THE UNIVERSITY OF MANITOBA

AN URBAN TRANSPORTATION DEMAND MODEL

by

KENNETH ARNOLD ROSIN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN CIVIL ENGINEERING

WINNIPEG, MANITOBA

OCTOBER, 1972

ACKNOWLEDGEMENTS

The author wishes to acknowledge the financial support of the National Research Council, Ottawa, and the Transportation Development Agency through the Center for Transportation Studies, University of Manitoba.

The author would also like to acknowledge his advisor, Dr. A. H. Soliman whose counsel and assistance were vital to the completion of this thesis. The author also wishes to thank Dr. Arvind Shah; and Messrs. Boris Hryhorczuk, and Ed. Geurtin of the City of Winnipeg, Streets and Traffic Department for their assistance.

Last but by no means least, the author wishes to thank his wife Dianne who typed the thesis.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

1.1 OBJECT OF STUDY

Most transportation studies are done in order to predict, in some measure, what the future will be with regard to the movement of people. This person movement is the result of various motives: the trip to and from the place of employment; the shopping trip; the social or recreational trip, etc. For each type of trip purpose there is generally a choice as to the means of travel: private auto, public transit, on foot, etc. If it is required to predict, with a reasonable degree of accuracy, the future patterns of a particular type of person movement, it is necessary to investigate the motivation behind the mode of travel. The reasons for travelling are usually fairly obvious, as mentioned previously: to reach one's place of work, to purchase certain goods, etc. However, the reasons for the choice of method of conveyance for these trips are not always quite as clearly defined.

The purpose of this study, therefore, is to obtain the most significant factors which determine the choice of mode for the type of person trip which accounts for a major portion of the total daily trips in virtually every major city in Canada - the peak hour work trip. The study will analyze the movement of workers from suburban zones to the central business district, which generally has the highest employment density of any area in any large Canadian city.

The results obtained will then be incorporated into a mathematical demand model which will enable the volume of work trips, from the suburban areas of a city to the central business district of that city, to be predicted for each mode of travel.

1.2 RELATIONSHIP TO OTHER STUDIES

Only within the last fifteen to twenty years has the concept of a "demand" model, based on the supply and demand principles of economics, been proposed as a method of predicting future travel patterns.

Several individuals, notably Maclynn(3), Domencich(1), and Kraft(2), among others have investigated the possibilities of such a "demand" model. Perhaps the most significant has been the work done primarily by G.A. Kraft.(7) He introduced concepts of predicting volume changes in travel as a result of changes in various factors, such as price of travel, level of income, etc. Simplistically, the demand model developed by Kraft predicts travel volumes or demand,

> "in terms of the price and service of the primary mode (or mode in question), of the price and service of each of the competitive modes, of the income status of the travellers, and in terms of certain population, attractiveness and employment characteristics of either the origin or destination..."(7)

for each mode and motivation for travel.

Some of the concepts developed by Kraft and others will be used in this study to develop a type of demand model which will predict future work trip volumes from suburban areas to the central business district for any middle-sized Canadian city.

1.3 ASSUMPTIONS AND LIMITATIONS

As stated in section 1.1, this study will concentrate on one specific trip purpose: the peak hour working trip. This trip purpose was chosen primarily because a very high percentage of all daily trips (as much as forty per cent or more) in many middle-sized Canadian cities occur in this time period. Therefore, the volume of trips represented by the peak hour volume is the highest hourly volume which is likely to be handled by any present or future transportation system. As a corollary for this study, the volume of trips represented by the peak hour volume from the suburban areas of any major Canadian city to the central business district of the same city represents the highest hourly volume which must be managed by the present suburban-central business district transportation network.

This trip purpose was also chosen because it is considered the most stable of any trip classification[12]. In other words, any prediction of future peak hour work trips would likely be more accurate than prediction of other trip types, such as the shopping trip, because of the relative haphazardness of the latter as to time of departure, as compared to the former.

CHAPTER 2: GENERAL DISCUSSION OF MODELS

2.1 EXPRESSION OF GENERAL EQUATION

The general equation of demand herein expressed is based on ideas postulated by Domencich.[1] It incorporates the ideas of relative travel times, relative travel cost and various socioeconomic and land use characteristics into an expression of urban transportation demand by any mode for the work trip purpose.

The general expression for the demand model may be stated thus:

$$W(i,j/M_o) = f\left[T(i,j/M_o),\ C(i,j/M_o),\ T(i,j/M_\alpha),\ C(i,j/M_\alpha),\ E(i)\right] \quad (1)$$

where:

$W(i,j/M_o)$ - the number of peak hour work trips between zones i and j by mode M_o.

$T(i,j/M_o)$ - the travel time components from zone i to zone j by mode M_o.

$C(i,j/M_o)$ - the travel cost components from zone i to zone j by mode M_o.

$T(k,j/M_\alpha)$ - the travel time components from zone i to zone j by mode $M_\alpha(=1, \ldots n)$.

$C(i,j/M_\alpha)$ - the travel cost components from zone i to zone j by mode $M_\alpha(=1, \ldots n)$.

$E(i)$ - socioeconomic and land use characteristics of workers in zone i.

2.2 VARIABLES TO BE CONSIDERED

As stated in the general expression the three basic variables which are to be used initially are: travel time, travel cost, and socioeconomic-land use factors. These will be discussed individually and further refined for use in a more specific equation.

2.2.1. Travel Time

The time involved in travelling from place of origin to place of employment in the CBD must be considered to be a contributing factor to any initial discussion of work trip volumes by either mode.

This time value may be further subdivided into the time spent in the vehicle (bus or auto) and the time spent outside the vehicle*. The former value will be hereafter referred to as the travel time, the latter value will be known as the excess time.

It will be postulated that these two values of time will each have a possible effect on the work trip volumes by either mode, and should both play a part in any initial demand model development.

Further to this, it will also be assumed that relative time values will be a better indicator of work trip volumes. By relative times is meant the travel time or excess time of the mode under consideration relative to the travel time or excess time of the other mode. These relative time ratios (hereafter known as travel time ratio and excess time ratio) are a measure of the travel

* Time spent outside the bus includes: walking time from home to bus stop, time spent waiting for bus, any transfer time and time spent walking from bus stop to place of employment.
Time spent outside the auto is the walking time from parking lot to place of employment.

time or excess time spent on one particular mode relative to the other mode. As such they may better illustrate the dependence of the volume of work trips on travel time or excess time because they relate one mode to the other as a sort of measure of travel time efficiency. This seems to be a truer indication of how a particular mode is chosen vis a vis time involved in travel, not by how long it takes by mode A, but how long it takes by mode A relative to mode B or vice versa.

In urban centers which have more than two basic modes of travel (such as auto, transit and rail for example) the relative values of the travel times for each mode could be considered separately by comparing A with B, A with C and B with C to obtain results analogous to the two mode system. A second approach would be to combine all public transit into one mode and treat the situation as a two mode system.

2.2.2. Travel Cost

The user cost of a work trip from the suburbs to the CBD is a second consideration when attempting to develop a discussion of work trip demand.

This user cost may be thought of as the amount spent on a transit fare, or the out of pocket costs of using a private auto (gas, oil, and lubrication) plus the terminal or parking costs.

Once again, as in the case of time, it will be assumed that the relative cost factors will be of more significance to any prediction of work trip demand than the straight cost by either mode.

The relative cost will be defined as the user cost for the mode under consideration in the demand equation divided by the user cost for the other mode. This ratio (hereafter known as the cost of travel ratio) is analogous to the time ratios discussed in section 2.2.1. That is, it is a measure of cost of using mode A relative to mode B or vice versa. This also appears to be a truer indication of an individual's thinking regarding choice of mode for the work trip - not necessarily how much it cost to utilize mode A but its user cost relative to mode B (for more than two modes, situation is analogous, see section 2.2.1.).

2.2.3. Socioeconomic - Land Use Factors

For purposes of analyzing the work trip from the suburban areas of a city to the CBD, several socioeconomic and land use factors may be considered.

Firstly, car ownership in the suburban zones of origin, usually expressed as cars per family. This factor is an indicator of income level in a particular zone, that is generally the higher the number of cars per family, the higher the income level. It seems fairly obvious that a profusion of autos in a particular suburban area will have a great bearing on the mode by which the work trip to the downtown area is made. Conversely, a dirth of cars in another suburban area might have quite a different effect on the modal choice. Both situations should be indicators of the number of working trips which utilize each mode in the respective zones.

A second factor to be considered is the population density of the origin zones, usually expressed in persons per acre. This value also may be thought of as an indicator of income level. Zones in which the population density is high are generally regarded as lower income zones as compared with the newer lower density (as a result of single family dwellings) and relatively higher income zones. The lower income, higher density zone working trips might possibly be more transit oriented than the lower density, higher income areas because these lower income areas will usually have less funds available for transportation to place of work.

A third possible factor, which may influence the work trip volumes by either mode, is the number of employed individuals who live in each suburban area (henceforth called the working population). One hypothesis regarding working population might be that a zone in which this value is quite high will produce more working trips destined for the CBD than zones of lower working population. This statement should be true for either mode.

Other factors which may be significant to the work trip volumes are: the frequency of transit service (this may be thought of as closely related to the excess time value for transit work trips. In effect, if the frequency of transit service is high it may be thought of as an indicator of higher transit work trip volumes; the distribution of persons as to age in the zone of origin (persons of an older age group may be thought to use the public transit system more frequently).

2.3 EQUATION POSSIBILITIES

The Kraft model as postulated in section 2.1 will be utilized in this section to develop equations which describe the volume of peak hour trips from the suburbs to the central business district for the auto and transit modes. The general equation will require slight alteration for this purpose. The two travel time components (T_i, j/M_o and T_i, j/M_α, representing mode M_o and mode(s) M_α) will be further subdivided (as described in section 2.2.1.) into a travel time ratio and excess time ratio for each mode. Also, the travel cost components (C_i, j/M_o and C_i, j/M_α for modes M_o and M_α) will be modified (as described in section 2.2.2.) to become cost of travel ratios for each mode.

These modifications to the general equation were instituted to relate the time and cost components of each mode more directly than is done in the general equation.

The ratios discussed above will be defined thus (see sections 2.2.1. and 2.2.2. for descriptions) travel time ratio (transit) - travel time by transit divided by travel time by auto.

travel time ratio (auto) - travel time by auto divided by travel time by auto.

excess time ratio (transit) - excess time by transit divided by excess time by auto.

excess time ratio (auto) - excess time by auto divided by excess time by transit.

cost of travel ratio (transit) - cost of travel by transit divided by cost of travel by auto.

cost of travel ratio (auto) - cost of travel by auto divided by cost of travel by transit.

2.3.1. The Transit Demand Equation

Based on the variables discussed in section 2.2 and the modified general demand equation, a relationship for the transit peak hour work trips originating in any suburban zone and destined to any central business district may be written:

$$W_T(i,j) = f(t_T, T_T, a, p_1, p_2, m_T, c, d, \ldots) \qquad (2)$$

where:

- $W_T(i,j)$ - volume of peak hour transit work trips originating in zone i (any suburban zone) and destined for zone j (the central business district).
- t_T - the travel time ratio (transit).
- T_T - the excess time ratio (transit).
- a - cars per family in zone i.
- p_1 - population density in zone i.
- p_2 - working population in zone i.
- m_T - cost of travel ratio (transit).
- c - frequency of transit service.
- d - age distribution in zone i.

2.3.2. The Auto Demand Equation

Analogous to the transit equation, an equation for the volume of peak hour work trips destined for any central business district and originating in any suburban zone may be expressed as:

$$W_A(i,j) = f(t_A, T_A, a, p_1, p_2, m_A, c, d, \ldots) \qquad (3)$$

where:

- $W_A(i,j)$ - volume of peak hour auto work trips originating in zone i (any suburban zone) and destined for zone j (the central business district).
- t_A - travel time ratio (auto).
- T_A - excess time ratio (auto).
- a - cars per family in zone i.
- p_1 - population density in zone i.
- p_2 - working population in zone i.
- m_A - cost of travel ratio (auto).
- c - frequency of transit service.
- d - age distribution in zone i.

2.3.3. Further Discussion

It should be noted that the above equations are merely a statement of a general situation. By no means is each variable applicable in every situation. The discussion as to whether or not a particular variable should be included in the expression or perhaps as to whether or not the variable should appear in an altered form (such as replacement of cars per family with the natural logarithm of cars per family, seeking a relative change due to this factor rather than an absolute change), depends upon the conditions extant in the city under study.

In other words, since each middle-sized city in Canada (and in North America, for that matter) has different existing transportation systems, different population distributions, different average income levels, etc., the particular equation which best predicts the volume of peak hour work trips from the suburbs to the central business district of any one city may not necessarily apply in another city.

Beginning with the basic demand equations postulated in sections 2.3.1. and 2.3.2., it is necessary to obtain, by investigation of the particular conditions in the city under study, the variables which will most accurately predict the required volumes. That is, the impact of each variable and its significance to the urban area being considered.

A test case which illustrates the impact of certain variables on a specific urban area is described in Chapter 3.

CHAPTER 3: TEST CASE: WINNIPEG

3.1 GENERAL DISCUSSION

This chapter will deal with the development of a pair of demand equations which will predict the peak hour work trip volumes, which originate in suburban areas and are destined for the central business district of the city of Winnipeg. This particular city was chosen as the test case for several reasons.

First, because of the author's close proximity to the study area. Second, because of the ready availability of all data which might be required for the equation development. Thirdly, because of its population (about 525,000), Winnipeg may be thought of as representative of an "average" sized Canadian city.

The peak hour chosen for Winnipeg was the A.M. peak hour. This choice was based on the fact that approximately forty percent of all daily trips were made in this time period[6] (7:30 A.M. to 8:30 A.M.), a much higher percentage than at any other time of the day.

It must be emphasized once again, that, despite any consideration of the "average" nature of Winnipeg, the equations of demand which will be developed in this section are for the specific conditions which exist in Winnipeg. This is not to say that similar equations will not apply to other Canadian urban centers, but a close investigation of the conditions which exist in these other cities should be completed before any decisions as to the validity of any equations developed in this chapter, are made.

3.2. DATA

The bulk of the data for the Winnipeg test case originated with the Winnipeg Area Transportation Study, an extensive study which included an origin and destination survey of all trip purposes for approximately 100% of the trips made.

3.2.1. Compilation and Assimilation

The transit data for the test case were obtained from the origin-destination survey conducted by the Winnipeg Area Transportation Study (as mentioned above). Because of the mass of survey forms collected, the aforementioned 100% of all trips, a twenty per cent sample of the A.M. peak hour transit suburb-to-CBD oriented trips, for the suburban areas chosen, was taken. The areas to be considered in the test case were divided into zones, which corresponded with the zones established for the Winnipeg Area Transportation Study, to facilitate analysis. These zones are illustrated on the map in Appendix I (the downtown destination zone is made up of four sub-zones which have been combined into one "super-zone" to represent the CBD for this test case).

The method of assimilation of the transit data was as follows: the origin point of each transit rider in a particular zone was located on a large scale map of the zone; the distance from point of origin to nearest appropriate transit route stop was then measured; assuming a walking rate of 1 mile in twenty minutes (source: Encyclopaedia Britannica)[16], it was then possible to calculate the time involved in walking to the particular bus stop; utilizing the time between buses or headway, the average amount of time spent in waiting for a bus was assumed to be half of the value of this headway up to a maximum time between buses of 10 minutes (or a waiting time of 5 minutes), beyond a 10 minute headway, it was assumed that the transit riders were familiar enough with the schedule of bus arrivals to wait only 5 minutes as opposed to one half the headway value. The amount of time spent on the vehicle itself was calculated from the distance between the point of boarding to the bus stop nearest the place of work and the average vehicle speed of 10.2 miles per hour*. Any transfer waiting time was assumed to be one half of the headway of the bus transferred to, regardless of the value of this headway. The transfer time did not consider a 5 minute maximum waiting period because of the fact that the schedules of interconnecting buses do not necessarily correspond and it would be difficult to time such interconnections to be not more than 10 minutes apart from the point of view of the transferee. The final time element, the walking time from bus stop to place of work, was calculated in a similar manner to that of the walking time from point of origin to

* This value was obtained from the records of the Winnipeg Transit Department.

nearest bus stop, with distance from bus stop to place of work substituting for distance from origin to bus stop. The only cost factor which was considered for the transit working trip was the actual fare ($.25) for the A.M. journey.

The data obtained for the transit work trips were, in summary: 20% of trip volumes from point of origin to place of work; the walking time from point of origin to nearest appropriate bus stop; the time spent waiting for the bus; the time spent on the bus; the transfer time, if applicable; the time spent walking from bus stop to place of work; and the cost to the transit user.

Much of the auto data utilized in this study were also obtained from the Winnipeg Area Transportation Study results. However, these results were in the form of matrices of origin and destination and, in the case of the vehicle travel time, represented an average value. Both the trip volumes and the travel times were for the approximate 100% sample. The other time element considered for auto trips was the walking time from parking lot to place of work. An average value of 6 minutes[(6)] was used for all work trips to the destination zones under consideration (the central business district). The cost of the auto work trip was broken down into operating costs and parking cost. The operating costs were further broken down into gasoline costs (an average of 12 miles per gallon at a rate of about .50¢ per gallon equals approximately $.042 per mile) and oil and lubrication costs (an average of 1000 miles per lubrication and oil change at a rate of about $5.00 equals $.005 per mile). These costs were calculated for an average distance from each zone of origin to each destination sub-zone. The average parking cost for the downtown destination area was found to be \$1.50 per day or \$.75 for the one way work trip considered.

In summary, the auto data obtained: 100% of trip volumes from origin to destination: average time spent in vehicle from each origin zone to each zone of destination; average time spent walking from the parking area to place of work; cost of the auto work trip.

Other data which may be necessary for the test case analysis included: population density in zone of origin, cars per family in zone of origin, age distribution in zone of origin, frequency of transit service, etc. Much of this data is obtainable from volume II of the Winnipeg Area Transportation Study.(6)

3.2.2. Assumptions

Several assumptions were made with regard to the data for the Winnipeg test case.

Since the volume of transit work trips collected from the survey forms represented a twenty percent sample, it was assumed that a multiplication factor of five could be used to obtain one hundred percent of the transit A.M. peak hour work trip passenger volumes from the selected suburban zones to the downtown sub-zones.

Since the exact location of each transit stop was not practically obtainable, upon discussion with Winnipeg Transit officials an interval of one quarter mile between transit stops was deemed as close to the actual stop intervals as was necessary for accuracy in obtaining walking time both to and from the stops.

With regard to potential transfer procedures, it was assumed that the person wishing to transfer did so at the first available point of interconnection between the appropriate transit routes, in order to facilitate the calculation of "in vehicle" transit travel time.

The transit work trip volumes in each zone of origin were averaged to be consistent with the auto trip data. Both sets of data were then averaged for the aforementioned amalgamated downtown sub-zones or "super-zone".

This then resulted in values which represent average time in vehicle (transit and auto travel time), average time out of vehicle (auto and transit excess times) and average user cost (transit and auto) along with population density, working population and cars per family for each suburban zone origin to the downtown "super-zone" (CBD).

3.3 DEVELOPMENT OF EQUATIONS

This section will deal with the development of a pair of equations to predict the A.M. work trip volumes by auto and transit from the suburban zones to the downtown or central business district of the Winnipeg test case.

The variables chosen for examination in the Winnipeg test case include (for each mode): travel time ratio, excess time ratio, cars per family in origin zone, working population in origin zone, population density in origin zone, cost of travel ratio.

It was decided that other variables previously mentioned, such as frequency of transit service and age distribution, were not as significant as the above variables, due to their repetitiveness (transit frequency is closely related to transfer and waiting time) or their lack of contribution to actual work trip considerations (most people over a certain age are retired and thus have no work trip as such).

3.3.1. Method of Solution

The method by which equations of the nature of those presented in sections 2.3.1. and 2.3.2., are solved most adequately is that of multiple regression analysis (together with the method of least squares).

Only a brief outline of the mathematical theory behind this method of solution will be presented in this section. No attempt at a more thorough coverage of this technique will be made in this paper. The reader is advised to consult the references in this section for a more detailed discussion of the topic.

An estimating regression function may be written as

$$Y = a + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + \ldots b_m(x_m - \bar{x}_m)$$

where:

Y - an estimate of the true expected value of the dependent variable, given the values of the m independent variable.

$x_1 \dots x_m$ - are the m independent variables.

$\bar{x}_1 \dots \bar{x}_m$ - are the means of the m independent variables.

$a, b_1 \dots b_m$ - are the estimated regression coefficients.

The primary purpose is to test the relationship between the dependent and independent variables, based on n sets of observed values (see references 4 and 5).

The method of least squares is used to estimate the regression coefficients, determined to that Z is minimized where

$$Z = \sum_{i=1}^{n} (y_i - Y_i)^2$$

and

y_i - the observed value of the dependent variable for the ith observation, which accompanies the observed values of $x_{1i}, x_{2i}, \dots x_{mi}$.

Y_i - the estimated value of the dependent variable (computed from the estimated regression) for the same values of $x_{1i}, x_{2i}, \dots x_{mi}$.

There are several computer programs available which utilize the multiple regression analysis procedure. The program chosen to analyze the test case data is entitled "Stepwise Multiple Regression Program (Stats 27)" and is available in the University of Manitoba Computer Center's Statistical Package.

Briefly, this program analyzes the relationship between the dependent variable and the independent variables to be considered.

It also selects, by "steps", the independent variables in order of "importance". The "importance" of the variable is based on the amount by which it reduces the sum of squares (the aforementioned Z value). Any variable may be deleted from the input file for purposes of testing the effect on the regression analysis of its deletion.

The procedure in this chapter, therefore, will be to utilize the data obtained for the Winnipeg area and the general equations postulated for auto and transit A.M. work trip demand (sections 2.3.1. and 2.3.2.) to attempt to develop demand equations for both modes, for the specific test case of Winnipeg.

3.3.2. Transit Demand Equation

The first stage in the development of the demand equation for transit involved the examination of several variables which might be used in the initial attempts to find the best fitting equation. It is to be noted at this point that all trial runs which led to the development of the final demand model for transit rider volumes utilized the same 20 suburban zones.

Consideration was first given to a purely linear equation for describing the volume of transit work trips. Also considered were a fully logarithmic relationship and several combinations of logarithmic and linear variables. Comparison of results of all types of equations were made and the "best fitting" equation chosen for testing. "Test runs" consisting of 15, 22, and 29 zones were done to examine the validity of the choice. The validity being established by a comparison of the results of the test runs with those of the equation chosen. The "success"of the comparison depends on certain statistical limitations. If these limitations are met, the chosen equation is considered the equation which most accurately predicts the volume of transit work trips.

Consideration was initially given to the relationship as postulated in section 2.3.1., equation (2), with deletions as noted in section 3.3.1.

The initial run of the "stepwise multiple regression analysis" procedure (see section 3.3.1.) produced a correlation matrix for the variables described. This matrix (Table 1) is an indicator of the intercorrelations between all variables both dependent and independent.

High intercorrelations between independent variables frequently results in a degeneration of the least squares regression procedure. The inclusion of both variables of a highly inter-correlated pair of independent variables should therefore be avoided.

High correlation between the dependent variable and any independent variable is an indicator of a significant relationship between same. (Note: a positive correlation coefficient means a direct relationship exists, a negative correlation coefficient means an inverse relationship between variables.)

TABLE I

Transit
Correlation Matrix Run 1

	1	2	3	4	5	6	7
1	1.00000	-.50587	-.17481	-.38875	.55835	.66741	-.00067
2	-.50587	1.00000	.54338	.04530	-.40046	-.19071	-.05241
3	-.17481	.54338	1.00000	-.04802	-.39266	.16381	-.60209
4	-.38875	.04530	-.04802	1.00000	-.47187	-.13804	-.02860
5	.55835	-.40046	-.39226	-.47187	1.00000	.09303	.50430
6	.66741	-.19701	.16381	-.13804	.09303	1.00000	-.47911
7	.00067	-.05241	-.60209	-.02860	.50430	-.47911	1.00000

LEGEND:

Variable 1 - Volume of trips (dependent)

Variable 2 - Excess time ratio (transit) (independent)

Variable 3 - Travel time ratio (transit) (independent)

Variable 4 - Cars per family (independent)

Variable 5 - Population density (independent)

Variable 6 - Working population (independent)

Variable 7 - Cost of travel ratio (transit) (independent)

Upon examination of Table 1, the matrix of correlations between the dependent variable (1) and each independent variable (2,3, 4,5,6,7) it was found that several dependent-independent pairs were correlated. These were: volume of trips with working population (.66741); volume of trips with population density (.55835); and volume of trips with excess time ratio (-.50587).

Also, there were several intercorrelated independent variables. These included: excess time ratio with travel time ratio (.54338); travel time ratio with cost of travel ratio (-.60209) and population density with cost of travel ratio (.50430).

Because of the previously mentioned degenerative effect of highly intercorrelated independent variables, runs 2 through 11 were an attempt to examine various combinations of variables while eliminating any possible degeneration due to high independent variable intercorrelations. A description of each run is as follows:

Run 2 - examined the effect of eliminating the travel time ratio from the input file.

Run 3 - elimination of the excess time ratio.

Run 4 - elimination of the cost of travel ratio.

Run 5 - elimination of population density.

Run 6 - elimination of excess time ratio and cost of travel ratio.

Run 7 - excess time ratio, travel time ratio and cost of travel ratio eliminated.

Run 8 - elimination of travel time ratio and cost of travel ratio.

Run 9 - elimination of population density and cost of travel ratio.

Run 10 - excess time ratio, travel time ratio and population density eliminated.

Run 11 - excess time ratio, travel time ratio, population density and cost of travel ratio eliminated.

A summary of the results of runs 1 through 11 is given in Table 2 (see page 27).

Examination of Table 2 indicates that the best combination of low standard error of estimate, high F - value and high multiple regression coefficient, R, appears to be run 8 (standard error = 26.474, F = 11.318, R = .839). This run eliminates the variables travel time ratio and cost of travel ratio.

The values of the regression coefficients, the standard error of the regression coefficients and the computed T* value for run 8 are reproduced, along with the intercept in Table 3 (page 29).

An examination of the computed T values of all four variables indicates that for a 95% confidence level, only variable 6, the working population, has a T value which falls outside the null hypothesis (that there is no correlation between the dependent variable and each independent variable) range of $T(15,.05), \pm 2.131$[8]. Therefore, since Table 3 represents the values of the linear variables which are the "best fit" for the analysis, it may be

*T is defined as the value utilized in the T - test for significance of a hypothesized observation.

TABLE 2

Transit
Summary of Results of Runs 1 - 11

Run	Variables Used	F*	Std. Error of Estimate	R*
1	6	7.561	28.761	.835
	5			
	2			
	4			
	7			
	3			
2	6	9.226	27.330	.840
	5			
	2			
	4			
	7			
3	6	7.006	30.315	.799
	5			
	3			
	4			
	7			
4	6	8.452	28.301	.828
	5			
	2			
	4			
	3			

* F is defined as : $\frac{\text{Explained (Regression) Variance}}{\text{Unexplained (Residual) Variance}}$

R is defined as : The multiple correlation coefficient

TABLE 2
(cont'd)

Run	Variables Used	F	Std. Error of Estimate	R
5	6	9.279	27.315	.840
	2			
	7			
	4			
	3			
6	6	9.303	28.444	.812
	5			
	3			
	4			
7	6	12.425	27.318	.815
	5			
	4			
8	6	11.318	26.474	.839
	5			
	2			
	4			
9	6	8.223	29.687	.793
	2			
	4			
	3			
10	6	9.730	29.662	.777
	7			
	4			
11	6	9.785	32.044	.714
	4			

TABLE 3

Transit
Results of Run 8

Variable Number	Regression Coefficient	Standard Error of Regression Coefficient	Computed T-Value
6	.02382	.00556	4.284
5	3.01648	1.45120	2.079
2	-30.06439	17.07196	-1.761
4	-52.56595	56.40726	- .932
	INTERCEPT	130.23848	

concluded that a strictly linear relationship between the dependent and independent variables does not exist for the variables chosen.

After consideration was given to the linear relationship, it was then proposed that a logarithmic, or perhaps a mixture of logarithmic and linear, relationship would more aptly describe the dependence of volume of transit work trips on the various independent variables.

Similar to the linear equation development, a series of runs were attempted as described below:

Run 12 - all natural log of variables.

Run 13 - natural log of excess time ratio, natural log of travel time ratio, the remainder linear.

Run 14 - natural log of travel time ratio, the remainder linear.

Run 15 - natural log of excess time ratio, the remainder linear.

The results of these runs are illustrated in Table 4, page 31.

Examination of the results of these runs indicates that run 15 has the best combination of low standard error of estimate, high F value and high multiple correlation coefficient, R (standard error = 24.753, F = 10.966, R = .881).

Therefore, run 15 was chosen for more detailed examination. The correlation coefficient matrix of the variables of run 15 is reproduced in Table 5, page 32.

Upon inspection of Table 5, it was found that correlation of fairly high magnitude existed between two dependent-independent pairs. They were: volume of work trips with working population (.78756); volume of work trips with population density (.55385).

Intercorrelations between the independent variables which were significant included: travel time ratio with ln excess time ratio (.54254); travel time ratio with cost-of-travel ratio (-.60209); population density with cost-of-travel ratio (.50430); and population density with cars per family (-.47187).

An examination of the step-wise regression procedure on the computer output revealed that the variables travel time ratio and population density reduced the sum of squares by 0.0% and 1.0% respectively. This coupled with the high intercorrelations of these two variables with at least two other independent variables, as

TABLE 4

Transit
Summary of Results of Runs 12 - 15

Run	Variables Used	F	Standard Error of Estimate	R
12	ln 6	9.490	.310	.865
	ln 2			
	ln 5			
	ln 7			
	ln 3			
	ln 4			
13	5	6.216	30.982	.806
	6			
	ln 2			
	4			
	7			
	ln 3			
14	6	8.947	26.907	.858
	5			
	2			
	7			
	4			
	ln 3			
15	6	10.966	24.753	.881
	ln 2			
	4			
	7			
	5			
	3			

TABLE 5

Transit
Correlation Matrix Run 15

	1	2	3	4	5	6	7
1	1.00000	-.47775	-.17481	-.38875	.55835	.78576	-.00067
2	-.47775	1.00000	.54254	.05033	-.39870	-.19885	-.06237
3	-.17481	.52454	1.00000	-.04802	-.39266	-.13853	-.60209
4	-.38875	.05033	-.04802	1.00000	-.47187	-.13996	-.02860
5	.55835	-.39870	-.39266	-.47187	1.00000	.44219	.50430
6	.78576	-.19885	-.13853	.13996	.44219	1.00000	.18198
7	-.00067	-.06237	-.60209	-.02860	.50430	.18198	1.00000

Variables identical to those of Table 1.

noted in the preceding paragraph, is adequate cause for the elimination of these two variables from further consideration.

From the standpoint of transportation planning, the fact that the travel time ratio has little or no significance with respect to the volume of transit work trips at first glance seems paradoxical. However, on delving deeper it may be stated that people do not mind spending more time on the vehicle when it is in motion; it is when they must increase their time outside of the vehicle (excess time) that the travel patterns may alter. In other words there may be certain "weights" on each of these times which reflect the travellers preference for in-motion time relative to out-of-vehicle time. This view reflects the philosophy of a number of transportation planners. With regard to population density, it appears that this variable has little significance with regard to the work trip volumes, perhaps because, with the recent advent of high-rise apartments, a high population density does not necessarily mean a lower income strata or fewer autos per family. Consequently, a high population density does not necessarily have any bearing whatsoever on the transit work trip volumes.

The next run, therefore, eliminated population density and travel time ratio from the regression analysis.

The results of this run (run 16) including the values of the regression coefficients, standard error of regression coefficient, computed T values, intercept, multiple correlation coefficient, F value, and standard error of estimate are shown in Table 6.

Upon inspection of the values of T computed, it was noted that all values fall outside the null hypothesis range for a 95% confidence level in which $T(15,.05)$, ± 2.131[8], except variable 7: the cost of travel ratio. This variable may be therefore eliminated.

TABLE 6

Transit
Results of Run 16

Variable Number	Regression Coefficient	Standard Error of Regression Coefficient	Computed T-Value
6	.03884	.00619	6.274
2	-62.21597	20.68925	-3.007
4	-104.27003	41.12988	-2.535
7	-751.98999	522.53979	-1.439

INTERCEPT 393.23438

Multiple Correlation Coefficient .890

F-Value 17.587

Standard Error of Estimate 22.248

An explanation of the elimination of the cost of travel ratio is a little difficult. It may be, however, that since the cost of travel ratio is an almost constant term (the numerator being a fixed transit fare), transit work trip riders do not directly consider the relative cost when chosing to ride public transit. This fact could be highly significant when the question of an alteration in the transit fee structure is considered. Perhaps the absolute cost of transit usage is of more significance, in any event, further study of the subject is warranted but beyond the scope of this analysis.

TABLE 7

Transit
Results of Run 17

Variable Number	Regression Coefficient	Standard Error of Regression Coefficient	Computed T-Value
6	.03479	.00780	4.463
2	-65.47389	21.30644	-3.073
4	-91.39723	39.17503	-2.333

INTERCEPT 170.69748

Multiple Correlation Coefficient .872

F-Value 19.473

Standard Error of Estimate 22.688

Run 17, therefore, eliminated the cost of travel ratio, as well as the aforementioned travel time ratio analysis. The results of this run are found in Table 7.

The results obtained from this run therefore represent the "best fit" for the partial logarithmic equation which relates transit work trip volumes to the variables selected for the Winnipeg test case analysis.

Since the results of attempting to develop an entirely linear equation were unsatisfactory and since the partially logarithmic equation succeeded in producing satisfactory mathematical results as well as reasonable transportation results, it was concluded that the partial logarithmic approach be accepted as the demand equation for the A.M. peak hour transit volumes of the Winnipeg area test case.

The equation may be written as:

$$W_T = 170.69748 - 65.47389 \ln T_T - 91.39723a + 4.03479p_2 \qquad (4)$$

where all variables are as described earlier in this section.

This equation was tested for analytical reliability in section 3.4.

3.3.3. Auto Demand Equation

The method of approach used in the auto work trips demand model for the Winnipeg test case was similar to the procedure used in developing the transit work trip demand model.

Again the linear equation was considered first, followed by a fully logarithmic and several mixed logarithmic and linear equations. Comparisons were made and the most satisfactory equation was chosen for testing. The testing procedure was identical to that used in the case of transit.

As stated in the previous paragraph, initial consideration was given to the relationship as previously stated in equation (3), section 2.3.2. (with deletions as described in section 3.3.1.).

The initial run of the regression procedure was examined for correlations between the dependent and independent variables and for intercorrelations between the dependent variables (run 18).

TABLE 8

Auto
Linear Correlation Matrix

	1	2	3	4	5	6	7
1	1.00000	-.05825	-.17421	-.00409	.30470	.76821	-.26470
2	-.05825	1.00000	.36343	-.22784	.38309	.05125	-.39834
3	-.17421	.36343	1.00000	.07626	.24360	-.29513	-.45773
4	-.00409	-.22784	.07626	1.00000	-.53393	-.32756	.43440
5	.30470	.38309	.24360	-.53393	1.00000	.34123	-.65951
6	.76821	.05125	-.29513	-.32756	.34123	1.00000	-.11138
7	-.26470	-.39834	-.45773	.43440	-.65951	-.11138	1.00000

LEGEND:

Variable 1 - volume of trips (dependent).

Variable 2 - excess time ratio (auto) (independent).

Variable 3 - travel time ratio (auto) (independent).

Variable 4 - cars per family (independent).

Variable 5 - population density (independent).

Variable 6 - working population (independent).

Variable 7 - cost of travel ratio (auto) (independent).

The correlation matrix for this initial run is reproduced in Table 8.

The most highly correlated independent-dependent pair was work trips and working population with a correlation of .76821. No other independent variable appeared to be highly correlated with the working trips.

Several pairs of independent variables were fairly highly correlated: cost of travel ratio and population density (-.65951); cars per family and population density (-.53393); cost of travel ratio and travel time ratio (-.45773).

Because of the degenerative effect of high intercorrelations between independent variables, both variables of a highly intercorrelated pair should not be included in the analysis.

Runs 19 through 26 were an attempt to examine various combinations of variables while eliminating any possible degeneration due to high independent variable intercorrelations:

Run 19 - examined the effect of eliminating population density from the input file.

Run 20 - elimination of travel time ratio.

Run 21 - elimination of cost of travel ratio.

Run 22 - elimination of cars per family.

Run 23 - elimination of travel time ratio and population density.

Run 24 - elimination of population density and cost of travel ratio.

Run 25 - elimination of travel time ratio and cost of travel ratio.

Run 26 - elimination of travel time ratio, population density and cost of travel ratio.

TABLE 9

Auto
Summary of Results of Runs 18 - 26

Run	Variables Used	F	Std. Error of Estimate	R
18	6	8.703	71.919	.854
	4			
	7			
	2			
	3			
	5			
19	6	11.098	67.311	.863
	4			
	7			
	2			
	3			
20	6	10.458	68.914	.856
	4			
	7			
	2			
	5			
21	6	6.798	80.995	.794
	4			
	5			
	2			
	3			

TABLE 9
(continued)

Run	Variables Used	F	Std. Error of Estimate	R
22	6	5.461	87.306	.755
	7			
	2			
	5			
	3			
23	6	13.904	64.651	.865
	4			
	7			
	2			
24	6	7.527	80.889	.778
	4			
	3			
	2			
25	6	9.909	75.780	.808
	4			
	5			
	2			
26	6	10.371	76.787	.788
	4			
	2			

Examination of this table indicates that run 23, which eliminates the variables travel time ratio and population density, is the linear equation of "best fit".

The values of the coefficients, the standard error of the regression coefficients and the computed T value for this run are reproduced in Table 10. The intercept point is also included.

TABLE 10

Auto
Results of "Best Fit" Run

Variable Number	Regression Coefficient	Standard Error of Regression Coefficient	Computed T-Value
6	.09485	.01376	6.895
4	375.03882	123.28131	3.042
7	-190.16187	63.45471	-2.997
2	-137.58734	102.79417	-1.338
		INTERCEPT 403.12329	

Examination of the computed T values indicates that for a 95% confidence level all values fall outside the null hypothesis range, which is: $T(15,.05)$, ± 2.131, except variable 2; the excess time ratio.

Therefore the actual best fit may be found by elimination of the excess time ratio. The results of this run can be found in Table 11.

TABLE 11

Auto
Results of Final Linear "Best Fit" Run

Variable Number	Regression Coefficient	Standard Error of Regression Coefficient	Computed T-Values
6	.09511	.01409	6.749
4	386.14990	126.00720	3.065
7	-161.22095	61.11493	-2.638

INTERCEPT 188.26660

Multiple Correlation Coefficient .857

F-Value 17.095

Standard Error of Estimate 64.254

It is therefore apparent that neither the travel time ratio (auto), nor the excess time ratio (auto) are significant in determining the number of auto working trips. The in-vehicle travel time lack of significance is analogous to the transit situation; people do not care how much time they spend actually in the vehicle. The auto situation then deviates from that of the transit since auto drivers also care little about the excess time they spend. This seems logical because the only excess time that one can consider for the automobile is the time it takes to walk from the parking lot to place of work, which is very frequently much less time than a person who rides the bus will spend walking to the bus, waiting for the bus, possibly transferring, and walking from the bus. The auto driver therefore sees one stretch of excess time coming at the end of his work trip to be of much less importance than the discontinuous series of excess time intervals which a transit rider must undergo.

The exclusion of the population density variable is a result of the fact that it is too closely related to other independent variables and also that the working population (by the high degree of correlation with the work trips) is a much better "population" indicator of auto work trips than is population density.

As in the development of the transit equation, a logarithmic and several mixed log-linear relationships were then investigated for the auto equation. A series of runs were attempted:

Run 27 - all logarithmic.

Run 28 - ln travel time ratio, ln excess time ratio.

Run 29 - ln travel time ratio.

Run 30 - ln excess time ratio.

The results of these runs are shown in Table 12.

An examination of the results illustrates the fact that none of the runs provided as good a "fit" as the linear run which utilized variables 6, 4, and 7. Further to this, it should be noted that in each run except for run 27, the first three variables in degree of importance were: 6, 4, and 7 respectively. Since an

TABLE 12

Auto
Summary of Results of Runs 27 - 30

Run	Variables Used	F	Std. Error of Estimate	R
27	ln 6	7.188	.492	.828
	ln 5			
	ln 4			
	ln 3			
	ln 7			
	ln 2			
28	6	8.103	73.990	.845
	4			
	7			
	ln 3			
	ln 2			
	5			
29	6	8.569	72.364	.852
	4			
	7			
	2			
	ln 3			
	5			
30	6	8.263	73.417	.847
	4			
	7			
	3			
	ln 2			
	5			

examination of the logarithmic and mixed log-linear runs as discussed previously did not produce a better fit than the initial run (run 18) and also since these runs (27 - 30) seemed to reinforce the final linear "best fit" choice (see above), it was concluded that the aforementioned final linear "best fit" (using variables 6, 4, and 7) should be chosen as the "basic" equation for the analysis. This equation may be written as:

$$W_A = 188.2660 + 386.14990\ a + .09511\ p_2 - 161.22095\ m_A \quad (5)$$

where all variables are as described earlier.

This equation along with equation (4), the transit work trip demand equation, was tested for analytical reliability in section 3.4.

3.4. TESTING OF DEMAND EQUATIONS

The equations of demand for auto A.M. peak hour work trips and transit peak hour work trips, for the Winnipeg test case, will now be examined for analytical reliability.

3.4.1. General Testing Procedure

Three types of test procedures will be considered to examine the reliability of the demand equations developed for the Winnipeg example. First, statistical inference tests: the F-value test of the variance, the confidence intervals for acceptability, and the T-test of acceptability. Secondly, "check" runs in which zones in the Winnipeg example are added or deleted to determine the stability of the parameters. The last test procedure will be an analysis of the residuals (the difference between the actual value of the work trip volume for a particular origin and the value predicted for this work trip volume by the demand equation), including a calculation of the root mean square error and the percent root mean square error for each residual, to determine the degree of accuracy of the estimated work trip volumes for the Winnipeg example.

3.4.1.1. Statistical inference tests

a) The F-Test[8]

This test is a measure which determines if the difference between a criterion, for two numerical distributions, are significant or simply a product of sampling error or random probability.

In other words:

$$F = \frac{\text{Explained (regression) variance}}{\text{Unexplained (residual) variance}}$$

for a confidence level of 95% (or an occurrence 95 times out of 100).

The minimum acceptable value for F will depend on the number of degrees of freedom, that is, the number of independent variables in the equation to be tested. These values of F are tabulated in all standard books of statistical tables.

If the F value obtained from the results of the chosen equation is larger than the minimum value established from the tables, the hypothesis that the explained variation is NOT SIGNIFICANT is rejected.

b) Calculation of Confidence Intervals (8)

The confidence interval, or the range of acceptability, for the population parameter being estimated by any regression coefficient β may be expressed as follows:

$$\text{UPPER AND LOWER LIMITS} = T(.05,\ n - \text{No. of Ind. Variables})\ (S.E.)$$

where

β = regression coefficient of some variable.

n - number of samples.

S.E. - standard error regression coefficient.

T(.05, n - No. of Ind. Var.) - value of T which corresponds to n-no. of individual variances degrees of freedom.

Therefore the confidence intervals for each equation in the Winnipeg test case may be determined enabling the acceptability of the equations to be examined during the "check" run procedure, since these limits represent the maximum and minimum values of the regression coefficients which are admissible for a confidence level of 95%.

c) The T-Test of Acceptability (7,8)

The T-test is an hypothesis test which is used to examine the significance of hypothesized observations.

In the case of the demand equations developed for the Winnipeg situation, a null hypothesis is postulated (for T-testing) that there is NO correlation between the dependent variable (volume of work trips by either mode) and the independent variables for each equation.

As mentioned previously (in section 3.4.1.1. (b) the value of T depends on the number of independent variables in the equation, for any chosen confidence level (in our case the confidence level is 95%).

3.4.1.2. "Check" runs

This reliability test is an examination of the stability of the parameters of the demand equation for the Winnipeg example.

Three "checks" were run for each demand equation, using the same computer program. The first check was an "updating", whereby nine new zones were added to the original twenty to test the effect of future updating and increasing the input data. The second "check" run was a "minor update", similar to the first "check" run except for the fact that only two new zones were added, a possible representation of a readjustment of the zones. The final "check" run was an attempt at observing the stability of the parameters of each demand equation with fewer than twenty zones (five zones being eliminated). This also might be a representation of a readjustment and/or amalgamation of zone.

The method of analyzing these "check" runs was to compare the variable means and regression coefficients of each "check" with the similar values obtained from the chosen demand equations. The confidence limits for the demand equations were then used to examine the stability and acceptability of the "check" runs as opposed to the demand equations.

3.4.1.3. Residual analysis

A statistical test of the estimation power of the regression equation is the calculation of the root mean square error and the percent root mean square error of the residuals.
The RMS error indicates the limits within which approximately 66.7% of the deviations between the observed value and the estimated value will fall. In other words, two thirds of the time, the estimated work trips will be within one RMS error of the observed values.

The percent RMS error is the percentage of the observed trips which is reprented by the RMS. A value of this percentage which is higher than 10 - 15% indicates a relatively poor degree of estimation for the particular zone considered.

3.4.2. Testing Transit Demand Equation

Utilizing the methods and techniques which have been illustrated in section 3.4.1., the demand equation for transit work trips in the Winnipeg example is now examined.

3.4.2.1. Statistical inference tests

a) F-Test

- for the transit demand equation chosen (see section 3.3.2.) the value of F was (see Table 7)

$$F = 19.473$$

From tabulated values of F(8) for 3 independent variables and 17 $\left(n - \frac{\text{no. of ind.}}{\text{variables}}\right)$ degrees of freedom, the minimum acceptable F value is

$$F(3,17) = 8.72$$

$\therefore$ the F value for the transit demand equation is within the range of acceptability.

b) Confidence Intervals

- utilizing the equation given in section 3.4.1.1.(b), the confidence limits for each regression coefficient of each input variable of the transit demand equation at a 95% confidence level are (with $T\left(n - \frac{\text{no. of ind.}}{\text{variables}}\right) = T(3,.05) = \pm 2.131$) :

Working Population

UPPER AND LOWER LIMITS = .03479 ± .01660

UPPER LIMIT = .05139

LOWER LIMIT = .01819

Excess Time Ratio (natural log)

UPPER AND LOWER LIMITS = -65.47389 ± 45.40000

UPPER LIMIT = -20.07389

LOWER LIMIT = -110.87389

Cars per Family

UPPER AND LOWER LIMITS = -91.39723 ± 83.35000

UPPER LIMIT = -8.04723

LOWER LIMIT = -174.74723

These limits will be referred to again when considering the check runs (section 3.4.2.2.).

c) The T-Test

In the case of the transit demand equation for the Winnipeg situation, the null hypothesis is that there is no correlation between the dependent variable (the volume of A.M. peak hour transit work trips) and the dependent variables (ln excess time ratio, cars per family and working population).

As mentioned in section 3.4.2.1.(b) above, the T value for 3 independent variables is ±2.131, for 95% confidence level. Bearing this in mind, an examination of the computed T values for the transit demand equation (see Table 7), reveals that all three T values fall outside the region of acceptance of the null hypothesis. Therefore the null hypothesis is rejected.

3.4.2.2. "Check" runs transit

As outlined in section 3.4.1.2., three check runs were done on the transit demand equation for Winnipeg: a 29 zone "update", a 22 zone "minor update" and a 15 zone "amalgamation and/or elimination". The results of these runs and their comparison with the original 20 zone equation follow.

(a) 29 zone check run results:

A comparison of the values of the variable means is illustrated in Table 13 (a), on the following page.

There is, as illustrated, very little effect on the means of variation in the sample.

TABLE 13(a)

Transit
Comparison of Means

Variable	Mean (20)	Mean (29)
6	1192.74487	1366.68604
ln 2	.42200	.45069
4	1.11400	1.13482

Table 13 (b) represents a comparison of the regression coefficients of the 29 zone run with the regression coefficients of the chosen (20 zone) equation. The variations noted were well within the previously established 95% confidence limits for variables 6, ln 2, and 4 (see section 3.4.2.1.(b)). No undue instability was noted in any of the coefficients.

TABLE 13(b)

Transit
Comparison of Regression Coefficients

Variable	Regression Coefficient (20)	Regression Coefficient (29)
6	.03479	.03926
ln 2	-65.47389	-54.37518
4	-91.39723	-81.91487

(b) 22 zone check run results:

produced little variation in the mean values and regression coefficients well within the 95% confidence limits established in section 3.4.2.1.(b).

(c) 15 zone "check" run results:

A comparison of the means of the variable is illustrated in Table 14(a).

TABLE 14(a)

Transit
Comparison of Means

Variable	Means (20)	Means (22)
Working Population	1192.74487	1277.45972
In Excess Time Ratio	.42200	.44580
Cars per Family	1.11400	1.13600

Once again there is little effect on the means of the variables if the sample is altered.

Table 14(b) is a comparison of the regression coefficients (see page 54).

The values obtained for the 15 zone check run are also well within the 95% confidence limits.

Table 15 is a summary of the results of all the check runs and the demand equation as well.

TABLE 14(b)

Transit
Comparison of Regression Coefficients

Variable	Regression Coefficient (20)	Regression Coefficient (15)
Working Population	.03479	.03829
In Excess Time Ratio	-65.47389	-74.36842
Cars per Family	-91.39723	-107.12245

TABLE 15

Transit
Summary of Results

Variable	Regression Coefficients			
	20 zone*	29 zone	22 zone	15 zone
Working Population	.03479	.03926	.03969	.03829
In Excess Time Ratio	-65.47389	-54.37518	-69.98245	-74.36842
Cars per Family	-91.39723	-81.91487	-95.72601	-107.12245

-Table 15 continued on following page

* this is the demand equation for transit as previously developed.

TABLE 15 continued

Variable	Means			
	20 zone*	29 zone	22 zone	15 zone
Working Population	1192.74487	1366.68604	1195.58618	1277.45972
In Excess Time Ratio	.42200	.45069	.46377	.44580
Cars per Family	1.14000	1.13482	1.12954	1.13600

* this is the demand equation for transit as previously developed.

As shown in this Table the parameters established for the demand equation of transit work trips are quite stable when subjected to an increase or decrease in the number of origin zones, for the sample case of Winnipeg.

<u>3.4.2.3. Residual analysis</u>

With reference to section 3.4.1.3., the root mean square error (RMS) and the percent root mean square error (%RMS) were now calculated for the observed and estimated A.M. peak hour transit work trips from the demand equation. These values are illustrated in Table 16.

The high %RMS values obtained for case 1 (35.00%), case 14 (16.35%), and case 21 (18.75%) are quite possibly due to the low trip volumes in these zones (10, 15, and 30 A.M. peak hour transit trips respectively). It is possible to eliminate these low volume zones

TABLE 16

Transit
RMS and %RMS of Residuals

Case	Y Value (volume)	Y Est. (Est. Vol)	Residual	RMS	%RMS
1	10	25.4	-15.4	3.5	35.0
2	140	139.8	.2	.1	.1
3	80	82.6	-2.6	.5	.6
4	35	51.5	-16.5	3.1	8.6
5	155	113.7	41.3	7.7	5.0
6	130	169.7	-39.7	7.4	5.7
7	45	46.5	-1.5	.9	1.9
8	85	70.3	14.7	2.7	3.2
9*	170	149.2	20.8	3.9	2.3
10	60	72.7	-12.7	2.7	3.9
11	105	52.0	53.0	9.8	9.4
12	110	121.3	-11.3	2.1	1.9
13	60	63.2	-3.2	.6	1.0
14	15	28.2	-13.2	2.5	16.4
15	75	79.7	-4.7	.9	1.2
16	55	75.7	-20.7	3.8	7.0
17	140	152.3	-12.3	2.3	1.6
18	65	69.8	-4.8	.9	1.4
19	75	82.3	-7.3	1.4	1.8
20	85	60.5	24.5	4.6	5.4
21	30	60.0	-30.0	5.6	18.8
22	170	172.4	-2.4	.5	.3
23	120	94.8	25.2	4.7	3.9
24	60	74.7	-14.7	.9	1.4
25	140	106.6	33.4	6.2	4.4
26	30	32.7	-2.7	.5	1.7
27	135	137.1	-2.1	.4	.3
28	85	74.7	10.3	1.9	2.2
29	40	45.8	-5.8	1.1	2.7

* Cases 1 - 9 were used as "check run" zones.

from the analysis without affecting the accuracy of the work trip demand equation (similar to the 15 zone situation in section 3.4.2.2.(c)), since such a small number of trips will have little effect on the predictive quality of the demand equation.

3.4.3. Testing the Auto Demand Equation

In a similar manner to section 3.4.2., the auto demand equation for the Winnipeg test case will now be examined.

3.4.3.1. Statistical inference tests

a) F-Test

The tabulated value of F is the same as that of the transit demand equation (see section 3.4.2.1.(a)), that is

$$F(3,17) = 8.72$$

The value of F obtained from the results of the auto work trip demand equation for the Winnipeg sample is

$$F = 17.095$$

∴ the auto demand equation value of F is within the range of acceptability.

b) Confidence Intervals

- similar to the transit equation, utilizing the method of section 3.4.1.1.(b), the regression coefficient confidence limits for each variable of the auto demand equation at a 95% confidence level are ($T = \pm 2.131$ again):

Working Population

UPPER AND LOWER LIMITS = .09511 ± .03000

UPPER LIMIT = .12511

LOWER LIMIT = .06511

<u>Cars per Family</u>

UPPER AND LOWER LIMITS = 386.14990 ± 268.50000

UPPER LIMIT = 654.64990

LOWER LIMIT = 117.64990

<u>Cost of Travel Ratio</u>

UPPER AND LOWER LIMITS = -161.22095 ± 130.15000

UPPER LIMIT = -31.07095

LOWER LIMIT = -291.37095

The above limits will again be referred to when considering the check runs (section 3.4.3.2.).

c) The T-Test

For the situation of the auto demand equation for the Winnipeg example, the null hypothesis can be stated: there is <u>no</u> correlation between the dependent variable (volume of A.M. peak hour auto trips) and the independent variables (working population, cars per family and cost of travel ratio).

Once again, the T-value for 3 independent variables is 2.131 for 95% confidence level. As illustrated in Table 11, the T values for the auto demand equation are outside the acceptability range of the null hypothesis. This then results in a <u>rejection</u> of the null hypothesis.

<u>3.4.3.2. "Check" runs auto</u>

Utilizing the same approach as in the transit check runs (see section 3.4.2.2.), a 29 zone "update", a 22 zone "minor update" and a decrease to 15 zones were all run for the auto demand equation for the Winnipeg sample.

a) 29 Zones

A comparison of the means of run of 29 zones with basic run was made

TABLE 17(a)

Comparison of 20 - 29 Means

Variable	Mean (20)	Mean (29)
6	1521.54980	1497.00000
4	1.15425	1.13362
7	3.71349	3.67034

Slight variation was observed for the 20 versus 29 zone means.

An examination of regression coefficients was now made.

TABLE 17(b)

Regression Coefficients 20 - 29

Variable	Regression Coefficient (20)	Regression Coefficient (29)
6	.09511	.09749
4	386.14990	349.21606
7	-161.22095	-158.38950

The variations were well within the previously established 95% confidence limits for variables 6, 4, and 7 (see section 3.4.3.1.(b)).

b) 22 Zones

There was little variation observed in the means and the variations in the regression coefficient were well within the 95% confidence limits.

c) 15 Zones

Comparison of means:

TABLE 18(a)

Comparison of 20 - 15 Means

Variable	Means (20)	Means (15)
6	1521.54980	1618.33325
4	1.15425	1.16333
7	3.71349	3.76866

Small variation in mean of variable 6 but not of significant magnitude for any concern.

The comparison of regression coefficients is on the following page.

TABLE 18(b)

Regression Coefficients 20 - 15

Variable	Regression Coefficient (20)	Regression Coefficient (15)
6	.09511	.09080
4	386.14990	419.98389
7	-161.22095	-208.54587

All variations again were well within 95% confidence limits.

Summarizing all test results:

TABLE 19

Auto
Summary of Test Results

Variable	Regression Coefficients			
	20 zone*	29 zone	22 zone	15 zone
6	.09511	.09749	.09334	.09080
4	386.14960	349.21606	321.75757	419.98389
7	-161.22095	-158.38950	-141.76591	-208.54587

Variable	Means			
	20 zone*	29 zone	22 zone	15 zone
6	1521.54980	1497.00000	1520.95435	1618.33325
4	1.15425	1.13362	1.15204	1.16333
7	3.71349	3.67034	3.69818	3.76866

* This is the demand equation for transit as previously established.

As shown in the above Table the auto work trip demand equation parameters established show reasonable stability when the number of origin zones is altered, for the Winnipeg test case.

3.4.3.3. Residual analysis

Calculation of RMS and %RMS are shown in Table 21 on the following page (for explanation see section 3.4.1.3.).

The "border line of acceptability" values of %RMS obtained in cases 3 (16.82), 10 (17.08), 13 (14.20), and 15 (13.64) cannot be explained by any quantitative means. Suffice to say that a far larger number of zones were well within the acceptable limits of about 10% than those which fell outside this figure. Therefore the predictive ability of the equation has not been drastically affected.

3.5 SENSITIVITY ANALYSIS

This section will deal with a further examination of the predictive ability of the auto and transit demand equations developed for the Winnipeg test case. This examination consists of altering any one of the chosen independent variables (while keeping the others constant) to gauge the effect of this alteration on the volume of A.M. peak hour work trips.

All three variables from each demand equation were altered and the results were plotted to obtain a "demand curve" for each particular zone of origin. The same five zones were plotted

TABLE 20

Auto
RMS and %RMS of Residuals

Case	Y Value (Volume)	Y Est. (Est. Vol.)	Residuals	RMS	%RMS
1	200	208.7	-8.7	5.1	2.7
2	190	143.8	46.2	8.6	4.5
3	55	94.8	-39.8	7.4	13.4
4	70	68.0	2.0	.4	.6
5	280	234.8	45.2	7.8	2.8
6	330	297.7	32.3	6.0	1.8
7	60	92.1	-32.1	5.9	9.9
8	175	250.3	-75.3	14.0	8.0
9*	160	143.5	16.5	3.1	1.9
10	160	169.4	-9.4	3.0	1.9
11	250	294.6	-44.6	8.3	3.3
12	85	140.0	-55.0	10.2	12.0
13	110	152.8	-42.8	8.0	7.2
14	80	81.8	-1.8	.3	.4
15	450	320.1	129.9	24.1	5.4
16	125	131.3	-6.3	1.2	.9
17	100	117.5	-17.5	1.0	1.0
18	350	291.0	59.0	11.0	3.2
19	60	108.6	-48.6	8.9	15.0
20	370	312.1	57.9	10.8	2.9
21	300	304.2	-4.2	.8	.3
22	20	25.6	-5.6	1.1	5.3
23	95	38.9	56.1	10.8	11.4
24	155	100.9	54.1	10.1	6.5
25	115	116.9	-1.9	1.1	1.0
26	180	112.4	67.6	12.6	7.0
27	200	254.0	-54.0	10.0	5.0
28	245	357.9	-112.9	21.0	8.6
29	150	156.4	-6.4	1.2	7.9

* Cases 1 - 9 were used as "check run" zones. None of the values calculated for the %RMS fell beyond the 15% range of acceptability. This indicates a reasonably good predictive ability for the auto demand equation.

for each alteration of variables in order to illustrate the zonal trend without a loss of clarity due to attempting to crowd too many zones onto a single graph (these 5 zones were chosen at random from the sample).

The independent variables were altered in the transit demand equation as follows. The excess time ratio was altered by decreasing the transit excess times by 3, 4, and 5 minutes and by increasing it by 3 and 5 minutes. The results of these alterations are shown in Graph 1. Alteration of the cars per family value was done by reducing the value by .1, .2, and .3 and increasing it by .2 and .3. These results are shown in Graph 2. The working population was increased by 300, 400, and 500 and decreased by 400 and 500. Results of these alterations are illustrated in Graph 3.

In the case of the auto demand equation, the cost of travel ratio and the cars per family were altered to examine the volume changes which would result. The cost of travel ratio with respect to auto was increased by .5 and 1.00 and decreased by .5, 1.00, and 1.50. The demand lines which resulted are illustrated in Graph 4. The cars per family values were altered in the same way as in the transit situation and these results are to be seen in Graph 5. Also the working population values were altered as in the transit case, the results are shown in Graph 6.

All of the volume trends which were established in the previously mentioned six graphs represent future predictions resulting from the alteration of a variable in the demand equation for transit or auto for the Winnipeg test case.

The graphs illustrate the stability of the demand equation variables when subject to alteration, with regard to prediction of volumes resulting from these alterations.

Figure 1 represents a comparison of average auto and transit work trip volume changes when the variables common to both demand equations (cars per family and working population in the zone of origin) are altered as described previously, with the alterations represented as a per cent of the average value.*

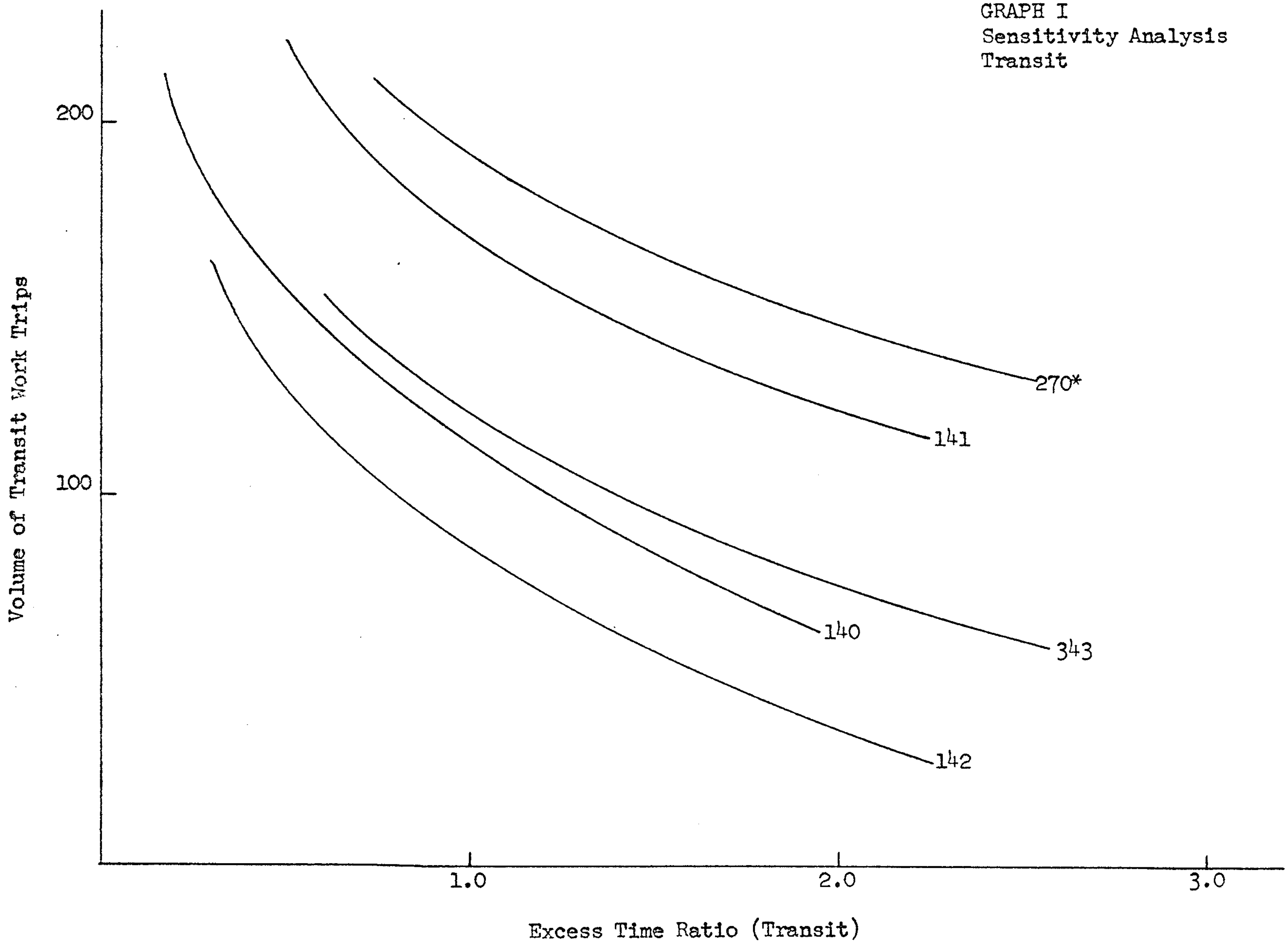
It appears that, with regard to the variable cars per family, a larger increase in auto trips occurs than the corresponding decrease in transit trips. Similar results are evident with a decrease in auto trips and the corresponding increase in transit trips.

With regard to the origin zone working population, it appears that although both trip volumes are increased when this variable increases and decreases correspondingly, the magnitudes of these changes differ.

Further discussion of the common variables will follow in the next section.

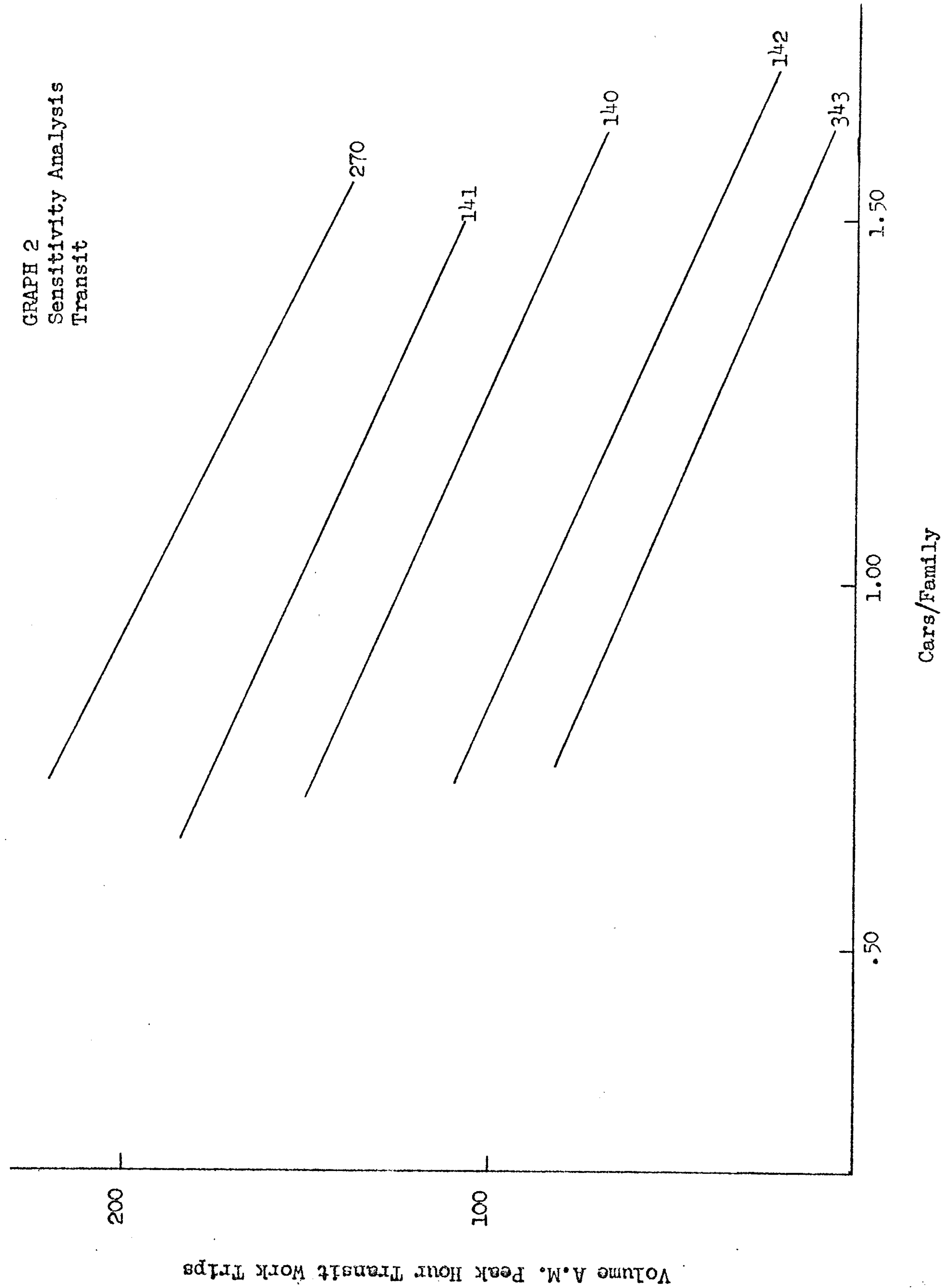
* The "base" line referred to in the figure is defined as the average estimated volume in the unaltered demand equation for either mode.

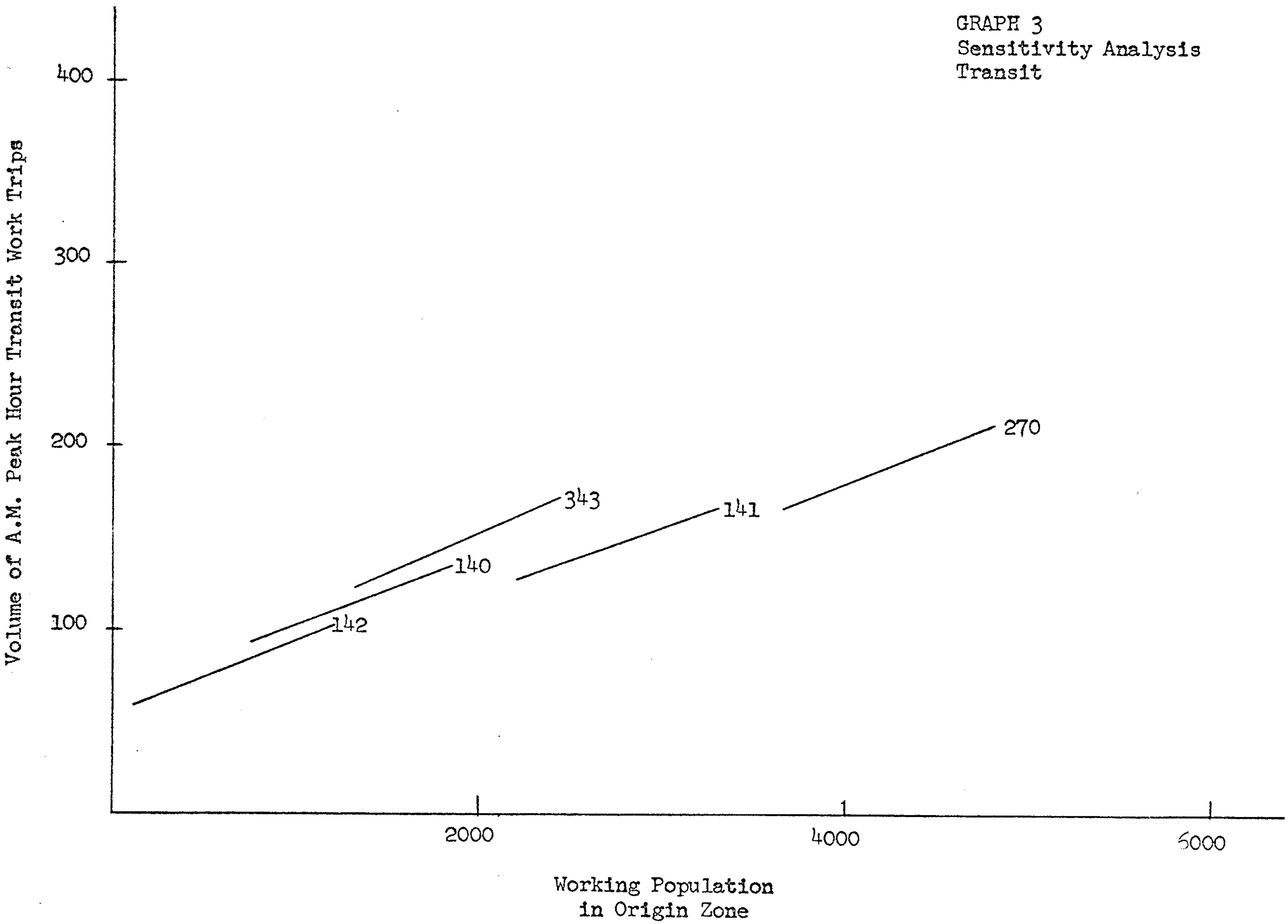
GRAPH I
Sensitivity Analysis
Transit



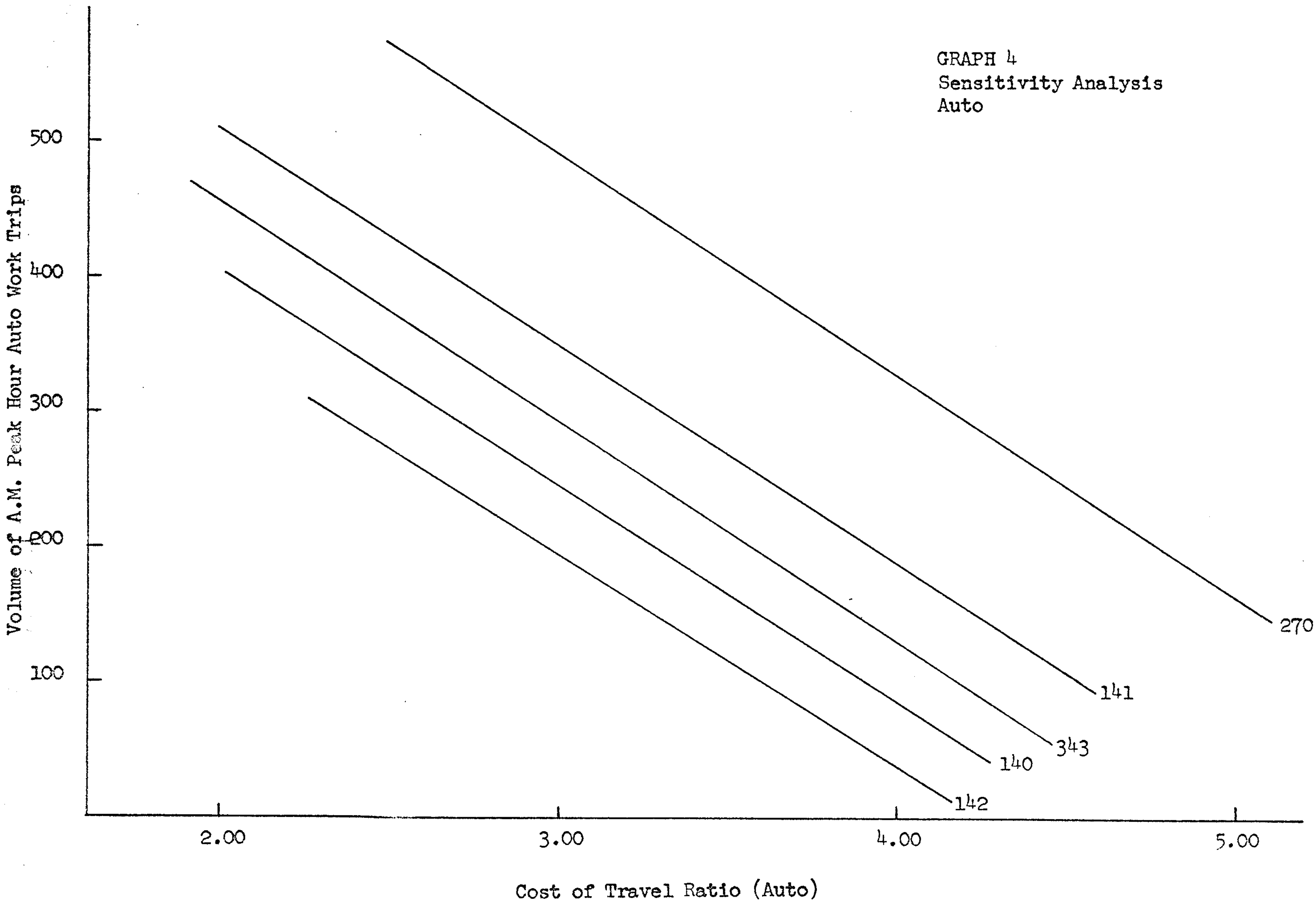
*Numbers refer to origin zones (see map Appendix I).

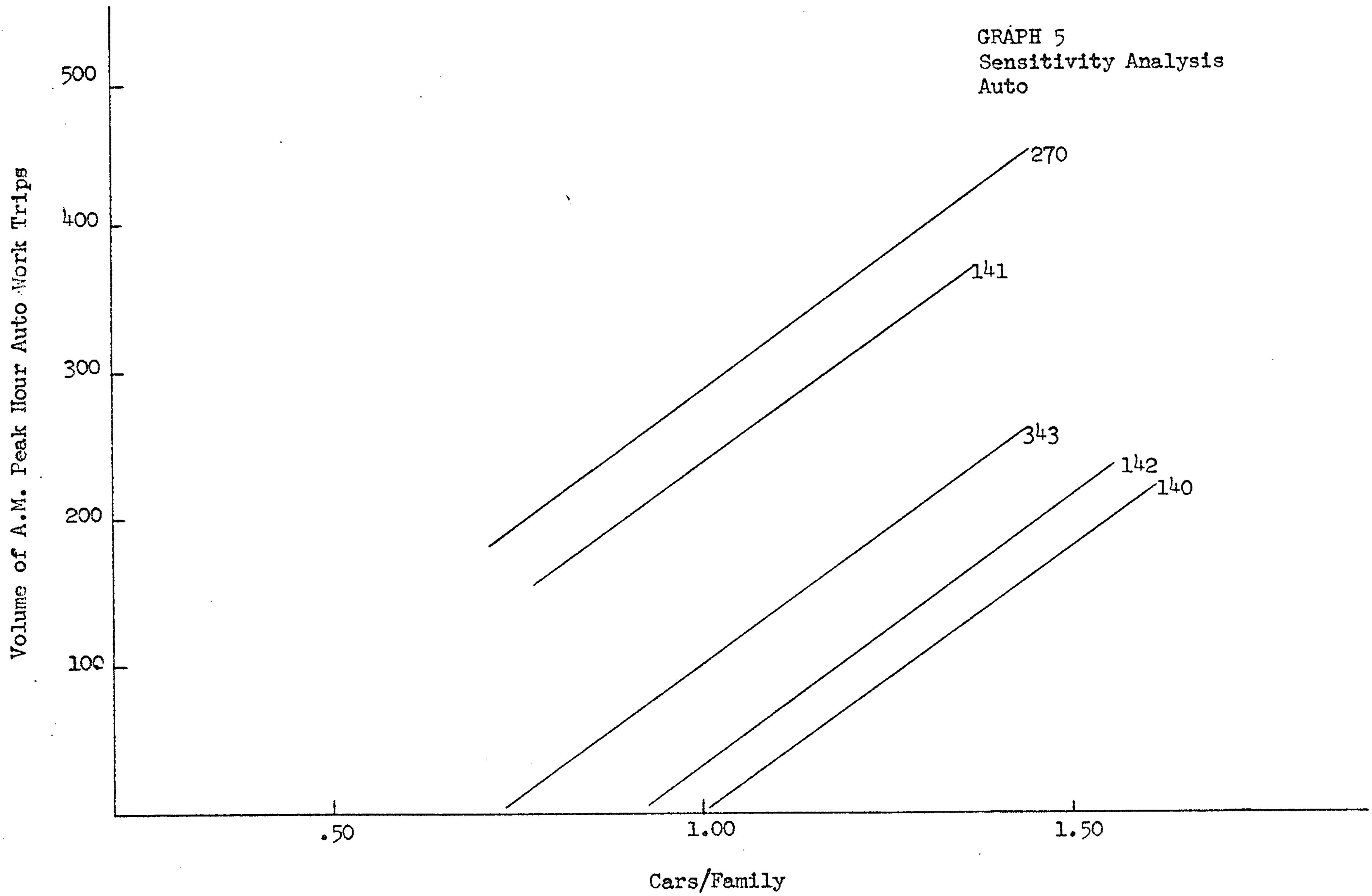
GRAPH 2
Sensitivity Analysis
Transit




GRAPH 3
Sensitivity Analysis
Transit
Volume of A.M. Peak Hour Transit Work Trips
400
300
200
100
343
140
142
141
270
2000
4000
6000
Working Population
in Origin Zone

GRAPH 4
Sensitivity Analysis
Auto





GRAPH 5
Sensitivity Analysis
Auto

GRAPH 6
Sensitivity Analysis
Auto

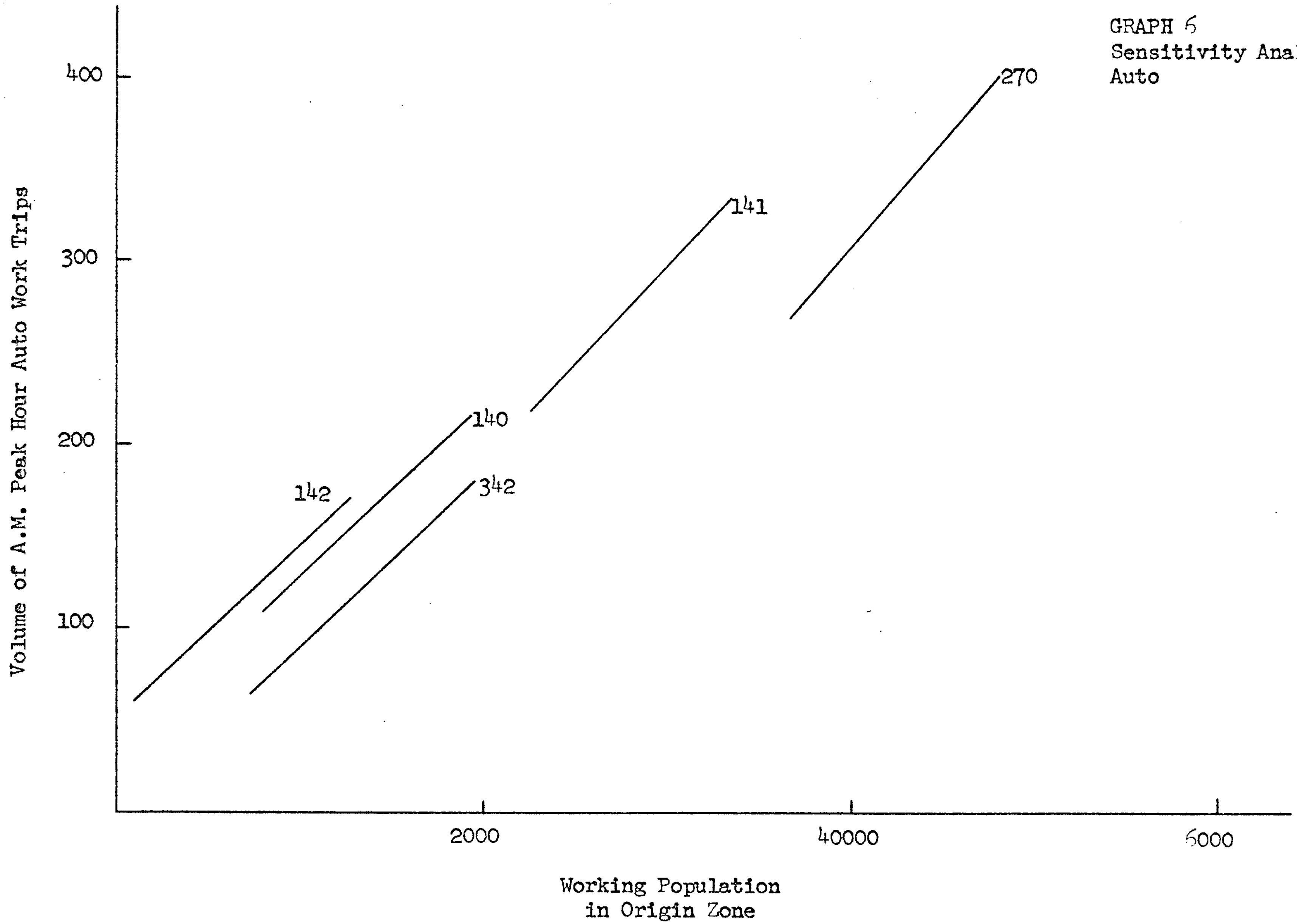
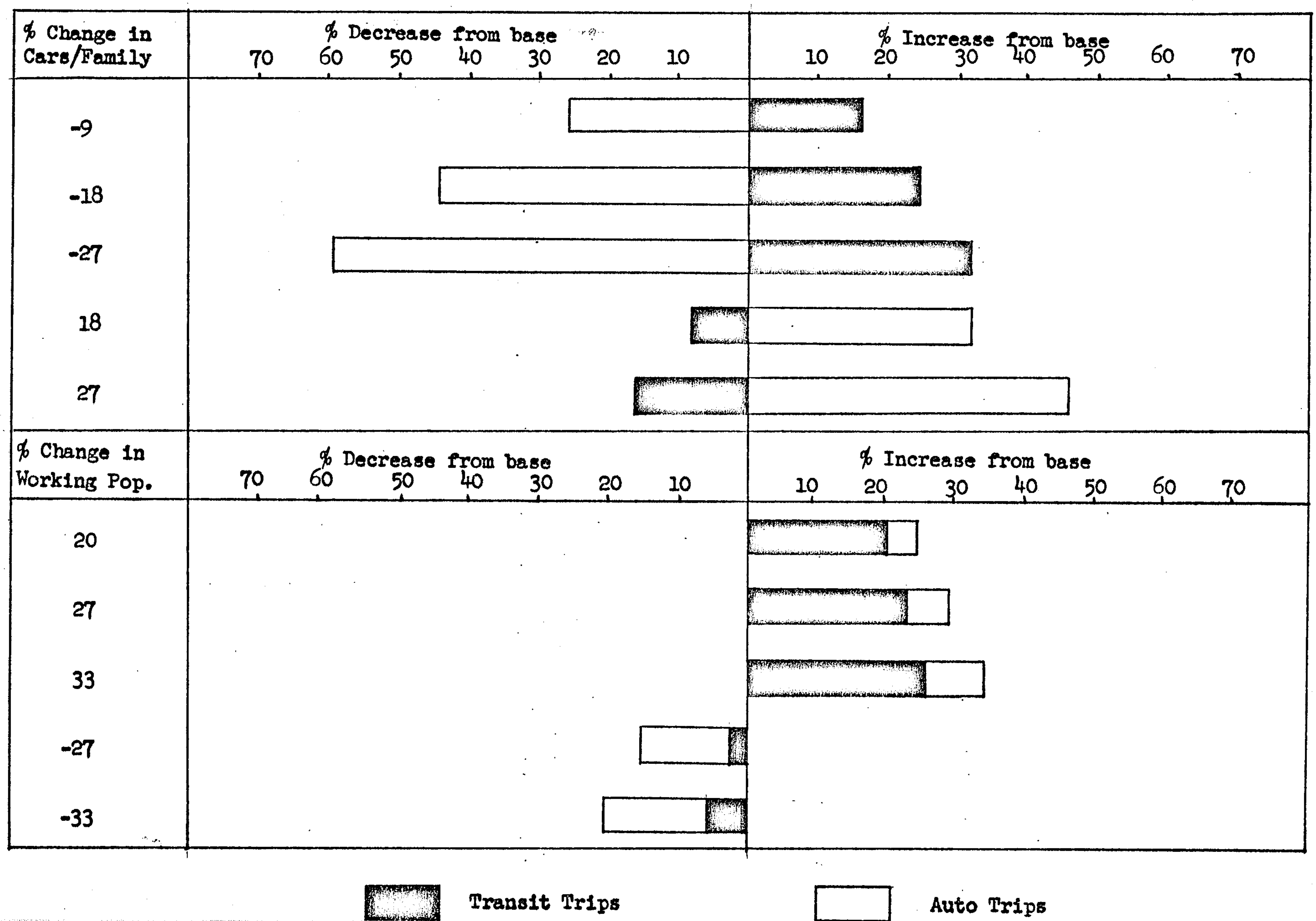


FIGURE I

Effect of Changes in Cars/Family and Working Population on Transit and Auto Trips



3.6 DISCUSSION OF RESULTS

To reiterate, the equations which best describe the demand for auto and transit A.M. peak hour work trips for the Winnipeg test case are

Transit:

$$W_T = 170.697 - 65.474 \ln t_T - 91.397\,a + .035 p_2$$

where, as before,

W_T - volume of A.M. peak hour transit work trips.

t_T - excess time ratio (transit).

a - cars per family in zone of origin.

p_2 - working population in zone of origin.

Auto:

$$W_A = 188.267 + 361.150\,a + .095 p_2 - 161.221\,m_A$$

where, as before,

W_A - volume of A.M. peak hour auto work trips.

a - cars per family in zone of origin.

p_2 - working population in zone of origin.

m_A - cost of travel ratio (auto).

Upon examination, it can be seen that two of the independent variables in both equations are the same: cars per family in the zone of origin (a) and working population in the zone of origin (p_2). However, in the case of cars per family, the sign is reversed. This may be thought to be consistent with the theory that an increase in the car ownership per family will cause a simultaneous increase in auto work trips and vice versa. However, if a particular origin zone is considered and there is an increase in car ownership per family, the decrease in transit riders is much less than the increase in auto trips according to the regression coefficients (-91.397 for transit equation versus 361.150 for auto equation). This suggests that not only is there a shift from transit to auto, but also possible shifts from car pools to individual drivers and/or, walking to driving for the Winnipeg sample. Further to this, the increase in auto trips may be because of the increased availability of that mode due to the increase in the cars per family ratio.

It is quite reasonable to assume that the origin zone working population variable will affect both the transit and auto A.M. peak hour working trips in the same manner. That is, by increasing or decreasing both auto and transit work trips with a corresponding increase or decrease in working population in the zone of origin. It should be noted, however, that an increase in origin zone working population will result in an effect on auto A.M. peak hour working trips which is almost three times as great as the corresponding effect on transit work trips (transit coefficient .035, auto coefficient .095) for the A.M. peak hour period, assuming all other variables are held constant. In other words, an increase in working population in a particular zone, caused perhaps by an influx of new families into a suburban sub-division for example, will produce almost three times the volume of auto trips as transit trips for the A.M. peak hour. For the particular case of Winnipeg, this situation is not too unreasonable when consideration is given to the fact that most suburban Winnipeg zones which may have the possibility of an increase in working population will accomplish this feat by an outward expansion, away from the central business district, which is the work place destination considered in this study. Usually the transit service to newer areas is not immediately available on a reliable basis. Therefore, it is to be expected that more people will rely on the private auto for their A.M. peak hour work trip than on the public transit system, even though the use of both modes would increase.

The third independent variable included in the transit equation is the natural log of the excess time ratio with respect to transit. Being a logarithmic function, this suggests that a relative change in the ratio of the excess time by transit to the excess time by auto will have a greater effect on the A.M. peak hour work trip volumes by transit than an absolute change, for the particular case of Winnipeg. Also, the effect of this relative change, if it is an increase, is to diminish the number of transit work trips. This seems to suggest that an increase in transit ridership would result if the excess time encountered in taking public transit were reduced. As shown in section 3.5, this trend to an increase in transit riders with a decrease in excess time ratio with respect to transit is borne out for the transit demand equation for the Winnipeg test case. Also the decrease in ridership appears to have a tendency to level off at a certain excess time ratio (which is a different value for each zone) and subsequently remain relatively constant (see graph 1, section 3.5.).

The third independent variable to appear in the auto demand equation is the cost of travel ratio with respect to auto. The negative sign suggests that any increase in this ratio will result in decreased auto A.M. peak hour work trips to the central business district of Winnipeg. It is significant that, although the cost of travel ratio with respect to transit was not a factor in the transit demand equation (discussed in section 3.3.2.) due, in part, to the constancy of the ratio, the cost of travel ratio with respect to auto was significant in the auto demand equation. It may be partly due to the greater variation in the cost of travel ratio with respect to auto from one origin zone to another, as compared to the very slight variation in the transit cost of travel ratio. This is the result of the fact that the denominator is now the constant (bus fare) value and the numerator is the out-of-pocket cost of auto operation plus the parking cost. It creates a value which is much larger in magnitude than the bus ratio and much less constant (the variation in the auto cost of travel ratio is from 3.52 to 4.33 while the transit ratio varies only from .23 to .28). Therefore, when applied to a relationship with volume of auto trips, the auto cost of travel will produce a more significant dispersion than a set of nearconstant values as in the case of the transit cost of travel ratio. Another possible explanation for the significance of the auto cost of travel ratio is the magnitude of the values. In no origin zone is the cost of operation and parking of an automobile less than 3.52 times as much as the cost of utilizing the public transit system. This could be of some importance regarding a discussion of choice of work trip mode.

The constant terms in each demand equation (170.697 for transit and 188.267 for auto), being of relatively large magnitude, require some explanation. The physical meaning of the constant term in the regression equation is that it represents the point of interception of the plane of the regression equation (in this case the hyperplane) with the plane of the Y axis (cartesian coordinate system). From the standpoint of the particular demand equations developed in this study, these constant values serve as a balance or check to the other terms of the equation, bringing the values estimated by the regression equation more in line with the actual observed values. In other words, enabling the number of A.M. peak hour work trips for Winnipeg to be estimated more precisely.

CHAPTER 4

CONCLUSIONS AND RECOMMENDATIONS

The avowed purpose of this study was, as stated in the introduction, to develop a mathematical demand model which will predict the volume of peak hour work trips which originate in suburban areas of any middle-sized city and are destined for the central business district of that city, for the transit and private auto modes of travel.

Such models were presented first in general form, then in a more specific form as multilinear equations with several independent variables. The latter were examined for the specific demand equations for auto and for transit A.M. peak hour work trips for Winnipeg were developed. These equations were able to predict the future travel demand by each mode utilizing the variables deemed significant for the Winnipeg study area. It must be emphasized that, for any other city, such demand equations will most likely be slightly different because of the different conditions extant in each Canadian city. However, the basic approach presented in this study is valid for any middle-sized Canadian city. Only the particular variables which appear in the demand equations may vary slightly from city to city.

The demand equations developed in this study may be used as an indicator of future travel demand in middle-sized Canadian cities. Such concerns to the urban milieu as improvement of public transit might possibly be answered by the use of the transit demand equation. It would enable a transit system analyst to gauge the effect on transit passenger peak hour volumes (to the CBD from suburban areas) of altering the transit fare structure, or decreasing the amount of excess time by improving the transit headways, or increasing the average speed of the transit system vehicles.

The auto trip and its future with regard to movement from suburb to CBD might also be more fully explored by use of the auto demand equation. The effect of placing certain constraints on the private auto trip, such as limiting parking space availability in the CBD or restricting the use of private vehicles in the downtown area might be examined by the auto demand equation.

These equations might also allow the transportation planner to examine the effect on one mode of altering the other mode. Any common variables among the equations will be an indicator of this effect and its magnitude. For example, what will be the result of an increase in car ownership with regard to both auto and transit passenger volumes? Or, would an increase in the tax on gasoline have a far reaching effect on transit ridership? Such questions and many others which relate the modes of travel might be answered in some measure by a consideration of the demand equations.

Future studies related to this particular area of travel demand should be concerned with refining the demand equations to enable prediction of not only A.M. peak hour CBD-bound suburban originating work trips, but also work trips to other areas of the city in question. Trips within the origin zones for purposes of work (in service industries) could also be investigated. The techniques used to obtain the demand equations could also be investigated in order to minimize the constant term in the regression equation, enabling a more reasonable prediction of actual working trips.

LIST OF REFERENCES

1. Domencich, T. A., G. Kraft, and J. P. Valette, "Estimation of Urban Passenger Travel Behaviour: An Economic Demand Model", Highway Research Record, No. 238, 1968.

2. Kraft, G. and T. A. Domencich, "Free Transit", Transportation and Poverty Conference, June 1968.

3. Mclynn, J. M. et al., "Passenger Demand and Modal Split Models", National Technical Information Service, United States Department of Commerce, 1969.

4. Sosslau, A. B., K. E. Heanue, and A. J. Balek, "Evaluation of a New Modal Split Procedure", Highway Research Record No. 261, 1969.

5. Canadian Council on Urban and Regional Research, "A Pilot Study of Relationships Between Socio-Economic Factors, User Attitude and Preferences and Urban Transportation System Attributes".

6. Transportation Planning Branch, Streets and Traffic Department, Metropolitan Corporation of Greater Winnipeg, Winnipeg Area Transportation Study, Vols. I and II, 1966.

7. Wohl, M. and B. V. Martin, Traffic Systems Analysis for Engineers and Planners, McGraw-Hill, 1967.

8. Steel, R. G. D. and J. H. Torrie, Principles and Procedures of Statistics, McGraw-Hill, 1960.

9. Zupan, J. M.,"Mode Choice: Implications for Planning" Highway Research Record, No. 251, 1968.

10. Wynn, F. H. and H. S. Levinson, "Some Considerations in Appraising Bus Transit Potentials", Highway Research Record,No. 197, 1967.

11. Larssing, J. B. and G. Hendricks, "How People Perceive the Cost of the Journey to Work", Highway Research Record, No. 197, 1967.

12. Shirnk, G. A., W. L. Grecco, and V. L. Anderson, "The Journey to Work: A Singular Basis for Travel Pattern Surveys", Highway Research Record, No. 240, 1968.

13. Creighton, R. Urban Transportation Planning, University of Illinois Press, 1970.

14. Meyer, J. R., J. F. Kain, and M. Wohl, The Urban Transportation Problem, Harvard University Press, 1965.

15. Irwin, N. A., and M. A. Wilson, "Information Requirements for Urban Transportation Planning", Proceedings of Seminar Series on Transportation, Center for Transportation Studies, University of Manitoba, Volume 4, June 1971.

16. Encyclopaedia Brittanica, 1970 edition.

A P P E N D I X   I

Map of

WINNIPEG

TEST

CASE