Dynamic Traffic Equilibrium with Discrete/Continuous Econometric Models

FRED L. MANNERING

University of Washington, Seattle, Washington 98195

SAMEER A. ABU-EISHEH

An-Najah National University, Nablus, West Bank, Israel

AUDUR T. ARNADOTTIR

University of Washington, Seattle, Washington 98195

A procedure is developed to simulate the spatial and temporal distribution of traffic flows on a simplified highway network for home-to-work commuter trips. The procedure is based on a discrete/continuous econometric framework in which travelers' choices of route and departure time are modeled. Equilibrium traffic flows are determined using the estimated econometric models in a classic supply/demand equilibrium. The potential usefulness of the approach is demonstrated through a number of simulation runs in which equilibrium traffic flows are determined under network capacity constraints and alternate signal timing strategies.

INTRODUCTION

Urban traffic congestion has consistently ranked as one of the most perplexing problems facing transportation analysts. At the heart of the problem is the dominance of single-occupant vehicle travel and the enormous social, economic and environmental costs associated with the construction of additional highway system capacity. Since single-occupant vehicle dominance and high construction costs are likely to persist well into the future, congestion can only be viewed as a problem of ever growing importance.

Researchers have long recognized that the analysis of urban traffic congestion must be predicated upon theoretically consistent models of travelers' response to traffic congestion. Traditionally, such modeling has focused solely on travelers' choice of route whereas other possible responses to congestion, such as changes in departure time and travel speed, were virtually ignored due, in large part, to the additional modeling complexity that their consideration warranted. Consequently, travelers' route choice was modeled with equilibrium assignment techniques, but the modeling approach was inherently static due to the fact that travelers' time-variant decisions of departure time and travel speed were not addressed. In recognition of the deficiencies of traditional modeling approaches, an entire body of literature has evolved to dispense with the static equilibrium assumptions of the past (BEN-AKIVA[4]). One branch of this literature has sought to incorporate traveler choice dynamics in an econometric framework by permitting the choice of departure times from a set of discrete time intervals (COSLETT,[10] ABKOWITZ,[1] HENDRICKSON and PLANK[17]). Related research has considered departure time within the context of rather rigorous definitions of user equilibrium (HENDRICKSON and KOCUR,[15] HENDRICKSON, NAGIN and PLANK,[16] FARGIER,[13] DE PALMA et al.[11]), and extensions of this approach, to include the choice of route, have provided fruitful results (MAHMASSANI and HERMAN[18]). More classic econometric-based approaches have also enjoyed some popularity and have provided a rich behavioral basis for departure-time/route-choice modeling (BEN-AKIVA, DE PALMA and KANAROGLOU,[6] ABU-EISHEH and MANNERING,[2] MANNERING[22]). Finally, exten-

Transportation Science
Vol. 24, No. 2, May 1990

0041-1655/90/2402-0105 $01.25

sive theoretical and empirical investigations of the dynamics of the thought process by which commuters arrive at a satisfactory choice of route and departure time have been successfully undertaken (MAHMASSANI and CHANG,[19,20] CHANG and MAHMASSANI,[9] MAHMASSANI and TONG[21]).

Overall, research on the varying aspects of traveler choices and dynamic equilibrium, as listed above, can best be characterized as slowly evolving, with the speed of evolution inhibited by the inherent complexity of the problem. The intent of this paper is to continue the advance of the field, but to do so by taking a theoretical and empirical approach that departs somewhat from the evolutionary course of the literature. Three key components of the current paper distinguish it from previous efforts: 1) departure time is viewed as continuous instead of discrete, 2) travelers are assumed to have some control over vehicle speed and consequently trip travel time, and 3) spatial and temporal equilibrium is defined in a classic economic sense and operationalized through the constraints imposed by the discrete/continuous econometric structure. While the development and demonstration of an operational dynamic equilibrium model are the primary objectives of the paper, the paper's departure from some of the accepted conventions of the field has the important side effect of encouraging researchers to rethink fundamental assumptions and to consider new aspects and dimensions of the dynamic equilibrium problem. This side effect may ultimately serve as the most significant contribution of this paper.

The paper begins with an overview of the empirical setting that will form the basis of subsequent model estimation and simulation experiments. This is followed by a presentation of the demand model structure, a discussion of specification issues, and estimation results. Equilibrium definitions are then specified and an appropriate equilibrium algorithm is outlined. The dynamic equilibrium model is then demonstrated through numerous simulation experiments and, finally, a summary of findings is made and appropriate conclusions are drawn.

1. EMPIRICAL SETTING

TO ENABLE the reader to better understand the econometric specification issues that will be addressed in the next section of this paper, we depart from the traditional theory-first format and provide a description of the empirical setting that will serve as the basis for model estimation and subsequent simulation experiments. To study travelers' choice of route and time-varying choices of departure time and travel speed, a survey of 151 morning commuters was conducted in State College, Pennsylvania in the spring of 1986. The survey was designed to collect a wide variety of data on the traveler's most recent trip to work including route choice, make, model and vintage of vehicle used, vehicle occupancy, departure time, arrival time, work start time, and preferred arrival time at work. In addition, general socioeconomic information was collected including income, age, sex, marital status, occupation and number of children.

The approach of the survey was to concentrate on a single origin–destination pair. The selected origin was a large residential development (Toftrees) located in suburban State College, and the destination was the highly concentrated central business district of State College and the adjacent campus of the Pennsylvania State University. As illustrated in Figure 1, three distinct and diverse routes connect the origin-destination pair; Atherton Street (a four-lane major arterial), Fox Hollow Road (a two-lane rural road), and the Route 322 by-pass (a four-lane expressway with a one-lane exit ramp). It is important to note that since the precise location of home and work was collected in the traveler survey, the actual point-to-point distances from work-to-home, on each of the three routes, can be computed and will vary from traveler to traveler. A summary of traveler point-to-point distances and basic route characteristics is given in Table I.

To supplement the data collected in the commuter survey, traffic-related data was gathered, via field observations, including traffic flow rates, peak hour volumes and traffic signal characteristics (phasings, cycle lengths, green times). The combination of the commuter survey and field observations provides a fairly rich and comprehensive data source and one that is well suited to the study of commuter responses to congestion.

2. DEMAND MODEL STRUCTURE AND SPECIFICATION ISSUES

WE BEGIN model development by restricting commuter options, in response to traffic congestion, to changing departure time and route. Thus, the options of changing mode or canceling the work trip entirely, are excluded. Since State College has abbreviated transit service and work trips are nondiscretionary in nature, the focus on only route and departure time choice is not unreasonable.

In modeling travelers' route and departure time choice, a discrete/continuous econometric structure is developed with the choice of route being discrete and the choice of departure time being continuous. In specifying such a structure, the utility provided by alternate routes is first defined as,

$$V_{kj} = \alpha ETT_{kj} + \beta OPCO_{kj} \tag{1}$$

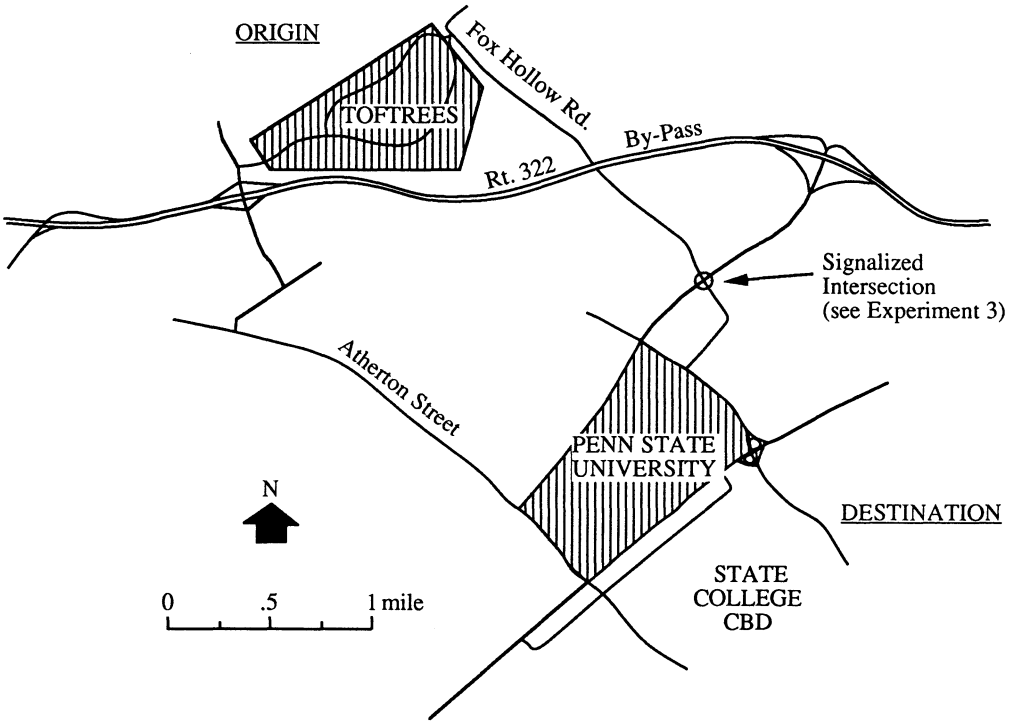

Fig. 1. Map of the study area.

TABLE I
Characteristics of Network Routes

Characteristic	Arterial	Rural Road	Expressway
Number of lanes (total, both directions)	4	2	4[a]
Hourly capacity of the most restrictive section (vph)	1600	933	467
Average point to point distance from origin to destination in miles (standard deviation in parentheses)	4.92 (0.86)	4.45 (0.81)	7.01 (0.87)
Average travel speed in miles per hour (standard deviation in parentheses)	26.9 (7.55)	24.5 (7.24)	37.9 (10.0)
Average travel time in minutes (standard deviation in parentheses)	11.57 (2.75)	11.69 (3.48)	11.66 (2.85)
Ratio of highest observed travel time to free flow travel time	1.77	2.33	1.88

[a] Two lanes at exit.

where V_{kj} is the observable utility provided to traveler k on route j, ETT_{kj} is the expected travel time for traveler k on route j, $OPCO_{kj}$ is the expected vehicle operating costs incurred by traveler k on route j, and α and β are estimable parameters. Note that expected travel time will be a function of route capacity, prevailing traffic flow at the time of departure (instantaneous flow), length of route, and the number of traffic signals on the route and their cycle lengths and effective green times. Vehicle operating costs will be a function of traffic flow at the time of departure and the fuel efficiency characteristics of the vehicle used.

Given this simple linear utility function, an estimable discrete probabilistic route choice model can be specified by adding a disturbance term to Equation 1 to account for unobserved influences on the route choice process (e.g., traveler preferences regarding scenery, pedestrian traffic, and so on) such that the total utility (observed and unobserved) is $U_{kj} = V_{kj} + \epsilon_j$. If the unobserved influences, ϵ_j's, are assumed to be generalized extreme value distributed, the standard multinomial logit form results (McFadden[27]),

$$P_{ki} = \exp[V_{ki}]/\sum_J \exp[V_{kj}] \tag{2}$$

where P_{ki} is the probability that traveler k will select route i from the set of available route choices J. Given this probabilistic form, the coefficients in the utility function (Equation 1) can be readily estimated by standard maximum likelihood methods.

For travelers' continuous choice of departure time, the following identity forms the basis of model estimation,

$$DT_k = WST_k - SD_k - TT_k - WAT_k. \tag{3}$$

Where DT_k is the home-to-work departure time of traveler k, WST_k is the work start time for travelers with fixed work start times and is the preferred arrival time for travelers without fixed work start times, SD_k is the schedule delay, defined as the amount of time between scheduled work start time and actual arrival time (schedule delay is, by definition, zero for travelers without fixed work start times), TT_k is the trip travel time, and WAT_k is the work access time (i.e., walking time from parking location to work). Although travelers could be viewed as having some influence over them, the terms WST_k and WAT_k are assumed to be exogenous to this study, as the data required to accurately model these factors would be extremely difficult to collect. This leaves travel time and schedule delay as the departure time determination factors that are assumed controllable by travelers.

Route travel time has traditionally been assumed to be beyond travelers' control. However, this assumption is only valid under extremely congested conditions as travelers have, in general, some control over their travel time because of their ability to alter driving speeds, risk taking behavior, and reaction times. For empirical convenience, a model of average speed is specified and, since distance is known, average travel time can be readily computed from this model. The speed model is,

$$SPEED_{kj} = \sigma ESP_{kj} + [1 - q_{kj}/c_j]\delta SES_k + \nu_{kj} \quad (4)$$

where $SPEED_{kj}$ is the average origin to destination speed of traveler k on route j, ESP_{kj} is the expected speed of traveler k on route j and is a function of route capacity, traffic flow at departure time (instantaneous flow), and the number of traffic signals on the route and their cycle lengths and effective green times, q_{kj}/c_j, is the volume to capacity ratio of route j at the traveler's departure time, SES_k is a vector of the socioeconomic characteristics of traveler k that affect driving speed, ν_{kj} is a disturbance term, and σ and δ are estimable parameters. The term $1 - q_{kj}/c_j$ in this equation reflects the assertion that as congestion increases, travelers will have less control over their own speeds. It should be mentioned that various nonlinear forms of the restriction that increasing congestion imposes on travelers' choice of speed were empirically tested, but were not found to be statistically superior to the form of Equation 4.

The final component needed to estimate traveler departure time, as determined by Equation 3, is schedule delay. Schedule delay is defined by the linear equation,

$$SD_{kj} = \pi + \psi ESP_{kj} + \Gamma SED_k + \Omega PREF_k + \eta_{kj} \quad (5)$$

where SD_{kj} is the schedule delay, which is defined only for travelers with fixed work start times, ESP_{kj} is the expected speed as defined in Equation 4, SED_k is a vector of socioeconomic characteristics influencing travelers' choice of schedule delay, $PREF_k$ is traveler k's preferred schedule delay assuming no traffic congestion, η_{kj} is a disturbance term, and π, ψ, Γ and Ω are estimable parameters.

The coefficients of Equations 4 and 5 can be estimated by standard regression techniques and those of Equation 2 by maximum likelihood, but two critical specification issues must be addressed. First, it has been mentioned that expected travel time and vehicle operating costs (in Equation 2), expected speed (in Equations 4 and 5), and the q_{kj}/c_j term (in Equation 5), are dependent on instantaneous traffic flow which is itself a function of travelers' departure time. If instantaneous traffic flows for observed traveler departure times were to be used, a potential for selectivity bias in model estimation would exist, since travelers are observed departing at only one time and it is not known what their behavior, with respect to route, speed, and schedule delay, would have been had they departed at some other time and faced different instantaneous traffic flows. Intuitively, there is reason to believe that their behavior may be different due to the nature of the departure time selection process. For example, commuters that choose to depart early will typically be risk averse travelers who may accept long schedule delays and have preferences for driving speeds and routes which differ from those of late departers.

To avert potential estimation bias resulting from this source, the instantaneous flow values used in the estimation of Equations 2, 4 and 5 (the precise use of which will be discussed in detail in the next section) are instrumented by regressing observed instantaneous traffic flows (defined as the traffic flow over a five minute time interval at the time of departure), for each of the three available routes, against all travelers' work start times (for those travelers with fixed work start times) and preferred arrival times (for those travelers without fixed work start times). The use of regression-predicted flow values, which are based on exogenous work start and arrival times, follows the popular indirect method of instrumental variable selectivity bias correction (MANNERING and HENSHER[24]). This type of selectivity bias correction has been shown to be theoretically and empirically valid by a number of studies (see DUBIN and MCFADDEN[12]).

The second specification issue is a more classic discrete/continuous selectivity bias problem. As with departure time, we only observe travelers making one route choice and, since it is unrealistic to assume that the speed and schedule delay behavior of travelers using the expressway, arterial, and rural road will be

identical, a selectivity bias will arise. For example, it might be expected that observed users of the expressway will tend to be faster drivers, in general, since the expressway route offers them the potential to drive at much higher speeds. Expressway drivers may also have an inherent preference for smaller schedule delays which may result from the lower variance in travel time (i.e., fewer traffic signals) offered by the expressway route choice. Given this, on the basis of observed route users, a censored sample exists since, for example, it is not clear as to how fast an expressway user would have driven, or what his/her schedule delay would have been, had he/she selected the arterial or rural route. If an appropriate econometric correction is not used, estimation of Equations 4 and 5 will be biased because users observed taking specific routes constitute a nonrandom sample that is formed from a systematic route selection process.

Although numerous correction techniques are available to avert potential estimation bias from this source (DUBIN and MCFADDEN,[12] MANNERING and WINSTON,[23] MANNERING and HENSHER[24]), the expected value method is the most suitable in this case. In applying this correction method, the expected speed (ESP_{kj}) variable used in Equations 4 and 5 is replaced by its expected value defined as,

$$\mathrm{ESP}_k = \sum_J \mathrm{P}_{kj}\mathrm{ESP}_{kj} \tag{6}$$

where P_{kj} is the probability of traveler k selecting route j as defined by Equation 2 and ESP_k is the expected speed as defined for Equations 4 and 5.

3. ESTIMATION RESULTS

GIVEN THE general econometric structure outlined in the preceding section, attention can now be directed to the specifics of model estimation. For estimation convenience, instantaneous flow is defined as a five-minute volume expanded to an equivalent hourly volume. With this definition of flow, we can precisely define the expected travel time term used in the route choice model (Equation 2) and the expected speed term used in the speed and schedule delay choice models (Equations 4 and 5). Expected travel time consists of two components: travel time on open sections of road and travel time resulting from intersection delay. The open road travel time is estimated using the Bureau of Public Roads (BPR) performance function (see BRANSTON[8]) and intersection delay is estimated by assuming uniform arrivals and departures. This results in the expected travel time term,

$$\mathrm{ETT}_{kj} = t_{j0}\left[1 + \zeta\left(\frac{q_{kj}}{c_j}\right)^{\xi}\right]d_{kj} + \frac{\sum_{\mathrm{M}_{kj}} r_{mj}^2}{2\mathrm{C}_m(1 - \rho_{mj})} \tag{7}$$

where ETT_{kj} is the expected travel time of traveler k on route j in minutes, t_{j0} is the free flow travel time at the prevailing speed limit, for route j, in minutes per mile, q_{kj} is the instantaneous five minute flow expanded to an equivalent hourly flow, faced by traveler k on route j, in vehicles per hour, c_j is the open road capacity of route j in vehicles per hour, d_{kj} is the point-to-point origin to destination distance faced by traveler k on route j, ζ and ξ are route specific parameters that are a function of route speed and capacity (Branston[8]), M_{kj} is the set of traffic signals faced by traveler k on route j, r_{mj} is the effective red time of signal m on route j per cycle in minutes, C_m is the cycle length of signal m in minutes, and ρ_{mj} is the traffic intensity of signal m defined as the average arrival rate (q_{kj}) divided by the saturation flow rate of the approach of route j, in vehicles per hour. Note that the traffic signal delay portion of this equation (i.e. the second term), as specified, is valid only when approach capacity exceeds arrivals. The standard modification of this term is used when arrivals exceed approach capacity (MANNERING and KILARESKI[26]). The values for expected speed used in the estimation of Equations 4 and 5 are obtained directly from Equation 7 by using distance (d_{kj}) and converting values to units of miles per hour.

Given the above definition of expected travel time along with vehicle operating cost (which is known from the price of fuel and the make, model and vintage of vehicle used to make the trip), the route choice model (Equation 2) can be readily estimated with instrumented values of instantaneous flow, $\hat{q}_{kj}$, used to avert possible departure time selectivity bias. The results of this estimation are given in Table II. The table indicates that both coefficient estimates are correctly signed and highly significant, statistically. The magnitudes of the coefficients are also reasonable, as the marginal rate of substitution indicates that the value of commuter time is \$6.32 per vehicle hour or, with an average sample vehicle occupancy of 1.1, \$5.75 per person-hour. It is also interesting to consider the route-specific elasticities of cost and time. Average elasticities, computed by sample enumeration, are presented in Table III. This table shows that expressway

TABLE II

Route Choice Coefficient Estimates (Corrected for Departure Time Selectivity Bias)

Variable	Coefficient	t-Statistic
Expected travel time (in minutes)	−0.667	−6.803
Vehicle operating costs (in dollars)	−6.328	−2.733
Number of observations		151
Log likelihood at zero		−165.89
Log likelihood at convergence		−101.36
Percent correctly predicted		74.17

TABLE III

Route Elasticities

Route Type	Route Elasticities with Respect to: Costs	Time
Arterial	−1.17	−5.71
Rural road	−0.476	−1.72
Expressway	−2.07	−6.04

TABLE IV

Speed Model (in MPH) Corrected for Route and Departure Time Selectivity Bias

Variable	Coefficient	*t*-Statistic
Expected speed (in mph)	0.642	6.86
Income (in thousands of dollars per year)	0.202	2.55
Number of vehicle passengers	4.036	2.29
Number of observations		151
R-squared		0.319

choice probabilities are most sensitive to cost and time changes while the rural road is least sensitive. This finding is largely the result of the longer distance that most travelers experience by using the expressway route alternative (see Table I).

The speed model (with the dependent variable being the average travel speed, in miles per hour, as specified in Equation 4) is estimated by ordinary least squares with corrections for route and departure time selectivity as previously discussed. Estimation results are given in Table IV. Individual coefficient estimates are all of plausible sign and statistically significant. The positive income coefficient suggests that higher income drivers tend to drive faster. This finding may be reflecting the fact that higher income travelers tend to drive higher performance vehicles and/or may tend to exhibit more aggressive driving behavior. The positive vehicle occupancy coefficient is likely capturing the tendency of travelers to increase driving speed in an attempt to compensate for the increased trip distance often encountered in multiple occupancy vehicle travel. Finally, the *R*-squared value of 0.319 is quite satisfactory when one considers the amount of variance in data of this type.

The estimation results of the schedule delay model (defined only for travelers with fixed work start times and with the dependent variable being the amount of time between scheduled work start time and actual arrival time, in minutes), with corrections for route and departure time selectivity bias, are presented in Table V. As found in previous research efforts (CHANG and MAHMASSANI[9]), travelers' preferred schedule delay (which is defined as the schedule delay assuming no traffic congestion) plays a dominant role in the

TABLE V

Schedule Delay Model (in Minutes) Corrected for Route and Departure Time Selectivity Bias

Variable	Coefficient	*t*-Statistic
Constant	12.104	1.05
Expected speed (in MPH)	−0.384	−1.71
Income (in thousands of dollars per year)	−0.108	−1.93
Age (in years)	0.118	1.69
Preferred schedule delay (in minutes)	0.821	7.75
Number of observations		90
R-squared		0.492

model. The positive coefficient associated with the age variable indicates that older people tend to be more conservative and have higher schedule delays. The opposite effect is true of income, with wealthier travelers reducing their schedule delay. This finding may reflect less significance being placed on on-time arrivals by higher income travelers (i.e., higher job security) and/or more risk seeking behavioral tendencies. The expected speed coefficient suggests that as congestion increases (and expected speed declines) schedule delays will increase. This is a reasonable finding since intuitively we expect travelers to increase schedule delays to account for the increased variability in travel times associated with congested, low expected-speed conditions. This finding is also consistent with the earlier work of Mahmassani and Tong[21] and others. Again, the overall *R*-squared is quite adequate considering the amount of variance inherent in the schedule delay data.

4. EQUILIBRIUM DEFINITIONS

GIVEN THE estimated demand models, as specified in Equations 2, 4 and 5, attention can now be directed toward operationalizing these models in a dynamic traffic equilibrium framework. The traditional approach to equilibrium on highway networks divorces itself from classic supply/demand equilibrium by using performance functions (such as the BPR function discussed earlier), which define a physical relationship between traffic flow and speed, as opposed to one that is some function of individual decision making (SHEFFI[28]). This paper departs from the performance function concept in that key variables (travel speed and schedule delay) are assumed to be an outgrowth of individual choices. As will be shown, this performance function departure permits the problem to be viewed as a supply/demand equilibrium in the classic economic sense.

To begin formalizing our approach, we adopt the definition of dynamic equilibrium as the spatial (route) and temporal distribution of traffic flow that

ensures that all travelers between all origins and destinations achieve utility maximizing behavior. This definition is consistent with the general notion of economic equilibrium as discussed in various sources (ARROW and HAHN[3]). In this study, with a single origin and destination pair, it is assumed that the number of morning work trips between the origin and destination is known and, as previously assumed for model estimation, that commuters' work start times (for fixed-time employees) and preferred arrival times (for travelers without fixed work starting times) are also known and exogenous to the equilibrium process. Given a series of discrete time intervals of a specified length (e.g., the five minute instantaneous flow intervals used in demand model estimation), equilibrium exists when the spatial and temporal distribution of traffic flows satisfy,

$$D_{tj}(\mathbf{Z}, \mathbf{R}, \mathbf{F}, \mathbf{V} \mid \mathbf{W}) = f_{tj} \quad \forall t, j. \tag{8}$$

Where t is the index of discrete time intervals of some specified length, j is an index of alternative routes between the origin and destination, $\mathbf{Z}$ is a matrix of traveler attributes, $\mathbf{R}$ is a matrix of route attributes that are not flow dependent, $\mathbf{F}$ is a matrix of instantaneous traffic flows for all routes and time intervals, $\mathbf{V}$ is a matrix of traveler-used vehicle attributes, $\mathbf{W}$ is a matrix of traveler's work-start and preferred-arrival times, $D_{tj}(.)$ is the conditional aggregate demand for route j in time interval t with prevailing traffic flows $\mathbf{F}$ (as specified by previously estimated Equations 2, 4 and 5), and f_{tj} is the instantaneous flow on route j in time interval t.

The equilibrium problem is one of finding flow values that ensure that Equation 8 is satisfied and that the summation of flows over all routes and time intervals is equal to the total number of trips made between the origin and destination. Intuitively, this equates to solving for the matrix of instantaneous flows ($\mathbf{F}$) that produces consistency between traveler departure times and desired route, speed, and schedule delay choices. It also is important to note that the flows for any given t' and j' are dependent on the flows of all other $t \neq t'$ and $j \neq j'$, as suggested by Equation 8. This implies a cross elasticity between different time periods. The existence of such a cross elasticity is consistent with recent developments in the dynamic equilibrium literature (BEN-AKIVA, CYNA and DE PALMA[5]).

With standard theorems relating to general equilibrium models, equilibrium flows satisfying Equation 8 can be shown to exist (ARROW and HAHN[3]). Further, with additional restrictions on travelers' route choice utility functions, equilibrium flows satisfying Equation 8 can be shown to be unique. Specifically, these restrictions relate to substitution and income effects. For proof of uniqueness it is necessary to assume that flows affect demand only through substitution effects and not income effects. While this assumption is potentially problematic in some cases, such as solving for equilibrium prices in durable goods markets, it is quite reasonable in the context of work trip choices which do not involve extensive capital investment. Unfortunately, while proof of existence and uniqueness is a straightforward adaptation of existing economic equilibrium proofs for the case of a single origin and destination, extensions of the modeling approach developed in this paper to multiple origins and destinations cannot be proven to be unique without unrealistic behavioral restrictions, due to the fact that the sole dependence between instantaneous route flows and specific origin-destination demand no longer exists. For additional details on general equilibrium models of this type see Arrow and Hahn,[3] BERKOVEC,[7] and MANNERING and WINSTON.[25]

5. EQUILIBRIUM ALGORITHM

FOR THE determination of instantaneous traffic flow, as indicated in Equation 8, discrete time intervals of a five minute duration are used so as to be consistent with the estimated demand models which have explanatory variables based on five minute instantaneous flows. Recall that since demand is now being aggregated, over the single origin–destination network, flow must be considered endogenous to the process, unlike the case of model estimation in which the effect of an individual traveler's choice on instantaneous flow could be assumed to be so small, that flow could be viewed as exogenous.

To arrive at equilibrium flows satisfying Equation 8, a simplistic heuristic algorithm is used. The algorithm begins by specifying an initial condition of the spatial and temporal distribution of traffic. The most obvious initialization is to distribute the total origin–destination traffic demand (assumed to be known) equally among all route and departure time interval combinations, thus producing equal values for all cells of the $\mathbf{F}$ matrix. With this initialization, estimated Equations 2, 4 and 5 are enumerated, through the sample, to arrive at a new distribution of flows over routes and departure times. This is illustrated as,

$$f_{tj} = \textstyle\sum_K \Lambda_{kj} P_{kjt} E \quad \forall j, t. \tag{9}$$

Where f_{tj} is the five minute flow at time interval t on route j for iteration n (i.e., elements of the $\mathbf{F}$ matrix), Λ_{kjt} is an indicator variable that is one if the departure time (as calculated in Equation 3) falls within time interval t for traveler k at iteration n, and zero otherwise, P_{ijt} is the probability of traveler k selecting route j with prevailing flows for time interval

t at iteration n, and E is a constant of expansion to permit the estimation sample to represent the total origin–destination demand (e.g., if the sample size is 151 and the total vehicle demand is 1510, then E is 10).

A convergence test is then undertaken to check whether or not the flow rates predicted by Equation 9 are acceptably close to those used in the demand calculations (i.e., either initially assumed flow rates or flow rates from a previous iteration). The convergence test is,

$$\sum_t |f_{tj}^n - f_{tj}^{n-1}| / f_{tj}^{n-1} < \kappa_j \quad \forall j \tag{10}$$

where κ is the convergence measure. A similar criteria has been defined by BEN-AKIVA, DE PALMA and KANAROGLOU,[6] and has been empirically proven to produce satisfactory results.

Finally, if convergence is not achieved in a given iteration, new flows are used and the process continues. To arrive at a new set of flows, a smoothing procedure has been found to provide the quickest convergence such that,

$$f_{tj}^{n+1} = 0.75 f_{tj}^{n-1} + 0.25 f_{tj}^n. \tag{11}$$

A similar procedure has been used with success in a modified capacity restraint traffic assignment by the Federal Highway Administration.[14]

6. SIMULATION EXPERIMENTS

IN ALL simulation experiments, a simplification of the actual area from which the survey data was collected, is used. The simplification is that it is assumed that all traffic on the network connecting the residential development with downtown State College (see Figure 1) is an outgrowth of this single origin–destination pair. Of course, in reality, many origin–destination pairs contribute volume to this street network. Total traffic demand, used to calculate E in Equation 9, was estimated from actual ground counts to be 3500 vehicles for the typical two hour period from 7:00 a.m. to 9:00 a.m., a period that covers the majority of work-trips in the State College area. Thus, using five minute intervals as previously discussed, the experiments include equilibrium over three routes and 24 time intervals.

6.1. Experiment 1: Existing Conditions (Base Case)

The results of the simulation of existing traffic conditions (the base case) are presented in Figure 2,

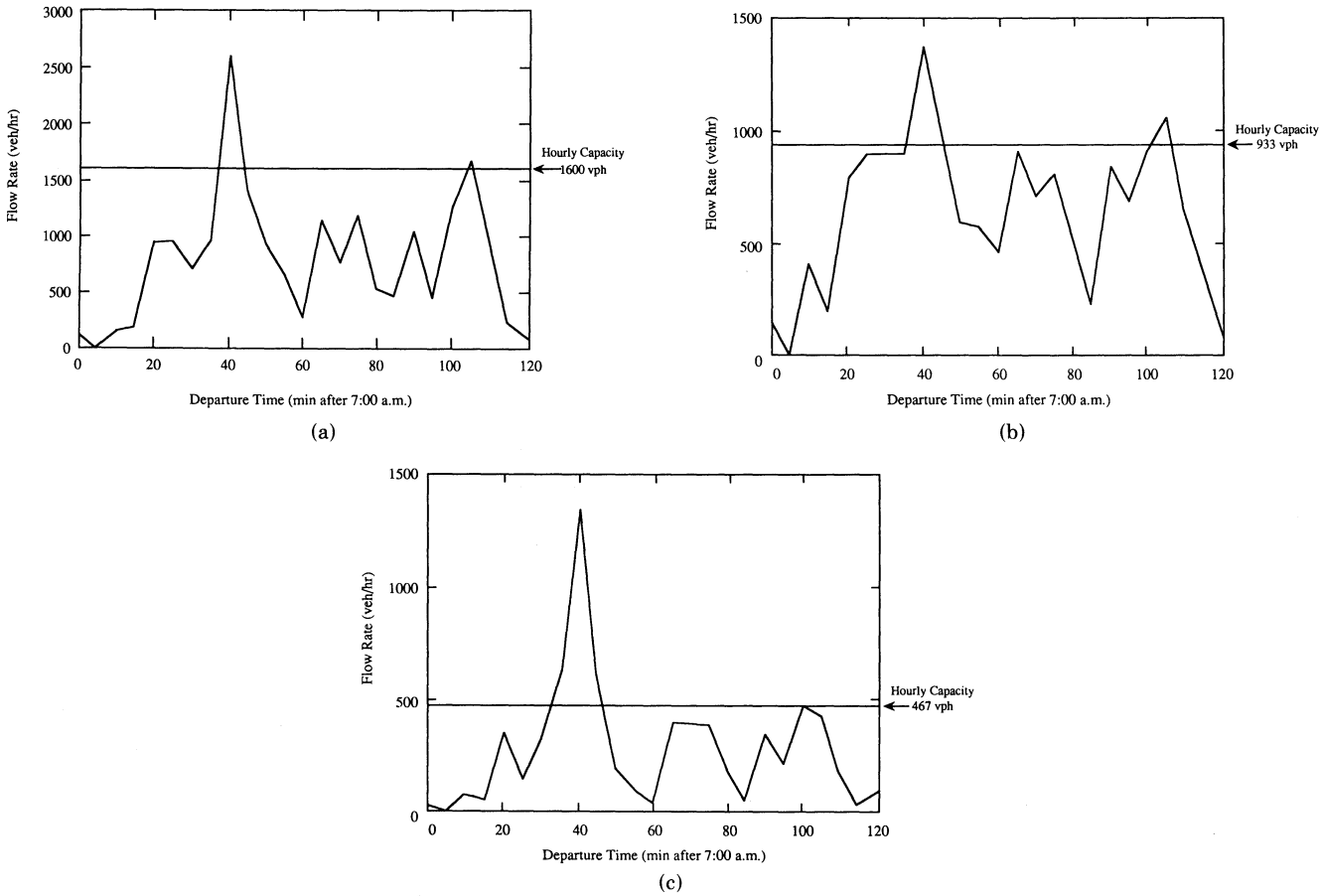

Fig. 2. Equilibrium traffic flow on (a) the arterial, experiment 1, (b) the rural road, experiment 1, and (c) the expressway, experiment 1.

a, b and c. This base case includes the vehicle demand, route characteristics, and the work start and preferred arrival times as they existed at the time of the commuter survey. The results shown in the Figures are a fairly typical representation of morning traffic peaking in a small metropolitan area.

In interpreting the results of Figure 2, it is important to note that when the instantaneous traffic flow rate exceeds the hourly capacity of the most restrictive section of the route (represented by the hourly capacity line in the Figures), queues will form and will not dissipate until some time after the flow rate falls below capacity. The length of queue can be determined by estimating the area of the flow rate curve above the hourly capacity line. The time for queue dissipation can be estimated by calculating the time needed for the area below the hourly capacity line and above the flow rate curve, following a queuing situation, to equal the area of the flow rate curve above the capacity line. For illustrative purposes, it is assumed that the restrictive capacity occurs at the point of trip origin, where the temporal distribution is readily computed. A driver's choice of departure time from his origin is therefore considered the same as the arrival time at the critical section. This is obviously a substantial simplification since, in reality, the restrictive capacity takes place at some point along the route.

Using this computational simplification, it is found that the longest-duration queue on the arterial begins at 7:37 a.m. and lasts 15 minutes, on the rural road begins at 7:35 a.m. and lasts 20 minutes and on the expressway off-ramp begins at 7:33 a.m. and lasts 30 minutes.

In addition to studying queue formation and dissipation, it is interestng to assess traveler welfare impacts. From economic literature on consumer welfare analysis (SMALL and ROSEN[29]) it is known that, with the logit route choice model, the total implicit cost of commuting is,

$$\mathrm{TIC} = \sum_{K} \left[-1/\lambda \ \ln \left[\sum_{J} \exp(\mathrm{V}_{kj}) \right] \right]. \qquad (12)$$

Where TIC is the total implicit cost of commuting, λ is the marginal utility of income, and V_{kj} is the flow dependent route utility as previously defined. The marginal utility of income is simply equal to the value of the coefficient associated with the operating cost variable in the route choice model, but opposite in sign (WINSTON and MANNERING[30]). Enumeration of Equation 12 through the sample of 151 travelers (expanded to approximate the total origin–destination vehicle demand of 3500) yields a total implicit cost of commuting, for the 7:00 a.m. to 9:00 a.m. morning peak, of $3,364. This value will be used as a basis for comparison in subsequent simulation experiments.

6.2. Experiment 2: Reduction of Route Capacity

This experiment assesses the impact of reducing the capacity of the arterial by 50% (to 800 vehicles per hour at the most restrictive section), which is the equivalent of a lane closure. The total travel demand is assumed to remain at the same level as in the base case. The results of the experiment are presented in Figure 3, a, b and c. As expected, the reduced arterial capacity results in a diversion of traffic to the two alternate routes and the peaking characteristics of traffic flow are less pronounced. Maximum queue durations are computed to be 24 minutes for the arterial, 39 minutes for the rural road, and over 1.5 hour for the expressway which has an unfavorable allocation of green time at the off-ramp intersection with the rural road.

Under this reduced capacity condition, the loss in total commuter welfare can be readily computed as the difference (between experiments 1 and 2) in the total implicit costs as defined in Equation 12. This difference in total implicit cost is the well-known economic concept of compensating variation, or the amount that would have to be paid to have travelers as well off, welfare-wise, after the capacity reduction as they were before the reduction. The welfare cost of the capacity reduction is computed to be $202 per each morning peak period, which can accumulate to a substantial welfare loss, even on this small and relatively lightly congested network, when one considers off-peak and afternoon peak costs of capacity reduction.

6.3. Experiment 3: Optimal Signal Timing

An interesting application of our dynamic model is one that computes optimal signal timing. From an economic perspective, optimal signal timing is that timing that provides the highest consumer welfare. In our case, this translates into the timing that results in the lowest total implicit cost as defined in Equation 12. The intersection of Fox Hollow Road and the Route 322 by-pass (rural road and expressway respectively) was chosen for the optimal signal timing experiment (see Figure 1). The intersection at the time of the commuter survey was a fixed-time signal with: 1) a saturation flow of 1400 vehicles per hour of effective green for both the rural road and the expressway off-ramp and 2) a cycle length of 60 seconds with 20 seconds of effective red allocated to the rural road and 40 seconds to the expressway off-ramp.

The procedure to determine optimal signal timings is to systematically compute dynamic equilibrium traffic flows for various cycle lengths and for the range of effective red allocated to the rural road for each cycle length. The combination of cycle length and effective red that produces the lowest total implicit

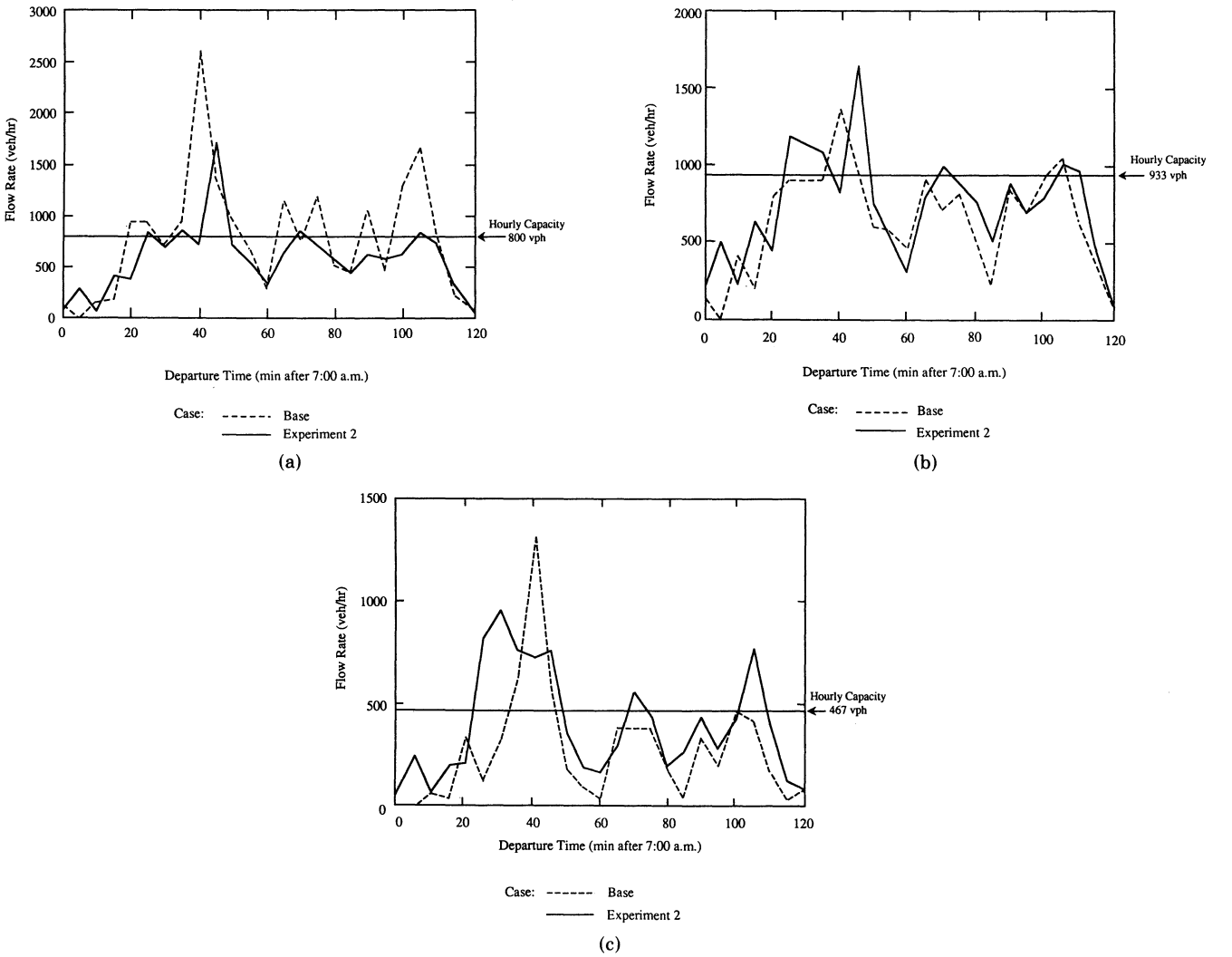

Fig. 3. Equilibrium traffic flow on (a) the arterial, experiment 2, (b) the rural road, experiment 2, and (c) the expressway, experiment 2.

cost will constitute optimal signal timing. Illustrations of effective red allocations and total implicit costs for 80 and 140 second signal cycle lengths are presented in Figures 4 and 5. Although both figures reveal fluctuations, resulting from algorithm convergence, that often making it difficult to pinpoint an exact optimal effective red allocation, the approximate optimal rural road effective red allocation is 28 seconds for the 80 second cycle and 44 seconds for the 140-second cycle. The selection of optimal effective red allocation for various cycle lengths yields the curve in Figure 6. Here it is seen that even at a 200-second cycle length, which is beyond practical recommendations, the total implicit cost of commuting continues to decline. This finding (i.e., the superiority of longer and longer cycle lengths) is consistent with the results obtained when using standard accepted signal optimization packages on the same intersection. It is believed that this is a

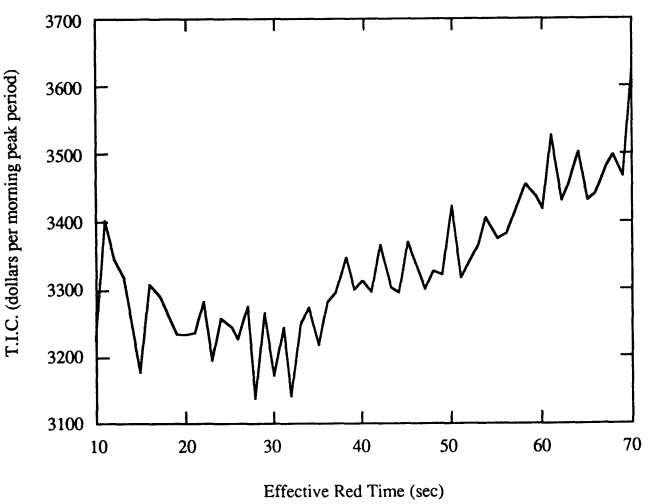

Fig. 4. Total implicit cost of commuting versus the length of effective red time allocated to the rural road for 80-second cycle length.

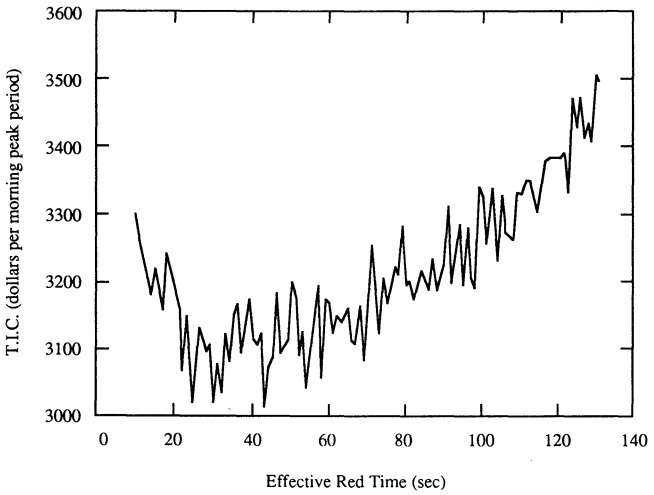

Fig. 5. Total implicit cost of commuting versus the length of effective red time allocated to the rural road for 140-second cycle length.

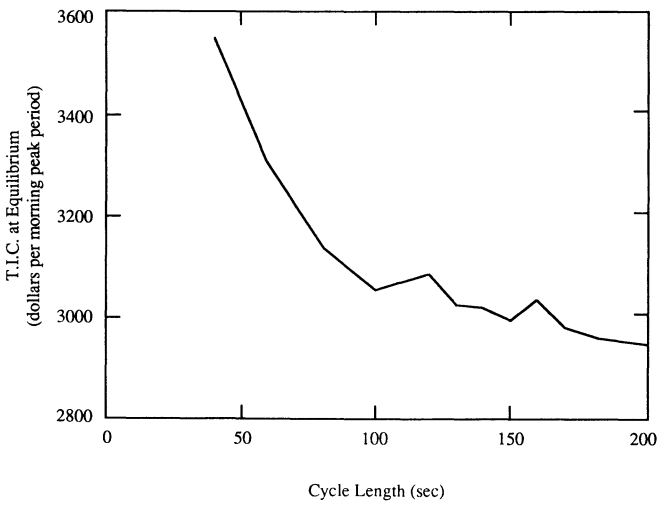

Fig. 6. Total implicit cost of commuting versus cycle length.

reflection of the simplistic two-phased, fixed-time nature of the signal being studied.

Although the minimization of total implicit cost is an exciting and promising approach to optimal traffic signal timing, the results contained herein must be viewed as exploratory for two reasons. First, the intersection being evaluated is too simplistic to be of value to most practical applications. Additional work on more complex intersections with multi-phase signals is needed to assess the potential of the approach. Second, the intersection queuing (see Equation 7) used in demand model estimation and subsequent equilibrium determination does not account for the randomness in vehicle arrivals. Thus, future work is needed to account for random vehicle arrivals so that a precise comparison between the total implicit cost approach and currently accepted signal optimization procedures can be made.

7. SUMMARY AND CONCLUSIONS

THIS PAPER provides a means of determining the spatial and temporal equilibrium distribution of traffic flows on a simplified network for home-to-work commuter trips. The procedure is based on the behavioral modeling of individual decisions in response to traffic congestion. Specifically, a discrete/continuous econometric framework is developed in which travelers' choice of route and departure time is modeled, and this is incorporated into an equilibrium framework from which a variety of traffic simulations can be undertaken.

The simulation experiments presented herein demonstrate the potential usefulness of the model in evaluating capacity restrictions and signal improvements. It is clear that the temporal considerations incorporated in this model, and other recently developed dynamic models that have appeared in the literature, greatly enhance the traffic analysis process relative to the traditional static equilibrium assumptions that are the current mainstay of traffic assignment algorithms.

In terms of future work, a number of important directions can be identified. First, although the expansion of our approach to multiple origins and destinations is an inherently difficult undertaking, it is obviously needed if widespread use of the model is to be achieved. Second, issues of the behavioral model's transferability among metropolitan areas must be addressed to lower the potential cost of model implementation. Finally, additional work on the optimal signal timing approach is warranted as previously discussed.

REFERENCES

1. M. ABKOWITZ, "An Analysis of the Commuter Departure Time Decision," *Transportation* **10** (1981).
2. S. ABU-EISHEH AND F. MANNERING, "A Discrete/Continuous Analysis of Commuters' Route and Departure Time Choice," *Trans. Res. Rec.* **1138,** 27–34 (1987).
3. K. ARROW AND F. HAHN, *General Competitive Analysis,* Holden-Day, San Francisco, 1971.
4. M. BEN-AKIVA, "Dynamic Network Equilibrium Research," *Trans. Res.* **19A,** 429–431 (1985).
5. M. BEN-AKIVA, M. CYNA AND A. DE PALMA, "Dynamic Model of Peak Period Congestion," *Trans. Res.* **18B,** 339–356 (1984).
6. M. BEN-AKIVA, A. DE PALMA AND P. KANAROGLOU, "Dynamic Model of Peak Period Traffic Congestion with Elastic Arrival Rates," *Trans. Sci.* **20,** 166–181 (1986).

7. J. BERKOVEC, "Forecasting Automobile Demand Using Disaggregate Choice Models," *Trans. Res.* **19B,** 315–329 (1985).
8. D. BRANSTON, "Link Capacity Functions: A Review," *Trans. Res.* **10,** 223–236 (1976).
9. G.-L. CHANG AND H. MAHMASSANI, "Travel Time Prediction and Departure Time Adjustment Behavior Dynamics in a Congested Traffic System," *Trans. Res.* **22B,** 275–290 (1988).
10. S. COSLETT, "Demand Model Estimation and Validation," Urban Travel Demand Forecasting Project, University of California, Berkeley, CA, 1977.
11. A. DE PALMA, M. BEN-AKIVA, C. LEFEVRE AND N. LITINAS, "Stochastic Equilibrium Model of Peak Period Traffic Congestion," *Trans. Sci.* **17,** 430–453 (1983).
12. J. DUBIN AND D. MCFADDEN, "An Econometric Analysis of Residential Appliance Holdings and Consumption," *Econometrica* **52,** 345–362 (1984).
13. P.-H. FARGIER, "Effects of the Choice of Departure Time on Road Traffic Congestion: A Theoretical Approach," *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory*, pp. 223–263, 1981.
14. Federal Highway Administration, "Urban Transportation Planning Package," Washington, DC, 1977.
15. C. HENDRICKSON AND G. KOCUR, "Schedule Delay and Departure Time Decisions in a Deterministic Model," *Trans. Sci.* **15,** 62–77 (1981).
16. C. HENDRICKSON, D. NAGIN AND E. PLANK, "Characteristics of Travel Time and Dynamic User Equilibrium for Travel-to-Work, *Proceedings of the Eighth International Symposium on Transportation and Traffic Theory*, pp. 323–347, 1981.
17. C. HENDRICKSON AND E. PLANK, "The Flexibility of Departure Times for Work Trips," *Trans. Res.* **18A,** 25–36 (1984).
18. H. MAHMASSANI AND R. HERMAN, "Dynamic User Equilibrium Departure Time and Route Choice on Idealized Traffic Arterials," *Trans. Sci.* **19,** 362–384 (1984).
19. H. MAHMASSANI AND G.-L. CHANG, "Dynamic Aspects of Departure-Time Choice Behavior in a Commuting System: Theoretical Framework and Conceptual Analysis," *Trans. Res. Rec.* **1037,** 88–101 (1985).
20. H. MAHMASSANI AND G.-L. CHANG, "Experiments with Departure Time Choice Dynamics of Urban Commuters," *Trans. Res.* **20B,** 297–320 (1986).
21. H. MAHMASSANI AND C. TONG, "Availability of Information and Dynamics of Departure Time Choice: Experimental Investigation," *Trans. Res. Rec.* **1085,** 33–49 (1986).
22. F. MANNERING, "Poisson Analysis of Commuter Flexibility in Changing Routes and Departure Times," *Trans. Res.* **23B,** 53–60 (1989).
23. F. MANNERING AND C. WINSTON, "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization," *Rand J. Econ.* **16,** 215–236 (1985).
24. F. MANNERING AND D. HENSHER, "Discrete/Continuous Models and Their Application to Transport Analysis," *Trans. Rev.* **7,** 227–244 (1987).
25. F. MANNERING AND C. WINSTON, "U.S. Automobile Market Demand," in *Blind Intersection: Policy and the Automobile Industry*, Brookings Institution, Washington, DC, 1987.
26. F. MANNERING AND W. KILARESKI, *Principles of Highway Engineering and Traffic Analysis*, John Wiley & Sons, New York, 1990.
27. D. MCFADDEN, "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. F. Manski and D. McFadden, MIT Press, Cambridge, MA, 1981.
28. Y. SHEFFI, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ (1985).
29. K. SMALL AND H. ROSEN, "Applied Welfare Economics with Discrete Choice Models," *Econometrica* **49,** 105–130 (1981).
30. C. WINSTON AND F. MANNERING, "Consumer Demand for Automobile Safety," *Am. Econ. Rev.* **74,** 316–319 (1984).

(Received March 1987; revisions received August 1988, September 1989; accepted September 1989