Travel Demand Forecasting with Dynamic Microsimulation

KONSTADINOS G. GOULIAS AND RYUICHI KITAMURA

A new travel demand forecasting system, based on microanalytic simulation and dynamic analysis, is discussed. The system consists of two components: a microsimulator of household socioeconomics and demographics and a dynamic model system of household car ownership and mobility. Each component comprises interlinked models formulated at the household level. Replicated in the socioeconomic and demographic microsimulator are interactions and causal paths that underlie life cycle evolution of individuals and households. Output from the sociodemographic component is then used by the dynamic model system of mobility to predict household car ownership, trip generation, and modal split. The parameters of the model system have been estimated using observations from five waves of the Dutch National Mobility Panel data, covering the period of 4 years from April 1984 through April 1988. Other sources of information, external to the panel data, were also used to estimate key parameters. The availability of the large-scale panel data has been essential for the development of the detailed demographic and mobility model components. The model system is a credible and flexible forecasting tool with which a wide range of future scenarios can be examined to answer a variety of "what if" questions. Issues related to the model structure, data requirements, estimation methods, assumptions, and forecasting performance are summarized.

In travel demand analysis and forecasting the recognition that time is an indispensable dimension of travel demand models is a recent phenomenon. A new forecasting method that explicitly accounts for the dynamic character of travel demand is described. The approach attempts to combine dynamic models of travel behavior with sociodemographic and economic microanalytic simulation to produce a flexible forecasting tool. The development of the Microanalytic Integrated Demographic Accounting System (MIDAS) is summarized, and its use in forecasting is discussed.

The use of cross-sectional models in travel demand forecasting involves some fundamental problems. First, it is based on the untested assumption that cross-sectionally observed variations in travel behavior can be used as valid indicators of behavioral changes over time. Second, future values of socioeconomic and demographic input variables are obtained using "allocation" methods, which "post-process" aggregate forecasts into "pseudo-disaggregate" data. As such, the methods fail to effectively and accurately capture the internal relationships among the input variables. And third, it does not properly represent response lags involved in long-term mobility decisions (e.g., residence location and car ownership).

K. G. Goulias, Department of Civil Engineering and Pennsylvania Transportation Institute, The Pennsylvania State University, University Park, Pa. 16802. R. Kitamura, Institute of Transportation Studies and Department of Civil Engineering, University of California at Davis, Davis, Calif. 95616.

An alternative travel demand forecasting system is presented in this paper. The system consists of two components: a microsimulator of household socioeconomics and demographics and a dynamic model system of household car ownership and mobility. Each component comprises interlinked models formulated at the household level. Replicated in the socioeconomic and demographic microsimulator are interactions and causal paths that underlie life cycle evolution of individuals and households. Simulation units evolve from year to year, experiencing marriages, divorces, births, deaths, and so forth. Employment, income, driver's license holding, education level, and household size and composition are among the variables that are internally generated in the simulation. User-defined parameters have been provided for modification to create any future growth path desired.

The parameters of the model system have been estimated using observations from five waves of the Dutch National Mobility Panel data, covering the period of 4 years from April 1984 through April 1988. Other sources of information, external to the panel data, were also used to estimate key parameters. The availability of the large-scale general purpose panel data has been essential for the development of the detailed demographic and mobility model components.

The model system is a flexible and credible forecasting tool with which a wide range of future scenarios can be examined to answer a variety of "what if" questions. It can replicate reality with accuracy comparable with other forecasting models and represents a new approach to forecasting travel demand. However, the method is complex, poses high demands in model estimation, and requires a large amount of data.

BACKGROUND AND STRUCTURE OF THE MICROSIMULATOR

The large systematic biases and the low predictive accuracy of long-range planning and forecasting motivated the adoption of strategic planning—the identification of preferable transportation policies as input to plan development (*1*). Policy strategies are identified beforehand, scenarios are developed for each strategy, and pertinent forecasts are derived. The strategic planning process requires forecasting tools to provide growth scenarios instead of point estimates. Moreover, forecasting creates the need for several alternative growth scenarios, each based on a different set of assumptions about economic development and demographic evolution.

In transportation the usual techniques are either problem-oriented descriptive analyses or forecasting procedures similar to those of the Urban Transportation Planning Process (UTPP).

The need for a system that can answer "what if" questions, under user-defined circumstances of the real world, is clear. This tool should allow the exploration of policies so that it could be used either as a decision insight system (a system that allows policy analysts to introduce subjective input and observe the output) or as an extrapolatory scenario-based system (a system that provides information about the future based on data and observed relationships from the past). The effort described in this paper is the starting point for the creation of a flexible and comprehensive "test bed" for alternative theories and methods of forecasting travel demand.

Socioeconomic and demographic information plays an important role in the four-stage UTPP procedure requiring input variables such as population, income, employment, and car ownership. These methods are extremely sensitive to the accuracy of the information provided externally. Travel demand models are usually cross-sectional individual- and household-based models, whereas the input to these models (i.e., sociodemographic and land use information) is obtained through approximate disaggregation techniques (*2*).

Whereas the techniques, models, and procedures used to obtain input to UTPP are disparate, they share one common characteristic: they are not at the same level of disaggregation as travel demand models. Most agencies transform regional information to the district level, and then from the district level to the traffic zone level. These allocation methods do not provide all the required information needed by the travel demand models. Additional detailed information is obtained using approximate post-processing procedures (disaggregation procedures). The provision of input at the zone level necessitates the application of travel demand forecasting models designed for households and persons at the traffic zone level, too. As expected, both the conversion of aggregate sociodemographic forecasts to zonal forecasts and the conversion of individual travel demands to zonal demands produce many errors throughout the process. Bajpai (*3*) observed that "techniques to project automobile ownership, household income, and household size from population and employment are highly recommended for future research." Travel demand forecasts based on existing techniques are questionable.

Most of the sociodemographic variables describe and attempt to replicate decisions made by individuals and households, so the need arises for models that predict just such variables at the elementary level of decision making. Aggregate responses to policy changes can be obtained by grouping households and individuals into the specific traffic zones or following any other aggregation scheme desired. This approach can be called a "bottom-up" procedure. It is well known that bottom-up approaches lead to more accurate results.

Microsimulation and Dynamic Analysis

Arrow (*4*) and Orcutt et al. (*5*) have shown that microsimulation is a particularly flexible approach in that it adopts a comprehensive system analysis to explain, predict, and compare the impacts of alternative transport policy strategies. The method enables the forecasting of direct and indirect effects of the simulated policies on the system analyzed. Microsimulation can help fill the gap in forecasting the input to travel demand models and provides the framework for designing a new dynamic forecasting tool. These compelling arguments in favor of microsimulation are examined here in view of the added complexity of the method and the increased data requirements.

When the data at hand are cross-sectional observations, the usual assumption is either that behavior does not change or that the changes are given by cross-sectional variations. Therefore, forecasts of changes over time are either nonexistent or are extrapolations from differences in the cross-sectional sample considered. Davies (*6*) notes that cross-sectional analyses fail to differentiate between age effects and cohort effects, fail to resolve ambiguities in causalities, cannot provide methods to consider observable or unobservable omitted variables, and exaggerate the behavioral effect of policy changes by not being able to incorporate phenomena such as inertial response to change.

One of the most promising research approaches to overcome these weaknesses is dynamic analysis. This is the procedure used to describe changes in behavior occurring over a period of time. Forecasts based on these estimates may prove better than cross-sectionally derived ones because models can be developed from dynamic hypotheses and tested with longitudinal information. Future behavior can be predicted by extrapolating observed changes that are reflected in the dynamic models.

Microsimulation and dynamic models need data for model estimation and the construction of microanalytic scenarios. The best source of data is a panel survey. In panel surveys the same information is collected on the same individuals over a period of time. Questionnaires and travel diaries are distributed at different times to the same individuals to collect detailed sociodemographic and travel data. Panel data enable us to develop models that relate behavioral changes to changes in contributing factors in dynamic context, specifying intertemporal causation properly (*7*).

Structure of the New System

The unique characteristic of the approach followed in this study is the combination of a dynamic model of travel behavior with dynamic microsimulation, which is motivated by the following. Since simulation in general implies modeling of a process that evolves over time, dynamic disaggregate models are the natural ingredient of the simulation. Hence, throughout the design of MIDAS, dynamic models at the level of the household and the household member are used to replicate real world changes in sociodemographic characteristics and mobility.

The forecasting tool is made of two components: the sociodemographic component and the mobility component. The sociodemographic component aims to realistically recreate the progression of a household through life cycle stages and simulate changes in the household members' socioeconomic and demographic attributes, such as employment status and driver's license holding. Then the mobility component uses these endogenously generated socioeconomic attributes to forecast household car ownership and mobility. The two components are integrated to form a comprehensive simulation system.

SOCIOECONOMIC AND DEMOGRAPHIC COMPONENT

In the simulation, a household member will age, form an independent household, gain employment, obtain a driver's license, marry, give birth, and so on. The size and composition of the household will change accordingly. A household member may be added to a household through a marriage, or a household may be split into two through a divorce. A child will leave his parents and form a new household. Such changes are probabilistically generated in the simulation. The model parameters that determine the probability of these events are obtained from the Dutch Panel data set.

Household Type Transition

In MIDAS the transition between household types is viewed as the fundamental element of household evolution representing household life cycle stage. Given a transition in household type, new household members are generated, or existing household members are eliminated, and member characteristics are altered in MIDAS. The transition in household types thus serves in MIDAS as a control that constrains the characteristics of household members.

Five household types are used: single-person households, households of a man-woman couple, nuclear family households, single-parent households, and other households. This classification, which is based on the major conclusion of the activity-based travel analysis that children of a household have an important influence on the travel patterns of its adult members, reflects the notion of life cycle (*8*).

For each household in the simulation, characteristics are first read from an input file comprising records of sample households from the Dutch Mobility Panel data set. Following this, the transition between household types is simulated for each time period (1 year is used as the time interval of the simulation). This process is based on a set of logit models that determine transition probabilities for each household as functions of attributes such as the presence of children by age group and the adult household members' age, education, and employment.

A set of subroutines has been developed to probabilistically change the attributes of household members, generate new members, or remove individuals from the household. For example, two subroutines are called in connection with the transition from family to family, or from single parent to single parent, when the number of children is two or more. Another routine is called in connection with the transition from single to single. It accounts for the possibility that the member of a single-person household passes away, and thus the household vanishes [a description of the routines is given elsewhere (*9*)].

Birth and Death

The probability that a woman in a household will give birth to a child in a given year is expressed as a function of the age and employment status of the woman and the number of children that already exist in the household. Observed frequencies obtained from the Dutch Panel data set are used to determine the probability that a woman in a household will give birth to a child.

A birth may be implied by a change in the household type (e.g., a couple to a family). In such cases, the logit models of household type transitions depict the probability of a birth. For example, the probability of a transition from couple to family is expressed as a function of the man's age and education and the woman's employment status. The event of birth is randomly generated in the simulation using these probabilities.

A single-person household is removed when a death takes place in the simulation. The possibility of death is also considered in connection with the transition from couple (or family) to single (or single parent). If a death does not take place in the simulation, the transition is regarded as a result of a divorce, and the household is split into two households.

Households Formed by Children

The event of "leaving the nest" (i.e., a child moving out and forming an independent household) is modeled as a function of the age, sex, and employment status of the child. Similar to the case of birth, this event is implied by household type transition from family to couple or from single parent to single. The probabilities of these transitions are represented by the logit models as functions of the number of children by age.

When the event of nest-leaving takes place in the simulation, a new household is added to the data file with a certain probability, representing the probability that the new household will remain in the same municipality. The evolution of this new household is simulated through the rest of the simulation period.

Employment and Income Models

The employment status of a person is determined using transition matrices developed by sex and age group. Each matrix contains the probability of change in employment from one status to another. For example, the two-by-two matrix for men in the 18-to-24 age bracket indicates that a person who is employed at Time t will also be employed at Time $t + 1$ with probability 0.929.

Given the employment status, the personal income is determined using a set of dynamic models. The personal income at Time t is assumed to be determined in part by the personal income at Time $t - 1$, called lagged dependent variable. It is also assumed that there is correlation between the unexplained effect of Time $t - 1$ and that of Time t, called serial correlation. The income models are developed for the four possible combinations of the employment status at Time $t - 1$ and Time t: (not employed, not employed), (employed, not employed), (not employed, employed), and (employed, employed).

The personal income of each household member is added in the simulation to obtain total household income. The employment transition matrices and the parameters of the income models are estimated using data obtained in a period of economic expansion (1984 through 1988). These parameters must be adjusted if the model is to be applied for a period of stable economy or economic recession. This adjustment requires examination of the impact of the regional and national economy on the parameters of these model components, which is outside the scope of this study.

Driver's Licenses and Education

The driver's license holding is determined using transition matrices similar to those for employment status. Compared with the transition matrices for employment status, the driver's license matrices in general have larger diagonal elements, which correspond to the transition from licensed to licensed or from nonlicensed to nonlicensed. This implies that license-holding status is less variable than employment status. Also notable is the stability in the transition probabilities across the age groups.

Education is among the explanatory variables used in the MIDAS mobility component, and it is necessary to determine education levels for those household members that are internally generated in the simulation process. This determination is not based on detailed modeling of education levels because it is clearly beyond the scope of this study.

The education levels of children that are generated in the simulation are determined randomly using the distribution of education levels by sex obtained for individuals 18 through 28 years old in the panel data. Education levels of new members that enter a household through a marriage are determined using the correlation between the education levels of married men and women. For example, the probability that a man has a given education level is determined by the education level of the woman who has been a member of the household in the simulation.

New Household Members

A set of personal attributes needs to be generated whenever a new household member is introduced in the simulation. When a new person enters a household through marriage, the person's age and education are determined on the basis of the existing member's age and sex. The new member's employment and income are then determined on the basis of age and sex.

For a newborn member of a household, only sex is determined at the time of birth; other attributes are determined when a person reaches the age of 18, using the probabilities of employment, license holding, and income.

The person attributes of "other" household members are determined as follows. First, the age and sex of the "other" individual are randomly generated on the basis of the age of the head of the household. Employment, license holding, education, and income are then randomly determined on the basis of the observed distribution of the attributes of "other" persons by age and sex.

Household Dissolution

A household is split into two or eliminated from the simulation after a divorce or other events that cause its dissolution. If children are present in the household, they are randomly assigned to the respective parents probabilistically. Only a fraction of newly formed households (formed through divorces or by children gaining independence) remain in the simulation. The value of 15 percent is chosen so that new households roughly replace households that disappear because of death and keep the total number of households in the simulation stable over the simulation years. This process replicates a demographically stable region.

Most model parameters are estimated using subsamples from the Dutch Panel data set. A subsample of Dutch Panel households is also used in the simulation. Observed household and person attributes of 1984, 1985, and 1986 are used as initial conditions; demographic and socioeconomic attributes and mobility levels of these and internally generated new households are simulated year by year to 2010 in MIDAS.

Input Parameters and Modifiers

The parameters in MIDAS can be classified into three categories. The first contains the coefficients of the dynamic models in the mobility component and the income models in the demographic component. These coefficients have been estimated from subsamples of the Dutch Panel data set using econometric methods and have been embedded in the MIDAS programming code. The second category contains 16 sets of parameters of the demographic components. Most represent transition probabilities associated with changes and are treated as input data. Their values have been estimated using the Dutch Panel data set. These parameters can be modified to represent a particular scenario of interest (e.g., an increase in women's labor force participation) or to incorporate external information. The third category is a set of input parameters that can be used for modifications of MIDAS settings. These are modifiers that can be used to change the annual growth of personal income, the birth probabilities, the male and female employment transition probabilities, the male and female license holding transition probabilities, and the household type transition probabilities.

Initial Sample Weighing

MIDAS stimulates the evolution of a subset of those Dutch Panel households that participated in Waves 1, 3, and 5. (The Dutch Panel is made of 10 contacts. The data used in this paper are from Waves 1, 3, 5, 7, 9, and 10, which correspond to March of 1984, 1985, 1986, 1987, 1988, and 1989, respectively. The data of Waves 1, 3, 5, 7, and 9 are used for estimation and the data of Wave 10 for validation.) Many models in MIDAS are dynamic, requiring observations from three time points in the simulation. Because of the initial sampling scheme and attrition, this subset of panel households does not represent the Dutch population. Two sets of weights have been developed for this subsample using available nationwide statistics. The weights are later used to duplicate households by Monte Carlo simulation [the derivation, use, and comparison between alternative weighing schemes are described elsewhere (*2*)].

MOBILITY COMPONENT

The MIDAS mobility component consists of a car ownership model, household motorized-trip generation models, a modal split model, car-trip distance models, and transit-trip distance models. All models are formulated for weekly totals. These mobility measures are obtained from the Dutch Panel survey in which only household members at least 12 years old were requested to report trips, and trips made by individuals less than 12 years of age are not reflected in the measures. Consequently, the MIDAS mobility component does not reflect trips made by individuals less than 12 years old.

Car Ownership Model

An ordered-response probit car ownership model is used to determine household car ownership in MIDAS. This model probabilistically describes the choice of an alternative from among a set of ordered discrete alternatives. A household's choice of the number of cars to own falls in this class of choice. The model assumes the presence of a latent variable that cannot be directly measured but is related to the observed choice—the number of cars owned in this case. Corresponding to a level of car ownership is a range of the latent variable value, which is defined by unknown threshold values. The model is a discrete choice dynamic model with serial correlation and was estimated in a five-stage maximum likelihood method (*10*).

The short-term MIDAS forecasting performance has been tested in a validation exercise. The models in the MIDAS mobility component are used to predict Wave 10 mobility measures using observed explanatory variable values from the Wave 10 data. Predictions thus obtained are then compared with observed measures in the Wave 10 data. The validation effort of this study is based on longitudinal data [i.e., a subset of observational time points (Wave 10 data) is set aside for validation]. If the models replicate Wave 10 observations well, evidence is offered that the models are capable of providing adequate short-term forecasting by replicating the sample closely.

The first part of Table 1 presents the average of five simulation runs. Car ownership levels are correctly forecast for approximately 90 percent of the sample households. The average number of cars per household is predicted to be 0.922, whereas the observed Wave 10 average is 0.945. The error is within 2.5 percent.

Dynamic Motorized-Trip Generation Models

Weekly household motorized-trip generation models, based on data from Waves 1, 3, 5, 7, and 9, have been developed separately for households with cars available and those without a car available. The variables used in the models are the number of diary keepers, number of women, number of men, number of workers, a set of income variables, car ownership, number of drivers, household types, residence area type, and a lagged dependent variable (number of trips a year ago).

Table 1 summarizes the validation results of the motorized-trip generation models. Two models have been formulated, separately for car-owning and carless households. The models are also dynamic with lagged dependent variables and serially correlated errors. Predictions are produced with two different methods: (*a*) using observed Wave 10 car ownership to classify sample households into car-owning and carless households and to exogenously determine the value of a multicar ownership dummy variable in the model for car-owning households; and (*b*) using simulated Wave 10 car ownership levels to classify households. The second method, which more closely represents MIDAS simulation forecasting, is subject to additional errors in household classification.

The results indicate that the models are performing well, in particular the one for car-owning households. The larger errors observed for the model for carless households are presumably due to smaller sample size.

Modal Split Model

Level-of-service data are not available to describe trip characteristics by alternative modes that connect given origin and destination zones. Modal split models that can be developed with this limitation are not trip-interchange (postdistribution) models that focus on modal competition at the disaggregate trip level. A new model structure, binomial logistic (BL), has been defined in this study to predict modal split.

Since land use and transportation network data for the 20 municipalities from which the Dutch Panel sample was initially drawn were not available, the only available measures on the supply side are a rough indicator of transit service level by municipality and accessibility measures by mode based on destination choice models (*11*).

The panel data set contains weekly travel information, which represents many travel mode choices repeated by the same household members. These repeated choices may be collectively explained by accessibility or other macroscopic level-of-service indicators.

Furthermore, mode choice may be made considering not each trip but a series of linked trips to be made as a whole by the individuals. Then the attributes of trips by alternative modes between a given origin and destination pair may not be as influential as one might think. To the contrary, household car ownership, the number of drivers in the household, overall level of transit development, and other sociodemographic attributes may be the major determinants of weekly household modal split. From this viewpoint, the appropriate measure of mode choice is the relative frequency of trips made by a particular mode rather than the mode chosen for each trip. These considerations motivated the new modeling effort reported by Goulias and Kitamura (*12*).

The BL model performed well in terms of data replication. The variables used were the number of diary-keepers in the household, number of cars available, number of drivers, and level of public transit availability. In particular, the results indicate that households without a car available and households in a large urban area with a regional transit district tend to have higher fractions of public transit trips.

The weekly household modal split model is validated similarly through simulation. The analysis here used Wave 10 observed explanatory variable values. The model's performance is evaluated in terms of the fraction of transit trips and the number of transit trips. The Wave 10 observed number of motorized trips is used together with a predicted fraction

TABLE 1 Mobility Component Validation with Wave 10 Observations

Car Ownership Model (five simulation runs)

	Predicted			
Observed	**Zero Cars**	**One Car**	**Two+ Cars**	**Total**
Zero Cars	217	12	0	229
(%)	17.2	0.9	0.0	18.1
One Car	21	816	39	876
(%)	1.7	64.5	3.1	69.2
Two+ Cars	0	59	101	160
(%)	0.0	4.7	8.0	12.6
Total	238	887	140	1265
	18.8	70.1	11.1	100

% of cases correctly classified = 89.7

Weekly Motorized-Trip Generation Models (five simulation runs)

	Car Owners		Non Car Owners	
	(a)	(b)	(a)	(b)
N	1036		229	
Trips Observed	32.1		12.1	
Trips Predicted	32.9	31.2	13.0	13.0
%Error	2.65%	-2.64%	7.51%	7.26%
MAE	9.2	10.6	5.7	5.9
MSE	134.6	182.6	51.1	58.0
R^2	0.725	0.620	0.648	0.597

Weekly Household Modal Split Model (five simulation runs)

	Proportion of Transit Trips	Number of Transit Trips
Observed	0.140	2.9
Pred (1)	0.146	
%Error	4.7%	
Pred (2)	0.134	2.7
% Error	4.5%	8.4%
MAE	0.121	2.9
MSE	0.040	22.7
Correlation	0.637	0.519

Notes: (a) Observed car ownership levels are used as input. (b) Simulated car ownership levels are used as input. MAE = Mean absolute error, average of the absolute difference between observed and estimated value. MSE = Mean square error, average of the squared difference between observed and estimated value. Pred(1) = Average of $(1/(1+\exp(-\beta'x))$ across observations. Pred(2) = Obtained by simulation.

of transit trips to obtain the latter measure. The model is performing well as indicated in Table 1.

In validation, the correlation coefficients between observed and predicted Wave 10 mobility measures are often as good as those obtained during model estimation; the models are not only replicating observed behavior well but also predicting future (i.e., Wave 10) behavior with comparable accuracy. The analysis of this section lends support to the simulation forecasting reported in the next section.

MIDAS LONG-TERM FORECASTING

The evolution of household demographics and socioeconomics, car ownership, and mobility is simulated with MIDAS using the expanded/weighed panel household samples. A simulation period of 25 years is used starting with 1986, when the Wave 5 survey was conducted, and ending in 2010. One year is used as the time increment in the simulation. Therefore, the characteristics of each sample household are updated 25 times in the simulation.

One of the objectives of this study is to examine whether dynamic microsimulation forecasting is practical and meaningful. Manipulation of the MIDAS parameters that have been estimated using the Dutch Panel data is kept to the minimum in this paper. In this section, the results of a baseline MIDAS run—Baseline Scenario—are compared with observed Dutch national mobility statistics (hereafter called the OVG mobility measures), car ownership forecasts by van den Broecke (hereafter called the VDB forecasts), and mobility forecasts by the national model.

The MIDAS baseline forecast represents an income growth of 57 percent by 2010. The results are presented in Table 2 for 1986 (the base year), 1995, 2000, 2005, and 2010. All MIDAS results presented in this section are averages of five simulation runs repeated for each simulation case using different seeds for random number generation.

Comparison with Observed 1986 OVG Mobility Measures

Dutch national mobility statistics (*13*) are used to examine the closeness to the Dutch population of the panel sample used in MIDAS. The results are summarized in Table 3. The survey years are exactly the same (i.e., 1986). The two sets of mobility measures are similar, in particular trip generation measures.

The MIDAS base year trip rates are consistently below the 1986 OVG trip rates. It is believed that the OVG mobility measures are averages over all days of the week, including Saturdays and Sundays. For example, the motorized-trip rate is 0.5 percent below the comparable OVG trip rate. This is a weak indication of underreporting in the Dutch panel survey (*14,15*).

Comparison with the VDB Forecasts

On the basis of a cohort model, van den Broecke produced driver's license holdings and car ownership forecasts for the Netherlands (*16–18*). His forecasts are compared with MIDAS forecasts in Table 3. The driver population and the national car ownership forecasts by VDB are close to the MIDAS forecasts.

Good agreement exists between VDB and MIDAS in the 2010 labor force participation forecasts, which are represented here as the percentage of employed persons in the total population. MIDAS assumes practically the same income growth rate as VDB. Considering the fundamental differences in data and methodology, the compatibility between the VDB forecasts and MIDAS results, including driver's license and car ownership, is striking.

Comparison with the National Model

The Dutch national model provides the only mobility forecasts available to this study (*19,20*). The results are summarized in Table 4 along with MIDAS forecasts. The differences in household size and labor force participation are similar to those seen earlier.

The 2010 driver's license holding in the national model forecasts is practically identical to the forecast by MIDAS. Driver's license holdings are forecast in the national mode using a set of discrete choice models formulated at the household level. Thus the forecast is not a simple extrapolation of observed trends. MIDAS forecasts are based on transition probabilities of license holdings, whereas van den Broecke's forecast relies on license ownership probabilities assumed for

TABLE 2 Baseline MIDAS Forecasts, 1986–2010

	Base Year 1986	MIDAS Forecasts 1995	2000	2005	2010	Growth
Population (x 10^6)	14.5				15.1	4.1%
Population, ≥ 12 Years Old (x 10^6)++	12.3				13.0	5.7%
Household Size	2.64	2.38	2.20	2.06	1.94	-26.5%
Labor Force Participation*	42.7%	49.8%	48.4%	45.0%	41.2%	
Average Income per Employed Person	100	127	134	143	157	
Number of Licensed Drivers (x 10^6)**	7.19				10.04	39.3%
Percent of Licensed Drivers	49.6%	58.2%	61.6%	65.2%	66.5%	
Number of Automobiles (x 10^6)**	4.50				7.10	57.8%
Automobiles per Person	0.31	0.39	0.42	0.45	0.47	51.6%
Automobiles per Household	0.82	0.92	0.92	0.92	0.90	9.8%
Automobiles per Driver	0.62	0.66	0.68	0.69	0.70	12.9%
Number of Motorized Trips per Week						
Per Person	9.35	11.64	12.29	12.91	12.86	37.5%
National Total (x 10^6)**	115.0				167.2	45.4%
Number of Car Trips per Week						
Per Person	8.28	10.38	10.95	11.54	11.47	38.5%
National Total (x 10^6)**	101.8				149.1	46.4%
Number of Transit Trips per Week						
Per Person	1.07	1.27	1.34	1.37	1.39	29.9%
National Total (x 10^6)**	13.2				18.1	37.1%

++Van den Broecke (*16*.)
The 2010 figure was adjusted to agree with the CPB forecast.
*Among individuals of 15 years old and over (CPB), or 18 years old and over (MIDAS).
**MIDAS forecasts are expanded using the national population (of individulas of 12 years old and over for mobility measures).

TABLE 3 Comparison of MIDAS Sample and MIDAS Forecasts with 1986 OVG Observations and VDB Forecasts

	OVG 1986*	MIDAS 1986
Population (x 10^6)		14.5
Population, ≥ 12 Years Old (x 10^6)++		12.3
Number of Motorized Trips per Week		
Per Person	9.73	9.35
National Total (x 10^6)**		115.0
Number of Car Trips per Week		
Per Person	8.47	8.28
National Total (x 10^6)**		101.8
Number of Transit Trips per Week		
Per Person	1.26	1.07
National Total (x 10^6)**		13.2
Vehicle-Kilometers Driven per Week		
Per Person	114.1	87.5
National Total (x 10^6)**		1076
Transit Passenger-Kilometers Trips per Week		
Per Person	28.7	23.0
National Total (x 10^6)**		283

	VDB[1]		MIDAS	
	1985	2010	1986	2010
Population (x 10^6)			14.5	15.1
Population, ≥ 12 Years Old (x 10^6)++			12.3	13.0
Labor Force Participation[2]	31%	38%	31.5%	38.6%
Average Income per Employed Person	100	170	100	157
Number of Licensed Drivers (x 10^6)**	6.90	9.30	7.19	10.04
Percent of Licensed Drivers in Population	48.0%	61.0%	49.6%	66.5%
Number of Automobiles (x 10^6)**	4.50	7.90	4.50	7.10
Automobiles per Person	0.31	0.52	0.31	0.47
Automobiles per Household			0.82	0.90

*CBS (*13*)
++Van den Broecke (*16*).
The 2010 figure was adjusted to agree with the CPB forecast.
**MIDAS forecasts are expanded using the national population of individulas of 12 years old and over.
[1]Van den Broecke (*17*)
[2]Percentage of employed persons in the total population.

respective population age cohorts. These three entirely different forecasting methods have produced 2010 driver population forecasts that are within 8 percent of each other.

Vehicle-kilometrage growth forecasts of MIDAS and the national model are again strikingly similar. The national model forecasts an increase of 72 percent by 2010. The corresponding MIDAS forecast is an increase of 80.5 percent.

The forecasts of public transit use are drastically different between the two. The national model predicts a slight decrease in public transit passenger-kilometers by 2010, and MIDAS forecasts an increase of 46 percent in car trips and an increase of more than 112 percent in vehicle kilometers. No changes in accessibility and levels-of-service are assumed in either method.

This discrepancy in public transit use between MIDAS and the national model is perhaps the single most important discrepancy. Unfortunately, there is no other comparable forecast available to this study to indicate which forecast is more likely. Both are based on elaborate model systems formulated at the household level. One important difference is that the national model is formulated using cross-sectional data, and longitudinal changes in population compositions are represented by weighting households (as in static microsimulation). MIDAS, on the other hand, is based on longitudinal data and simulates household evolution over time.

CONCLUSIONS

This study, representing an entirely new approach to travel demand forecasting, is based on the recognition that no external demographic and socioeconomic forecasts are furnished at levels that meet the data requirements of sophisticated discrete choice models currently used in transportation planning. Specifically, no external forecasts are produced to provide a multivariate distribution of the array of explanatory variables typically used in travel choice models at the levels where these models are formulated (i.e., households or individuals).

TABLE 4 Comparison of MIDAS Forecasts with National Model Forecasts

	National Model#		MIDAS		
	1986	2010	1986	2010	Growth
Population (x 10^6)	14.3	15.1	14.5	15.1	
Population, ≥ 12 Years Old (x 10^6)++			12.3	13.0	
Household Size	2.70	2.29	2.64	1.94	
Total Workforce (x 10^6)	4.6	6.1			
Labor Force Participation	39.2%	48.5%	42.7%	41.2%	
Number of Licensed Drivers (x 10^6)**	6.6	10.4	7.19	10.04	
Percent of Licensed Drivers	46.2%	68.9%	49.6%	66.5%	
Number of Automobiles (x 10^6)**	4.3	7.9	4.50	7.10	
Automobiles per Person	0.30	0.52	0.31	0.47	
Automobiles per Household	0.81	1.20	0.82	0.90	
Change in Weekday Vehicle-Kilometers[2]	+72%				
Vehicle-Kilometers Driven per Week					
Per Person			87.5	149.4	70.7%
National Total (x 10^6)**			1076	1942	80.5%
Change in Weekday BMT Passenger-Kilometers[2]	-7%				
Change in Weekday Rail Passenger-Kilometers[2]	-2%				
Transit Passenger-Kilometers per Week					
Per Person			23.0	46.3	101.3%
National Total (x 10^6)**			283	602	112.7%

Vrolijk, Gunn and van der Hoorn (*19*), Gunn, van der Hoorn and Daly (*20*)
++Van den Broecke (*16*).
**MIDAS forecasts are expanded using the national population (of individulas of 12 years old and over for mobility measures).
[1]Estimated using the total population and the number of households used in a National Model study.
[2]Read from a graph in (*20*)

The use of dynamic microsimulation is motivated by its flexibility and its ability to forecast direct and indirect effects of the simulated policies on the system analyzed. Microsimulation helped to fill the gap in forecasting the input to travel demand models and provided the framework for designing the new dynamic forecasting tool MIDAS. It generates demographic and socioeconomic, as well as car ownership and mobility forecasts internally through microsimulation. A system of dynamic models estimated using the Dutch National Mobility Panel data set is applied in this simulation.

The primary objective of the study—to determine whether long-range travel demand forecasting can be practically and meaningfully performed using microsimulation with a system of dynamic models and parameters estimated using a panel data set—has been met, along with the secondary objective—to design a flexible tool for building scenarios based on alternative policy strategies. The forecasting exercise reported here is the evidence that a dynamic microsimulator is a credible forecasting model system. In addition, a large number of parameters can be modified by the user to represent scenarios of interest; the microsimulator can automatically simulate the repercussions that follow and reflect them in its mobility forecasts.

Dynamic microsimulation offers many advantages over the traditional cross-sectional models with externally produced sociodemographic variables. However, it is complex and requires a large amount of data. The estimation of a dynamic mobility model requires more data than does a corresponding cross-sectional model. The estimation of dynamic models using panel data requires additional attention because of panel attrition, conditioning, and fatigue.

The dynamic microsimulator described in this paper is the first step toward a full-fledged dynamic microsimulation forecasting system in the transportation planning field. Despite meeting the study objectives, the dynamic microsimulator is not yet a completed tool. Its current version needs to be improved in a number of ways (*2,9*).

ACKNOWLEDGMENTS

This research was supported by the U.S. Department of Transportation Region 9 Transportation Center. The comments on an earlier version of this paper provided by the anonymous reviewers are gratefully acknowledged.

REFERENCES

1. C. R. Fleet. Workshop on Long-Range and Strategic Forecasting Techniques. In *Special Report 201: Travel Analysis Methods for the 1980s*, TRB, National Research Council, Washington, D.C., 1983.
2. K. G. Goulias. *Long-Range Forecasting with Dynamic Microsimulation*. Ph.D. dissertation. University of California-Davis, Davis, 1991.
3. J. N. Bajpai. *National Cooperative Highway Research Program Report 328: Forecasting the Basic Inputs to Transportation Planning at the Zonal Level*. TRB, National Research Council, Washington, D.C., 1990.

4. K. Arrow. Microdata Simulation: Current Status, Problems, Prospects. In *Microeconomic Simulation Models for Public Policy Analysis*, (R. Haveman and K. Hollenbeck, eds.), Vol. 2, Academic Press, New York, 1980.
5. G. Orcutt, A. Glazer, R. Harris, and R. Wertheimer II. Microanalytic Modeling and the Analysis of Public Transfer Policies. In *Microeconomic Simulation Models for Public Policy Analysis*, (R. Haveman and K. Hollenbeck, eds.), Vol. 1, Academic Press, New York, 1980.
6. R. B. Davies. The Limitations of Cross Sectional Analysis. In *Longitudinal Data Analysis*, (R. Crouchley, ed.), Gower, Aldershot, England.
7. R. Kitamura. Panel Analysis in Transportation Planning: An Overview. *Transportation Research*, Vol. 24A, No. 6, 1991, pp. 401–415.
8. L. P. Kostyniuk and R. Kitamura. Life Cycle and Household Time-Space Paths: Empirical Investigation. In *Transportation Research Record 879*, TRB, National Research Council, Washington, D.C., 1982.
9. R. Kitamura and K. G. Goulias. *MIDAS: A Travel Demand Forecasting Tool Based on a Dynamic Model System of Household Demographics and Mobility*. Projectbureau Integrale Verkeer-en Vervoerstudies, Ministerie van Verkeer en Waterstaat, the Netherlands, 1991.
10. R. Kitamura. A Dynamic Model System of Household Car Ownership, Trip Generation, and Modal Split: Model Development and Simulation Experiments. *Proc., 14th Australian Road Research Board Conference*, Part 3, Australian Road Research Board, Vermont South, Victoria, Australia, 1988, pp. 96–111.
11. J. Geinzer and A. Daly. *Zuidvleugel Study—Report 9, Models of Car Ownership and License Holding*. Cambridge Systematics Europe, b.v., The Hague, the Netherlands, 1981.
12. K. G. Goulias and R. Kitamura. Analysis of Binary Choice Frequencies with Limit Cases: Comparison of Alternative Estimation Methods and Application to Weekly Household Mode Choice. *Transportation Research* (forthcoming).
13. *De Mobiliteit va de Nederlandse Bevolking in 1986*. Centraal Bureau voor de Statistiek, The Hague, the Netherlands, 1987.
14. T. F. Golob and H. Meurs. Biases in Response over Time in a Seven-Day Travel Diary. *Transportation*, Vol. 13, 1986, pp. 163–181.
15. H. Meurs, L. van Wissen, and J. Visser. Measurement Biases in Panel Data. *Transportation*, Vol. 16, No. 2, 1989, pp. 175–194.
16. A. A. J. van den Broecke. *Demografische Ontwikkelingen en Verkeers- en Vervoerproblematiek in de Komende 25 Jaar*. Amsterdam, the Netherlands, 1987.
17. A. A. J. van den Broecke. *De Mogelijke Groei van het Personenautobezit tot 2010*. Amsterdam, the Netherlands, 1987.
18. A. A. J. van den Broecke. Long-Term Forecasting of Car Ownership with the Cohort Processing Model. *PTRC 16th Summer Annual Meeting, Proceedings of Seminar D*. PTRC Education and Research Services, Ltd., London, 1988, pp. 183–191.
19. P. Vrolijk, H. Gunn, and T. van der Hoorn. Waar komt de groei vandaan? Presented at the 14th Colloquium Vervoersplanologisch Speurwerk, Den Haag, the Netherlands, 1987.
20. H. Gunn, T. van der Hoorn, and A. Daly. Long-Range, Country-Wide Travel Demand Forecasts from Models of Individual Choice. Presented at the Fifth International Conference on Travel Behavior Analysis, Aix-en-Provence, France, 1987.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.