0.20	0	0.01	0.15	0.49	4.70
Shopping (8)	0.46	0.66	0.04	0.14	0.90	0.01	0.16	2.37	8.16	12.90
Home (0)	13.19	3.12	0.54	3.54	7.48	0.18	0.48	6.72	0.01	35.26
Total	22.79	6.52	0.80	4.57	12.50	0.25	4.91	12.87	34.77	

Persons classified as non-gainfully employed workers (including housewives and other unpaid home-workers, retired workers, persons permanently incapacitated for any gainful employment, and students) were not included in the employed persons totals. Also placed in each dwelling unit card was the net area of residential land use per dwelling unit. This was determined by dividing the area of residential land in a given zone less the vacant land zoned residential by the number of dwelling units in the zone. The area was recorded in 1,000's of square feet per dwelling unit.

The driving time from the centroid of each zone to the central business district (CBD) was determined to the nearest 0.1 min from the minimum driving time paths developed from a travel time study in Hutchinson.

Before the development of estimating equations for work trip attractions, it was decided, as was found in an Iowa study (4), that the number of jobs in a zone would be expected to be a potent indicator of home-work trip attractions. Information on various categories of employment in each zone was therefore collected. An attempt was made to determine the employment in the various zones as it existed at the time the O-D survey was made, but only major changes in employment could be determined. However, a very good correlation of total number of jobs in the survey area with the 1960 Census data was obtained.

The employment study was made in the Hutchinson office of the Kansas State Employment Service. The excellent cooperation of the employment service personnel enabled completion of the survey within 3 days with two persons collecting most of the data. The number of self-employed persons and the number of employees in each of the following types of business and industry were tabulated by zone: agricul-

TABLE 4

DISTRIBUTION OF AUTO DRIVER TRIPS BY PURPOSE OF TRIP

Purpose	Initial Category No.	Initial Category %	Final Category No.	Final Category %
Home-work	18,183	25.53	18,183	25.53
Home-social-recreation	11,010	15.45	-	-
Home-shopping	10,599	14.88	-	-
Home-miscellaneous[a]	10,092	14.16	-	-
Home-other[b]	-	-	31,701	44.49
Non-home	21,358	29.98	21,358	29.98
Total	71,242	100.00	71,242	100.00

[a]Includes business, medical-dental, school, change travel mode and eat meal.
[b]Includes social-recreation, shopping and miscellaneous.

ture and forestry; mining; construction; manufacturing and processing; transportation, communications and public utilities; wholesale, retail, finance and insurance; personal services; amusement and recreation; professional; and government.

Selection of Zones for Reduced Sample Study

Zones were selected for the reduced sample survey to reflect a range in zonal characteristics such as residential density, car ownership rate, population density, distance from CBD and distance to the nearest large employment center. An investigation of the Hutchinson data showed that the principal CBD zone, 12, was, in fact, the large employment center in the city. A zonal map (Fig. 1) was used in selecting the zones. The number of dwelling units, cars owned by residents of the zone, total number of persons residing in the zone, cars per dwelling unit, and persons per dwelling unit were noted on a zone map for each zone having any significant number of dwelling units.

The selection of zones was further based on opinions of those familiar with the nature of residential areas in Hutchinson so that zones of varying economic status or value of residence would be chosen. After consultation with statisticians it was decided that a minimum of 14 zones would be required if an estimating equation with as many as seven terms were to be used. Data from 14 zones, giving 14 pieces of information, were judged to give a satisfactory statistical estimate of the predictive power of the estimating equation. A minimum of seven degrees of freedom was judged to be required. Zones 12, 14, 16, 24, 33, 51, 53, 57, 59, 61, 62, 64, 75, and 77 were chosen for the study.

Development of Equations for Estimating Zonal Productions

Early attempts were made to develop estimating equations based on groups of samples within each of seven initially selected zones. It was hoped that interviews obtained in a single zone could be grouped into small subsamples that would give good estimates of the productions and attractions of that particular zone.

Information concerning the assessed evaluation per dwelling unit in zone 75 (Fig. 1) was obtained in a two-day study of the city records in Hutchinson. The evaluation data were tested for value as an economic indicator in the production of trips. The results were inconclusive since the trip production based on the subsamples showed great variability. Examination of the evaluation study indicated that such a study for each zone in the area would probably not produce data of significant aid in the estimation of zonal trip productions and attractions.

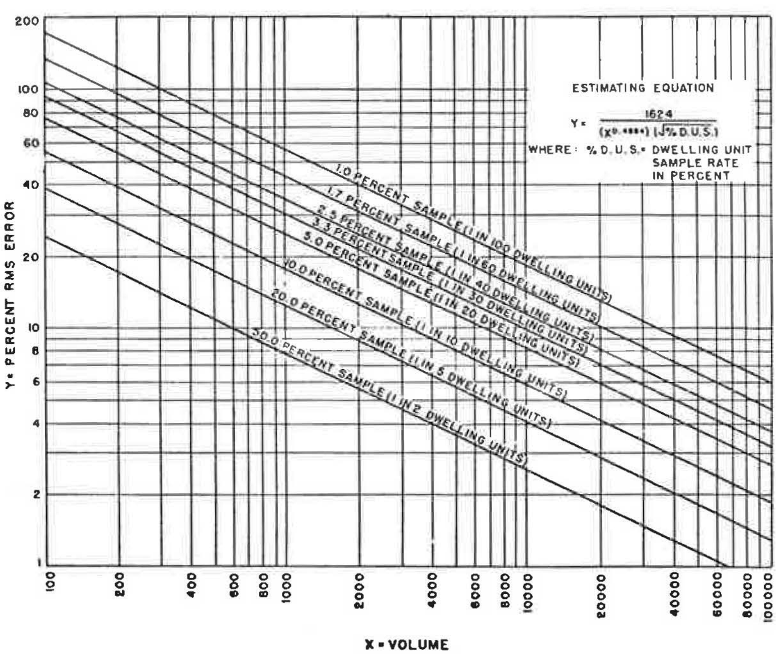
It was also believed that the number of employed persons per dwelling unit might be a better indicator of work trip production than persons per dwelling unit. To test this hypothesis, four sets of regression equations were developed using, among the variables, either persons per dwelling unit or employed persons per dwelling unit. The data used were from the comprehensive survey of the 22 zones containing more than 125 dwelling units. There was less than one percent difference between the value of R^2, the coefficient of determination, for each pair of equations. It was concluded, therefore, that employed persons per dwelling unit was no better indicator of work trip production than persons per dwelling unit.

Due to the great variability among subsamples as well as the problems anticipated in gathering data on zonal characteristics in such small areas, the estimation of productions was carried out using zonal averages of such information as cars and persons per dwelling unit, area of residential land per dwelling unit and distance to the CBD in minutes. The development of the estimating equations was carried out using the multiple regression technique in which the form of an equation is estimated and the coefficient of each term is obtained from the least squares best fitting curve. The measure of fit was obtained as an output of the computer program used (7). The SCRAP "Sixteen-twenty Card Regression Analysis Progam" used in this research is one of a number of such programs available in most computing centers.

Two sets of estimating equations of trip production were developed. One was based on zonal characteristics obtained from each of 22 zones as a result of sampling 20 percent of the dwelling units in each zone. The 22 zones were those in which over 25 dwelling units were sampled. In essence, the 22 zones represented the universe of zones of substantial size. The second set of equations was based on zonal characteristics obtained from each of the previously mentioned 14 selected zones. However, in this case, a sampling rate was established for each zone.

To estimate the total number of non-home-based trip productions, non-home-based trips were treated precisely as home-based trips. These productions represented only the number of non-home trips made by the residents of each zone and did not distribute the trips according to location of trip end. A regression analysis, similar to that used for home-based productions, was made on these trips with the resulting equation, as expanded to the entire area, giving the total number of non-home trip productions or attractions. This number was later used in expanding the non-home-based productions to this estimated total.

Selection of Reduced Samples.—The selection of the reduced sample size in the 14 zones was made in accordance with research conducted by Sosslau and Brokke (8). To use Figure 3, it is necessary to estimate the number of trips per zone and to find an acceptable root-mean-square (RMS) error before selecting the appropriate rate of sampling. A level of accuracy yielding an expected RMS error of zonal trip production of 15 percent or less was acceptable in this case. The estimation of trips produced per zone was made and the citywide average was found to be about five auto-driver trips per dwelling unit. Where an estimate of this average is not available for a city, studies



$$\text{Dwelling unit sample rate in percent} = \left[\frac{1{,}624}{\left(\text{percent RMS error}\right)\left(\text{volume}^{0.4884}\right)}\right]^2$$

Figure 3. Relation of percent root-mean-square error and volume for various dwelling unit sample rates (8).

of trip-making characteristics of similar cities should suffice for this estimate.

Figure 3 was entered with X, the estimated volume of trips in a given zone, and the 15 percent RMS error line indicated the minimum percentage of dwelling units to be sampled in that zone. The subsample was drawn from the O-D sample and selected by a computer program.

The subsample in zone 12 was selected as follows:

1. Estimated trips = 5 x 270 dwelling units = 1,350.
2. Enter Figure 3 with 1,350 and intersect the horizontal 15 percent RMS error line.
3. Read 15 ± percent sample.
4. A 15 percent sample is equivalent to three-fourths of a 20 percent sample (the existing sample size).
5. The computer program was devised to select the dwelling unit and trip cards for every n th sampled dwelling unit and in this case a random selection of a number from 1 to 4 was made to indicate the starting sample and thereafter every fourth dwelling unit sample was selected. This group was discarded and the remaining three-fourths of the original 20 percent sample was taken. The original expansion factors were multiplied by the ratio of the original number of sampled dwelling units to the reduced number of sampled dwelling units and this new expansion factor was placed in the dwelling unit and trip cards for the reduced samples.

Table 5 shows the sample size selected in each of the 14 zones. There were 402 dwelling units in the reduced samples in the 14 selected zones. The comprehensive O-D survey consisted of 1,359 interviews obtained from a 20 percent sample of dwelling units in the same 14 zones.

Development of Equations for Estimating Zonal Attractions

The SCRAP regression analysis program (7) was used in developing two sets of estimating equations for trip attractions by trip purpose. One set was obtained using the attractions as distributed according to the data from the reduced sample in the 14 selected zones. The second set was obtained using data from the comprehensive O-D survey for all zones having 40 or more trips attracted for each of the three purpose groupings.

For the first set of equations, even though trips produced by only 14 zones were used, the attraction ends of the trips were distributed to many zones. For example, assume that 10,000 trips of a given purpose were produced by the 14 zones and of these 200 were attracted to zone 1, 300 to zone 4, 1,000 to zone 12, 2,000 to zone 20, 1,500 to zone 50, etc. The total of all trips attracted would be 10,000. The trips attracted

TABLE 5

REDUCED SAMPLE SELECTION IN 14 SELECTED ZONES

Zone	Estimated Auto Driver Trips per Zone[a]	% Sample Size to Obtain 15% RMS Error	Fraction of Orig. 20% Sample	No. of Interviews Selected	O-D Expansion Factor	No. of Dwelling Units, Orig. Sample	Expansion Factor, Reduced Sample[b]
12	1,350	15	3/4	35	5.32	46	6.99
14	3,760	5	1/4	32	5.21	126	20.51
16	2,160	7	1/3	25	5.11	77	15.74
24	1,350	11	1/2	24	5.11	49	10.43
33	1,490	10	1/2	29	4.98	57	9.79
51	7,360	2.5	1/8	29	5.34	233	42.90
53	4,820	3.3	1/6	27	5.18	164	31.46
57	1,570	10	1/2	26	5.19	53	10.58
59	3,010	5	1/4	26	5.29	102	20.75
61	1,785	9	1/2	29	5.03	58	10.06
62	2,060	7	1/3	25	5.05	76	15.35
64	2,870	6	1/3	35	5.13	105	15.39
65	3,490	5	1/4	30	5.21	123	21.36
77	2,470	7	1/3	30	5.16	90	15.48

[a] 5 × No. of dwelling units per zone. [b] O-D expansion factor × No. of dwelling units in original sample/No. of interviews selected.

to each zone were divided by 10,000, the total of trips produced, giving the proportion of all trips attracted to a given zone. Thus zone 1 would attract 0.02 trips per trip produced, zone 4 would attract 0.03, zone 12 would attract 0.10, zone 20 would attract 0.20, and zone 50 would attract 0.15.

Based on such zonal characteristics as various types of land use in 1,000's of square feet, numbers of jobs of various classifications and total number of persons and dwelling units in the zone, an equation was developed for each trip purpose that estimated the trip attractions in the various zones. The dependent variable was the number of trips attracted per trip produced.

A similar procedure was followed in developing equations based on the 20 percent sample obtained in the O-D survey. The estimating equation for the non-home productions was developed in the same manner as that for trip attractions.

In the case of home-work attractions, it was discovered that the best predictor of trips attracted was the total employment in a zone. Since the SCRAP program would not handle a problem with a single independent variable, the equation was developed using a computer program which fitted a polynomial to the data by the method of least squares.

The equation for trip attractions, along with non-home productions, had to be multiplied by the total number of trips produced in the study area to give total attractions per zone. Thus, zonal estimates of trip attractions (and non-home productions) were the product of two estimates and, in general, were less satisfactory than estimates of home-based productions.

Development of Travel Time Factors

The calibration of the gravity model (Eq. 2) was carried out using the three trip purposes (home-work, home-other and non-home) and consisted principally of the determination of travel time factors which resulted in a trip length frequency distribution comparing satisfactorily to that of the surveyed population. Two sets of travel time factors for each trip purpose were determined, one from zonal productions and attractions from the comprehensive O-D survey and the other from those obtained from the reduced sample size in the 14 selected zones.

The gravity model formula as used requires input parameters of zonal trip productions, attractions, travel time factors, and zone-to-zone adjustment factors. The zone-to-zone adjustment factors were used as unity throughout this study because of no apparent effect on travel patterns of defined zonal characteristics.

The travel times used for determination of the corresponding travel time factors for use in the gravity model were made up of the terminal time on each end of the trip plus the zone-to-zone driving time. The zone-to-zone minimum driving time was obtained from the "time trees" or minimum driving time paths developed from travel time study data. The driving time for intrazonal trips (those trips with both ends in a given zone) was not available from the time trees and was estimated at 1 min for each zone after inspection of interzonal times for all adjacent zones. The interzonal times of adjacent zones were, in all cases, slightly less than 2 min. The maximum intrazonal time was also about 2 min and a reasonable average time was believed to be 1 min. Other methods of determining intrazonal times have been previously discussed (1).

The terminal time of one end of a trip may be made up of the time spent in looking for a parking space, the time spent waiting before a vehicle can be parked and the time spent walking from the parking place to the actual destination. The terminal time of the other end of the trip may consist of the time spent walking from home to garage or parking lot and the driving time from garage or parking lot to the street. The initial estimates of terminal times for the zones in Hutchinson were made after consultation with personnel who were familiar with Hutchinson. The CBD, zones 12, 13 and 50 (Fig. 1), were each given terminal times of 3 min and each of the zones adjacent to the CBD was given a terminal time of 1.5 min. Some changes in these terminal times resulted in better trip end balance for some zones and some trip purposes. Table 6 shows the final sets of terminal times used in the study.

TABLE 6

FINAL TERMINAL TIMES BY ZONE AND TRIP PURPOSE

Terminal Time (min)	Trip Purpose		
	Home-Work	Home-Other	Non-Home
3.0	12, 13, 50[a]	None	12, 13, 50
2.5	14, 15, 16, 17, 18, 51, 52, 53[b]	None	14, 15, 16, 17, 18, 51, 52, 53
2.0	19, 58, 59, 60, 61, 63[c]	12, 13, 14, 15, 16, 17, 18, 19, 50, 51, 62, 53, 58, 59, 60, 61, 63	19, 58, 59, 60, 61, 63
1.5	All others[d]	All others	All others

[a]CBD zones. [b]Zones adjacent to CBD. [c]Other highly developed zones. [d]Relatively undeveloped zones.

The interzonal travel time between any two zones was made up of the terminal time of the production zone plus the driving time between the zones plus the terminal time of the attraction zone. Intrazonal travel time for a given zone was made up of twice the zonal terminal time plus the intrazonal driving time for that zone.

A set of travel time factors using the comprehensive O-D zonal productions and attractions was developed. An initial set of travel time factors was assumed and the trip interchanges between all zones were computed. The trip length frequency distribution of the trip interchanges was determined by finding the number and the percentage of trips falling in each 1-min increment of driving time. The estimated trip length frequency was then compared to the actual trip length frequency distribution obtained from the O-D data. The comparison was made in three ways: (a) both distributions expressed as percent trips for each 1-min driving time should, when plotted, be relatively close to one another; (b) the average trip length, in minutes, for both sets of data should be within ±5 percent of each other; and (c) the person hours of travel for both sets of data should be within ±5 percent of each other (1). The average trip length was determined by multiplying the number of trips of each incremental trip length by the length of trip (driving time) in minutes and dividing this product by the total number of trips. The vehicle-hours of travel were obtained by multiplying the number of trips of each incremental trip length by the length of trip in minutes and dividing the product by 60. Computer programs were written to determine the trip length distributions as well as the average trip length and vehicle-hours of travel.

If the comparisons were not within the limits cited, an adjustment was made in the initially assumed set of travel time factors for each trip purpose. The travel time factors were adjusted manually by a procedure which follows from the question: "What must be done to the travel time factor at each travel time increment to bring the gravity model estimated percentage of trips, in each travel time increment, into closer agreement with the surveyed trips at each increment?" The actual adjustment was made for each travel time increment by multiplying the initial travel time factor for each increment by the ratio of the percentage of surveyed trips to the percentage of estimated trips for the respective time increments. The adjusted travel time factors (for each 1 min of travel time) were then plotted against the respective travel time increments on log-log graph paper in most cases and straight-line graph paper in others. The second set of travel time factors was then determined from a hand-fitted line of best fit to the adjusted factors. The gravity model was then run using the second set of travel time factors and the comparisons of trip length frequency, etc., were repeated. This process was continued until satisfactory agreement among the comparisons was reached.

In the case of home-work trips, 12 sets of adjustments were required before acceptable agreement was reached. Better estimates of initial travel time factors would have resulted in fewer iterations being required. This was graphically illustrated by home-other trips when the Iowa travel time factors (4, 9) were used for the initially estimated factors; only four iterations were required. In addition, much time was spent in adjusting to the trip length frequency curve. The Iowa travel time factors (4, 9) are shown in Table 7.

The second set of travel time factors was developed in a similar manner except that productions and attractions obtained from data from the 14 zones with the reduced sample size were used. The trip length frequency, average trip length and vehicle-hours of travel, against which comparisons were made, were those resulting from the O-D data obtained from the reduced sample size in the 14 zones. The Iowa travel time factors (4, 9) were used as the first estimate of the factors used for each trip

TABLE 7

CEDAR RAPIDS, IOWA, TRAVEL TIME FACTORS[a]

Travel Time (min)	Travel Time Factors by Purpose			Travel Time (min)	Travel Time Factors by Purpose		
	Work	Non-Home Based	Other-Home Based		Work	Non-Home Based	Other-Home Based
1	2.00	3.00	5.00	28	0.33	0.07	0.10
2	2.00	2.25	3.66	29	0.31	0.06	0.09
3	2.00	1.80	2.20	30	0.29	0.05	0.08
4	1.50	1.40	1.45	31	0.27	0.04	0.06
5	1.25	1.15	1.20	32	0.25	0.03	0.04
6	1.10	1.00	1.00	33	0.23	0.02	0.03
7	1.00	0.90	0.90	34	0.21	0.01	0.02
8	0.93	0.80	0.80	35	0.19	0.01	0.01
9	0.87	0.70	0.70	36	0.17	-	-
10	0.84	0.62	0.62	37	0.15	-	-
11	0.80	0.56	0.56	38	0.14	-	-
12	0.76	0.49	0.50	39	0.13	-	-
13	0.72	0.43	0.45	40	0.12	-	-
14	0.68	0.38	0.41	41	0.11	-	-
15	0.64	0.34	0.38	42	0.10	-	-
16	0.61	0.30	0.35	43	0.09	-	-
17	0.58	0.27	0.32	44	0.08	-	-
18	0.55	0.24	0.30	45	0.07	-	-
19	0.52	0.22	0.27	46	0.06	-	-
20	0.49	0.20	0.25	47	0.05	-	-
21	0.47	0.18	0.23	48	0.04	-	-
22	0.45	0.16	0.21	49	0.04	-	-
23	0.43	0.14	0.19	50	0.04	-	-
24	0.41	0.12	0.17	51	0.03	-	-
25	0.39	0.10	0.15	52	0.03	-	-
26	0.37	0.09	0.13	53	0.02	-	-
27	0.35	0.08	0.11				

[a]Source: (4, 9).

TABLE 8

TRAVEL TIME FACTORS BY TRIP PURPOSE

Travel Time (min)[a]	Travel Time Factors					
	Home-Work		Home-Other		Non-Home	
	F.S.[b]	R.S.[c]	F.S.[b]	R.S.[c]	F.S.[b]	R.S.[c]
4	2.78	2.12	1.74	2.70	1.40	3.00
5	2.40	1.70	1.22	1.84	1.15	2.25
6	2.20	1.41	0.94	1.23	1.00	1.80
7	1.97	1.20	0.78	0.91	0.90	1.40
8	1.78	1.06	0.65	0.68	0.80	1.15
9	1.58	0.93	0.56	0.51	0.70	1.00
10	1.43	0.84	0.49	0.41	0.62	0.90
11	1.32	0.76	0.43	0.34	0.56	0.80
12	1.20	0.70	0.39	0.28	0.49	0.70
13	1.12	0.65	0.35	0.23	0.43	0.62
14	1.04	0.60	0.32	0.19	0.38	0.56
15	0.96	0.56	0.30	0.17	0.34	0.49
16	0.90	0.53	0.28	0.14	0.30	0.43
17	0.85	0.49	0.25	0.12	0.27	0.38
18	0.80	0.46	0.23	0.11	0.24	0.34
19	0.76	0.44	0.21	0.10	0.22	0.30
20	0.72	0.43	0.20	0.09	0.20	0.27

[a]No travel time of less than 4 min was possible for any trip.
[b]F.S. - Model input: full sample O-D productions-attractions, all zones.
[c]R.S. - Model input: reduced sample O-D productions-attractions, 14 zones.

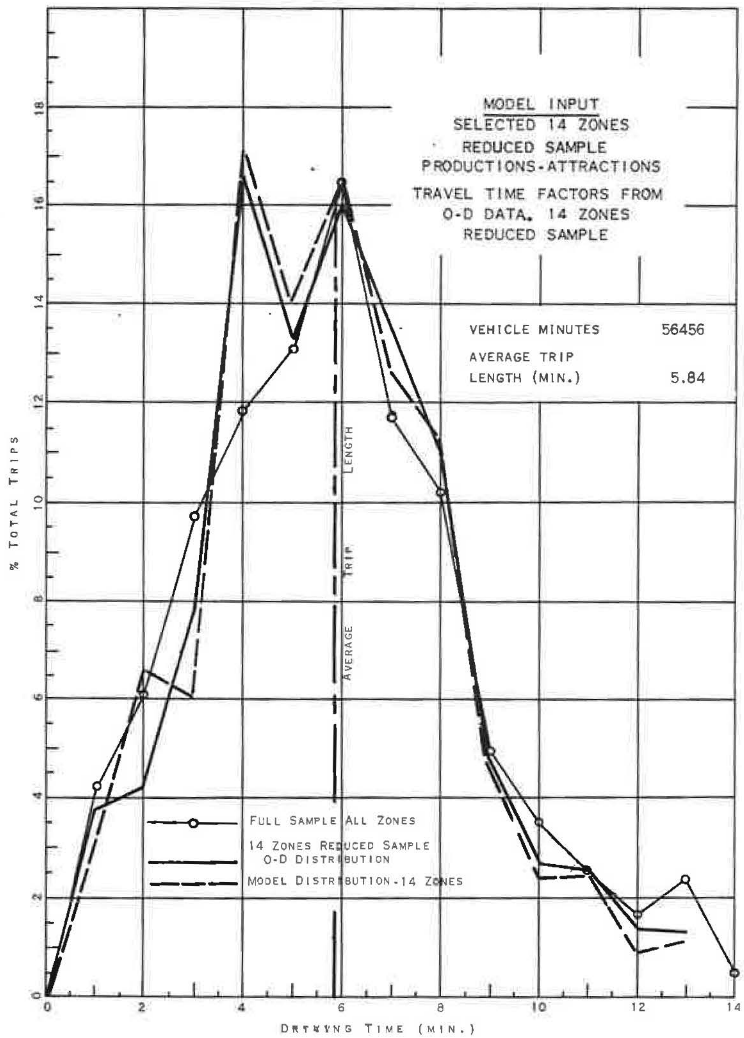


Figure 4. Comparison of trip length frequency using O-D and model data, home-work trips.

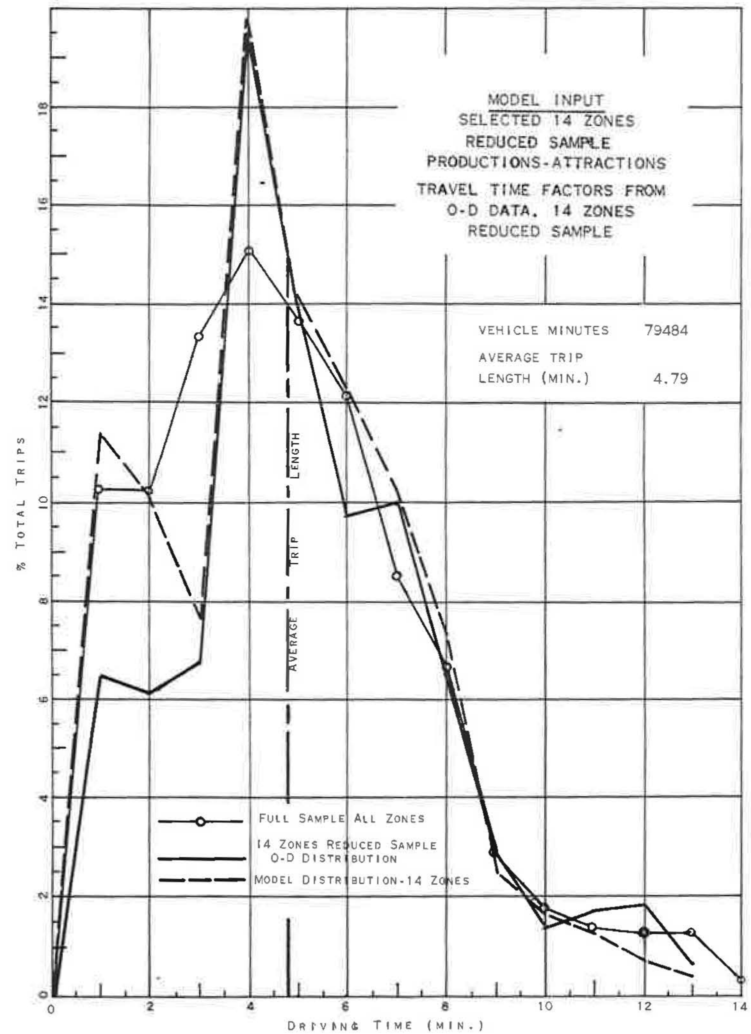


Figure 5. Comparison of trip length frequency using O-D and model data, home-other trips.

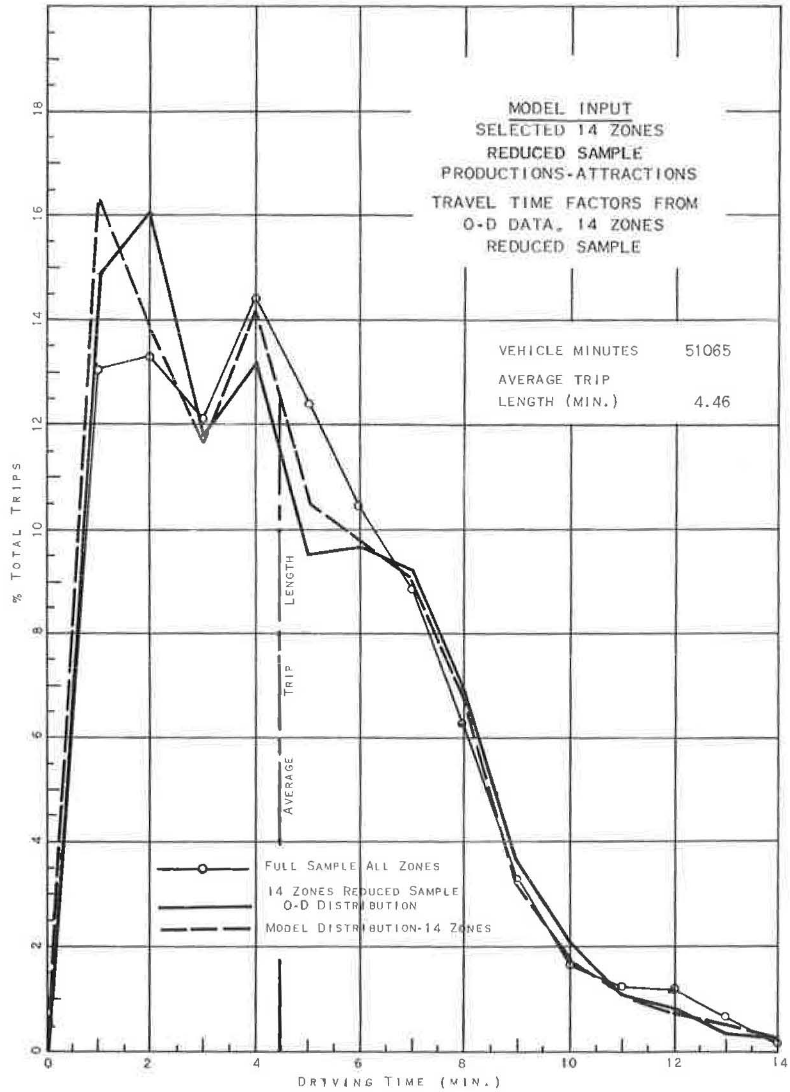


Figure 6. Comparison of trip length frequency using O-D and model data, non-home trips.

purpose. From three to seven iterations were necessary to arrive at acceptable factors. The computer time was much reduced because of the great reduction in number of trips to be distributed. Figures 4, 5, and 6 show the comparison of trip length distributions as obtained for the 14 zones, reduced sample O-D data vs the model distribution using the second set of travel time factors. Table 8 shows the best set of travel time factors developed in each case.

The trip-length frequency data were developed on the basis of driving time rather than travel time. An examination of the computational procedures indicates that, with little difficulty, the distribution could have been made on the basis of travel time if terminal times were introduced as input data. The results of the calibration process probably would have been more satisfactory had this been done.

ESTIMATING EQUATIONS

The following estimating equations for trip productions and attractions were developed using the multiple regression analysis technique.

Trip Production

OSP 22030--home-work, full sample:

$$Y = -5.78775 + 10.25491\,(\text{cars/DU}) + 1.52715\,(\text{pers./DU}) - 0.70627\,(\text{CBD dist.}) - 2.66103\,(\text{cars/DU})\,(\text{pers./DU}) + 0.30996\,(\text{pers./DU})\,(\text{CBD dist.}) - 0.039634\,(\text{CBD dist.})^2$$

Y = trips/dwelling unit. (3)

RSP 14030--home-work, reduced sample:

$$Y = -0.54297 - 0.96297\,(\text{cars/DU}) + 0.79424\,(\text{pers./DU}) + 0.13594\,(\text{CBD dist.}) + 0.31954\,(\text{cars/DU})\,(\text{pers./DU}) - 0.10496)\,(\text{pers./DU})\,(\text{CBD dist.}) + 0.018626\,(\text{CBD dist.})^2$$

Y = trips/dwelling unit. (4)

OSP 22039--home-other, full sample:

$$Y = -5.92767 + 11.60937\,(\text{cars/DU}) + 1.39224\,(\text{pers./DU}) - 1.20225\,(\text{CBD dist.}) - 2.26609\,(\text{cars/DU})\,(\text{pers./DU}) + 0.29830\,(\text{pers./DU})\,(\text{CBD dist.}) + 0.007200\,(\text{CBD dist.})^2$$

Y = trips/dwelling unit. (5)

RSP 14039--home-other, reduced sample:

$$Y = 4.56907 - 6.09284\,(\text{cars/DU}) - 1.69056\,(\text{pers./DU}) + 0.58893\,(\text{CBD dist.}) + 2.98149\,(\text{cars/DU})\,(\text{pers./DU}) - 0.29073\,(\text{pers./DU}\,(\text{CBD dist.}) + 0.042162\,(\text{CBD dist.})^2$$

Y = trips/dwelling unit. (6)

NHP 04514--non-home, full sample:

$$Y = -2.398094 - 0.0051391\,(LU_1) - 0.017017\,(LU_5) + 0.054498\,(LU_6) + 0.058424\,(\text{jobs}) + 0.048989\,(\text{pers./zone}) - 0.084438\,(\text{tot. DU/zone}) - 0.017802\,(\text{cars/zone}) + 0.0000005793\,(LU_1)^2 + 0.000003317\,(LU_5)^2 + 0.00052392\,(LU_6)^2 - 0.000023462\,(\text{jobs})^2 - 0.000023616\,(\text{pers./zone})^2 + 0.000089245\,(\text{tot. DU/zone})^2 + 0.000060598\,(\text{cars/zone})^2$$

Y = trips/zone/1,000 trips produced in 14 selected zones. (7)

NRP 04514--non-home, reduced sample:

$$Y = -6.22386 - 0.0046915\,(LU_1) + 0.043900\,(LU_5) + 0.045839\,(LU_6) + 0.070971\,(\text{jobs}) + 0.058497\,(\text{pers./zone}) - 0.15106\,(\text{tot. DU/zone}) + 0.038259\,(\text{cars/zone}) + 0.0000005283\,(LU_1)^2 - 0.00020223\,(LU_5)^2 + 0.00064132\,(LU_6)^2 - 0.00002833\,(\text{jobs})^2 - 0.00002883\,(\text{pers./zone})^2 + 0.00013757\,(\text{tot. DU/zone})^2 + 0.00002562\,(\text{cars/zone})^2$$

Y = trips/zone/1,000 trips produced in 14 selected zones. (8)

Trip Attraction

Work trip attraction--home-work, full sample (adj. jobs):

$$Y = 1.109 + 0.0624\,(\text{jobs})$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.} \qquad (9)$$

RS 460--home-work, reduced sample (zone 12 omitted):

$$Y = 1.092802 + 0.058113\,(\text{jobs})$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.} \qquad (10)$$

MRA 51409--home-other, full sample:

$$Y = -0.62306 - 0.048951\,(LU_6) + 0.005561\,(LU_9) - 0.0052420\,(\text{pers./zone}) + 0.035644\,(\text{tot. DU/zone}) - 0.050611\,(\text{who.-ret. jobs}) + 0.06504\,(\text{pers. serv. jobs}) + 0.064090\,(\text{prof. jobs}) - 0.012982\,(\text{grouped jobs}) + 0.32256\,(LU_{220}) + 1.95827\,(LU_{240}) + 1.63904\,(LU_{250}) + 0.39525\,(LU_{280}) + 0.00017289\,(LU_6)^2 - 0.00000037908\,(LU_9)^2 - 0.0000001466\,(\text{pers./zone})^2 + 0.000018689\,(\text{tot. DU/zone})^2 - 0.0000094125\,(\text{who.-ret. jobs})^2 - 0.00061319\,(\text{prof. jobs})^2 + 0.0001369\,(\text{grouped jobs})^2 - 0.0032360\,(LU_{220})^2 - 0.054781\,(LU_{240})^2 - 0.0055373\,(LU_{280})^2 \qquad (11)$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.}$$

RSA 51309--home-other, reduced sample:

$$Y = -0.021097 - 0.080372\,(LU_6) + 0.0045021\,(LU_9) + 0.0078854\,(\text{pers./zone}) - 0.013214\,(\text{tot. DU/zone}) - 0.049754\,(\text{who.-ret. jobs}) + 0.26185\,(\text{pers. serv. jobs}) - 0.018923\,(\text{prof. jobs}) - 0.0041486\,(\text{grouped jobs}) + 0.35568\,(LU_{220}) + 0.88353\,(LU_{240}) + 2.63884\,(LU_{250}) + 0.35530\,(LU_{280}) + 0.00017385\,(LU_6)^2 - 0.0000003212\,(LU_9)^2 - 0.000003921\,(\text{pers./zone})^2 + 0.000059441\,(\text{tot. DU/zone})^2 - 0.000024614\,(\text{who.-ret. jobs})^2 - 0.00003059\,(\text{prof. jobs})^2 + 0.000078437\,(\text{grouped jobs})^2 - 0.0031158\,(LU_{220})^2 - 0.032500\,(LU_{240})^2 - 0.0043047\,(LU_{280})^2$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.} \qquad (12)$$

NHA 04714--non-home, full sample:

$$Y = -0.722068 - 0.004496\,(LU_1) - 0.055532\,(LU_5) + 0.052999\,(LU_6) + 0.045612\,(\text{jobs}) + 0.02664\,(\text{pers./zone}) - 0.0314071\,(\text{tot. DU/zone}) - 0.009572\,(\text{cars/zone}) + 0.0000005513\,(LU_1)^2 + 0.00012975\,(LU_5)^2 + 0.00045254\,(LU_6)^2 - 0.000019178\,(\text{jobs})^2 - 0.000016506\,(\text{pers./zone})^2 + 0.000061364\,(\text{tot. DU/zone})^2 + 0.000044775\,(\text{cars/zone})^2$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.} \qquad (13)$$

NRA 04714--non-home, reduced sample:

$$Y = -3.03559 - 0.0073974\,(LU_1) - 0.024895\,(LU_5) - 0.0027639\,(LU_6) + 0.056735\,(\text{jobs}) + 0.062327\,(\text{pers./zone}) - 0.11725\,(\text{tot. DU/zone}) - 0.014188\,(\text{cars/zone}) + 0.0000014912\,(LU_1)^2 + 0.000048958\,(LU_5)^2 + 0.00063131\,(LU_6)^2 - 0.00002446\,(\text{jobs})^2 - 0.00002917\,(\text{pers./zone})^2 + 0.00014079\,(\text{tot. DU/zone})^2 + 0.000049897\,(\text{cars/zone})^2$$
$$Y = \text{trips/zone/1,000 trips produced in 14 selected zones.} \qquad (14)$$

OSP 14034--non-home, full sample:

$$Y = 3.34575 - 6.39670\,(\text{cars/DU}) - 0.52092\,(\text{pers./DU}) + 0.000120\,(\text{CBD dist.}) + 2.52203\,(\text{cars/DU})\,(\text{pers./DU}) - 0.35194\,(\text{pers./DU})\,(\text{CBD dist.}) + 0.13654\,(\text{CBD dist.})^2$$

$$Y = \text{trips/dwelling unit.} \tag{15}$$

RSP 14034--non-home, reduced sample:

$$Y = 3.75602 - 6.31798\,(\text{cars/DU}) - 1.19446\,(\text{pers./DU}) + 0.46274\,(\text{CBD dist.}) + 2.76797\,(\text{cars/DU})\,(\text{pers./DU}) - 0.45012\,(\text{pers./DU})\,(\text{CBD dist.}) + 0.13799\,(\text{CBD dist.})^2$$

$$Y = \text{trips/dwelling unit.} \tag{16}$$

In these equations, LU_x indicates 1,000's of square feet of land use; x, if a single digit, indicates major group land use; and x, if three digits, indicates land-use categories within major group land uses. The land-use codes have been published previously (10, Appendix A).

In Eqs. 3, 4, 5, 6, 15, and 16:

- DU = dwelling units which responded to the O-D interview;
- pers. = persons; and
- CBD dist. = distance from the zone centroid in question to the centroid of zone 12, in minutes.

In Eqs. 7 through 14:

- tot. DU/zone = total number of dwelling units per zone;
- grouped jobs = total jobs in wholesale, retail, finance, personal services, amusement, recreation, professional government, and self-employed;
- jobs = total of all jobs;
- who.-ret. jobs = total jobs in wholesale, retail, finance and insurance;
- pers. serv. jobs = total jobs in personal service; and
- prof. jobs = total jobs in professional area.

Equations 15 and 16 were used only for the estimation of numbers of non-home productions or attractions.

ANALYSIS OF RESULTS

Estimates of Trip Production and Attraction

The coefficients of correlation, R, for the estimating equations were obtained from the SCRAP regression analysis program and are shown in Table 9. The squared correlation coefficient or coefficient of determination, R^2, is a measure of the amount of variation about the mean that the estimating equation explains. Although many of the R^2 values were quite high, this coefficient did not necessarily indicate the predictive power of the various equations. For the same data, however, higher values of R^2 did indicate better predictive power of the form of equation being used.

A more meaningful statistical test of the estimating power of the equations was felt to be the calculation of RMS errors. The RMS error for each equation was computed by summing the squares of the differences between the estimated and surveyed values of production or attraction and dividing the total squared differences by the number of zoned productions or attractions and finding the square root of the quotient:

$$\text{RMS error} = \sqrt{\frac{(Y - Y_{est})^2}{N}} \tag{17}$$

TABLE 9

COEFFICIENTS OF CORRELATION AND DETERMINATION FROM REGRESSION ANALYSIS

Eq. No.	R	R^2
3	0.936	0.876
4	0.952	0.907
5	0.918	0.843
6	0.912	0.831
7	0.995	0.990
8	0.991	0.982
9	-[a]	-[a]
10	-[a]	-[a]
11	0.994	0.989
12	0.992	0.984
13	0.995	0.990
14	0.986	0.972
15	0.968	0.937
16	0.961	0.923

[a]Not determined, equation developed in "polynomial best fit" program (see Table 10 for RMS error comparisons).

TABLE 10

RMS ERRORS OF ESTIMATING EQUATIONS

Eq. No.	Avg. Trips per Zone	RMS Error	% RMS Error[a]
3	302	65	21
4	297	118	40
5	545	137	25
6	545	171	31
7	306	71	23
8[b]	306	218	71
8[c]	307	146	48
9	219	39	18
10[b]	219	88	40
11	452	208	51
12[b]	411	208	51
13	297	65	22
14[b]	306	220	72
14[c]	306	162	53

[a]% RMS error = 100 (RMS error)/average trips per zone.
[b]Using estimated total non-home productions to expand to zonal totals.
[c]Using O-D total non-home productions to expand to zonal totals.

where

Y = surveyed value,

Y_{est} = value estimated from regression equation, and

N = number of values.

The RMS error indicates the limits within which about two-thirds of the deviations between the observed and the estimated values will fall. The RMS errors, shown in Table 10, for the developed estimating equations were smallest when the regression equation was based on data obtained from the comprehensive O-D survey. The recorded RMS error, in most cases, appears to be reasonable when one considers that this is equivalent to stating that two-thirds of the time the estimated zonal productions or attractions can be expected to be within one RMS error of the actual value. The estimating power of Eqs. 8, 10, 12, and 14 were much improved when O-D survey productions were used to expand to zonal values. A plot of the estimated values vs O-D values of zonal productions or attractions provided an excellent graphical portrayal of the goodness of fit of the estimating equations. Figures 7 through 12 show the comparison of O-D trips per zone by purpose to the estimated zonal trips as obtained from Eqs. 4, 6, 8, 10, 12, and 14. Comparisons for all equations have been previously published (10). If the estimated value was equal to the O-D value, the plotted point fell on the 45° line. The plot of the O-D value ± RMS error vs the O-D value indicates a band within which one would expect the estimated values to fall about two-thirds of the time.

Gravity Model Distribution

As noted, two sets of travel time factors were developed. One set, F_1, was based on the O-D productions, attractions and trip length frequency distribution as obtained from the comprehensive O-D data in all zones; the other, F_2, was based on productions,

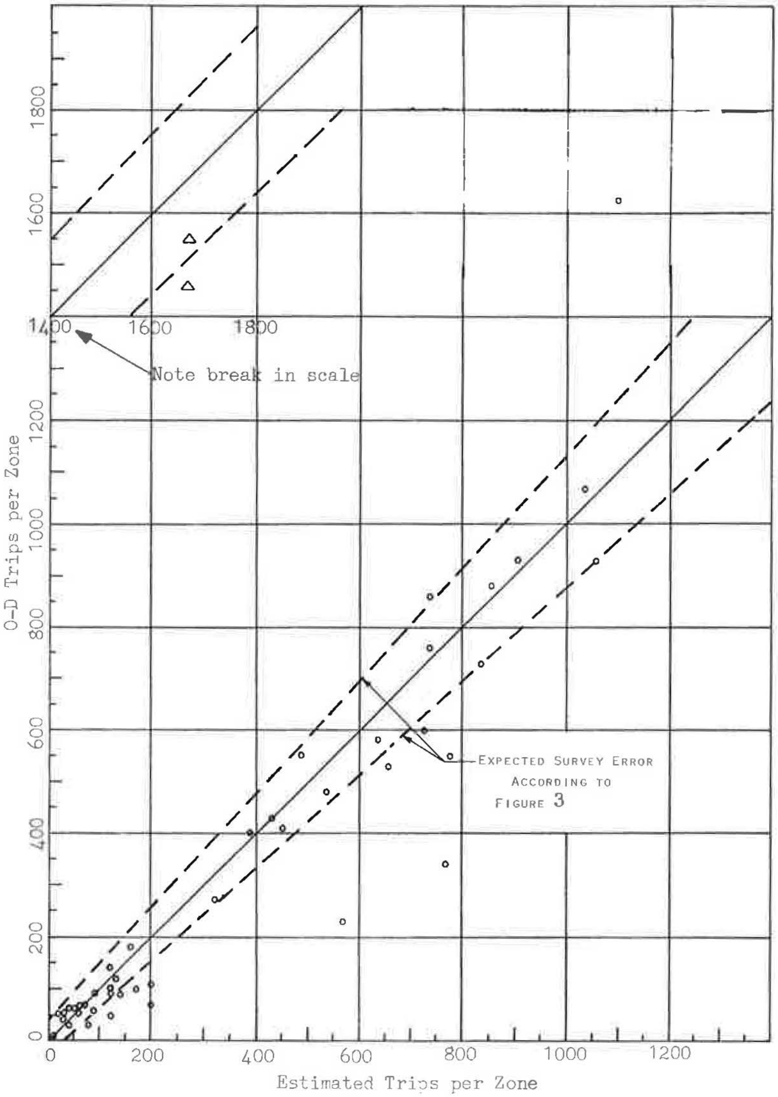


Figure 7. Comparison of home-work trips produced (Eq. 4).

attractions, and the trip length frequency distribution obtained from the reduced sample in the 14 selected zones. Table 8 shows the developed travel time factors. The trip distribution of the model was analyzed using the following four combinations of model parameters:

Combination 1—O-D productions, attractions, and travel time factors, F_1;
Combination 2—O-D productions, attractions, and travel time factors, F_2;
Combination 3—estimated productions, attractions, and travel time factors, F_1; and
Combination 4—estimated productions, attractions, and travel time factors, F_2.

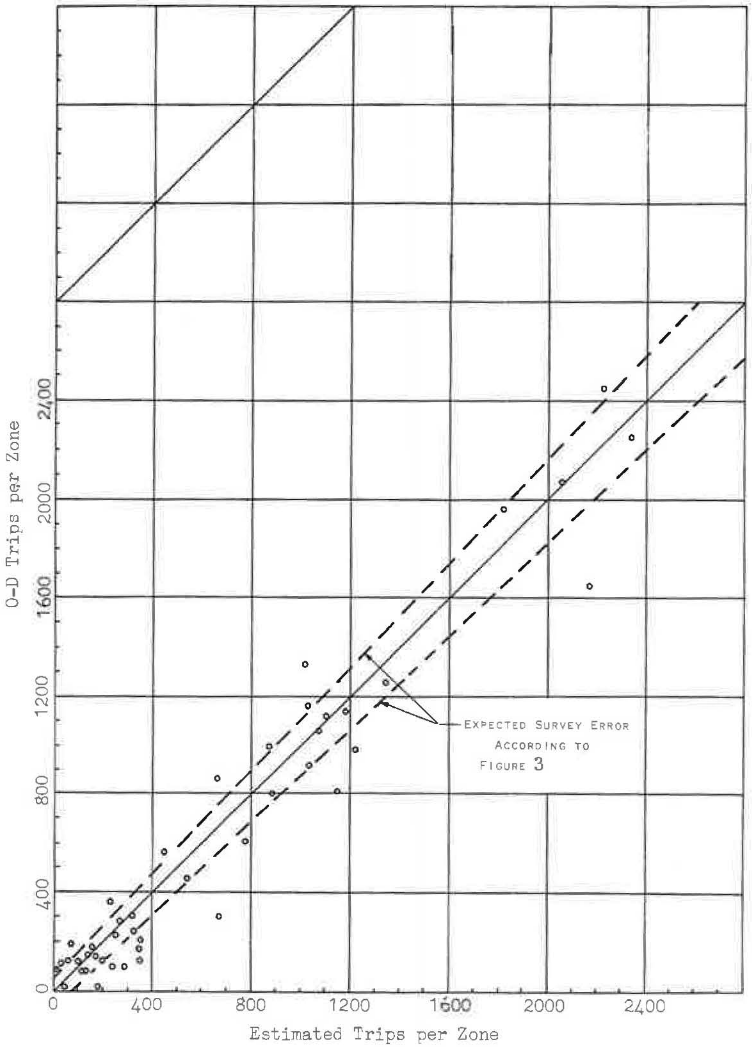


Figure 8. Comparison of home-other trips produced (Eq. 6).

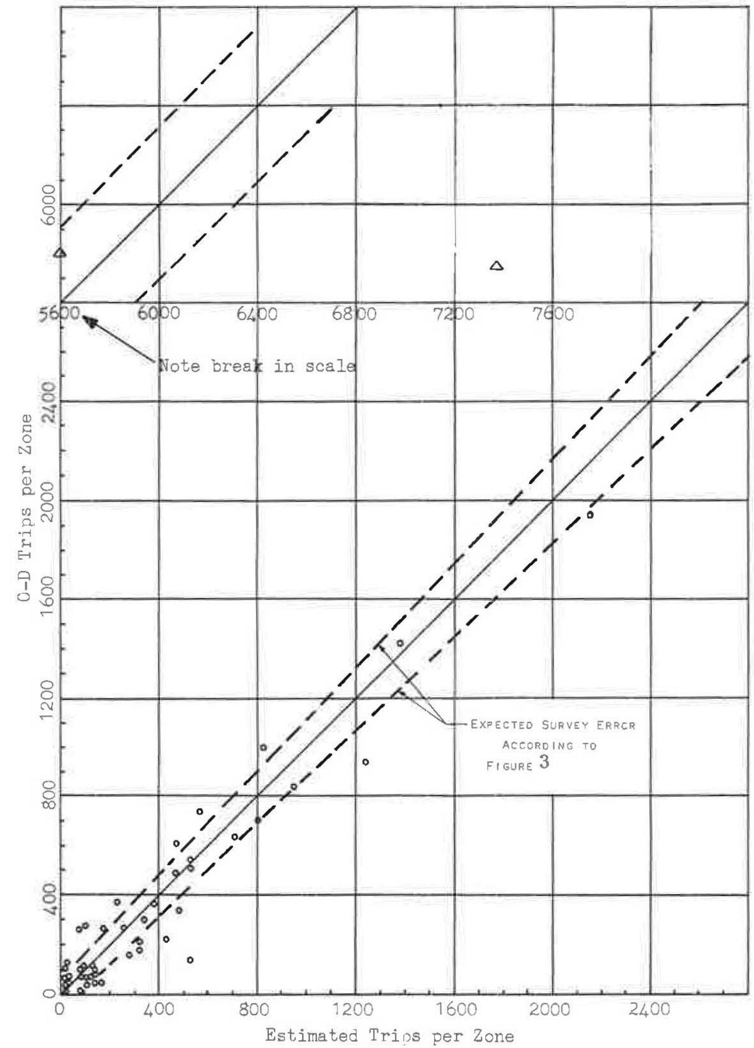


Figure 9. Comparison of non-home trips produced (Eq. 8).

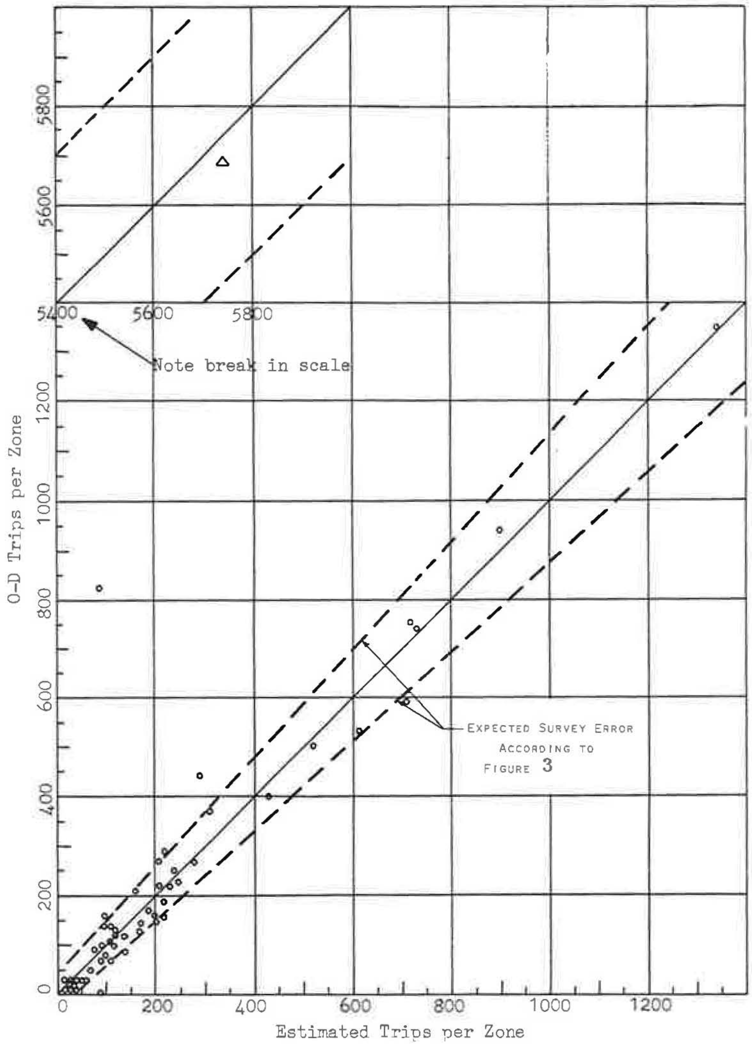


Figure 10. Comparison of home-work trips attracted (Eq. 10).

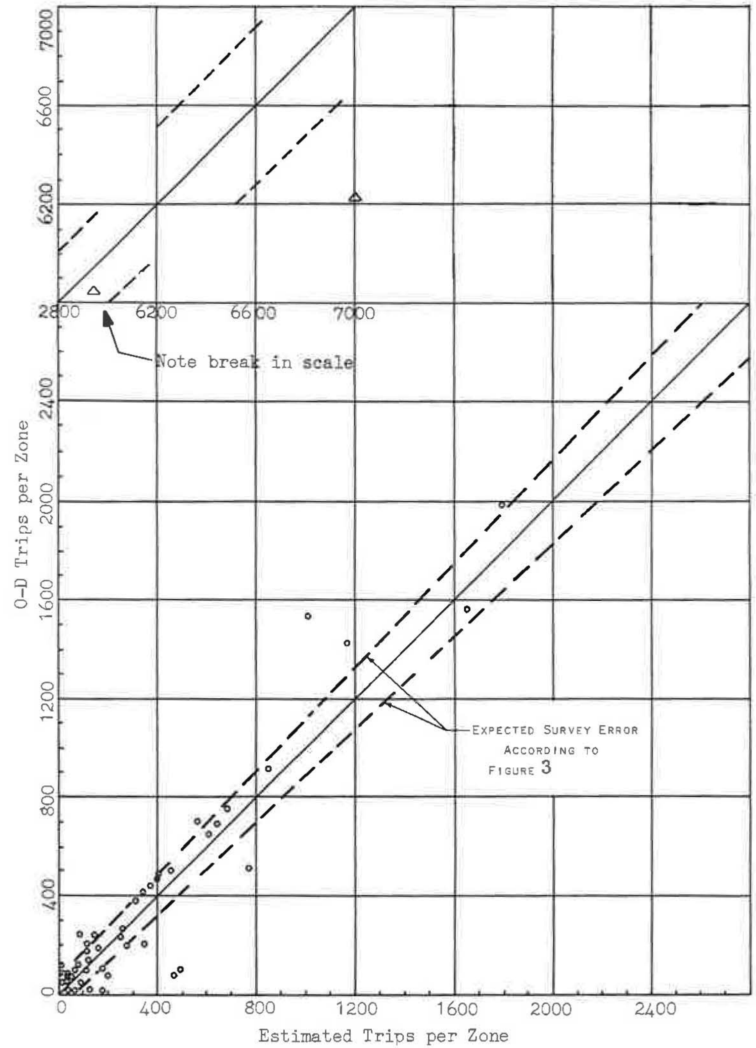


Figure 11. Comparison of home-other trips attracted (Eq. 12).

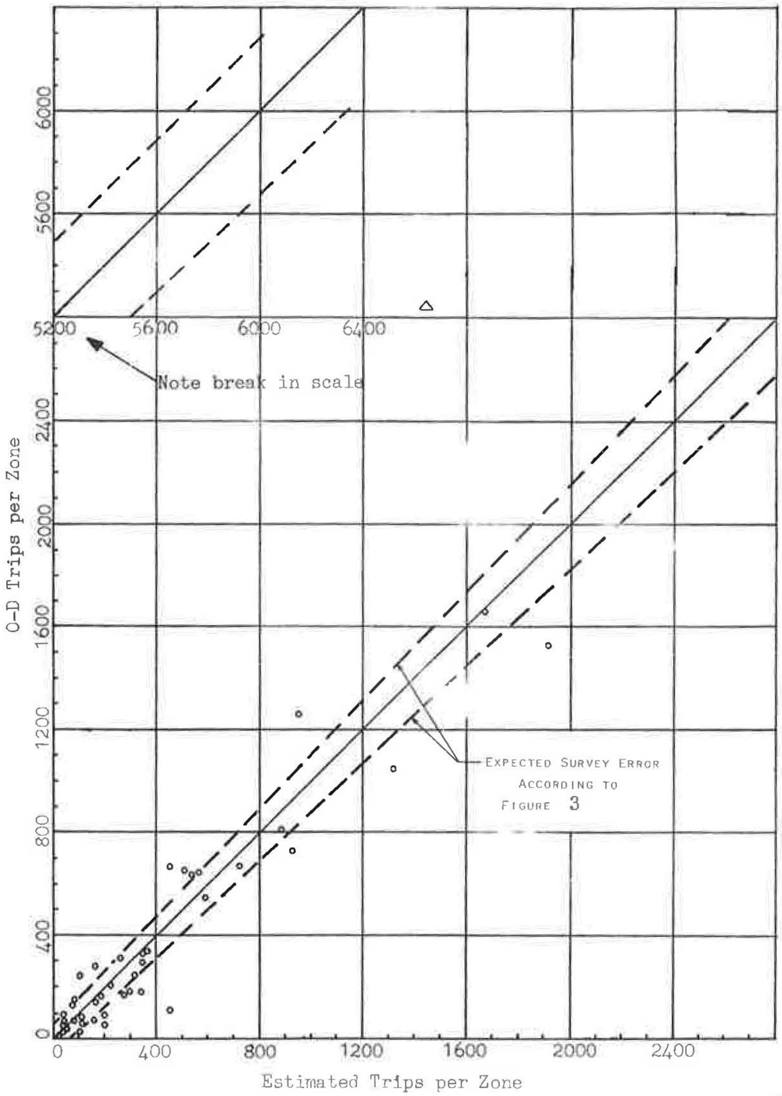


Figure 12. Comparison of non-home trips attracted (Eq. 14).

Screenline Comparison.—Seven screenlines were chosen for a comparison of crossings using the O-D data and those obtained from the gravity model with the various combinations of parameters. The location of the various screenlines is shown in Figure 1. Crossings of screenline 6 showed the largest percent difference; however, the number of trips crossing the line was very small, making it difficult to obtain a close agreement in percent. It is believed, however, that this did indicate some geographical bias in the model and could probably have been remedied by increasing the terminal times in the zones south of screenline 6 or by applying zone-to-zone k factors. The total number of trips involved, however, did not appear to warrant such adjusting procedures. The various screenline crossings are compared in Table 11.

Trip Length Comparison.—It was felt that the comparison of the total amount of travel and average trip length as obtained from the various model distributions would serve as measurements of the adequacy of the model. Figures 13 through 15 show the

TABLE 11

COMPARISON OF SCREENLINE CROSSINGS USING O-D AND MODEL DATA

Screen-line[a]	Crossings from Com-plete O-D Survey	Full Sample[b]				Reduced Sample[e]			
		Model F_1[c]	Model/O-D (%)	Model F_2[d]	Model/O-D (%)	Model F_1[c]	Model/O-D (%)	Model F_2[d]	Model/O-D (%)
(a) All Trips									
1	27,564	27,478	99.7	29,924	108.6	27,861	101.1	29,884	108.4
2	20,557	20,946	101.9	21,930	106.7	21,157	102.7	22,212	108.1
3	27,656	26,439	95.6	28,248	102.1	26,678	96.5	28,505	103.1
4	24,530	24,008	97.9	25,299	103.1	24,240	98.8	25,473	103.8
5	26,828	24,765	92.3	28,643	106.8	26,419	98.5	28,536	106.4
6	4,202	5,156	122.7	5,720	136.1	5,095	121.3	5,704	135.7
7	10,028	9,682	96.5	11,302	112.7	9,734	97.1	11,755	117.2
Total	141,365	138,474	98.0	151,066	106.9	141,184	99.9	152,069	107.6
(b) Home-Work Trips									
1	8,507	8,731	102.6	9,537	112.1	8,793	103.4	9,438	110.9
2	6,140	6,229	101.4	6,762	110.1	6,328	103.1	6,990	113.8
3	8,908	8,245	92.6	8,427	94.6	8,320	93.4	8,565	96.1
4	7,714	7,535	97.7	7,566	98.1	7,609	98.6	7,642	99.1
5	7,844	7,644	97.5	8,339	106.3	7,642	97.4	8,007	102.1
6	1,235	1,427	115.5	1,838	148.8	1,450	117.4	1,839	148.9
7	3,596	3,440	95.7	3,672	102.1	3,522	97.9	4,194	116.6
Total	43,944	43,251	98.4	46,141	105.0	43,664	99.4	46,675	106.2
(c) Home-Other Trips									
1	9,959	9,916	99.6	10,275	103.2	10,007	100.5	10,332	103.7
2	9,002	9,182	102.0	9,784	108.2	9,292	103.2	9,840	109.3
3	11,694	11,289	96.5	12,518	107.0	11,450	97.9	12,634	108.0
4	10,509	10,342	98.4	11,464	110.8	10,500	99.9	11,569	110.1
5	9,785	9,388	95.9	9,840	100.6	9,605	98.2	10,073	102.9
6	1,731	2,130	123.1	2,274	131.4	2,128	122.9	2,256	130.3
7	4,239	4,077	96.2	4,971	117.3	4,045	95.4	4,902	115.6
Total	56,919	56,324	99.0	61,126	107.4	57,027	100.2	61,606	108.2
(d) Non-Home Trips									
1	9,098	9,055	99.5	10,112	111.1	9,061	99.6	10,114	111.2
2	5,415	5,535	102.2	5,385	99.4	5,538	102.3	5,383	99.4
3	7,054	6,905	97.9	7,304	103.5	6,908	97.9	7,306	103.6
4	6,309	6,131	97.2	6,271	99.4	6,131	97.2	6,262	99.3
5	9,198	9,181	99.8	10,465	113.8	9,172	99.7	10,456	113.7
6	1,235	1,515	122.7	1,608	130.2	1,518	122.9	1,609	130.3
7	2,193	2,165	98.7	2,660	121.3	2,167	98.8	2,659	121.2
Total	40,502	40,487	100.0	43,805	108.2	40,495	100.0	43,789	108.1

[a]See Figure 1 for the locations of screenlines. [b]Using productions and attractions from the complete O-D survey.
[c]Using travel time factors developed with model input productions and attractions from the complete O-D survey.
[d]Using travel time factors developed with model input productions and attractions from the 14 zones with a reduced sample size. [e]Using productions and attractions obtained from the regression estimating equations developed from the reduced sample in the 14 selected zones.

trip length frequency distribution, by purpose, of the O-D data vs the model with input from the small sample data. These comparisons show good agreement of the model and O-D data. Additional comparisons are found in the earlier report (10). The average lengths of trip in minutes and total vehicle-hours of travel as compared in Table 12 are also in close agreement.

Since there were few diagonal streets in Hutchinson, the distance, in miles, between zone centroids was measured by determining the L distance (sum of map coordinate differences); the total vehicle-miles of travel was taken as the product of the zone-to-zone interchange and the L distance between the zone centroids in question with the summation being made over all zones. The average trip length in miles was taken as the total vehicle-miles of travel divided by the total number of trips. The comparisons of the average length of trip in miles and total vehicle-miles of travel are given in Table 13 and again indicate close agreement.

Comparison of District-to-District Movements.—A comparison was also made of the district-to-district movements given by both the model and O-D data. The comparison was originally made using zone-to-zone movements but the small numbers of trips gave results having little stability. The RMS errors for the various volume groups and trip purposes, as shown in Tables 14 through 16, indicate that two-thirds of the time the difference between district interchanges, as given by the model and O-D data, is expected to be equal to, or less than, the value of the RMS error.

District-to-district trip interchanges can be used in determining interchange volumes between the CBD and various corridors. This can provide a check on the geographical bias of the model. However, in this analysis, it was felt that the screenline checks gave a dependable test for bias.

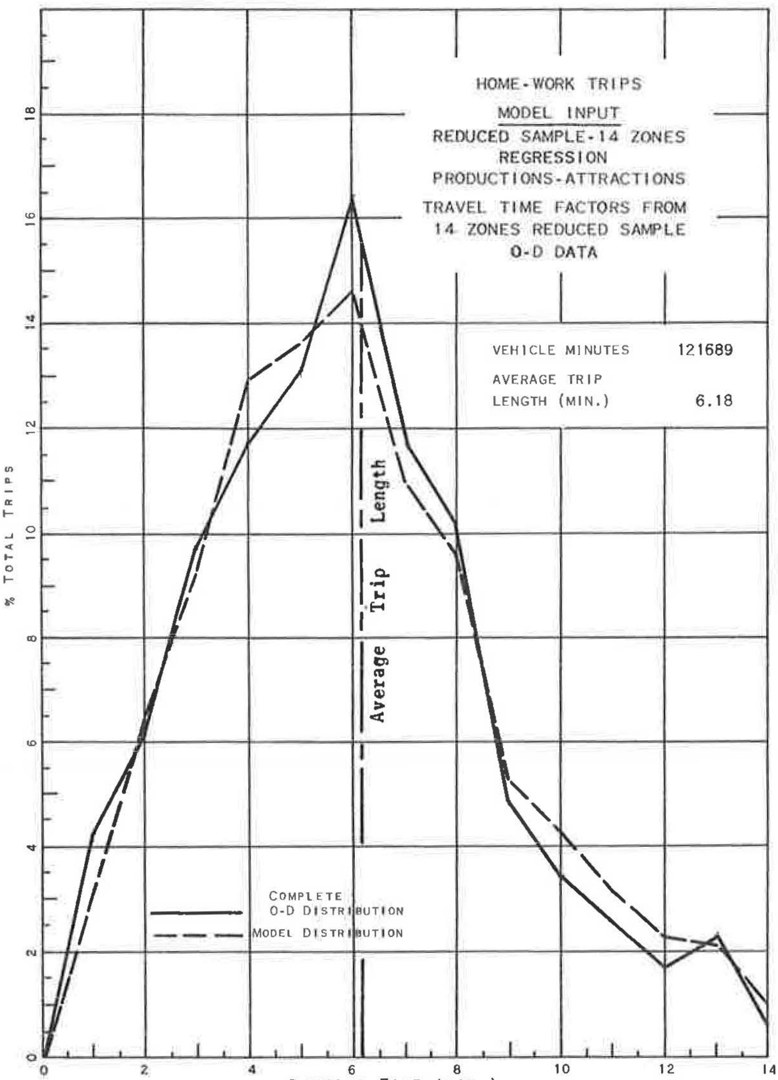


Figure 13. Comparison of trip length frequency using O-D and model data, home-work trips, combination 4 parameters.

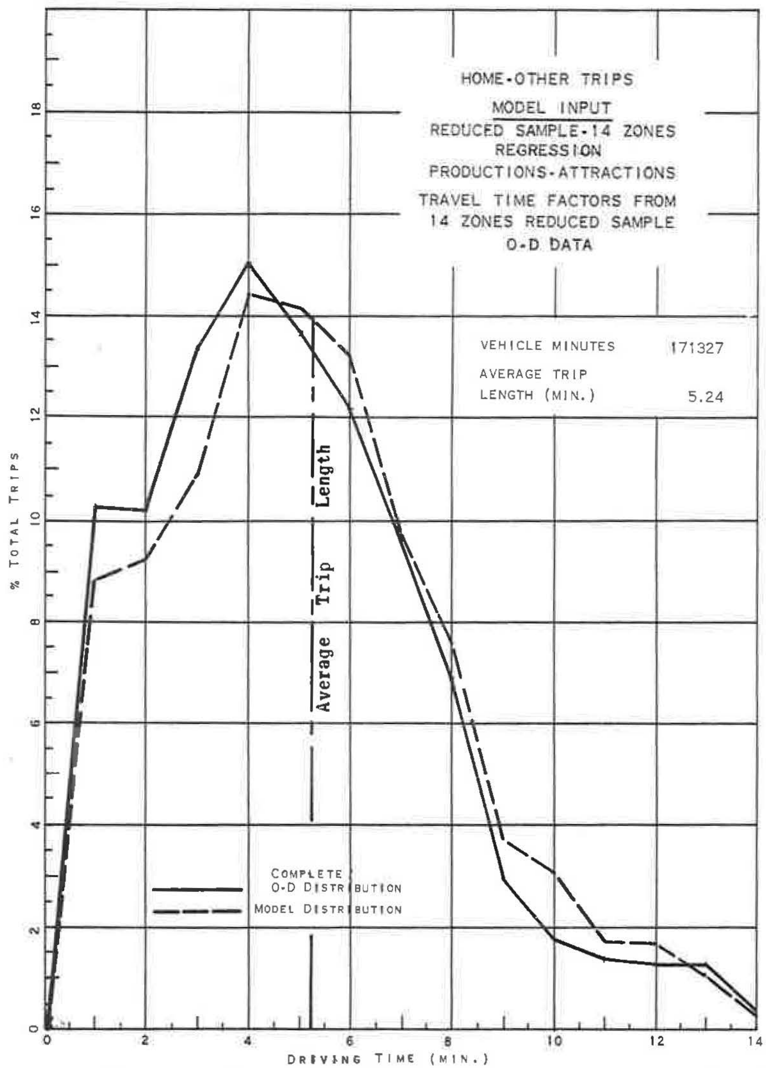


Figure 14. Comparison of trip length frequency using O-D and model data, home-other trips, combination 4 parameters.

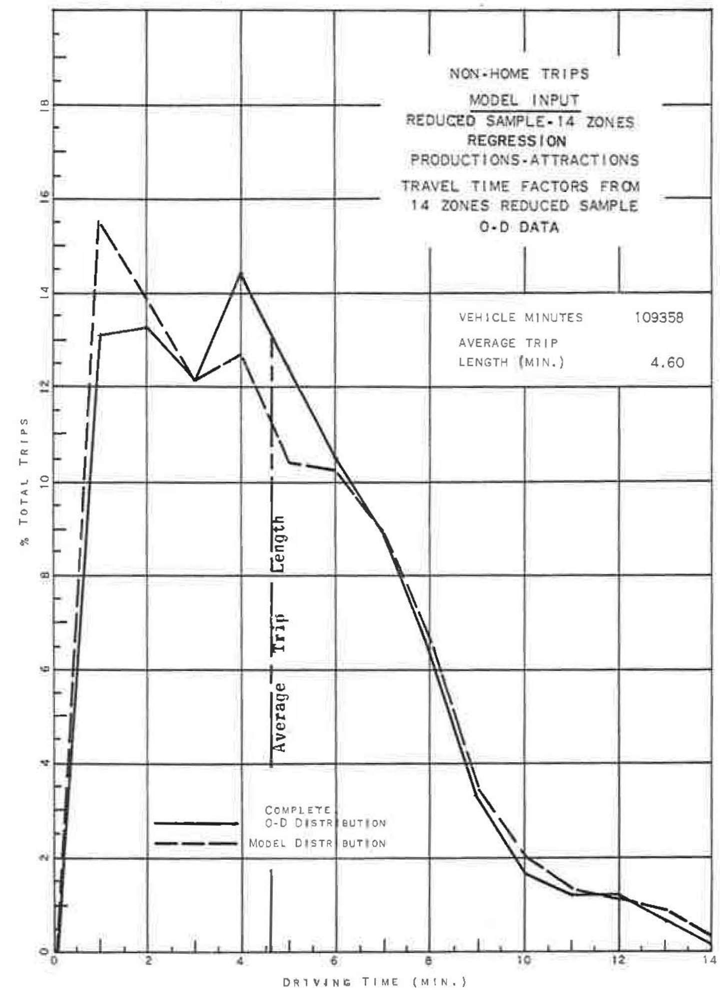


Figure 15. Comparison of trip length frequency using O-D and model data, non-home trips, combination 4 parameters.

TABLE 12

VEHICLE-HOURS OF TRAVEL AND AVERAGE TRIP LENGTH FROM O-D AND MODEL DATA

Purpose	Complete O-D Survey Data		Full Sample				Reduced Sample			
			O-D Prod.-Attract.		Regression Prod.-Attract.		O-D Prod.-Attract.		Regression Prod. Attract.	
	Veh-Hr	Avg. Trip Length (min)	Veh-Hr	Avg. Trip Length (min)	Veh-Hr	Avg. Trip Length (min)	Veh-Hr	Avg. Trip Length (min)	Veh-Hr	Avg. Trip Length (min)
Home-work	1,831	6.06	1,823	6.03	1,970	6.01	1,842	6.09	2,028	6.18
Home-other	2,584	4.91	2,610	4.95	2,845	5.22	2,627	4.98	2,855	5.24
Non-home	1,597	4.48	1,657	4.64	1,823	4.60	1,657	4.64	1,823	4.60

TABLE 13

VEHICLE-MILES OF TRAVEL AND AVERAGE TRIP LENGTH FROM O-D AND MODEL DATA

Purpose	Complete O-D Survey Data		Full Sample				Reduced Sample			
			O-D Prod.-Attract.		Regression Prod.-Attract.		O-D Prod.-Attract.		Regression Prod.-Attract.	
	Veh-Mi	Avg. Trip Length (mi)	Veh-Mi	Avg. Trip Length (mi)	Veh-Mi	Avg. Trip Length (mi)	Veh-Mi	Avg. Trip Length (mi)	Veh-Mi	Avg. Trip Length (mi)
Home-work	35,331.6	1.946	35,325.6	1.947	38,327.3	1.948	35,765.0	1.971	40,007.6	2.033
Home-other	48,248.2	1.527	49,193.3	1.552	54,894.9	1.680	49,454.6	1.563	54,882.0	1.679
Non-home	30,755.6	1.438	30,630.7	1.431	33,731.9	1.419	30,636.7	1.431	33,717.9	1.418

TABLE 14

ANALYSIS OF DISTRICT-TO-DISTRICT MOVEMENTS[a]

Volume Group	Freq.[b]	Total Trips		Mean Diff.	Std. Dev.	RMS Error	% RMS Error
		O-D	Model				
0-99	33	1,858	2,491	19	49	52	93
100-199	32	4,648	6,219	-49	76	90	62
200-299	32	7,875	8,831	-29	89	93	38
300-399	14	4,891	6,075	-84	132	157	44
400-499	8	3,476	4,024	-68	109	129	29
500-599	2	1,089	582	253	158	298	54
600-699	4	2,576	2,883	-76	234	247	38
700-799	4	2,943	2,608	83	143	165	22
800-899	2	1,641	1,800	-79	162	180	22
900-999	2	1,896	1,735	80	77	111	11
1,000-1,499	10	12,750	12,558	19	328	329	25
1,500-1,999	4	6,753	6,659	23	289	290	17
2,000-2,999	4	9,581	10,739	-289	411	503	21
3,000-3,999	1	3,549	4,837	-1,288	-	1,287	36
4,000-4,999	-	-	-	-	-	-	-
5,000-5,999	1	5,199	3,795	1,404	-	1,403	27
6,000-6,999	-	-	-	-	-	-	-
7,000-7,999	-	-	-	-	-	-	-
8,000-8,999	-	-	-	-	-	-	-
9,000-9,999	-	-	-	-	-	-	-
10,000-999,999	-	-	-	-	-	-	-
Total		70,725	75,836				

[a]All trips, full sample, regression productions-attractions, travel time factors from reduced sample O-D data.

[b]Number of district-to-district movements within volume group.

TABLE 15

ANALYSIS OF DISTRICT-TO-DISTRICT MOVEMENTS[a]

Volume Group	Freq.[b]	Total Trips		Mean Diff.	Std. Dev.	RMS Error	% RMS Error
		O-D	Model				
0-99	33	1,858	2,554	-21	48	52	93
100-199	32	4,648	5,907	-39	83	92	63
200-299	32	7,875	8,752	-27	88	92	37
300-399	14	4,891	6,034	-81	137	160	45
400-499	8	3,476	3,923	-55	116	129	29
500-599	2	1,089	611	239	164	290	53
600-699	4	2,576	2,970	-98	250	269	41
700-799	4	2,943	2,653	72	164	180	24
800-899	2	1,641	1,815	-87	118	147	17
900-999	2	1,896	1,725	85	78	116	12
1,000-1,499	10	12,750	12,603	14	309	309	24
1,500-1,999	4	6,753	6,714	9	303	303	18
2,000-2,999	4	9,581	10,923	-335	374	502	20
3,000-3,999	1	3,549	4,846	-1,297	-	1,296	36
4,000-4,999	-	-	-	-	-	-	-
5,000-5,999	1	5,199	3,780	1,419	-	1,418	27
6,000-6,999	-	-	-	-	-	-	-
7,000-7,999	-	-	-	-	-	-	-
8,000-8,999	-	-	-	-	-	-	-
9,000-9,999	-	-	-	-	-	-	-
10,000-999,999	-	-	-	-	-	-	-
Total		70,725	75,810				

[a]All trips, full sample, regression productions-attractions, travel time factors from full sample O-D data.

[b]Number of district-to-district movements within volume group.

TABLE 16

ANALYSIS OF DISTRICT-TO-DISTRICT MOVEMENTS[a]

Volume Group	Freq.[b]	Total Trips		Mean Diff.	Std. Dev.	RMS Error	% RMS Error
		O-D	Model				
0-99	33	1,858	1,965	-3	39	39	70
100-199	32	4,648	5,041	-12	60	61	42
200-299	32	7,875	8,389	-16	99	100	40
300-399	14	4,891	4,915	-1	90	90	25
400-499	8	3,476	3,638	-20	78	81	18
500-599	2	1,089	506	291	85	303	55
600-699	4	2,576	2,517	14	91	92	14
700-799	4	2,943	3,001	-17	196	197	26
800-899	2	1,641	1,757	-58	29	65	7
900-999	2	1,896	1,662	117	85	145	15
1,000-1,499	10	12,750	12,433	31	294	296	23
1,500-1,999	4	6,753	6,399	88	251	256	15
2,000-2,999	4	9,581	10,498	-229	159	279	11
3,000-3,999	1	3,549	3,868	-319	-	318	8
4,000-4,999	-	-	-	-	-	-	-
5,000-5,999	1	5,199	4,018	1,181	-	1,180	22
6,000-6,999	-	-	-	-	-	-	-
7,000-7,999	-	-	-	-	-	-	-
8,000-8,999	-	-	-	-	-	-	-
9,000-9,999	-	-	-	-	-	-	-
10,000-999,999	-	-	-	-	-	-	-
Total		70,725	70,607				

[a]All trips, full sample, complete O-D productions-attractions, travel time factors from full sample O-D data.

[b]Number of district-to-district movements within volume group.

CONCLUSIONS

1. Current zonal trip productions and attractions were adequately estimated from mathematical models developed from a small sample of home interviews (8) taken in a sample of the origin-destination zones. Best estimates resulted for home-based trip productions but estimated non-home-based trip productions and all trip attractions appeared to be adequate for planning purposes.

2. Mathematical models developed from current comprehensive O-D or reduced sample data should be of great value in estimating future zonal trip productions and attractions.

3. Only three trip purposes (home-work, home-other, and non-home) were found to be practical divisions of all trips for prediction of zonal trip productions and attractions from mathematical models based on comprehensive or reduced sample O-D data.

4. For home-work trip attractions, the number of jobs in a zone was the only important factor.

5. For home-work trip productions, the number of employed persons per dwelling unit was not found to be a more important factor than persons per dwelling unit.

6. For all trip productions, the number of persons and the number of cars per dwelling unit were found to be very important factors. Other factors of importance for trip productions were distance to the CBD for home-work and home-other trips, area of various land uses, and number of jobs for non-home trips.

7. For trip attractions other than the home-work trip, the number of persons per zone, the number of types of jobs in the zone, and the areas devoted to various land uses were found to be important factors.

8. Travel time factors for distribution of trips by the gravity model were satisfactorily estimated by calibrating the gravity model with trip length frequency data developed from a small sample of home interviews taken in a sample of the O-D zones.

9. The travel time factors which were developed varied in value for the different trip purposes for the same travel time separation.

10. The gravity model using trip productions and attractions and travel time factors developed from a small sample of home interviews taken in a sample of O-D zones distributed trips among all zones to give an adequate reproduction for planning purposes of the trip distribution obtained in a comprehensive O-D survey.

11. The gravity model using trip production and attractions and travel time factors developed from a comprehensive O-D survey distributed trips among all zones to give a good reproduction of the trip distribution obtained in the comprehensive survey.

ACKNOWLEDGMENT

This research project was conducted under an agreement between the Kansas Highway Commission and the Civil Engineering Department and Engineering Experiment Station of Kansas State University. The research was financed with Highway Planning Survey ($1\frac{1}{2}$ percent) Federal Highway funds.

REFERENCES

1. Calibrating and Testing a Gravity Model for Any Size Urban Area. U. S. Bureau of Public Roads, July 1963.
2. Voorhees, A. M. A General Theory of Traffic Movement. Proc. Inst. of Traffic Eng., 1955.
3. Voorhees, A. M. Forecasting Peak Hours of Travel. Highway Research Board Bull. 203, pp. 37-46, 1958.
4. Wiant, R. H. A Simplified Method for Forecasting Urban Traffic. Highway Research Board Bull. 297, pp. 128-145, 1961.
5. 1960-1980 Traffic Patterns in Ottumwa. Report of Traffic and Highway Planning Dept., Div. of Planning, Iowa State Highway Commission, in coop. with U. S. Bureau of Public Roads, Mar. 1961.

6. Voorhees, A. M., and Morris, Robert. Estimating and Forecasting Travel for Baltimore by Use of a Mathematical Model. Highway Research Board Bull. 224, pp. 105-114, 1959.
7. Sixteen-twenty Card Regression Analysis Program (SCRAP). IBM Program Library File No. 6.0.003, IBM, New York.
8. Sosslau, A. B., and Brokke, G. E. Appraisal of Sample Size Based on Phoenix O-D Survey Data. Highway Research Board Bull. 253, pp. 114-127, 1960.
9. Gravity Model Used by Iowa Planning Survey. Report of Div. of Planning and Research, Region 5, U. S. Bureau of Public Roads, July 1962.
10. Smith, B. L. An Application of Gravity Model Theory to a Small City in Kansas Using a Small Sample of Origin-Destination Data. Kansas Highway Commission, Dept. of Highway Planning, Topeka, Sept. 1963.