# **OPERATIONS RESEARCH CENTER**

working paper



# MASSACHUSETTS INSTITUTE OF TECHNOLOGY

# A COMBINED TRIP GENERATION, TRIP DISTRIBUTION, MODAL SPLIT AND TRAFFIC ASSIGNMENT MODEL

K. Nabil Ali Safwat\*

Thomas L. Magnanti\*\*

OR 112 82

March 1982

revised and submitted to <u>Transportation Science</u> February 1985

\* Supported by U.S. AID through the Technology Adaptation Program at M.I.T.

\*\* Supported in part by NSF Grant No. 79-26225-ECS. This paper was completed, in part, while this author was visiting the Graduate School of Business at Harvard University.

# ABSTRACT

We introduce a transportation equilibrium model that simultaneously predicts trip generation, trip distribution, modal split, and traffic assignment by algorithms that are guaranteed to converge to an equilibrium and are computationally efficient for large-scale systems. The model is formulated as an equivalent optimization problem, yet it allows realistic, flexible and behaviorally acceptable demand models.

# 1. INTRODUCTION

During the last ten years, much of the research in transportation planning has focused on ways to improve predictive modelling. One of the most predominant themes of this research has been an effort to develop comprehensive models, and related computational procedures, for computing short-run transportation equilibria. These integrated models recognize that user decisions concerning trip frequency, destination, mode and route choices are inherently interrelated. By combining these user decisions, the models aim to provide better predictions of transportation systems' performance (delay times, costs) and user travel behavior (demand patterns).

This trend toward integrated modelling contrasts sharply with earlier methods for predicting traffic equilibria. The earlier procedures, which have been applied to hundreds of transportation studies throughout the world for the past 30 years and still are in use today, have viewed transportation planning as a sequential process, often with four stages--trip generation, trip distribution, mode, and route choice. The Detroit Metropolitan Area Traffic Study [1955], the Chicago Area Transport Study [1960] and the Cairo Urban Transport Study [1981] illustrate this practice, as do guidelines prepared by the U.S. Federal Highway Administration [1970, 1972] and the U.S. Urban Mass Transit Authority [1976]. Unfortunately, the sequential approach has an inherent weakness; its solution need not be internally consistent. That is, because each stage in this type of sequential planning depends upon the others, the performance or demand levels that one needs to assume as given inputs at any one stage need not agree with those that one determines as outputs from the other stages. This deficiency has precipitated attempts to model all of the stages simultaneously.

Research intended to meet this objective of model integration has proceeded in three directions. One of these lines of investigation has significant computational advantages; the others permit richer modelling of user behavior. Regrettably, to date none of these approaches has generated models that are both behaviorally acceptable and computationally tractable for large-scale applications.

The first of the simultaneous approaches, which originates with the early and seminal research of Beckman et al [1956], views the equilibrium model as an equivalent optimization problem that when solved yields the desired equilibrium solution. The primary advantage of this formulation is that the equilibrium problem becomes a convex optimization problem (assuming monotonicity of demand and performance) that can be solved efficiently by any of several convergent algorithms (Bruynooghe, Gibert, and Sakorovitch [1968], Bertsekas and Gafni [1981], Dembo and Klincewicz [1981], Leblanc [1973], Nguyen [1974, 1976a, 1976b], Golden [1975] and Florian and Nguyen [1974]). The main disadvantage of this formulation is behavioral. It requires strong modelling assumptions that frequently are unrealistic, particularly an assumption that demand between each origin-destination (0-D) pair depends solely upon the performance between that O-D pair.

The basic equivalent optimization formulation has several modelling enrichments. Evans [1976] extended the formulation to include trip distribution, assuming fixed trip generation and an entropy model for trip distribution. Using the fact that an entropy distribution model implies a logit mode-split model, Florian and Nguyen [1978] further extended the formulation to include modal split. Each of these extensions shares the computational advantages of the equivalent optimization formulation. Again,

-2-

the deficiencies are behavioral; the entropy model is not based upon any behavioral principles. Moreover, those modelling extensions are rigid. Because the formulations incorporate entropy distribution and fixed trip generation, the models are not flexible enough to accommodate situations in which a goodness-of-fit test with observed data shows that the entropy model is not a correct functional form.

The second simultaneous approach views the equilibrium conditions as a system of equations and inequalities to be solved directly. In this form, the equilibrium conditions can be interpreted as describing a nonlinear complementarity problem (Aashtiani and Magnanti [1981]), a stationary point problem (Asmuth [1978]), or a variational inequality problem (Smith [1979], Dafermos [1980]).

This approach has substantial behavioral advantages, but is limited computationally. It permits general demand or performance functions and yet insures existence and uniqueness of an equilibrium, even with only mild continuity and/or monotonicity assumptions imposed upon the data. In principle, this general model can be solved by convergent fixed point algorithms (Hearn and Kuhn [1977], Asmuth [1978]) or, by projection algorithms (Dafermos [1980, 1981], Pang and Chan [1981]). The fixed point algorithms are limited, however, to very small problems. Similarly, computational experience has suggested that the proposed projection algorithms are inefficient for this type of application (see Fisk and Nguyen [1980]). The general model can also be solved by an efficient Newton type algorithm (Aashtiani [1979]), but this algorithm only guarantees local convergence (Pang and Chan [1981]).

A third line of investigation enriches the modelling of user behavior by permitting user perception of performance to be stochastic. Sheffi and Daganzo [1980] view this stochastic equilibrium problem as a traffic

-3-

assignment problem on an extended network and cast the model as an equivalent optimization problem. They use a disaggregate probit model for demand and combine both deterministic and stochastic assignment of trips to paths on the extended network. Although their algorithm is convergent (with some restrictions imposed upon the probit model specification), the procedure is limited in practice because it requires substantial computational effort for even modestly-sized problems.

This summary of previous studies illustrates the tradeoffs between the realistic (behavioral fidelity), technical (convergence), and practical (computational efficiency) aspects of modelling the equilibrium problem. None of the previous models has been successful in addressing all these issues.

Our goal in this paper is to develop a model that comes closer to achieving all three objectives. It is intended to strike a balance among the realistic, technical, and practical considerations of the problem. We propose an equilibrium model that is behaviorally acceptable; moreover, it has a unique equilibrium that can be computed efficiently by a convergent algorithm.

In spite of prevailing views of many researchers concerning the behavioral limitations of the usual optimization approach, we have formulated an equivalent optimization problem that relaxes two of the major behavioral restrictions imposed upon transportation demand. In our formulation, trip generation can depend upon the system's performance through an accessibility measure that is based on the random utility theory of users' behavior; in addition, trip distribution is given by a logit model. The formulation is a convex program that can be solved efficiently by convergent algorithms (see Safwat [1982]).

-4-

In the next two sections, we describe our equilibrium model and formulate an equivalent optimization problem. In the fourth section, we prove that these two forms of the problem are equivalent by deriving yet another equivalent optimization formulation which is stated in a form that is more likely to be familiar to the reader. The fifth section shows that singly and doubly constrained gravity models are, respectively, special and limiting versions of our demand models. The sixth section contains discussion and conclusions.

-5-

# 2. A SIMULTANEOUS TRANSPORTATION EQUILIBRIUM MODEL (STEM)

In this section we present the underlying theory and the basic assumptions of an equilibrium model that describes users' travel behavior in response to system's performance on a transportation network. We first introduce some notations:

- (N,A), a directed graph (i.e., any transportation network) consisting
   of a set N of nodes and a set A of links;
- i, an origin node in the set N;
- j, a destination node in the set N;
- ij, an origin-destination pair;
- p, a simple (i.e., no node repeats) path in the network (N, A);
- a, a link in the set A;
- I, the set of origin nodes  $(I \subseteq N);$
- $D_i$ , the set of destinations that are accessible from a given origin i  $(D_i \subseteq N)$ :

R, the set of origin-destination pairs;

- P<sub>ij</sub>, the set of simple paths from origin i to destination j;
- P, the set of simple paths in the network (P =  $U \{P_{ij}:i \in I, j \in D_i\}$

Now let us describe the basic assumptions for the different components of our STEM model.

#### 2.1 USER UTILITY FUNCTFONS

We assume that a typical user travelling from a given origin i associates a utility  $v_{ij}$  with each destination j in the set  $D_i$  of destinations perceived to be accessible from i. Because users do not usually have perfect information concerning the system and analysts cannot quantify all the factors that influence users' utilities, we assume that utility functions are random and may be decomposed into a measured (observed) utility component plus an additive random (error) term; that is,

$$v_{ij} = V_{ij} + \varepsilon_{ij}$$
, for all  $ij \in \mathbb{R}$  (2.1)

where

 $v_{ij}$  = utility of travel from i to j;  $V_{ij}$  = measured (observed) utility of travel from i to j; and  $\varepsilon_{ij}$  = random (unobserved) utility of travel from i to j.

We further assume that the measured utility is a function of socioeconomic characteristics of both the destination (e.g., consumption levels, population) and the user (e.g., income, profession, education) as well as the system's performance, and may be expressed as follows:

$$V_{ij} = -\theta u_{ij} + \sum_{w=1}^{W} \theta_{w} g_{w}(A_{wj})$$
  
=  $-\theta u_{ij} + A_{j}$ , for all ijcR. (2.2)

In this expression,

u<sub>ij</sub> = the "perceived" cost of travel from i to j,
A<sub>wj</sub> = the value of the w<sup>th</sup> socio-economic variable that influences
 trip attraction at destination j;

- $g_w(A_{wj}) = a$  given function specifying how the w<sup>th</sup> socio-economic variable,  $A_{wj}$ , influences trip attraction;
  - $A_j$  = the composite effect that the socio-economic variables, which are exogenous to the transport system, have on trip attraction at destination j.

The quantities  $\theta$  and  $\theta$  for  $w = 1, \dots, W$  are coefficients to be estimated.

Notice that  $\theta$  is a positive coefficient; the negative sign associated with it reflects the behavioral assumption that, everything else being equal, the utility decreases as travel cost increases. During the time period required to achieve short-run equilibrium, which we are predicting, the socio-economic activities in the system will remain essentially unchanged. Consequently, we assume that the composite effect of these activities,  $A_j$ , is a fixed constant. That is, for a given specification of the socio-economic system, we assume that the observed utility of travel from i to j depends solely on the perceived travel cost,  $u_{ij}$ , that is,

$$V_{ij} = V_{ij}(u_{ij})$$
, for all ijcR.

We will also assume that the perceived cost of travel from i to j on any route is the sum of travel costs on the links that comprise that route. We will elaborate on how transportation policies and the system's usage influence perceived travel costs as we present the basic assumptions concerning link cost functions, modal split, and traffic assignment.

### 2.2 ACCESSIBILITY

Accessibility is a term that is widely used, but rarely defined (and measured) rigorously and satisfactorily [Dalvi and Martin (1976)]. In order to overcome this deficiency, Ben-Akiva and Lerman (1977) have defined accessibility as "some composite measure which describes the characteristics of a group of travel alternatives as they are preceived by a particular individual". They also have considered accessibility measures in the context of the random utility theory of users' behavior, which assumes that utility functions are random and that users are utility maximizers. Based on this theory, they have suggested that accessibility may be appropriately measured by the expected maximum utility to be obtained from a particular travel choice situation [other researchers such as Williams (1977) and Daganzo (1979) have also suggested and studied this measure].

-8-

Following this same line of thought, we define accessibility as a composite measure of the transportation system's performance and the socioeconomic system's attractiveness as perceived by a typical user travelling from a given origin. Accessibility of an origin will then be the value of the expected maximum utility obtained by travelling from that origin; that is,

 $S_{i} = E [\max_{j \in D_{i}} v_{ij}], \text{ for all } i \in I$  (2.3)

where

E is the expectation operator,

and the maximization is taken over all destinations D accessible from origin i.

Recall that the utility (as defined in section 2.1) has a random error term. In order to obtain an operational measure of accessibility, we must assume some probabilistic distribution for the random terms in the utility functions. A well-known and often used assumption in travel demand analysis is that the error terms are independent and identically distributed as a type-I extreme value distribution (we will elaborate on this assumption when discussing trip distribution). Making this assumption, the references cited earlier show that accessibility is given by the natural logarithm of the sum of exponentials of measured utilities to all accessible destinations; that is,

$$S_{i} = \ln \sum_{j \in D_{i}} \exp(V_{ij}), \text{ for all } i \in I$$
where  $V_{ij}$  is given by (2.2).
(2.4)

-9-

# 2.3 TRIP GENERATION

We assume that trip generation is a function of socio-economic activities, socio-economic characteristics of the users, and the transport system's performance. Specifically, we assume that trip generation is given by a general linear model with the measure of accessibility as one of its variables. That is,

$$G_{i} = \alpha S_{i} + \sum_{\ell=1}^{L} \alpha_{\ell} f_{\ell}(E_{\ell i})$$
  
=  $\alpha S_{i} + E_{i}$ , for all icI (2.5)

where

G<sub>i</sub> = the number of trips generated from i;

- E<sub>li</sub> = the value of the l<sup>th</sup> socio-economic variable that influences
   trip generation from origin i;
- $f_{l}(E_{li}) = a$  given function specifying how the  $l^{th}$  socio-economic variable  $E_{li}$ , influences trip generation; and
  - E<sub>i</sub> = the composite effect that the socio-economic variables, which are exogenous to the transport system, have on trip generation from origin i.

The quantities  $\alpha$  and  $\alpha_{\ell}$  for  $\ell = 1, ..., L$  are coefficients to be estimated.

As noted earlier, since the socio-economic activities are essentially unchanged in the short run, we assume that their composite effect,  $E_i$ , is a fixed constant. That is, for a given specification of the socio-economic system, we assume that trip generation is dependent solely on the system's performance as measured by the accessibility variable; that is,

 $G_i = G_i(S_i)$ , for all icI.

Since the accessibility variable S, in our model is a natural logarithm (expression (2.4)), its value may vary, in theory, between  $-\infty$  and  $+\infty$ . In practice, however, accessibility has some finite upper limit (i.e., the systems attractiveness when travel costs are zero throughout the system); we argue that it also has some finite lower limit. Specifically, we assume that our specification of the network, and particularly our definition of origins, implies that each accessibility variable is nonnegative. A sufficient, though not necessarily required, condition for S, to be nonnegative is that the measured utility of travel from i to at least one destination j in the set D, is nonnegative (i.e.,  $V_{ij} \ge 0$  for some  $j \in D_i$ ). That is, at least one destination in the system is "attractive" to users at any given origin, an assumption that should be satisfied in many, if not all, realistic systems. Suppose to the contrary, that the minimum travel costs to all destinations in the set D, are sufficiently large to give negative values for all measured utilities. Then either (i) no trips will be generated from i and thus, we might as well have deleted that origin from the analysis, or (ii) some trips must be generated from origin i regardless of the system's performance. In the later case, we assume that when accessibility in (2.4) becomes negative it no longer affects the number of trips generated; instead, the exogenous socio-economic composite variable  $E_i$  in (2.5) becomes predominant. That is, E, trips must be generated due to socio-economic forces. Hence, we assume that accessibility is nonnegative and specified as follows.

$$S_{i} = \max \left\{ 0, \ln \sum_{j \in D_{i}} \exp(-\theta u_{ij} + A_{j}) \right\}, \text{ for all it}.$$
 (2.6)

-11-

### 2.4 TRIP DISTRIBUTION

Adopting the random utility theory of users' behavior, we say that the probability  $(PR_{ij})$  that a typical user at any given origin i chooses to travel to any given destination j in the set  $D_i$  is equal to the probability that the utility of travel to j is greater than (or equal to) that of any other destination k in the set  $D_i$ . That is,

 $PR_{ij} = Probability [v_{ij} \ge v_{ik} \text{ for all } k \in D_i].$ 

Different assumptions on the probabilistic distribution of the random (error) terms of the utility functions lead to different trip distribution models. Since we are assuming that the error terms are independent and identically distributed as type-I extreme value (Gumbel) distribution, trip distribution is given by the well-known "logit" model:

$$^{T}_{ij} = G_{i} - \frac{\exp(-\theta u_{ij} + A_{j})}{\sum_{\substack{k \in D_{i}}} \exp(-\theta u_{ik} + A_{k})}, \text{ for all ij} \in \mathbb{R}.$$
(2.7)

Here T<sub>ij</sub> equals the number of trips travelling from i to j.

The type-I extreme value distribution describes the limiting distribution of the largest value of n independent and identically distributed random variables as n becomes large, assuming that the common distribution has an upper tail that falls off "in an exponential manner" as in the normal distribution [see Gumbel (1958) for more details].

These assumptions are invoked frequently in travel demand analysis and the resulting "logit" model is known to be very robust, practical and analytically tractable. These desirable features account for the model's popularity. In addition, as we will demonstrate later, our logit distribution

<sup>\*</sup> See, for example, Domencich and McFadden (1975) for the derivation of the logit model.

model is quite flexible and general, compared to other gravity models which may be viewed as special cases.

# 2.5 MODAL SPLIT AND TRAFFIC ASSIGNMENT

Several alternative assumptions on both modal split and traffic assignment may be considered within the framework presented in this paper [see Safwat (1982)]. However, to simplify the development and notation in this paper, we assume that each user chooses the mode and route combination that minimizes his total perceived cost from node of origin to node of destination. Implied in this assumption is the possibility of transferring from one mode to another in the middle of any given trip.

We assume that perceived travel costs are represented at the link level by a set of link cost functions. Each link cost function is assumed to depend upon the flow over that link and to be continuous and nondecreasing. These frequently invoked assumptions reflect congestion effects on perceived costs.

The above assumptions on modal split, traffic assignment, and system's performance imply a Wardrop user equilibrium model of path choice, That is, the perceived costs on all used paths (i.e., mode-route combinations) between any given O-D pair are equal and not greater than those on unused paths:

$$C_{p} = \sum_{a \in A} \delta_{ap} \cdot C_{a}(F_{a}) \begin{cases} = u_{ij} \text{ if } H_{p} > 0 \\ & , \text{ for all } p \in P_{ij} \text{ and } ij \in \mathbb{R} \end{cases} (2.8)$$
$$\geq u_{ij} \text{ if } H_{p} = 0$$

where

C<sub>p</sub> = the total perceived travel cost on a path p joining origin i and destination j;

$$C_a$$
 (F<sub>a</sub>) = the perceived travel cost on link a as a function of the link flow  $F_a$ ;

$$H_p =$$
the flow on path p; and

$$\delta_{ap} = \begin{cases} 1 & \text{if link a belongs to path p} \\ 0 & \text{otherwise} \end{cases}$$

# 2.6 THE STEM MODEL AND THE EQUILIBRIUM PROBLEM

Combining the modelling ingredients described so far gives the following simultaneous transportation equilibrium model.

(STEM):

$$G_{i} = \alpha S_{i} + E_{i}, \text{ for all } i$$

$$S_{i} = \max \left\{ 0, \ln \sum_{j \in D_{i}} \exp(-\theta u_{ij} + A_{j}) \right\}, \text{ for all } i \in I$$

$$T_{ij} = G_{i} \quad \frac{\exp(-\theta u_{ij} + A_{j})}{\sum_{k \in D_{i}} \exp(-\theta u_{ik} + A_{k})}, \text{ for all } i j \in R$$

$$C_{p} = u_{ij} \text{ if } H_{p} > 0$$

$$C_{p} \geq u_{ij} \text{ if } H_{p} = 0 \right\} \text{ for all } p$$

where

$$C_{p} \equiv \sum_{a \in A} \delta_{ap} \cdot C_{a}(F_{a}) \cdot$$

The equilibrium problem now becomes one of predicting  $G_i$  and  $S_i$  for all icI,  $T_{ij}$  and  $u_{ij}$  for all ijcR, and  $F_a$  and  $C_a$  for all acA

- (1) simultaneously,
- (2) with a procedure that is guaranteed to converge to an equilibrium that is proven to exist and to be unique, and
- (3) efficiently (in the computational sense).

As noted in the introduction to this paper, in spite of prevailing views of many researchers concerning the behavioral limitations of the usual optimization approach, we propose to model and solve this equilibrium problem by formulating an optimization problem (ECP) and showing that under mild assumptions on demand and performance the (ECP) problem has a unique solution that is equivalent to the (STEM) model.

In the next section, we introduce the (ECP) formulation and state a theorem which shows the equivalence between (ECP) and (STEM). In the fourth section, we prove this equivalency as well as the existence and uniqueness of equilibrium, by deriving yet another equivalent optimization formulation which is stated in a form that is more likely to be familiar to the reader. The fifth section describes some limiting and special cases of the (STEM) model. Safwat (1982) develops and tests a convergent algorithm for solving the STEM model.

-15-

# 3. AN EQUIVALENT CONVEX PROGRAM (ECP)

Consider the following (convex) optimization problem (ECP):

Minimize  $Z(S,T,H) = J(S) + \psi(T) + \phi(H)$ 

subject to:

$$\sum_{j \in D_{i}} T_{ij} = \alpha S_{i} + E_{i}, \text{ for all } i \in I$$
(3.1)

$$\sum_{p \in P_{ij}}^{N} H_p = T_{ij}, \quad \text{for all } ij \in I \quad (3.2)$$

$$S_{i} \geq 0, \quad \text{for all } i \in I$$

$$T_{ij} \geq 0, \quad \text{for all } i j \in R$$

$$H_{p} \geq 0, \quad \text{for all } p \in P$$

$$(3.3)$$

where

$$J(S) = \frac{1}{\theta} \sum_{i \in I} \left[ \frac{\alpha}{2} S_i^2 + \alpha S_i - (\alpha S_i + E_i) \ln(\alpha S_i + E_i) \right],$$
  

$$\psi(T) = \frac{1}{\theta} \sum_{i \in I} \sum_{j \in D_i} \left[ T_{ij} \ln T_{ij} - A_j T_{ij} - T_{ij} \right],$$
  

$$\phi(H) = \sum_{a \in A} \int_{o}^{Fa} C_a(w) dw, \text{ and}$$
  

$$F_a = \sum_{p} \delta_{ap} H_p.$$
(3.4)

The constraints (3.1) and (3.2) are the flow conservation equations on the transport network, stating that the number of trips distributed from a given origin to all possible destinations should equal the total number generated from that origin and that the number of trips on all paths joining a given origin-destination pair should equal the total number distributed from that origin to that destination. The constraints (3.3) state that all the decision variables should be nonnegative as postulated earlier. The expression (3.4) defines the link-path incidence relationships stating that the flow on a given link equals the sum of flows on all paths sharing that link.

The objective function Z has three sets of terms. The last of these,  $\phi(H)$ , corresponds to the familiar transformation introduced by Beckmann et al (1956). The second set of terms,  $\psi(T)$ , is similar to those used by Evans (1976) and by Florian and Nguyen (1978), as well as in other related models. The first set of terms, J(S), is new. In fact, what distinguishes our formulation from other models is the definition of the accessibility measure S<sub>i</sub>, its introduction as a decision variable in the optimization problem, and the specification of the first set of terms J(S) in the objective function of (ECP).

The importance of the (ECP) optimization problem is that even with very mild assumptions imposed upon the problem data, it is a convex program which has a unique solution that is equivalent to the (STEM) equilibrium model. Formally, the equivalence theorem may be stated as follows:

# THEOREM 3.1 (EQUIVALENCY):

Suppose that  $\theta > 0$ ,  $0 < \alpha \leq E_i$  for all ieI, and that  $C_a$  is continuous and a nondecreasing function of  $F_a$  for all aEA. Then (ECP) is a convex program whose optimality conditions are equivalent to the simultaneous transportation equilibrium model (STEM).

Because of this equivalency, it is possible to study the qualitative characteristics of the STEM model (i.e., existence and uniqueness) and compute an equilibrium by studying and solving a nicely structured optimization problem (i.e., ECP). We pursue the first of these objectives in the next section.

-17-

# 4. DERIVATION OF RESULTS

In this section, we prove that the optimization problem ECP and the STEM equilibrium model are equivalent and show that the STEM model has a solution which is unique.

Rather than establishing these results, as is possible, directly from the formulation of ECP given in section 3 [see Safwat (1982) for these proofs], we derive and study another optimization problem ECP\*, whose form is more likely to be familiar to the reader. In order to formulate ECP\*, we first derive several properties of our demand functions. More specifically, we show, even with very mild assumptions imposed upon the problem data, that the demand function has an inverse, and that the Jacobian matrix of the inverse demand function is symmetric and negative definite.

For notation, let  $u^{i} = (u_{ij}; j \in D_{i})$  be the vector of travel costs from origin i to its destination  $D_{i}$ . Similarly let  $T^{i}(u^{i}) = (T_{ij}(u^{i}); j \in D_{j})$  denote the vector of trips distributed from origin i as a function of travel costs and let  $t^{i} = (t_{ij}; j \in D_{i})$  denote a given vector of trip distributions.

PROPOSITION 4.1 For each origin i, the demand function  $T^{i}(u^{i})$  with components

$$T_{ij}(u^{i}) = (\alpha S_{i} + E_{i}) \frac{exp(-\theta u_{ij} + A_{j})}{\sum_{\substack{k \in D_{i}}} (-\theta u_{ik} + A_{k})}$$
(4.1)

has an inverse  $u^{i}(t^{i}) = (u_{ij}(t^{i}); j \in D_{i})$  over the domain  $t^{i} = (t_{ij}; j \in D_{i}) > 0$ . The components  $u_{ij}(t^{i})$  of  $u^{i}(t^{i})$  are given by

$$u_{ij}(t^{i}) = \frac{1}{\theta} \left[ A_{j} - \ln t_{ij} + \ln \sum_{k \in D_{i}} t_{ik} - \frac{1}{\alpha} \sum_{k \in D_{i}} t_{ik} - E_{i} \right] \cdot$$
(4.2)

-18-

PROOF:

Let  $T_{ij}(u^i) = t_{ij}$  be given. Adding equations (4.1) over the destinations accessible from origin i and substituting for S<sub>i</sub> gives

$$\sum_{j \in D_{i}} t_{ij} = \alpha S_{i} + E_{i} = \alpha \ln \sum_{j \in D_{i}} \exp(-\theta u_{ij} + A_{j}) + E_{i}$$

Using the leftmost equality to substitute for  $\alpha S_i + E_i$  in (4.1) and using the equality of the outermost expressions to substitute for the denominator of (4.1) in terms of the  $t_{ij}$ , gives

$$t_{ij} = \begin{pmatrix} \sum t_{ik} \\ k \in D_i \end{pmatrix} \frac{\exp(-\theta u_{ij} + A_j)}{\exp\left[\frac{1}{\alpha} \begin{pmatrix} \sum t_{ik} - E_i \\ k \in D_i \end{pmatrix}\right]}$$

Taking the natural logarithm of both sides of this equality and solving for  $u_{ij}$  gives (4.2).

PROPOSITION 4.2 For each *i*, the Jacobian  $\forall u^{i}(t^{i})$  of the inverse demand function  $u^{i}(t^{i})$  is symmetric.

PROOF: If  $\ell \neq j$ , then

$$\frac{\partial \mathbf{u}_{ij}}{\partial \mathbf{t}_{ik}} = \frac{1}{\sum_{\substack{k \in D_i \\ ik}} t_{ik}} - \frac{1}{\alpha}$$
(4.3)

Therefore,  $\frac{\partial u_{ij}}{\partial t_{i\ell}} = \frac{\partial u_{i\ell}}{\partial t_{ij}}$  for all j and  $\ell$ .

PROPOSITION 4.3 For each *i*, the Jacobian  $\forall u^{i}(t^{i})$  of  $u^{i}(t^{i})$  is negative definite at any point  $t^{i} = (t_{ij}) > 0$  satisfying  $\sum_{\substack{k \in D_{i}}} t_{ik} \geq \alpha$ .

PROOF: First note that for any fixed origin i

$$\frac{\partial^{\mathbf{u}}\mathbf{i}\mathbf{j}}{\partial^{\mathbf{t}}\mathbf{i}\mathbf{j}} = -\frac{1}{\mathbf{t}\mathbf{i}\mathbf{j}} + \frac{1}{\sum_{\substack{k \in \mathbf{D}_{\mathbf{i}}}} \mathbf{t}\mathbf{i}\mathbf{k}} - \frac{1}{\alpha}.$$

which combined with (4.3) shows that

$$\nabla u^{i}(t^{i}) = \left(\frac{1}{\sum\limits_{k \in D_{i}} t_{ik}} - \frac{1}{\alpha}\right) E - D$$

where E is a matrix of ones and D is a diagonal matrix with diagonal entries  $d_{jj} = \frac{1}{t_{ij}}$ .

Thus for any  $|D_i|$ -dimensional column vector y,

$$y^{T} \nabla u^{i}(t^{i})y = \left(\frac{1}{\sum_{j \in D_{i}} t_{ij}} - \frac{1}{\alpha}\right) \left(\sum_{j} y_{j}\right)^{2} - \sum_{j} \frac{1}{t_{ij}} y_{j}^{2}.$$

This expression is negative whenever  $y \neq 0$ ,  $\sum_{j \in D_i} t_{ij} \geq \alpha$ , and each  $t_{ij} > 0$ .

Since each inverse demand function  $u^{i}(t^{i})$  defined by (4.2) has a symmetric Jacobian, it can be integrated as a line integral. Also, since the Jacobian is negative definite, if mild assumptions are imposed upon the data, the integral is a concave function.

Based on these properties of the demand functions, we can formulate the following optimization problem.

ECP\*: Minimize 
$$\phi(H,t) \equiv \sum_{a o} \int_{a}^{F_{a}} C_{a}(w) dw - \sum_{i o} \int_{a}^{t^{i}} u^{i}(\tau^{i}) d\tau^{i}$$
 (4.4)

Subject to: 
$$\sum_{p \in P_{ij}} H_p = t_{ij}$$
 for all i,j (4.5)

$$\sum_{j \in D_i}^{t} t_{ij} \ge E_i \quad \text{for all } i \tag{4.6}$$

$$t_{ij} > 0$$
 for all i,j (4.7)

$$H_p \ge 0$$
 for all p (4.8)

$$\mathbf{F}_{\mathbf{a}} \equiv \sum_{\mathbf{i},\mathbf{j}} \sum_{\mathbf{p} \in \mathcal{P}_{\mathbf{i},\mathbf{j}}} \delta_{\mathbf{a}\mathbf{p}} \mathbf{H}_{\mathbf{p}} \cdot$$

This formulation comes closer than ECP to the formulation stated by Dafermos (1980) and, in somewhat different form, by Aashtiani (1979), which are essentially generalizations of the usual optimization problem stated by Beckmann et al (1956). However, the ECP\* problem differs from that stated by Dafermos (1980) in several ways. First, the inverse demand function u(t) decomposes into a separate inverse  $u^{i}(t^{i})$  for each origin i; as a result, the single line integral  $\int_{0}^{t} u(\tau) d\tau$  in her formulation becomes a sum of line integrals, one defined for each origin. Second, this formulation contains the additional constraint (4.6) imposed upon the trips  $t_{ij}$  made from origin i. Third, the variables  $t_{ij}$  are restricted to be positive, rather than merely nonnegative. Moreover, our formulation uses a specific functional form for  $u^{i}(t^{i})$ , rather than a more general and unspecified form used by Dafermos.

<sup>&</sup>lt;sup>†</sup> The inverse demand functions  $u^{i}$  are not defined if any component of  $t^{i}$  equals zero. Therefore, let the lower limit on the line integral in (4.4) be defined by setting each component of  $\tau$  at value  $\varepsilon > 0$  and interpret the line integral as the limit as  $\varepsilon$  approaches zero.

In fact, our main purpose in introducing the line integral formulation (ECP\*) is to show that the symmetry conditions required on demand functions are not nearly as restrictive as previously thought.

Below we establish equivalency between ECP\* and STEM, and prove existence and uniqueness of equilibrium on our STEM model. We then note that ECP and ECP\* are indeed equivalent.

THEOREM 4.4 Assume that  $\theta$  and  $\alpha$  are positive. Then the Kuhn-Tucker conditions for ECP\* are identical with the STEM model.

PROOF: Let  $v_{ij}$  and  $\lambda_i \ge 0$  be Kuhn-Tucker multipliers for constraints (4.5) and (4.6). Let  $k_i = |D_i|$  and let  $v^i = (v_{i1}, v_{i2}, \dots, v_{ik_i})$ . Also, let e be a vector of ones with |I| components. Then the Kuhn-Tucker conditions for ECP\* are:

$$\frac{\partial \Phi}{\partial H}_{\mathbf{p}} - \mathbf{v}_{\mathbf{ij}} \geq 0 \qquad \text{for all } \mathbf{i}, \mathbf{j}, \mathbf{P} \qquad (4.9)$$

$$\begin{bmatrix} \frac{\partial \Phi}{\partial H_p} & -v_{ij} \end{bmatrix} H_p = 0 \qquad \text{for all } i,j,P \qquad (4.10)$$

$$\nabla_{t^{i}} \Phi(\mathbf{H}, t) + \mathbf{v}^{i} - \lambda_{i} \mathbf{e} = 0 \quad \text{for all } i \tag{4.11}$$
$$\lambda_{i} \left[ \sum_{j \in D_{i}} t_{ij} - E_{i} \right] = 0 \quad \text{for all } i \tag{4.12}$$

and (4.5) - (4.8).

Since each t<sub>ij</sub> > 0, equation (4.5) implies that H<sub>p</sub> > 0 for some path P joining origin i to destination j. This fact, and the fact that  $\frac{\partial \phi}{\partial H_p}$  is given by

$$\frac{\partial \Phi}{\partial H_{p}} = \sum_{a} \delta_{ap} C_{a}(F_{a}),$$

implying from (4.9) and (4.10) that  $v_{ij} = \min \left[\sum_{a} \delta_{ap} C_a(F_a)\right]$  where the minimum is taken over all paths p joining origin i and destination j. These are the mode split and traffic assignment conditions of STEM.

Since 
$$\nabla_t i\phi(H,t) = -u^i(t^i)$$
, conditions (4.11) become  
 $u^i(t^i) = v^i - \lambda_i e$ .

Letting  $u_{ij} = v_{ij} - \lambda_i$  and inverting  $u^{i}(t^{i})$ , we have from (4.1)

$$t_{ij} = (\alpha S_i + E_i) \frac{\exp(-\theta v_{ij} + \theta \lambda_i + A_j)}{\sum_{\substack{k \in D_i \\ k \in D_i}} \exp(-\theta v_{ik} + \theta \lambda_i + A_k)}$$
(4.13)

where

$$S_{i} = \ln \sum_{j \in D_{i}} \exp \left(-\theta v_{ij} + \theta \lambda_{i} + A_{j}\right)$$
$$= \theta \lambda_{i} + \ln \sum_{j \in D_{i}} \exp \left(-\theta v_{ij} + A_{j}\right). \qquad (4.14)$$

Eliminating the common factor  $\exp(\theta \lambda_i)$  from its numerator and denominator, we see that expression (4.13) reduces to the trip distribution function (2.7) of STEM with  $v_{ij}$  in place of  $u_{ij}$ . Summing (4.13) over  $j \epsilon D_i$  and invoking (4.6) gives

 $\alpha S_{i} + E_{i} = \sum_{j \in D_{i}} t_{ij} \ge E_{i}.$ 

Since  $\alpha > 0$ , this inequality implies that  $S_i \ge 0$ . Moreover, if  $\lambda_i > 0$ , the complementarity conditions (4.12) implies that the last inequality holds as an equality and thus  $S_i = 0$ . But then, these facts, the hypothesis  $\theta > 0$ , and (4.14) imply that whether  $\lambda_i > 0$  or  $\lambda_i = 0$ ,

$$S_i = \max\{0, \ln \sum_{j \in D_i} \exp(-\theta v_{ij} + A_j)\}$$
.

Consequently, after we have eliminated the variables  $\lambda_i$ , the Kuhn-Tucker conditions (4.5)-(4.12) become the conditions of the STEM model.

This theorem shows that any solution to the STEM model satisfies the Kuhn-Tucker conditions of ECP\*. We have yet, however, to show that ECP\* has an optimal solution or that any solution to STEM corresponds to an optimal solution to ECP\*. The second of these properties requires some additional hypothesis on the performance functions  $C_a$  and the data  $E_i$  that will insure that ECP\* is a convex program.

# THEOREM 4.5 (Equivalency, Existence, and Uniqueness)

Suppose that  $\theta > 0$ , that  $E_i \ge \alpha > 0$  for all *i*, and that each performance function  $C_a(F_a)$  is real valued and nondecreasing over the domain  $F_a \ge 0$ . Then STEM and ECP\* have a solution and solutions to STEM correspond in a one-to-one fashion with the optimality conditions to ECP\*. The performance costs  $u_{ij}$  and trip distributions  $t_{ij}$  in the STEM model are unique. If  $C_a(F_a)$  is strictly increasing, the arc flows  $F_a$  of STEM are unique as well.

# PROOF:

# Equivalency:

Since each  $E_i \geq \alpha$ , proposition 4.3 shows that  $\nabla u^i(t^i)$  is negative definite over the feasible region to ECP\*. Consequently, if  $C_a(F_a)$  is nondecreasing, the objective function of ECP\* is convex over the feasible region. But then, ECP\* is a convex program and its Kuhn-Tucker conditions are sufficient as well as necessary. Therefore, Theorem 4.4 demonstrates the one-to-one correspondence between solutions to STEM and the Kuhn-Tucker conditions to ECP\*.

-24-

# Uniqueness:

Since  $\nabla u^{i}(t^{i})$  is negative definite, its line integral in (4.4) is strictly concave as a function of  $t^{i}$ , implying that the optimal values of  $t^{i}$  in ECP\* are unique. Since the Jacobian of the constraints (4.5) and (4.6) of ECP\* has full row rank, the Kuhn-Tucker variables associated with any given optimal solution to the problem are unique. But since the Kuhn-Tucker variables for any convex program are independent of its optimal solutions since they are subgradients for the perturbation function of that problem (see Rockafellar [1970]), the Kuhn-Tucker variables  $v_{ij}$  to (4.5) and  $\lambda_{i}$ to (4.6) are unique and hence so are the  $u_{ij}$ 's given by  $u_{ij} = v_{ij} - \lambda_{j}$ .

If  $C_a(F_a)$  is strictly increasing, then the objective function in ECP\* is strictly concave as a function of the arc flows  $F_a$  and so these arc flows are unique.

# Existence:

To complete the proof, we must show that ECP\* has a solution. First, let us evaluate the line integral in (4.4).

Let  $L_{ik} = \{\tau_{ij}: \tau_{ij} = t_{ij} \text{ for } j < k, 0 \leq \tau_{ik} \leq t_{ik}, \text{ and } \tau_{ij} = 0 \text{ for } j > k\}.^{\dagger}$ For each i, we evaluate the line integral over the path defined by  $L_{i1}, L_{i2}, L_{i3}, \dots$ In the region  $L_{ik}$ ,

$$\int_{\mathbf{L}_{j \leq k}} u^{\mathbf{i}} (\tau^{\mathbf{i}}) d\tau^{\mathbf{i}} = \frac{1}{\theta} \left[ A_{\mathbf{k}} \tau_{\mathbf{i}\mathbf{k}} - \tau_{\mathbf{i}\mathbf{k}} \ln \tau_{\mathbf{i}\mathbf{k}} + \tau_{\mathbf{i}\mathbf{k}} + \tau_{\mathbf{i}\mathbf{k}} + \tau_{\mathbf{i}\mathbf{k}} \right]$$

$$+ \left( \sum_{j \leq k} t_{\mathbf{i}j} + \tau_{\mathbf{i}\mathbf{k}} \right) \ln \left( \sum_{j \leq k} t_{\mathbf{i}j} + \tau_{\mathbf{i}\mathbf{k}} \right) - \tau_{\mathbf{i}\mathbf{k}} - \frac{1}{2\alpha} \left( \tau_{\mathbf{i}\mathbf{k}}^2 - 2\left[ \sum_{j \leq k} t_{\mathbf{i}j} - E_{\mathbf{i}} \right] \tau_{\mathbf{i}\mathbf{k}} \right) \right] \right]$$

$$= \frac{1}{\theta} \left[ A_{\mathbf{k}} t_{\mathbf{i}\mathbf{k}} - t_{\mathbf{i}\mathbf{k}} \ln t_{\mathbf{i}\mathbf{k}} - \frac{1}{2\alpha} \left[ t_{\mathbf{i}\mathbf{k}}^2 + 2 \left( \sum_{j \leq k} t_{\mathbf{i}j} - E_{\mathbf{i}} \right) t_{\mathbf{i}\mathbf{k}} \right] + \left( \sum_{j \leq k} t_{\mathbf{i}j} \right) \ln \left( \sum_{j \leq k} t_{\mathbf{i}j} \right) \right]$$

$$= \left( \sum_{j \leq k} t_{\mathbf{i}j} \right) \ln \left( \sum_{j \leq k} t_{\mathbf{i}j} \right) \right] .$$

<sup>†</sup> More precisely define L with  $\tau_{ij} = \varepsilon$  for j > k and later take the limit as  $\varepsilon_{ik}$  approaches zero in (4.15). See the previous footnote.

Summing these integrals for k = 1, 2, ..., we see that the last two terms telescope and, therefore, that the objective function  $\phi(H,t)$  in (4.4) becomes

$$\phi(\mathbf{H},\mathbf{t}) = \sum_{\mathbf{a}} \int_{0}^{\mathbf{F}_{\mathbf{a}}} \mathbf{C}_{\mathbf{a}}(\mathbf{w}) d\mathbf{w} + \frac{1}{\theta} \left[ \sum_{\mathbf{i}} \left\{ \sum_{j}^{\Gamma} \mathbf{A}_{j} \mathbf{t}_{\mathbf{i}j} + \sum_{j}^{\Gamma} \mathbf{t}_{\mathbf{i}j}^{2n} \mathbf{t}_{\mathbf{i}j} - \left(\sum_{j}^{\Gamma} \mathbf{t}_{\mathbf{i}j}\right)^{2n} \left(\sum_{j}^{\Gamma} \mathbf{t}_{\mathbf{i}j}\right) + \frac{1}{2\alpha} \left( \sum_{j}^{\Gamma} \mathbf{t}_{\mathbf{i}j} - \mathbf{E}_{\mathbf{i}}^{2} \right)^{2} - \mathbf{E}_{\mathbf{i}}^{2} \right] .$$

$$(4.15)$$

Now define  $\phi(H,t)$  when any  $t_{ij} = 0$  by setting 0ln0 = 0 and let  $t_{ij} \ge 0$  replace the constraints  $t_{ij} > 0$  in ECP\*. Then the feasible region becomes closed and  $\phi(H,t)$  is continuous on the feasible region. First note that this modified problem has an optimal solution. For if  $(H_p^k, t_{ij}^k)$  for k = 1, 2, ... is any sequence of feasible solutions whose norms approach  $+\infty$ , then some  $t_{ij}$  approaches  $+\infty$ . But then since the first term in the definition of  $\phi(H,t)$  is nonnegative and since the quadratic term in (4.15) is asymptotically dominant in the  $t_{ij}$ terms,  $\phi(H^k, t^k)$  approaches  $+\infty$ . But this norm condition implies that the modified problem has an optimal solution (for example, see Ortega and Reinboldt [1970, theorem 4.3.3]).

Thus far we have established existence of a solution with  $t_{ij} \ge 0$ . We next show that each  $t_{ij} > 0$  in any optimal solution to the modified problem. Let  $(H_p^*, t_{ij}^*)$  be any feasible solution to the modified problem with some  $t_{ij}^* = 0$ . Let  $(H_p^*, t_{ij}^*)$  be any other feasible solution with all  $t_{ij}^* > 0$ . Consider the solutions

$$(H_{p}(z), t_{ij}(z)) = (1-z)(H_{p}^{*}, t_{ij}^{*}) + z(H_{p}^{'}, t_{ij}^{'})$$

for 0 < z < 1. Then letting  $\phi(z) = \phi(H(z),t(z))$ , we see that

$$\frac{d\phi(z)}{dz} = \sum_{a} \delta_{ap} C_{a}(F_{a}) (H_{p}' - H_{p}'') - \sum_{i j} \sum_{\theta} \frac{1}{\theta} \left[ A_{j} - \ln t_{ij}(z) + \ln \sum_{j \in D_{i}} t_{ij}(z) \right] (t_{ij}' - t_{ij}'')$$
$$- \sum_{i} \left\{ \frac{1}{\alpha} \left[ \sum_{j \in D_{i}} t_{ij}(z) - E_{i} \right] \sum_{j} (t_{ij}' - t_{ij}') \right\}.$$

Since some  $t_{ik}^* = 0$  this derivative approaches  $-\infty$  (via the term  $[\ln t_{ik}(z)][t_{ik}^* - t_{ik}^*] = t_{ik}^* [\ln t_{ik}(z)])$  as z approaches 0 and thus  $t_{ik}(z)$ approaches  $t_{ik}^* = 0$ . (Note that no term in this derivative approaches  $+\infty$ because  $\sum_{j \in D_i} t_{ij}(z) \ge E_i$ ). But then, since  $\phi$  is continuous,  $\phi(H(z), t(z)) < \phi(H^*, t^*)$  for sufficiently small but positive values of z. Consequently,  $t_{ij} > 0$  in any optimal solution to the modified problem and thus ECP\*, and so STEM, has a solution. []

We conclude this section by noting that ECP\* and ECP are alternate forms of the same problem; thus, as a byproduct of this section, we have established the equivalence between ECP and STEM. To see that ECP and ECP\* are equivalent, use the constraint  $\sum_{j \in D_i} T_{ij} = \alpha S_i + E_i$  in ECP to solve for  $S_i$  in terms of  $T_{ij}$ . Substituting this value for  $S_i$  in the objective function to ECP gives (4.15) plus the constant  $\frac{1}{2\alpha} \sum_{i} E_i^2 - \sum_{i} E_i$ . Consequently, the two objective functions differ only by a constant. Since  $\alpha > 0$ , the constraints  $\sum_{j \in D_i} T_{ij} = \alpha S_i + E_i$  and  $S_i \ge 0$  are equivalent to the constraints (4.6) of ECP\*. Therefore, ECP and ECP\* are equivalent optimization problems.

# 5. SPECIAL AND LIMITING CASES

In this section we illustrate the generality and the range of applications of the STEM model. We first show that a singly constrained gravity model with an exponential delay function may be used within the STEM model to describe trip distribution. This trip distribution model is a special case of the more general logit model. We also show that the STEM model can be used to approximate as closely as desired any given doubly constrained gravity model with fixed productions and attractions.

Let  $D_j > 0$  be the number of trips attracted to destination j. Also let  $A_i = lnD_i$ . Then the distribution model (2.7) becomes

$$T_{ij} = G_{i} \frac{\frac{D_{j}e^{-\theta u}ij}{\sum_{k} D_{k}e^{-u}ik}}{\sum_{k} D_{k}e^{-u}ik}$$

This is a gravity model with an exponential delay function.

Now suppose that the number of trips generated at an origin i,  $0_i > 0$ , is fixed, the number of trips  $D_j$  attracted to any j is fixed, and that  $\sum_{i} 0_i = \sum_{j} D_j$ .

We show that by a judicious choice of the data  $A_j$ ,  $\alpha$  and  $E_i$ , the STEM model approximates these productions and attractions as  $\theta$  approaches 0.

First note that if all costs C are nonnegative, then all  $u_{ij}$  are nonnegative. Thus, if  $\theta > 0$ ,

$$S_{i} = \ln \sum_{j \in D_{i}}^{-\theta u_{ij}} e^{+A_{j}} \leq \ln \sum_{j \in D_{i}}^{A_{j}} e^{-\beta u_{ij}}$$

Therefore,

$$\sum_{i} G_{i} \leq K \equiv \sum_{i} (\alpha \ln \sum_{j \in D_{i}} e^{A_{j}} + E_{i}).$$

Assuming that  $C_{a}(F_{a})$  is continuous implies that

 $K' = \max \max \min C_{p}(F)$ i,j  $0 \le F_a \le K p \in P_{ij}$  exists. Here  $P_{ij}$  denotes the set of available paths joining origin i and destination j and  $C_p(F) = \sum_{a} \delta_{ap} C_a(F_a)$ . Since  $u_{ij} \leq C_p(F)$  for any  $p \in P_{ij}, 0 \leq u_{ij} \leq K'$ . Therefore,

$$\begin{array}{c} - \theta u_{ij} + A_{j} \leq A_{j}, \\ \theta u_{ij} + A_{j} \geq -\theta K' + A_{j}, \end{array}$$

and as  $\theta$  approaches 0,

$$-\theta u_{ij} + A_{j} \rightarrow A_{j}$$

Consequently,

$$\exp (-\theta u_{ij} + A_{j}) \text{ approaches } \exp (A_{j}) \text{ and } S_{i} = \ln \sum_{j \in D_{i}} e^{-\theta u_{ij} + A_{j}} \text{ approaches}$$

$$\ln \sum_{j} e^{j} \text{ as } \theta \text{ approaches zero. Thus}$$

$$T_{ij} = (\alpha S_{i} + E_{i}) \frac{\exp(-\theta u_{ij} + A_{j})}{\sum_{k} \exp(-\theta u_{ij} + A_{k})}$$

approaches

$$T_{ij}^{*} = (\alpha \ln \sum_{j} e^{A_{j}} + E_{i}) \frac{\exp(A_{j})}{\sum_{k} \exp(A_{k})}$$

Now let  $A_j = \ln D_j$ , let  $\alpha > 0$  be chosen sufficiently small so that  $\alpha \ln \sum_j D_j < O_j$  for all i, and let  $E_i = O_j - \alpha \ln \sum_j D_j$ . Then  $\sum_j T_{ij}^* = \alpha \ln \sum_j D_j + E_i = O_j$  for all i and  $\sum_i T_{ij}^* = (\sum_i O_i) \frac{D_j}{\sum_k D_k} = D_j$  for all j.

Therefore for  $\theta > 0$ , but sufficiently small, the STEM model approximates the doubly constrained gravity model as closely as desired.

# 6. SUMMARY AND CONCLUSIONS

Review of previous studies illustrates the tradeoffs between the realistic (behavioral), technical (convergence), and practical (computational efficiency) considerations in modelling the equilibrium problem. None of these studies has been successful in addressing all aspects of the problem.

In this paper, we have presented an equilibrium model, STEM, that permits trip generation, trip distribution, modal split and traffic assignment on transportation networks to be predicted simultaneously by convergent and computationally efficient algorithms. In the STEM model, trip generation can depend upon the system's performance through an accessibility measure that is based on the random utility theory of users' behavior, and trip distribution is given by a logit model. Modal split and traffic assignment are user optimized; however, several alternative assumptions can be considered within the framework presented in this paper. We have proven existence and uniqueness of equilibrium by formulating an equivalent optimization problem and studying its qualitative characteristics. We have demonstrated the richness of the STEM model by showing that it can approximate as closely as desired other commonly used demand models which may be thought of as limiting or special cases.

Safwat (1982) studies computational aspects of this equilibrium problem and the formulation ECP. Currently, a joint project between M.I.T. and Cairo University, in cooperation with the Egyptian Ministry of Transport, is applying an extended version of the model to the intercity multimodal transport system of Egypt.

-30-

#### REFERENCES

- 1. Aashtiani H. Z. (1979) The multi-modal traffic assignment problem. Ph.D. Thesis, Operations Research Center, Massachusetts Institute of Technology.
- Aashtiani H. Z. and Magnanti T. L. (1981) Equilibria on a congested transportation network. SIAM Journal on Algebraic and Discrete Methods, 2(3), 213-226.
- 3. Asmuth R. L. (1978) Traffic Network Equilibrium. Technical Report SOL-78-a, Stanford University, Stanford, California.
- 4. Beckmann M., McGuire C. B. and Winston C. B. (1956) Studies in The Economics of Transportation. Yale University Press, New Haven, Connecticut.
- 5. Ben-Akiva M. and Lerman S. R. (1977) Disaggregate travel demand and mobility choice models and measures of accessibility. Presented at the 3rd International Conference on Behavioral Demand Modelling, Australia.
- 6. Bertsekas D. P. and Gafni E. M. (1981) Projected Newton Methods and Optimization of Multicommodity Flows. Working Paper No. LIDS-P-1140, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Mass.
- 7. Bruynooghe M., Gibert A. and Sakorovitch M. (1968) Une methode d'affectation du traffic. Persented at the 4th International Symposium on the Theory of Traffic Flow, Karlsruhe.
- 8. Cairo Urban Transport Project (1981) Technology Adaptation Program, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- 9. Chicago Area Transport Study (1960), Final Report.
- 10. Dafermos S. (1981) The general multimodal network equilibrium problem. Submitted to *Networks*.
- 11. Dafermos S. (1980) Traffic equilibrium and variational inequalities. Transportation Science 14(1), 42-54.
- 12. Daganzo C. F. (1979) Multinomial Probit, Academic Press, New York.
- 13. Dalvi M. Q. and Martin K. M. (1976) The measurement of accessibility: some preliminary results. *Transportation* 5(1), 17-42.
- 14. Dembo R. S. and Klincewicz J. G. (1981) A scaled reduced gradient algorithm for network flow problems with convex separable costs. *Mathematical Programming Study* 15, 125-147.
- 15. Detroit Metropolitan Area Traffic Study (1955).
- 16. Domencich T. and McFadden D (1975) Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam.
- 17. Evans S. P. (1976) Derivation and analysis of some models for combining tr ρ distribution and assignment. *Transpn. Res.* 10, 37-57.

- 18. Fisk C. and Nguyen S. (1980) Solution algorithms for network equilibrium models with asymmetric user costs. Publication No. 167, Centre de recherche sur les transports, Universite de Montreal, Montreal, Canada.
- 19. Florian M. and Nguyen S. (1978) A combined trip distribution mode split and trip assignment model. *Transpn. Res.* 12, 241-246.
- 20. Florian M. and Nguyen S. (1974) A method for computing network equilibrium with elastic demands. *Transportation Science* 8(4), 321-332.
- 21. Golden B. (1975) A minimum-cost multi-commodity network flow problem concerning imports and exports. *Networks* 5, 331-356.
- 22. Gumbel E. J. (1958) Statistics of Extremes. Columbia University Press, New York.
- 23. Hearn D. W. and Kuhn H. (1977) Network Aggregation in Transportation Planning--Final Report. Mathtech, Inc., Princeton, New Jersey.
- 24. Leblanc L. J. (1973) Mathematical programming algorithms for large scale network equilibrium and network design problems. Ph.D. Thesis, Department of Industrial Engineering and Management Sciences, Northwestern University.
- 25. Nguyen S. (1976a) A Mathematical Programming Approach to Equilibrium Methods of Traffic Assignment with Fixed Demands. Publication #17, Centre de Recherche sur Les Transports, Universite de Montreal, Montreal, Canada.
- 26. Nguyen S. (1976b) Equilibrium Traffic Assignment Procedures with Elastic Demands. Publication #39, Centre de Recherche sur Les Transports. Universite de Montreal, Montreal, Canada.
- 27. Nguyen, S. (1974) An algorithm for the traffic assignment problem. Transportation Science 8(3), 203-216.
- 28. Ortega J. M. and Reinboldt W. C. (1970) Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York.
- 29. Pang J. S. and Chan D. (1981) Iterative Methods for Variational and Complementarity Problems. Technical Report GSIA, Carnegie-Mellon University, Pittsburgh, Pa.
- 30. Rockafellar, R. T. (1970) The simultaneous prediction of equilibrium on Princeton, New Jersey,
- 31. Safwat K. N. A. (1982) The simultaneous prediction of equilibrium on large-scale networks: a unified consistent methodology for transportation planning. Ph.D. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology (in progress).
- 32. Sheffi Y. and Daganzo C. F. (1980) Computation of equilibrium over transportation networks: the case of disaggregate demand models. *Transportation Science* 14(2), 155-173.

- 33. Smith M. J. (1979) The existence, uniqueness and stability of traffic equilibria. *Transpn. Res.* 13B, 295-304.
- 34. U. S. Federal Highway Administration (1972) Urban Transportration Planning: General Information. U. S. Department of Transportation.
- 35. U. S. Federal Highway Administration (1970) Urban Transportation Planning: General Information and Introduction to System 360. U. S. Department of Transportation.
- 36. U. S. Urban Mass Transit Authority (1976) Urban Transportation Planning System--Reference Manual. U. S. Department of Transportation.
- 37. Williams H. C. W. L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A* 9, 285-344.