EQUILIBRIUM CONCEPTS IN MODAL SPLIT MODELLING

By James Beattie Morison[1]

(Reviewed by the Urban Transportation Division)

ABSTRACT: The potential usefulness of extrapolating Wardrop's first principle regarding traffic equilibrium on road systems to modal split modelling is evaluated. A conceptual framework based upon equilibrium theory is presented, a model for peak hour home based work trips is developed and the price equality prediction of the equilibrium model is tested. The test, which used data from the 1971 City of Winnipeg work trip survey, provides support for the equilibrium theory. It is noted that a logit model can be derived from this model. It is concluded that this model and research based upon it can lead to a better understanding of modal choice behavior and to a reliable and stable modal choice model.

INTRODUCTION

The purpose of this paper is to evaluate the possibility of applying the concept of equilibrium in transportation systems to the analysis of modal choice behavior.

The concept of traffic equilibrium has formed the basis of many trip assignment models since it was first proposed by Wardrop (8). Recent work by Florian and Nguyen (4) has shown the reliability of equilibrium modelling in trip assignment. Abdulaal and LeBlanc (1) have done some work on a model that applies the equilibrium concept to modal choice.

Currently modal choice models consist of either a set of diversion curves or some type of 'generalized cost' diversion curve (logit model). These have the advantage of being simple to develop, but there are some doubts as to their reliability. The study of equilibrium models will provide fresh insight into this problem.

The paper is organized into seven sections. The first section describes the theory behind the equilibrium model. The second section looks at the concept of equilibrium in more detail. The third section provides an example of a work trip model. The fourth section presents a test of the price equality prediction of the equilibrium model using peak hour home based work trips in the city of

---

[1]Transportation Planner, Transportation Planning Div., City of Calgary, Calgary, Alberta, Canada.

Note. -Discussion open until February 1, 1983. To extend the closing date one month, a written request must be filed with the Manager of Technical and Professional Publications, ASCE. Manuscript was submitted for review for possible publication on September 17, 1981. This paper is part of the Transportation Engineering Journal of ASCE, Proceedings of the American Society of Civil Engineers, (c) ASCE Vol. 108, No. TE5, September, 1982. ISSN 0569-7891/82/0005-0457/$01.00.

Winnipeg (Canada). The fifth section deals with the advantages of the model and looks at some of the implications. The sixth section will present some conclusions, and the seventh section suggests some further research.

EQUILIBRIUM THEORY IN MODAL CHOICE

The model developed here is based upon three main concepts. The first is the concept of equilibrium. The basic idea of equilibrium is that any system will eventually end up in a state where all the forces will balance each other. This is known as the equilibrium state. It was first proposed for traffic on a road system by Wardrop (8). The second concept is 'Travelers Expected Price' (TEP). The TEP is a combination of all the 'costs' that a traveler expects to pay when he makes a trip. The third concept is the 'Marginal Utility' (MU) idea. The MU is the additional 'Utility' which a traveler obtains from making a trip. Utility is a concept from economics that represents the benefits that a Person obtains from any activity.

In this paper, three assumptions have been made regarding the behavior of individuals in a transportation system. The first is that an individual in a transportation system will choose the mode that he perceives to be the best. The second is that all individuals are perfectly aware of all the characteristics of all modes and routes available to them in making their choice. The third assumption is that the users being considered are homogeneous with respect to the way they evaluate and perceive the characteristics of the transportation system.

TEP is assumed to be a linear combination of all the 'cost' factors affecting the choice of mode. This includes in-vehicle travel time, access time (walk time, wait time, and transfer time), and the monetary cost of travel (bus fare, vehicle costs, and parking). Each mode will have its own TEP equation where each factor is weighted according to how it is perceived by travelers.

The marginal utility is assumed to be a combination of all the 'benefit' factors affecting the choice of mode. It would seem reasonable that most of the MU of the majority of trips is associated with reaching the destination and only a small portion would be derived from the trip itself. Therefore for most types of trips the MU can be ignored when analyzing the choice of mode.

In this model it will be assumed that the best option is the one which the tripmaker perceives as having the highest MU/TEP ratio. This is somewhat analogous to the benefit /cost analysis used in an economic evaluation.

Wardrop's first principle states that travellers on a road system will tend to settle into an equilibrium state where no individual can improve his travel time by a change in route. In this model this principle is extended to a multiple mode system by introducing the MU/TEP ratio to replace travel time. Explicitly this is: Travelers in a transportation system will tend to settle into an equilibrium state where no change in mode or route by an individual will improve that individual's MU/TEP ratio.

In order to simplify working with the model, all relationships are assumed to be linear. Over short ranges, a straight line is a reasonable approximation of most curves.

In summary, the three major components of this model are the concepts of traveller's equilibrium, the Traveler's Expected Price (TEP), and the Marginal Utility (MU). The model is derived to a large extent from Wardrop's first principle which in turn can be derived from economic concepts of optimality and equilibrium.

The concept of equilibrium may be better understood if a few examples are considered. A system at equilibrium is not always a static system, because there may be some sort of change occuring. A good indication of this is a salt water solution. Once the water has dissolved all the salt it can accommodate, any additional salt remains as a precipitate. At this point there are two processes taking place. The first is the ongoing precipitation of the sodium and chlorine atoms out of solution. The rate at which this occurs is a function of the number of atoms in the solution. The second is the disolving of the sodium and chlorine atoms, in the precipitate, into the water. The rate at which this occurs is a function of the 'free space' in the water into which the atoms can be dissolved. The equilibrium state occurs when these two rates are equal. It is most important to note here that although many atoms are continually moving from one state to the other (solution /precipitate) the system remains at equilibrium.

Equilibrium in a traffic system is similar to this. At any point in time a small percentage of travellers will be reconsidering the route that they have chosen. Regardless of how they make the decision, the travel times at equilibrium by any two alternate routes will tend to be equal. If more of the travellers who reconsider their route choice on a particular day choose route one over route two, then route one would have a higher travel time as compared to route two. On the following day because route two has a lower travel time, the travellers who reconsider their route will be more likely to pick that route. This will bring the two travel times closer to being equal. Thus, over a period the system will tend to move toward the equilibrium state. The important thing to note here is that rather than staying at the equilibrium point the system will actually fluctuate around it. Because of the other factors that affect personal behavior the deviation from the equilibrium point on a given day may be quite substantial. This is not noticeable in the case of a saturated salt solution because the number of individual atoms is large. In transportation systems this fluctuation has been noted (3).

EXAMPLE OF MODEL FOR PEAK HOUR HOME BASED WORKED TRIP

Demonstrating how this model would work in practice, a model for peak hour home based work trips has been developed. This model will be used in the empirical test in the next section.

In the last section it was stated that for most trips the marginal utility is largely a function of reaching the destination and is not affected significantly by the mode taken. This is likely to be true of peak hour home based work trips. Therefore, disregarding the marginal utility should not greatly affect the model validity while making it easier to work with. It should be remembered that this simplification would not necessarily apply to all trip purposes and the concept of marginal utility remains an integral part of the conceptual model. In this example the "transportation system" consists of a single origin, a single destination, and two competing modes between the origin and the destination.

TEP equations can be written for each mode:

$$TEP_1 = A_0 \sum_{i=1}^{n} A_i \times X1_i \quad (1)$$

$$TEP_2 = B_0 \sum_{i=1}^{m} B_i \times X2_i \quad (2)$$

in which TEP_1 = Travelers Expected Price for mode 1; TEP_2 = Travelers expected price for mode 2; A_0, A_i, B_0, and B_i = weighting coefficients; $X1_i$ and $X2_i$ = the variables affecting the TEP; n = the number of variables affecting the TEP of mode 1; and m = the number of variables affecting the TEP of mode 2.

Even before the exact variables in the equation are known it can be assumed that one or more of the variables are affected by the volume of trips using that mode. This is because of the relationship between volumes and speeds on sections of the road system. Generally, the volume on one mode can affect the TEP of the other mode. So both volumes are indicated in these two rewritten TEP equations:

$$TEP_1 = F'(V_1, V_2) \quad (3)$$

$$TEP_2 = F''(V_1, V_2) \quad (4)$$

in which F' and F" represent two different arbitrary functions, V_1 = volume of trips on mode 1, and V_2 = volume of trips on mode 2.

At equilibrium no individual can decrease his TEP by a change in mode. For this simple system occurs only when the TEP's alternate modes are equal. This is called "price equality" and can be written as:

$$TEP_1 = TEP_2 = TEP_e \quad (5)$$

in which TEP_e = the equilibrium value of the TEP.

This concept of TEP or "price" equality will be the basis for the test in the next section. If the TEP's of alternate modes can be shown to be equal in a real life situation this will lend support to the model.

The total volume in the system is equal to the sum of the volumes using the two modes. This can be written:

$$V_t = V_1 + V_2 \quad (6)$$

in which V_t = total volume of trips.

Finally an equation accounting for the generation and distribution of trips can be written:

$$V_t = F'''(TEP_e) \quad (7)$$

in which F''' represents an arbitrary function.

Fig. 1 shows the bimodal equilibrium state described by Eqs. 3, 4, 5, 6, and 7. In this example the TEP of one mode is independent of the volume of the other. A numerical example could be developed that would give the actual volumes on the two modes but this approach is interesting for demonstrating only, and could not be used for a practical application. A iterative assignment technique

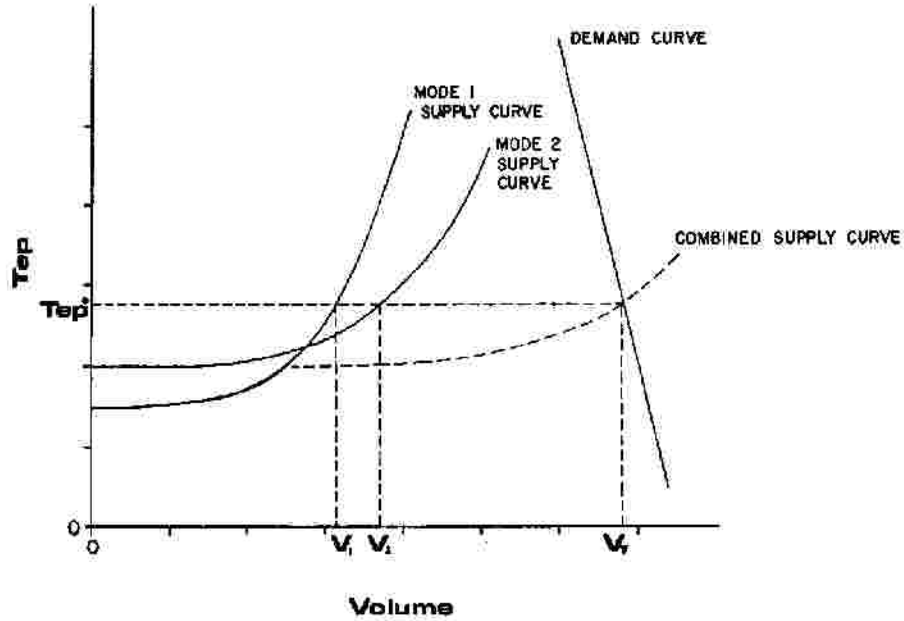

FIG. 1.—Demonstration of Bimodal Equilibrium State

such as that used in many traffic assignment programs would have to be used in practice.

The same approach can be applied to trips made for other purposes. In these cases however the marginal utility may have to be considered.

TESTING PRICE EQUALITY PREDICTION

In this section the equilibrium model will be tested by looking at the price equality prediction. This prediction will be tested by calculating TEP values for auto and transit cases where the situation described in the previous section is believed to hold. The resulting TEP estimates will be compared by: firstly, a visual examination of a scatter plot of the two values; secondly, a calculation of Pearson's 'r'; and thirdly, a calculation of Theil's inequality coefficient 'U'. Theil's inequality was calculated using the formula:

$$U = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Z1_i - Z2_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}Z1_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}Z2_i^2}} \quad (8)$$

in which $Z1_i$ = the value of variable 1 for observation *i*; and $Z2_i$ = Value of variable 2 for observation *i*.

Data for this test came from the City of Winnipeg (Canada), which in 1971 had a population of 529,000 and was Canada's fourth largest city. In 1971 the City of Winnipeg conducted a work trip survey using a sample of about 20%. In conjunction with the work trip survey, transit and road travel times, and other data were collected. 92 O-D pairs were selected for inclusion in this test. A number of variables were obtained for each O-D pair, and the data base was transferred to a computer file where Statistics Package for the Social Sciences (SPSS) was used for the analysis.

The following variables were selected for inclusion in the TEP equations:

1. In-vehicle travel time [TT_a (auto), TT_b (transit)], is the time spent in the vehicle during the trip.
2. Terminal time [AT_a (auto), AT_b (transit)] is the time spent walking to or from the vehicle at either the origin or the destination.
3. Fare [BF_b (transit)], is the average bus fare. This is converted to minutes by dividing by the value of time, which is assumed to be 25% of the rate of income.
4. Wait time [WT_b (transit)] is the time spent waiting at the origin terminal or at a transfer point for a bus.
5. Parking Cost [PC_a (auto)] is the cost of parking a car at the destination. It is converted to units of time by dividing by the value of time.
6. Operating Cost [OC_a (auto)] is the cost of operating a car for the trip. It is expressed in terms of time by dividing by the value of time.
7. Auto Availability [AA_a (auto)] is an estimate of the cost of becoming an auto driver, and is equal to $1 times (1.0 minus the number of cars divided by the total population). It is expressed in terms of time by dividing by the value of time.

The weighting factors to be used in the TEP equations should be obtained from value of time studies, which could use some form of discriminant analysis. The disaggregate data base needed to do this was not available. Therefore, values were assumed based upon data from Shunk and Bouchard (7). It would be expected that the weighting factors would differ from city to city as a function of climate and other factors. However, for this preliminary analysis these values should be adequate.

The Transit TEP equation used was:

$$TEP_b = 1.0 \times TT_b + 2.5 \times AT_b + 2.5 \times WT_b + 1.0 \times BF_b \qquad (9)$$

The Auto equation has a constant included to allow the means of the two equations to be equal. The constant was determined by trial and error, and it may represent the cost of becoming a private vehicle user. The auto equation used was:

$$TEP_a = 1.0 \times TT_a + 2.5 \times AT_a + 0.5 \times PC_a + 1.0 \times OC_a + 1.0 \times AA_a + 7.7 \qquad (10)$$

The values for TEP, and TEP, are computed in Fig. 2.

If two variables are equal a scatterplot of corresponding values would lie along a 45° line. In Fig. 2 this pattern appears.

Pearson's 'r,' or the correlation coefficient, is used to measure the strength of a relationship between two variables. In the comparison the value of Y is about 0.77, which indicates a strong relationship.

Theil's inequality coefficient, 'U,' is often used to determine how accurate an equation predicts the actual value. It ranges in value from 0-1.0; a value of

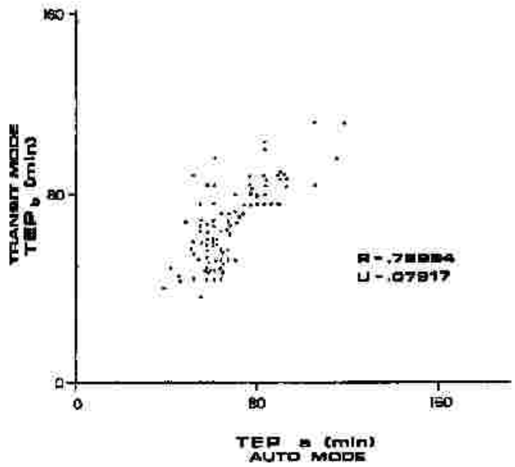

FIG. 2.—Comparison of Predicted TEP Values for Auto and Transit Modes

less than 0.01 is considered very good, 0.01-0.05 is good, 0.05-0.10 is acceptable, and over 0. 10 is not considered acceptable. For the comparison in Fig. 2 the value of U falls just within the acceptable range.

REVIEW OF ASPECTS OF EQUILIBRIUM MODEL

The review in this section will address four areas. The first is a comparison of the equilibrium model and the logit model. The second is a description of some advantages of the equilibrium model. The third is an examination of the problems of the equilibrium model, and the fourth is an investigation of the implications of the equilibrium model.

If the equilibrium model is a good representation of the real world it should be possible to explain existing modal split models. The most common type of modal split model being used now is the logit model. The logit model theory states that a plot of the 'Utility difference' between two modes vs. the modal split produces a logit curve. This amounts to a residual analysis of the TEP differences in the equilibrium model. If the equilibrium model were valid and the population were homogeneous, the plot of the residuals vs. the modal split would be a straight line at the point where the residuals are equal to zero. However, if the population were not homogeneous then there would be some deviation from this line, and this deviation would be a function of the modal split. Using Eqs. 9 and 10, a plot of the difference between the predicted TEP values for the two modes vs. the modal split was made. This is presented in Fig. 3. In Fig. 3 there is some deviation from the zero difference line, but there does not seem to be any trace of a logit type curve. The failure to produce such a pattern shows some support for the equilibrium model over the logit model. However, at low modal splits there is a much better chance of a "wild" datum occurring since a single individual could affect the estimate of the modal split. This could result in a logit curve fitting the data quite well. This would not compromise the validity of an equilibrium model.

There are two major advantages of the equilibrium model. The first is the usefulness of the model in evaluating policies. By seeing how the various factors are weighted policies can be formulated and evaluated.

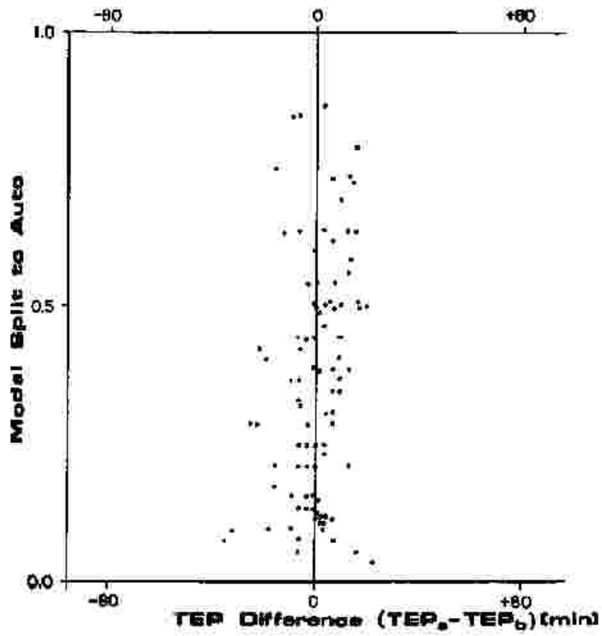

FIG. 3.—Observed Modal Split Plotted Against Difference in TEP Values

A second advantage of this model is that it is "behavioral" and the weighting factors used in the TEP equations have a real meaning which may allow them to be predicted under different conditions. These factors can change from place to place and from time to time, and further research could determine how these changes occur and possibly allow planners to effect a change on their own.

Although this study supports the prediction of price equality, there may still be problems with the equilibrium model. Some of these problems are considered in this section.

The first problem is that there is a considerable time lag before equilibrium is reached in a transportation system. As a result the system may not be at equilibrium at any given time. To get some idea of the magnitude of the time lag the effects of a transit strike should be considered. A transit strike will throw a system out of equilibrium. It has been noted (5) that it takes about a year before transit ridership reaches prestrike levels. Since in an area that is growing rapidly the equilibrium point will be constantly changing it is possible that a system may never reach equilibrium. This does not invalidate the equilibrium model, but it could mean that the modal split may be inherently unpredictable except on a highly aggregate level. This would apply to all modal split models.

The second problem is that the assumption of a homogeneous population is not valid. This would mean that not all people would evaluate their TEP's the same. For example, the weighting factors may vary as a function of social or economic characteristics. A person who already owns a car does not have the capital cost factor in his equation for TEP.. To summarize this, the population as a whole can be divided into three market segments. In the first, the TEP for car is much lower than the TEP for transit. In the second, the TEP for transit is much lower than the TEP for the car. In the third, the two TEP's are about the same. The model has in effect been constructed on the assumption that the third segment accounts for 100% of the population. If the third group is sufficiently large, no problem is encountered.

The concept of three market segments has been noted by other researchers and has been referred to as 'captive' and 'choice' riders. The problem with this approach is that the term captive implies that those people can never change to the other mode. Historically, during the 1940's and 1950's, the large captive market for transit almost disappeared. Significant changes in the TEP for alternate modes can affect some of the captive riders. This is what occurred during the 1940's and 1950's with a rapid increase in the availability of automobiles.

The third problem is that the assumption that all individuals in the system are perfectly aware of the options open to them is not true in the strictest sense. In many models this assumption has been made without any major difficulties.

There are five important implications of the model on the practice of transportation planning which should be pointed out. The first implication is in the evaluation of the fare elasticity of transit demand. Most studies show that the demand is relatively inelastic with respect to fare changes (2). This fits in well with the idea that most transit users are captive. However, if the concept of TEP replaces fare it is obvious the demand is not inelastic but extremely elastic. This is because the so called captive transit rider is not really captive in the long run. This means that a transit system cannot count on a captive market for revenue, but also it means that a major shift from the auto mode to transit is possible as a result of changes in service.

The second implication is the effect of the auto availability variable of the TEP equation. A person who does own a car does not perceive this as a factor while a person who does not own a car does. Thus a person who changes from a 'captive' transit user to an auto driver 'loses' that part of the TEP for the auto mode. It would therefore be expected that an increase in the transit TEP would have a much higher elasticity than a decrease would. A transit user who is lost to the auto mode by an increase in the transit TEP will not shift back unless the transit TEP can be reduced by an amount greater than the original increase, plus contribution of auto availability to the TEP. The catastrophy theory may be useful in understanding this.

The third implication concerns the use of rapid transit lines to increase transit use. The basis of this approach is if the overall travel time is reduced then there will be an increase in the transit ridership. Unfortunately some recent studies have suggested that changes in travel time don't affect the transit ridership greatly. This is based upon interpretation of the weighting factors obtained. But the equilibrium model indicates this may be misleading, and that the shape of the TEP curve is important. If the shape of the TEP equation is not greatly affected by the volume then relatively small changes in travel time can result in substantial changes in the volume.

The fourth implication also affects the potential of rapid transit. In value of time studies which form the basis upon which the TEP is developed it has been noted that 'access' times are often valued at as much as 2.5-3.0 times the invehicle time. Thus a one minute reduction in the wait time at a bus stop has as great an effect on the modal split as a reduction of 2.5 or 3.0 in the in-vehicle time. This is an important point since in developing rapid transit lines, planners will trade off an increase in the access time to produce an overall reduction in the total travel time. But since the access time is valued more highly, an overall reduction in the travel time can result in an increase in the TEP. This increase will, in turn, cause a drop in the modal split to transit, the exact opposite of its intention.

The fifth implication noted is the time lag in reaching equilibrium. In the course of this examination it was suggested that a transportation system may take a long time to reach equilibrium. Because of the rapid change in the factors affecting the location of the equilibrium point the system may for all practical purposes never reach equilibrium and may in fact deviate substantially from it. This would mean the modal split may in fact be impossible to predict.

CONCLUSIONS

A number of conclusions were drawn from the analysis and review in this paper.

The general idea of the equilibrium modal split model was supported by the study presented here, although more research will be needed before the model can be widely accepted and applied.

The equilibrium model is general in its nature, and it is possible to explain the logit model in terms of the equilibrium model.

The effects of parking, vehicle cost, and access times on the modal split are substantial and should be considered in any modal split model.

Further research on the equilibrium model should eventually lead to a better understanding of the modal choice process and the development of a reliable and stable modal split model.

SUGGESTIONS FOR FURTHER RESEARCH

The first project in any further research effort would be to test the price equality prediction in as many situations as possible. This would involve developing the weighting factors in the TEP equation from discriminant analysis of disaggregate data, then plotting estimates of TEP's for alternate modes between O-D pairs and testing for equality.

The second project would involve the development of a computer program that would apply the equilibrium modal split model. This would then be used to test the predictive ability of the model.

A number of other projects can be obtained from further theoretical studies, in particular other trip purposes. It may also be possible to extend the equilibrium concept to trip generation and distribution.

ACKNOWLEDGMENTS

The work presented here originated as the author's M.Sc. thesis (6) for the University of Manitoba Department of Civil Engineering under the direction of A. H. Soliman, with the financial support of the RTAC IAC Award. Further work was done by the author at the City of Calgary Transportation Department. The author would like to thank his family and the many other people who gave their advice.

APPENDIX I.- REFERENCES

1. Abdulaal, M., and Leblanc, L. J., "Methods for Combining Modal Split and Equilibrium Assignment Models," *Transportation Science, Vol.* 13, No. 4, Nov., 1979, p. 292.
2. Baum, H. J., "Free Public Transport," *Journal of Transportation Economics and* Policy, Jan., 1973, p. 3.
3. Bendtsen, P. H., "Errors Affecting the Different Components of Traffic Predictions," *Transportation Research, Vol.* 9, 1975, p. 111.
4. Florian, M., and Nguyen, S., "An Application and Validation of Equilibrium Trip Assignment Methods," *Publication No.* 28, University of Montreal, Centre de Recherche sur les Transports, Montreal, Quebec, Canada, Aug., 1975.
5. Harmatuck, D. J., "The Effects of Service Interruption on Bus Ridership Levels in a Middle-sized Community," *Transportation Research, Vol.* 9, 1975, p. 43.
6. Morison, J. B., "A Proposed Modal Choice Model for Urban Transportation Systems," thesis presented to the Department of Civil Engineering, of the University of Manitoba, at Winnipeg, Manitoba, Canada, in Oct., 1978, in partial fulfillment of the requirements for the degree of Master of Science.
7. Shunk, G. A., and Bouchard, R. J., "An Application of Marginal Utility to Travel Mode Choice," *Record* 322. *HRB*, Highway Research Board, Washington, D.C., 1970, p. 30.
8. Wardrop, J. G., "Some Theoretical Aspects of Road Traffic Research," *Proceedings*, Institute of Civil Engineering, Part II, Vol. 1, 1952, p. 325.

APPENDIX II.-NOTATION

The following symbols are used in this paper:

A_i	=	arbitrary weighting factor for mode *a;*
AA_a	=	Automobile Availability for auto;
AT_a	=	terminal time (access time) for auto;
AT_b	=	terminal time (access time) for transit;
B_i	=	arbitrary weighting factor for mode *b;*
BF_b	=	Bus Fare (transit);
F( )	=	indicates an arbitrary or unknown function;
MU	=	Marginal Utility;
OC_a	=	Operating Cost (auto);
PC_a	=	Parking Cost (auto);
‘r'	=	Pearson's correlation coefficient;
TEP	=	Traveler's Expected Price;
TEP_e	=	TEP at equilibrium;
TT_a	=	In-Vehicle Travel Time (auto);
TT_b	=	In-Vehicle Travel Time (transit);
U	=	Theil's Inequality Coefficient; Vi = volume of trips;
WT_b	=	Wait Time Bus; and
X_i	=	arbitrary variable affecting TEP.