A dynamic model system of household car ownership, trip generation, and modal split: model development and simulation experiment

Ryuichi Kitamura

ABSTRACT A dynamic model of household car ownership and mode use is developed and applied to demand forecasting. The model system consists of three interrelated components: car ownership, mechanized trip generation, and modal split. The level of household car ownership is represented as a function of household attributes and mobility measures from the preceding observation time point using an ordered-response probit model. The trip generation model predicts the weekly number of trips made by household members using car or public transit, and the modal split model predicts the fraction of trips that are made by public transit. Household car ownership is a major determinant in the latter two model components. A simulation experiment is conducted using sample households from the Dutch National Mobility Panel data set and applying the model system to predict household car ownership and mode use under different scenarios on future household income, employment, and drivers' license holding. Policy implications of the simulation results are discussed.

KEYWORDS Automobile ownership · Dynamic model · Mode split ·
Ordered probit · Panel data · Simultaneous equations

1. INTRODUCTION

This paper presents the development of a simultaneous equations system of household car ownership, mechanized trip generation, and modal split, and its application to simulate future car ownership and mode use. The analysis is motivated by the concern that the increasing household car ownership in the Netherlands may lead to excessive car trip demand that cannot be reasonably met by expanding the infrastructure. Underlying this

---

R. Kitamura
University of California, Davis, CA 95616, USA

concern is the commonly held belief that an increase in car ownership leads to an increase in mobility and car use, and a decrease in public transit use. It is also believed that an increase in household income or licensed drivers will lead to an increase in car ownership. However, there are indications that an increase in income tends to lead to an increase in public transit use. A question that then arises is whether or not public transit patronage will diminish as income and driver's license holding increase and, consequently, car ownership increases, or will public transit prosper with increasing car ownership?

This issue can be best addressed by adopting an analytical framework in which household car ownership and travel behavior are treated as interactive elements: car ownership influences mobility, and mobility in turn leads to a decision to acquire or dispose of a car. Because this interactive relationship evolves over time, incorporating dynamism into the analysis is of extreme importance. Recent findings from dynamic analyses of travel behavior support this viewpoint by showing that inferences made using cross-sectional observation may lead to inaccurate characterization of travel behavior (e.g., Goodwin 1986 1987; Kitamura and van der Hoorn 1987). It is believed that a model system estimated using a panel data set will be capable of representing dynamic aspects of travel behavior more accurately and offer better prediction of travel behavior.

There exist many studies that have addressed the issue of car ownership and car sales (e.g., Ben-Akiva and Lerman 1974; Mogridge 1983), or car acquisition and utilization (Train and Lohrer 1983; Mannering and Winston 1985; de Jong and Cramer 1987). The effort in this present study takes on the approach of modeling car ownership at the household level on the basis of demographic and socioeconomic characteristics of households. This is based on the belief that the major factors that influence the acquisition or disposal of a family car are sociodemographic in nature, e.g., young household members forming a separate household, a member of the household gaining employment, etc. (see Town 1983; Onnen et al. 1987).

A panel data set consisting of repeated observations of the same households over time is used in this study to develop a dynamic simultaneous equations system. The objective of this research is (1) to identify dynamic relationships among household car ownership, mode usage, and sociodemographic characteristics, and further (2) to show that such a dynamic model system can be practically used to forecast future travel demand. Throughout this study, the household is used as the unit of analysis.

In the model system, trip generation and modal split are assumed to be functions of household car ownership. In addition to this conventional assumption, it is assumed that car ownership is influenced by trip generation and modal splits of the past. These three interactive measures of household travel behavior are endogenous to the model system. Accordingly, the system may be viewed as a dynamic model of mechanized trip generation and modal split in which household car ownership is an endogenous variable.

The modeling effort assumes that observed behavior at time t depends on the behavior at time $t-1$. One motivation for this formulation is to capture the dynamism among household characteristics, car ownership and mobility. In addition, an analysis that preceded the effort reported here (Kitamura 1987c) has indicated that a lagged dependent variable significantly contributes to the model's fit. Because of the inclusion of a lagged dependent variable, the models of this study behave differently from conventional cross-sectional models in which observed behavior at t, say $Y(t)$, is expressed as a function of contributing factors also observed at t, $X(t)$. A change in X will be immediately reflected as a change in Y and behavior is depicted to be entirely independent of the past history of either X or Y in a cross-sectional model. With lagged dependent variables, a change in X

will cause a change in behavior that continues over a certain period and the same change in X may show a different impact in the short run depending on the past history of Y.

The structure of the model system of this study is similar to that in Hensher, Bernard, Smith, and Milthorpe (1987). The study emphasis, however, differs importantly between the two studies. Hensher, et al. propose a dynamic model system to describe vehicle type choice (fleet size, body type, and model) and utilization using a system of nested logit and linear regression models. In the model system of this present study, the car ownership component is concerned only with the number of cars accessible to the household and is formulated using the ordered-response probit. In contrast, the modal split between private car and public transit is incorporated in the present model system. Hensher, et al. consider a range of variables representing vehicle characteristics, while this present study is concerned with medium- to long-term impact of economic and sociodemographic changes on car ownership and use. The two studies take different approaches to account for household-specific effects. An error component scheme is used by Hensher et al. in the linear vehicle utilization model, while in the present study, serially correlated errors are used. Smith et al. (1986) address the issue of initial condition. In contrast, the present model system incorporates correlations among the error term of the car ownership model and those of the trip generation and modal split components.

The effort reported in this paper toward a dynamic model system of household car ownership and mobility is ongoing. The model components, estimation methods, and forecasting procedures are continuously improved and refined. The results presented in this paper may not necessarily be the final outcome of the effort. Nonetheless, reporting the results available thus far is useful because they establish that a dynamic model system can be a practical planning tool.

The rest of this paper is organized as follows. The structure of the model system is presented in the next section, together with a brief discussion of the procedure for parameter estimation. Next, the results of model estimation are presented by model component. This is followed by a discussion of the results of simulation experiments. In the experiments, the model system is applied to predict the impact of increases in household income, number of licensed drivers, and number of workers in households upon car ownership and trip generation by mode. The last section is a summary of the study.

2. MODEL SYSTEM

The three endogenous variables of the model system are household car ownership, mechanized trip generation, and modal split (defined in this study as the fraction of transit trips among the total mechanized trips made by household members). Figure 1 presents the interrelationships assumed among these three elements. As indicated by the vertical arrows, each of the three elements is assumed to be dependent upon its own value at the preceding time point. This dependence is represented in the model by lagged dependent variables, i.e., by introducing the dependent variable from the preceding time point to the model as an explanatory variable.

Mechanized trip generation and modal split are assumed to be dependent upon car ownership at the same time point. This is a conventional assumption used in most trip generation and mode choice models. Household car ownership is assumed to be influenced by trip generation and modal split from the preceding time point. This reflects the hypothesis that an experience at one time point may lead to a decision to acquire or dispose of a car, thus influencing car ownership at later time points. Whether such an effect of experience exists or not can be statistically tested by estimating the model system.
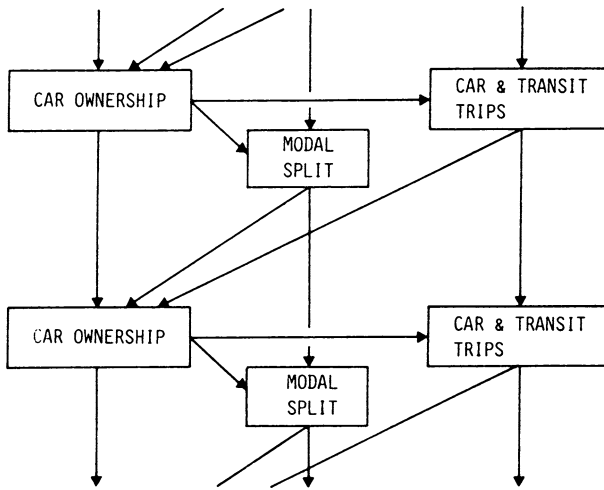
Fig. 1 Structure of the model system

The car ownership model is formulated using the ordered-response probit model. The linear regression model is applied to mechanized trip generation, while the logistic response curve is used to represent modal split. The logistic curve is estimated using a weighted least-squares procedure. Each of these model components is described in the rest of this section.

2.1 Car Ownership Model

Let the level of household car ownership be expressed by three categories: no car, one car, and two or more cars, and define A(i,t) as

$$A(i,t) = \begin{cases} 0 \text{ if household i has no car available at time t;} \\ 1 \text{ if household i has one car available at time t;} \\ 2 \text{ if household i has two or more cars available at time t.} \end{cases}$$

Consider a latent variable, I(i,t), for household i and time point t such that
$I(i,t) = \alpha'X(i,t) + \theta'D(i,t-1) + a_1G(i,t-1) + a_2F(i,t-1) + \varepsilon(i,t)$, and

$$A(i,t) = \begin{cases} 0 \, if \, I(i,t) \leq h(t); \\ 1 \, if \, h(t) < I(i,t) \leq k(t); \\ 2 \, if \, k(t) < I(i,t); \end{cases}$$

where
α and θ are coefficient vectors;
$X(i,t)$ is a vector of exogenous variables;
$D(i,t-1)$ is a vector of dummy variables representing the car ownership level at time $t-1$; a_1 and a_2 are scalar coefficients;
$G(i,t-1)$ is the number of mechanical trips made by household members at time $t-1$;
$F(i,t-1)$ is an indicator of modal split at time $t-1$;
$\varepsilon(i,t)$ is a random error term; and

h(t) and k(t) are unobserved threshold values for car ownership at time t.

If $\varepsilon(i,t)$ has a standard normal distribution, household car ownership can be formulated as an ordered-response probit model,

$\Pr[A(i,t) = n]$

$$= \begin{cases} \Phi(h(t) - B(i,t)), & \textit{if } n = 0; \\ \Phi(k(t) - B(i,t)) - \Phi(h(t) - B(i,t)), & \textit{if } n = 1; \\ 1 - \Phi(k(t) - B(i,t)), & \textit{if } n = 2; \end{cases}$$

where

$B(i,t) = \alpha'X(i,t) + \theta'D(i,t-1) + a_1G(i,t-1) + a_2F(i,t-1);$ and

Φ = the standard normal cumulative distribution function.

This formulation is used in the analysis to represent household car ownership in a dynamic context.

2.2 Mechanized Trip Generation Model

The standard least squares regression analysis is applied to develop a model of mechanized trip generation. The dependent variable of the model, G(i,t), is the weekly total number of trip segments made by mechanized modes (car driver, car passenger, bus, tram, metro, and train) as reported by diary-keepers in the household. The model has the form

$G(i,t) = \beta'Z(i,t) + q'D(i,t) + aG(i,t-1) + U(i,t);$

where

β and q are coefficient vectors;

a is a scalar coefficient;

Z(i,t) is a vector of explanatory variables; and

U(i,t) is a random error term.

The inclusion of D(i,t) reflects the assumption that mechanized trip generation is dependent on the level of car ownership. Also note that the model includes the lagged dependent variable, $G(i,t-1)$.

The model is estimated without applying any transformation to this dependent variable because a weekly total number of trips is likely to lead to residuals with desirable properties due to the law of large numbers. Previous model estimation results using weekly measurements agree with this expectation. No correction for possible heteroscedasticity is conducted for this model, also because such correction applied to similar weekly measurements in previous efforts produced no appreciable difference in estimated coefficients or standard errors.

2.3 Modal Split Model

Let p(i,t) be the fraction of public transit trips among the total mechanized trips.

$p(i,t) = H(i,t)/G(i,t),$

where H(i,t) is the total number of transit trips and we assume that $G(i,t) > 0$. Suppose this ratio can be expressed using a logistic function as

$p(i,t) = 1/[1 + \exp(F(i,t))],$

where

$F(i,t) = \gamma'W(i,t) + r'D(i,t) + bF(i,t-1) + V(i,t)$

and γ and r are coefficient vectors, b is a scalar coefficient, W(i,t) is a vector of explanatory variables, and V(i,t) is a random error term. After a frequently used transformation, the left-

hand side of the above equation can be expressed as the logit, $\ln[C(i,t)/H(i,t)]$, where $C(i,t) = G(i,t) - H(i,t)$. Assuming that $V(i,t)$ is normal, this model can be estimated using weighted least squares, with the weight (see, e.g., Theil 1971, p. 635)

[H(i,t) C(i,t)/G(i,t)] $^{1/2}$.

2.4 Estimation of the Model System

The inclusion of the lagged dependent variables implies that model estimation will encounter added difficulty if serial correlation is present among the errors. Correlated errors across the models will also impose problems because of the inclusion of endogenous variables from other model components. Consistent estimators exist if the model system is linear. The present model system, however, includes the ordered-response probit model to which these estimators are not applicable. Model estimation for this study is therefore conducted using correction terms to account for correlated errors. The method is an extension of that in Heckman (1979) and discussed in Kitamura (1987a; also see Kitamura and Bovy 1987, for related discussions). Note that the linear models (trip generation and modal split after transformation) can be estimated using generalized least squares, together with correction terms to account for correlations involving errors of the car ownership model. This, however, is not performed in this study. Also note that the use of the correction terms introduces heteroscedasticity under certain conditions, leading to inconsistent estimation of standard errors. Correction for this added heteroscedasticity remains a future task.

3. DATA SET

The observations from waves 1, 3, 5 and 7 of the Dutch National Mobility Panel survey are used in this study. The four waves were conducted in the spring of 1984, 1985, 1986, and 1987, respectively. A unique feature of this panel data set is that, in addition to the conventional set of demographic and socioeconomic information on the household and its members, a weekly travel diary was collected from each member of the sample households who was at least 12 years old. The weekly trip information available from the data set makes trip frequency information more reliable and allows analysis of weekly travel patterns and activity scheduling. The variables that appear in the models presented in this report are summarized in Table 1.

The numbers of households in the data file are 1,774, 1,687, 1,849, and 1,926 in wave 1, 3, 5, and 7, respectively. In each wave, new households were introduced into the panel while some households left the panel. Because of this sample refreshment and attrition, the size and composition of the sample varies substantially from analysis to analysis in this study; 1,031 households are in waves 1 and 3, 1,333 in waves 3 and 5, and 1,391 in waves 5 and 7; 853 in waves 1, 3 and 5, and 1,035 in waves 3, 5 and 7. 668 households are in all four waves. The number of households available for analysis is usually smaller than these figures because of missing variables.

Additional discussions of the aim, format and characteristics of this panel data set can be found in Golob et al. (1986a), Golob and Meurs (1986), Kitamura and Bovy (1987), Meurs et al. (1987a), and Meurs et al. (1987b). A sample of analyses using the Dutch Panel data set includes Golob (1986, 1987), Golob and Meurs (1987, 1988), Golob et al. (1986b), Goodwin (1987), Kitamura (1987a–c, 1988), Kitamura and van der Hoorn (1987), Recker et al. (1987), and van der Hoorn and Kitamura (1987).

Table 1 Definition of the Variables Used in the Model System

Variable	Definition
CARTRIPS	Weekly total number of car trip segments reported by household members
TRANTRIPS	Weekly total number of transit trips (bus, tram, metro, train) reported by household members
MODALSPLIT	ln[CARTRIPS + 0.5)/(TRANTRIPS + 0.5)]
MECHTRIPS	Sum of CARTRIPS and TRANTRIPS
NCARS	Number of cars available to the household
ONECAR	1 if the household has one car available; 0 otherwise
TWOCARS	1 if the household has two or more cars available; 0 otherwise
NDIARIES	Number of household members ($\geq$12 years old) who filled out diaries
NDRIVERS	Number of licensed drivers in household
ONEDRIVER	1 if NDRIVERS is 1; 0 otherwise
TWODRIVERS	1 if NDRIVERS is 2; 0 otherwise
MLTDRIVERS	1 if NDRIVERS is greater than 2; 0 otherwise
NONDRIVERS	Number of household members ($\geq$12 years old) who did not hold driver's licenses
NWORKERS	Number of employed household members
MLTWORKERS	1 if NWORKERS is greater than 1; 0 otherwise
CHLD0-6	Number of children between 0 and 6 years old
CHLD7-11	Number of children between 7 and 11 years old
CHLD12-17	Number of children between 12 and 17 years old
SINGLE	1 for a single-person household; 0 otherwise
ELDERLY	1 for a household without children whose representative member is over 65 years old; 0 otherwise
YNGNOCHLD	1 for a household of a couple without children whose representative member is under 35 years old;
DINCS	1 if YNGNOCHLD is 1 and both members of the couple work; 0 otherwise
YNGHIED	1 if YNGNOCHLD is 1 and at least one member has a college degree; 0 otherwise
YUPPY	1 for a household in a large metropolitan area for which YNGHIED is 1; 0 otherwise
$\sqrt{\text{INCOME}}$	Square-root of the net annual household income (in Dfl) divided by 100
PPINCOME	Net annual income per household member (in Dfl) divided by 10,000
LOEDUC	1 if the household member with the highest education had completed only primary school; 0 otherwise
HIEDUC	If at least one person in the household has a college degree; 0 otherwise
BOV-LARGE	1 if the household resides in a metropolitan area with highly developed transit systems; 0 otherwise
BOS-MEDIUM	1 if the household resides in a community which is served by rail; 0 otherwise
FOR-SMALL	1 if the household resides in a community which is not served by rail; 0 otherwise

4. RESULTS OF MODEL ESTIMATION

The results of model estimation are presented in this section by model component. Because the models include lag terms, their estimation requires that sample households be in every wave of the survey used for the estimation. This substantially reduces the sample size as the number of households in all odd numbered waves is less than 670. In order to increase the sample size, observations are pooled to create a data base with three time points:

records of 853 households in waves 1, 3 and 5 and those of 1,035 households in waves 3, 5 and 7 are pooled to form a data file of 1,888 records. Each record in the pooled data file contains observations from three time points, either waves 1, 3, and 5, or 3, 5, and 7. Implicit in this pooling is the assumption that the model parameters remain stable across the waves. The observations from the first time point are used to obtain initial correction terms and those from the last two time points are used to estimate the model coefficients.

4.1 Car Ownership Model

The estimated coefficients and associated statistics of the household car ownership model are presented in Table 2. Overall, the model is highly significant with a likelihood ratio chi-square value $(= -2\left[L(\hat{C}) - L(\hat{\beta})\right])$ of 1966.2 with degrees of freedom (df) 15, and a pseudo-coefficient of determination $(= 1 - L(\hat{\beta})/L(\hat{C}))$ of 0.635.

Table 2 Ordered-Response Probit Car Ownership Model

	$\hat{\beta}$	t
ONECAR(t − 1)	2.773	15.00
TWOCARS(t − 1)	5.284	17.76
MECHTRIPS(t − 1)/100	.817	1.79
MODALSPLIT(t − 1)	.017	.43
Q[D(t − 1)]	−.461	−5.98
U(t − 1)/100	−.639	−1.13
V(t − 1)	.052	.87
ONEDRIVER(t − 1)	1.061	4.58
TWODRIVERS(t − 1)	1.390	5.42
MLTDRIVERS(t − 1)	1.735	5.16
DDRIVERS(t,t − 1)*	.845	8.07
NWORKERS(t)	.095	1.51
SINGLE(t)	.180	1.04
PPINCOME(t)/10,000	.060	.96
BOV-LARGE(t)	−.051	−.48
h(t)	2.546	10.62
k(t)	6.348	23.24
N	1832	
L(0)	−2012.7	
L($\hat{C}$)	−1548.3	
L($\hat{\beta}$)	−565.2	
−2[L(0) − L($\hat{\beta}$)]	2895.0 (df = 17)	
−2[L($\hat{C}$) − L($\hat{\beta}$)]	1966.2 (df = 15)	
1 − L($\hat{\beta}$)/L(0)	.719	
1 − L($\hat{\beta}$)/L($\hat{C}$)	.635	

L(0) = log-likelihood with all model coefficients set to 0
L($\hat{C}$) = log-likelihood with h(t) and k(t) alone
L($\hat{\beta}$) = maximum log-likelihood with coefficients as shown in the table
*DDRIVERS(t,t − 1) = NDRIVERS(t) − NDRIVERS(t − 1)

Household car ownership at time t is heavily dependent on that at time t − 1 (ONE-CAR(t − 1) and TWOCARS(t − 1)). The lagged mobility indicators, number of mechanized trips at time t − 1 (MECHTRIPS(t − 1)) and modal split (MODALSPLIT(t − 1)), on the other hand, are not as influential; the coefficient of MECHTRIPS is positive but relatively small and not significant at α = 0.05. The coefficient of MODALSPLIT is insignificant. The effect of past trip making and mode usage upon car ownership appears to be only marginal.

The number of drivers in the household is a significant determinant of car ownership. This factor is represented by a set of four variables: ONEDRIVER(t − 1), TWODRIVERS(t − 1) and MLTDRIVERS(t − 1), representing the value at time t − 1, and DDRIVERS(t,t − 1), the change between t − 1 and t. The first three dummy variables are introduced to represent a possible non-linear effect, the existence of which is evident from the estimated coefficients.

If there has been no change over time in the number of drivers, the expected difference in the latent variable value between no-driver households and one-driver households is, *ceteris paribus*, 1.061 − 0 = 1.061, that between one-driver households and two-driver households is 1.390 − 1.061 = 0.329, and that between two-driver households and multi-driver households is 1.735 − 1.390 = 0.345. Note the large coefficient (0.845) of the difference in the number of drivers (DDRIVERS), which indicates that an increase or decrease in the number of drivers has a one-time impact on car ownership and suggests that such a change tends to trigger a change in the car ownership level.

Other indicators of household characteristics – number of workers (NWORKERS), single-person household (SINGLE), income per person (PPINCOME) and transit service level (BOV-LARGE) – have mostly anticipated signs, but their effects are minor. Quite notably, the transit service level indicator (BOV-LARGE) is insignificant. This is presumably because of the infrequent change in this variable, which in turn is due to the relatively small number of households that changed residence locations between surveys. Accordingly, the effect of transit service levels is already represented by the lagged dependent variable and the insignificance of BOV-LARGE does not necessarily imply that transit service levels are not associated with household car ownership.

Three correction terms are introduced into the model to account for possible correlations among the error terms. The negative and significant coefficient of Q[D(t − 1)] implies the presence of negative serial correlation in the error term of the car ownership model. This implies, for example, that a household may not acquire a car immediately after its number of drivers increases (a negative error term), but it is likely to do so in the next period in which a positive error is likely to realize. Therefore the negative serial correlation may be interpreted as an indication of lags in car ownership change.

The coefficients of the error terms of the trip generation model (U(t − 1)) and modal split model (V(t − l)) are not significant; car ownership is significantly influenced by neither past mobility indicators nor their error terms. (The correction terms have been developed for the case which involves up to three errors, while the model here involves four error terms. However, the insignificance of two of the correction terms implies that correlations among errors are properly accounted for.)

4.2 Mechanized Trip Generation Model

Household car ownership, change in number of drivers, and income are the major determinants of mechanized trip generation (Table 3). The car ownership dummy variables are multiplied by the number of diary-keepers such that the coefficients of these variables represent the difference in trip generation per diary-keeper. The coefficients of ONECAR and TWOCARS indicate that a diary-keeper from a one-car household on average makes

Table 3 Household Weekly Mechanized Trip Generation Model

	$\hat{\beta}$	t
MECHTRIPS(t − 1)	.881	27.39
Q[D(t)]	.060	.08
U(t − 1)	−.379	−9.64
V(t − 1)	−.368	−1.51
ONECAR(t − 1) × NDIARIES(t)	.955	3.15
DONECAR(t, t − 1) × NDIARIES(t)*	2.711	3.94
TWOCARS(t − 1) × NDIARIES(t)	1.490	3.30
DTWOCARS(t, t − 1) × NDIARIES(t)	4.615	4.48
DNDIARIES(t, t − 1)	3.653	4.09
DDRIVERS(t, t − 1)	3.207	4.13
DNWORKERS(t, t − 1)	1.649	2.94
DCHLD0-6(t, t − 1)	−.992	−1.22
$\sqrt{INCOME(t - 1)}$	1.732	2.51
DINCOME(t, t − 1)	3.617	4.20
Constant	−1.837	−1.61
N	1832	
R^2	.673	
F	267.6	
df	(14, 1817)	

Dependent variable = MECHTRIPS(t)

*DONECAR(t, t − 1) = ONECAR(t) − ONECAR(t − 1)

DTWOCARS(t, t − 1) = TWOCARS(t) – TWOCARS(t − 1)

DNDIARIES(t, t − 1) = NDIARIES(t) – NDIARIES(t − 1)

DDRIVERS(t, t − 1) = NDRIVERS(t) – NDRIVERS(t − 1)

DNWORKERS(t, t − 1) = NWORKERS(t) – NWORKERS(t − 1)

DCHLD0-6(t, t − 1) = CHLD0-6(t) – CHLD0-6(t − 1)

DINCOME(t, t − 1) = $\sqrt{INCOME(t)}$ – $\sqrt{INCOME(t - 1)}$

0.955 more mechanized trips than does a diary-keeper from a no-car household, and that a diary-keeper from a multi-car household makes 1.490 more trips.

The use of the difference terms reflects the hypothesis that a change in these attributes has a short-term impact on trip generation that is different from the long-term effect that will realize after behavior reaches a new equilibrium. This short-term impact is assumed to be symmetric in the model formulation, namely, both an increase and a decrease in car ownership have effects of the same magnitude (see Kitamura 1987b, for alternative assumptions). For example, the number of mechanized trips per diary-keeper will increase by 2.711 following the acquisition of a first car (from 0 to 0 + 2.711), and it will decrease by the same 2.711 when a one-car household disposes of its car. Acquiring a second car will lead to 1.904 more trips per diary-keeper (from 0.955 to 0.955 − 2.711 + 4.615), while disposing of a second car implies 1.904 fewer trips.

The difference in trip rate implied by the difference terms is much larger than the cross-sectional difference depicted by the coefficients of ONECAR and TWOCARS. Short-term response to car ownership change tends to be an over-reaction. The income coefficients

offer a similar indication of over-reaction in mechanized trip generation following an income change.

The lagged dependent variable is extremely significant with a coefficient of 0.881, indicating strong stability in trip generation over time. The estimated coefficient of $U(t-1)$ indicates the presence of negative serial correlation in the error term of this model. Previous analyses (Kitamura and Bovy 1987; Kitamura 1987a) have indicated that error terms of trip generation models are positively correlated over time. However, the models used in the previous analyses did not contain lagged dependent variables. The reason for the contradictory results is currently under investigation.

The coefficients of the other two correction terms indicate that the error term of this trip generation model is not correlated with that of the car ownership model and its correlation with the error term of the modal split model is, at most, very weak. The model's overall fit is very good ($R^2 = 0.673$ with 1,832 observations) and its residuals exhibit desirable properties.

4.3 Modal Split Model

The modal split model was estimated using the logit defined earlier as the dependent variable. The estimated coefficients of ONECAR, TWOCARS, DONECAR, and DTWOCARS (Table 4) show a positive association between car ownership and the

Table 4 Household Weekly Modal Split Model

	$\hat{\beta}$	t
SPLIT(t − 1)	.228	10.60
Q[D(t)]	−.405	−4.63
U(t − 1)/100	.513	2.84
V(t − 1)	.044	1.40
ONECAR(t − 1)	1.249	16.41
DONECAR(t, t − 1)*	1.983	9.80
TWOCARS(t − 1)	1.950	16.60
DTWOCARS(t, t − 1)	3.005	8.39
CHLD7-11(t)	.076	1.70
CHLD12-17(t)	−.161	−5.65
$\sqrt{INCOME}$(t − 1)	−.500	−12.37
DINCOME(t, t − 1)	−.278	−4.08
BOV-LARGE(t)	−.276	−4.99
BOS-MEDIUM(t)	−.165	−3.19
Constant	1.660	28.77
N	1832	
R^2	.487	
F	123.4	
df	(14, 1817)	

Dependent variable = MODALSPLIT (t)

*DONECAR(t, t − 1) = ONECAR(t) − ONECAR(t − 1)
DTWOCARS(t, t − 1) = TWOCARS(t) − TWOCARS(t − 1)
DINCOME(t, t − 1) = $\sqrt{INCOME}$(t) − $\sqrt{INCOME}$(t − 1)

fraction of trips made by car, either as driver or passenger (a positive coefficient of this model indicates a positive contribution of the variable to car use). The coefficients of the level variables (ONECAR and TWOCARS) and the difference variables (DONECAR and DTWOCARS) suggest short-term over-reaction in modal split similar to that found earlier for mechanized trip generation.

The negative coefficient of number of children between 12 and 17 years old (CHLD12-17) who are also diary-keepers, can be linked to the frequent use of public transit by this group. Higher transit use in larger metropolitan areas with well-developed public transit systems is reflected in the coefficients of the transit service level indicators (BOV-LARGE and BOS-MEDIUM).

The highly significant, negative coefficients of the income variables indicate that households with higher incomes tend to use public transit more frequently. A possible explanation of this result is the use of public transit for commuting by high-income groups found in other studies of the Dutch Panel data set (e.g., Meurs, et al., 1987b). Also note that the coefficient of income difference (DINCOME) is about one-half of that of the income level variable ($\sqrt{\text{INCOME}}$); change in modal split that follows income change involves time lags.

Modal split at the household level changes more spontaneously in response to changes in contributing factors. This is evidenced by the coefficient of the lagged dependent variable (0.228), which is significant, but much smaller than that for trip generation (0.881). The coefficient of $V(t-1)$ is positive, but not significant. The error term of the modal split model has only weak serial correlation.

The coefficient of $Q[D(t)]$ is negative and significant; a negative correlation exists between the unobserved elements influencing modal split and those influencing car ownership. A household that has fewer cars than expected (a negative error for car ownership) tends to make heavier use of the cars that they have (a positive error for modal split), and vice versa. The positive coefficient of $U(t-1)$ suggests that a household that was more mobile than expected in the previous period tends to make heavier use of cars. At the same time, the result points to the need for accounting for a possible correlation between the errors of trip generation and modal split in the same period.

During the model estimation process, it was found that the residual distribution of this modal split model diverts substantially from the normal distribution. It was also found that there exists a substantial number of households with no transit trips or no car trips, requiring the introduction of correction terms when computing the logit (0.5 is added to the observed frequencies of car and transit trips). This may be the reason for the over-prediction of public transit use found in the simulation experiments reported in the next section. The deficiencies of this modal split model need to be resolved in a future effort. It is also noted that the significance of the coefficients of the correction terms, $Q[D(t)]$ and $Q[D(t-1)]$, implies added heteroscedasticity and inconsistent estimation of t-statistics. The estimation results point to the need for implementing procedures to account for the added heteroscedasticity.

5. SIMULATION EXPERIMENT

In order to demonstrate how the model system developed in this study can be applied to forecasting, simulation experiments are conducted to predict car ownership and mode usage by the panel households under hypothetical scenarios. Simulation was conducted for

a period of 12 years with 1,002 households that were in waves 3, 5 and 7. The scenarios examined are:

a. Base case, no change in any contributing factors.
b. Income increase of 100% during a 10-year period.
c. Random increase in licensed drivers.
d. Random increase in employed household members.

In each simulation run, observations from wave 5 (simulation time t = 0) are used as the initial (lagged) values. The simulation starts using exogenous variable values at wave 7 (t = 1) and simulating wave-7 car ownership, trip generation, and modal split. The normal pseudo-random number generator of the SPSSx package is used to generate error terms that have the same standard deviations as obtained during the model estimation. The values of the exogenous variables are held constant throughout the simulation run unless otherwise mentioned. The four scenarios are described below.

5.1 Scenario a, Base Case

All exogenous variable values are held constant at their wave-7 values throughout the experiment with wave-5 endogenous variable values used as the lagged terms. This base case run serves as a control against which results from other experiments can be compared.

5.2 Scenario b, Income Increase at a 7.18% Annual Rate

Household income increases by 7.18% each year starting at t = 3 through t = 12. This deterministic increase doubles the household income at t = 12.

5.3 Scenario c, Increase in Licensed Drivers

In each simulation year starting at t = 3, each non-driver adult ($\geq$ 18 years old) is converted to a driver randomly with probability 0.25. Once a person becomes a driver, he or she remains a driver in the rest of the simulation. Non-drivers of elderly households, however, are exceptions and remain as non-drivers throughout the simulation.

5.4 Scenario d, Increase in Workers

Non-working adult members of a household are randomly converted to workers using the same rule: each non-worker has a probability 0.25 of becoming a worker; once a non-worker becomes a worker, he or she remains a worker throughout the simulation. Again, the working status of the members of elderly households is not changed during the simulation. Household income is increased as the number of workers increases, assuming arbitrarily that a new worker earns 60% of the average personal income of the other workers in the household. In case the new worker is the only worker in the household, he or she is assumed to earn 60% of the sample-wide average personal income.

These scenarios are purely hypothetical; they do not represent anticipated growths in income, licensed drivers, workers or income. For example, the income increase assumed here is much greater than that indicated by a forecast (Centraal Plan Bureau 1985). The intent of this study is to create exaggerated changes in contributing factors and amplify changes in behavior as depicted by the model, such that clear indications of model performances can be obtained. Currently, the simulation model is being extended and refined

to incorporate external forecasts of key variables and also to incorporate dynamic changes in household characteristics, its evolution through lifecycle stages, aging of its members, retirement, etc.

5.5 Results of Experiments

Because wave 5 observations are used as input to the simulation, the simulated results can be compared against observed wave-7 behavior to assess the model performance. Observed and simulated transitions in car ownership from wave 5 to 7 are presented in Table 5. The simulation in general replicates the observed transition quite well. For example, 197 households had no car available in both wave 5 and wave 7; the average frequency of five simulation runs is 193.6. The only slight discrepancy is the absence of multi-car to no-car transition in the simulated table. Overall, however, it is evident that the model is capable of replicating observed changes remarkably well.

The number of mechanized trips is also well replicated by the model system (Table 6). The average of five base-case simulation runs is 26.9 trips for period t = 1, and the corresponding wave-7 observed mean trip rate is 26.7. Modal split, however, is not equally well replicated; the simulation runs resulted in a mean transit trip rate of 3.3, which is 12.2% of the total mechanized trips. The corresponding observed values are 2.5 trips and 9.4%.

This over-prediction of transit ridership propagates itself in the first few rounds of the simulation experiments, leading to an approximately 15% increase in the transit trip rate over the 12-year simulation period. Judging from the discrepancy observed at t = 1 against wave-7 observations, it is likely that the model's forecast offers an inflated ridership figure. Possible reasons for this result are, as pointed out earlier, the skewed distribution of residuals (to which symmetric normal random numbers are applied in the simulation) and

Table 5 Observed and Simulated Transitions in Car Ownership from Wave 5 (t = 0) to Wave 7 (t = 1)

Wave 5 (t = 0)	Wave 7 (t = 1)			
	0	1	2+	Total
a. Observed Transitions				
0	197	18	0	215
1	18	635	31	684
2+	2	25	76	103
Total	217	678	107	1002
b. Simulated Transitions (Realization from a Sample Run)				
0	193	22	0	215
1	19	632	33	684
2+	0	27	76	103
Total	212	681	109	1002
c. Simulated Transitions (Average of Five Runs)				
0	193.6	21.4	.0	215
1	22.2	630.2	31.6	684
2+	0	31.0	72.0	103
Total	215.8	682.6	103.6	1002

Table 6 Results of Five Base-Case Simulation Runs

Simulation Run	1.	2.	3.	4.	5.	Avg.	Observed (t = 1)
No. of Mechanized Trips	27.3	26.4	26.6	27.3	26.7	26.9	26.7
No. of Car Trips	23.8	23.2	23.5	23.9	23.5	23.6	24.2
No. of Transit Trips	3.5	3.2	3.1	3.4	3.2	3.3	2.5
Percent of Transit Trips	12.8	12.0	11.6	12.5	12.1	12.2	9.4

the large number of households without either car trips or transit trips, which makes the application of logit linear regression less accurate. This model component is currently being reformulated to account for these limitations.

The results of base-case simulation runs are summarized in Fig. 2a. Five runs are repeated for each scenario and the averages are used as the forecast values. The number of car trips shows a slight increase of 7% over the 12-year period and the number of transit trips increases by approximately 15%, as noted earlier. Both no-car households and multi-car households increase slightly in this base case. The changes in the endogenous variables are very minor, however. The results presented in this figure are used as the basis for comparison of the other three scenarios.

The effect of income increasing is shown in Fig. 2b. With an annual rate of 7.18%, household income increases exponentially (starting at t = 3) and has doubled by t = 12. The consequence of this increase is evident from the figure; the number of car trips, number of transit trips, and multi-car households all increase, while the number of no-car households remains unchanged and the number of one-car households decreases slightly. Overall, however, the change in car ownership is modest despite the substantial change in income.

Quite notable is the increase in transit usage. The positive contribution of an income increase to transit usage more than compensates for the negative contribution of car ownership, which also increases as a result of an income increase. This model system, in which car ownership is one of the endogenous variables, thus offers an indication that the commonly held belief that an income increase will lead to an increase in car ownership, which in turn will lead to a decline in transit patronage, may not hold true in the Netherlands. An income increase is not detrimental to public transit patronage, despite its contribution to car ownership.

With the hypothetical rule that 25% of adult non-drivers become drivers each year, the average number of drivers in the sample households increases from 1.47 at t = 1 (wave 7) to 2.03 at t = 12. This increase of drivers has a drastic impact on household car ownership and mode use (Figure 2c). The number of multi-car households nearly doubles, while the number of no-car households decreases by more than one-third. Car trips steadily increase by 15% during the simulation period, and transit trips show an unstable pattern around the original value. Relatively speaking, public transit usage is suppressed under the increase of licensed drivers.

Simulation runs with scenario d (Fig. 2d) resulted in an increase in the average number of workers per household from 0.99 to 2.00, accompanied by a 32% income increase from 29,600 Dfl to 39,200 Dfl over the simulation period. These changes lead to slight decreases in one-car and no-car households and an increase in multi-car households. The number of car trips increases and the number of transit trips even increases at a large rate. Presumably,
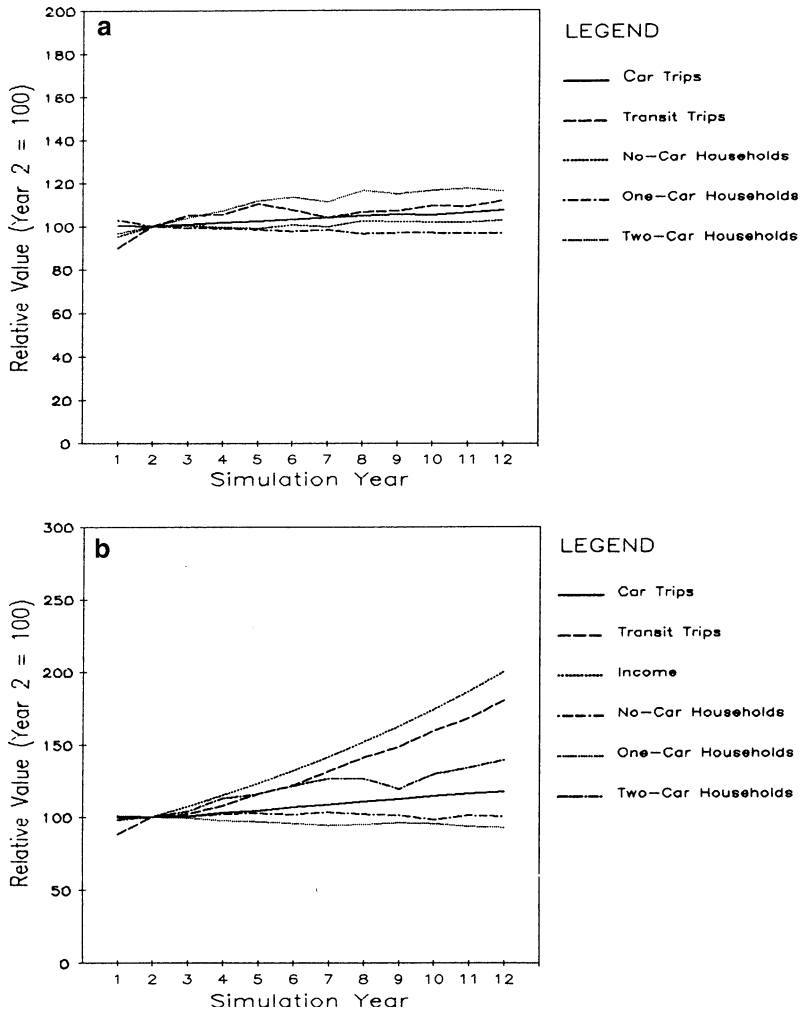
Fig. 2 a Base-case simulation results (average of 5 simulation runs with 1,002 households). **b** 7.18% annual increase in household income (average of 5 simulation runs with 1,002 households). **c** 25% of non-drivers becoming drivers each year (average of 5 simulation runs with 1,002 households). **d** 25% of non-workers gaining employment each year (average of 5 simulation runs with 1,002 households)

increased commuting trips due to the increased number of workers leads to increased transit use.

Simulated car ownership and mode use are compared across the scenarios in Figures 3a–d. It is evident from Figs. 3a and 3b that an increase in income, drivers, or workers leads to a more rapid increase of multi-car households than in the base-case scenario. Also no-car households decrease more under these three scenarios. The impact of these changes on car ownership is evident.

The increase in drivers has the most substantial impact on car ownership, especially on the decrease of no-car households, while income increase has only a moderate effect. The result suggests that the major determinants of household car ownership are
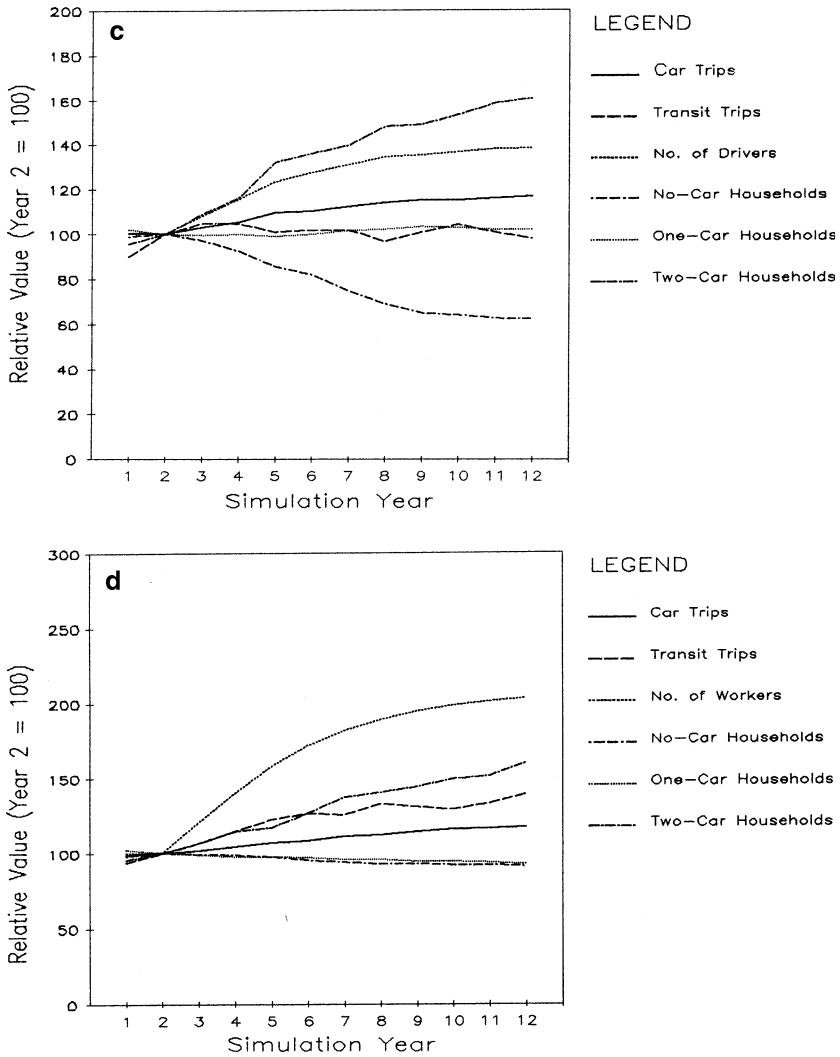
Fig. 2 continued

sociodemographic in nature, and that household income plays only a minor role. It may be the case that the diverse price range of new and used cars available in the market makes a family car affordable to households with a wider income range, leading to the weak correlation between household income and car ownership.

The car trip rate increases slightly with household income, number of drivers, and number of workers (Fig. 3c). This modest increase, approximately 10% more than the base case at $t = 12$, is quite surprising in light of the large increases observed in car ownership. The result suggests that the car trip rate is not proportional to the number of cars available; indeed it is family members, not family cars that make trips. Also, note that the effects of income, drivers, and workers on the car trip rate are approximately the same, despite their markedly different impact on car ownership.
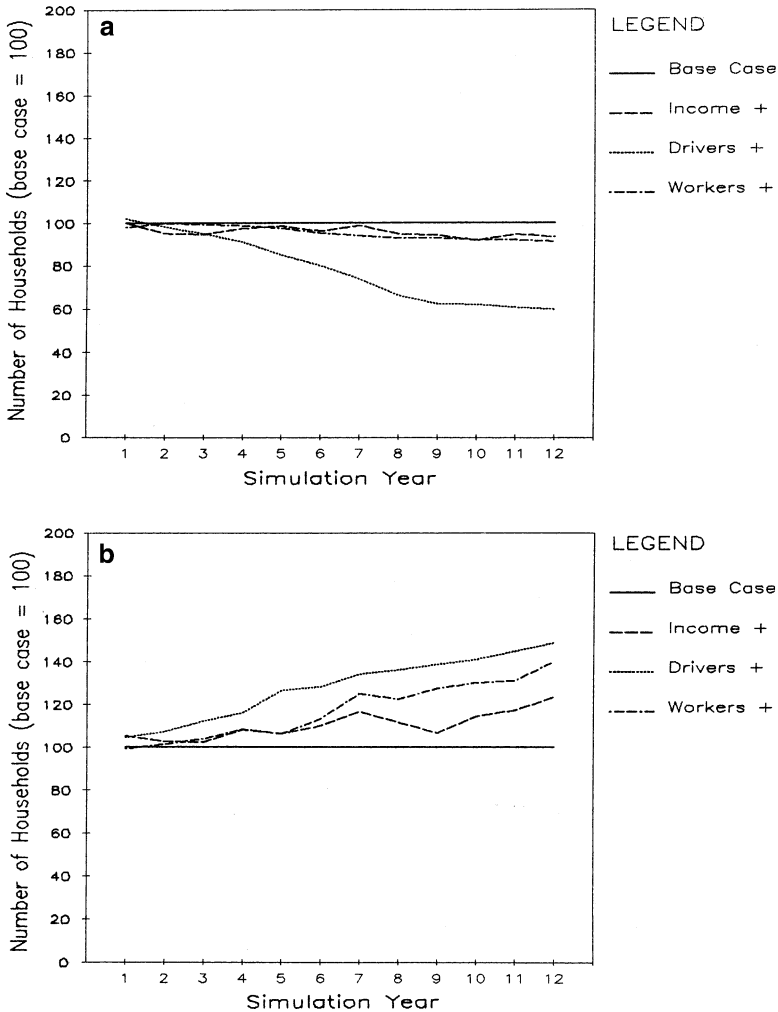
Fig. 3 a Comparison of simulation results across scenarios, number of no-car households (base case = 100). **b** Comparison of simulation results across scenarios, number of multi-car households (base case = 100). **c** Comparison of simulation results across scenarios, number of car trips (base case = 100). **d** Comparison of simulation results across scenarios, number of transit trips (base case = 100)

This is not the case for the public transit trip rate (Fig. 3d). An increase in income or number of workers leads to an increase in transit trips, while an increase in the number of drivers results in transit trip rates below those in the base-case scenario. Income or employment increases lead to increased household car ownership, but not to a decline in transit patronage. An increase in licensed drivers, however, leads to suppressed transit ridership.

The simulation experiment reported in this section is based on hypothetical scenarios for changes in income, number of drivers, and number of workers. The assumed rate of change is perhaps much larger than what will actually be observed in the next 12 years. Accurate forecasting of these contributing factors is beyond the scope of this paper; the purpose of
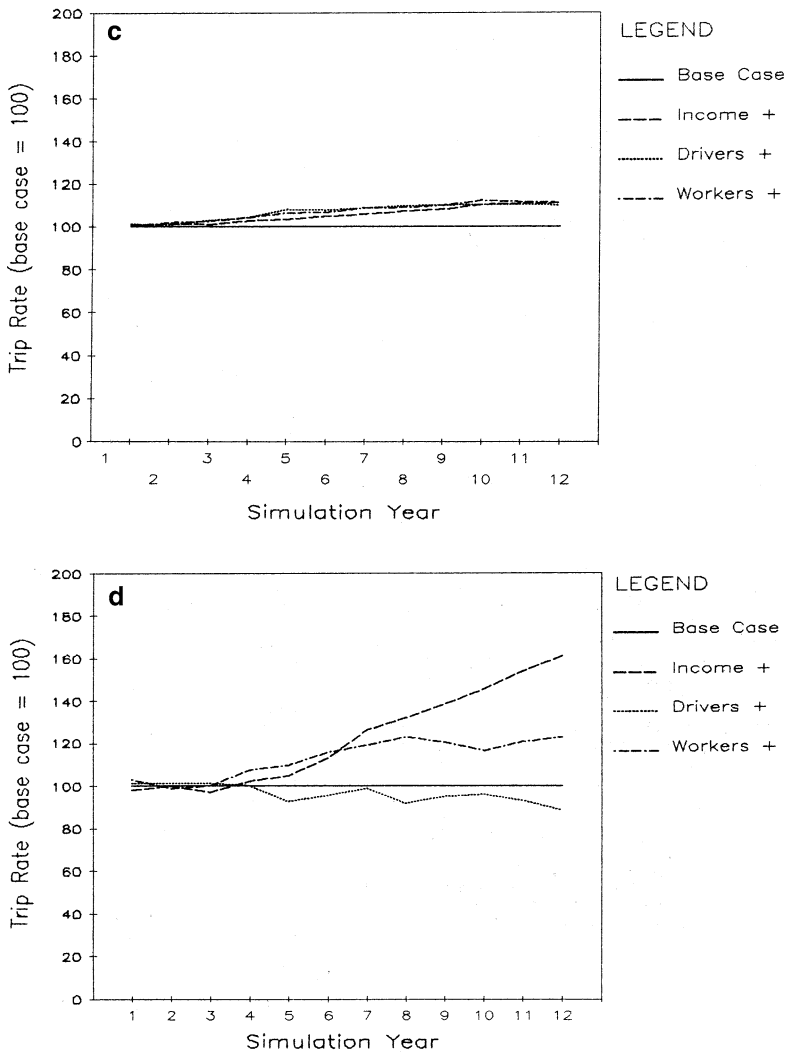
Fig. 3 continued

this simulation has been to demonstrate that the dynamic simultaneous equations system can be utilized for practical forecasting and policy evaluation.

Nevertheless, the results offer a number of behavioral insights. The simulation results have shown that an income increase has a relatively minor impact upon household car ownership. Major determinants appear to be sociodemographic, rather than economic in nature. The results have also shown that income increase and the resulting moderate increase in car ownership do not imply decreased public transit use. To the contrary, transit trip rates increase with income. An increase in licensed drivers, on the other hand, impacts car ownership substantially and, unlike an income increase, suppresses public transit use. The result is in line with the warning that the emergence of new cohorts of young drivers who were brought up with a family car and for whom the automobile has been a way of

life, may substantially impact future public transit usage (Goodwin 1987). Finally, the study has shown that simulation based on a dynamic model system can produce useful forecasts for transportation planning.

6. CONCLUSION

A dynamic, simultaneous equations system is developed in this study to represent household car ownership, generation of mechanized trips, and the modal split between private car and public transit. The model system is estimated using the Dutch National Mobility Panel data set and used in simulation experiments to assess the impact of increases in income, number of drivers, and number of workers upon car ownership and mode use. The model components are formulated with lagged dependent variables and correction terms to account for possible correlations among the error terms.

The number of drivers in the household is the predominant and most significant factor influencing car ownership. Household income is one of the factors that significantly affect the generation of mechanized trips. Other influential variables include car ownership, number of drivers, and number of workers. Modal split is influenced by car ownership, presence of school-age children, and transit service level. Importantly, household income has a very significant and positive effect upon transit use. As income increases, the total number of trips made by car or public transit increases, and the fraction of trips made by public transit also increases. The lagged dependent variables of these models are always very significant and. as anticipated, have positive coefficients. The coefficient of the lagged term is the smallest for the modal split model, suggesting that mode choice is less stable and more responsive to changes in household characteristics and other contributing factors.

The simulation experiments using hypothetical scenarios indicated that the effects of income, drivers, and workers on the car trip rate are approximately the same, despite their markedly different impacts on car ownership. The moderate increase in household car ownership caused by an income increase does not lead to decreased transit ridership; to the contrary, transit usage increases as income increases over time. Similarly, an increase in employed household members leads to increased car ownership and transit use. An increase in licensed drivers has the most pronounced effect on household car ownership and, unlike the case of increased income or workers, is detrimental to public transit patronage. This result warns that the ever increasing driver population may seriously deter public transit use, especially as cohorts of young drivers, who were brought up in motorized environments, progress through life-cycle stages.

Being one of the first attempts to develop a dynamic model system of travel behavior and to apply it to forecasting, the exercise of this study is subject to certain limitations. Nonetheless, the study has shown that a dynamic model can be conveniently estimated and practically applied to forecast future demand. The joint forecasting of car ownership, trip generation, and modal split obtained from the simulation experiment has offered valuable insights for policy development.

The research effort reported here is currently being extended in several important directions. Better estimation procedures are being sought for the dynamic model system that consists of both discrete and continuous endogenous variables with correlated errors and lagged and cross-lagged endogenous variables. A long-range forecasting simulation procedure is also being developed using the dynamic model system together with the Dutch Panel sample. Crucial in this extension is the introduction of a component that performs realistic forecasting of demographic and socio-economic attributes of

households. The model system offers dynamic forecasts along a semi-continuous time axis, not a forecast for an isolated future time point based on observed cross-sectional variations in behavior. Therefore, the simulation requires continuous updating of socio-demographic variables. In order to best achieve this, a lifecycle simulator that replicates the evolution of a household through lifecycle stages is being developed as an extension of MIDCAT (Goodwin et al. 1987). Finally, incorporation of level-of-service variables is being considered for enhanced usefulness of the model system in forecasting and policy evaluation.

References

Ben-Akiva, M., and S.R. Lerman (1974) Some estimation results of a simultaneous model of auto ownership and mode choice to work. Transportation, 3, 357–376.

Centraal Plan Bureau (1985) De Nederlandse Economie op Langere Termijn. The Hague, The Netherlands.

de Jong, G.C., and J.S. Cramer (1987) A Censored Regression Model of Private Car Use. Report AE 4/87, Faculty of Actuarial Science and Econometrics, University of Amsterdam, Amsterdam, The Netherlands.

Golob, J.M., L.J.M. Schreurs, and J.G. Smit (1986a) The design and policy applications of a panel for studying changes in mobility over time. In Behavioural Research for Transport Policy, pp. 81–95. VNU Press: Utrecht, The Netherlands.

Golob, T.F. (1986) A non-linear canonical correlation analysis of weekly trip chaining behaviour in the Netherlands. Transportation Research, 20A, 5, 385–399.

Golob, T.F. (1987) Effects of Income and Car Ownership on Trip Generation: A Structural Equations Model and Its Temporal Stability. UCI-ITS-WP-87-7, Institute of Transportation Studies, University of California, Irvine, CA.

Golob, T.F., and H. Meurs (1986) Biases in response over time in a seven-day travel diary. Transportation, 13, 163–181.

Golob, T.F., and H. Meurs (1987) A structural model of temporal change in multimodal travel demand. Paper presented at the 66th Annual Meeting of the Transportation Research Board, Washington, DC.

Golob, T.F., and H. Meurs (1988) Modeling the dynamics of passenger travel demand by using structural equations. Environment and Planning A, 20, 1197–1218.

Golob, T.F., L. van Wissen, and H. Meurs (1986b) A dynamic analysis of travel demand. Transportation Research, 20A, 5, 401–414.

Goodwin, P.B. (1986) A panel analysis of changes in car ownership and bus use. Traffic Engineering and Control, 27 October, 519–525.

Goodwin, P.B. (1987) Family Changes and Public Transport Use 1984–1987: A Dynamic Analysis Using Panel Data. Report prepared for the Project Bureau of the Netherlands Ministry of Transport, Bureau of Goudappel Coffeng, Deventer, The Netherlands.

Goodwin, P.B., M.C. Dix, and A.D. Layzell (1987) The case for heterodoxy in longitudinal analysis. Transportation Research A, 21A, 363–376.

Heckman, J.J. (1979) Sample selection bias as a specification error. Econometrica, 47, 153–161.

Hensher, D.A., P.O. Bernard, N.C. Smith, and F.W. Milthorpe (1987) Modelling the dynamics of car ownership and use: a methodological and empirical synthesis. Paper presented at the 5th International Conference on Travel Behavior, Aix-en-Provence, France.

Kitamura, R. (1987a) A panel analysis of household car ownership and mobility. Proceedings of the Japan Society of Civil Engineers, No. 383/IV-7, 13–27.

Kitamura, R. (1987b) Determinants of household car ownership and utilization: A panel analysis. Paper presented at the 5th International Conference on Travel Behavior, Aix-en-Provence, France.

Kitamura, R. (1987c) An analysis of weekly activity patterns and travel expenditure. In Behavioral Modeling Approaches in Geography and Planning (edited by R.G. Golledge and H.J.P. Timmermans), pp. 399–423 (Croom Helm: London).

Kitamura, R. (1988) Formulation of trip generation models using panel data. Transportation Research Record, 1203, 60–68.

Kitamura, R., and P.H.L. Bovy (1987) Analysis of attrition biases and trip reporting errors for panel data. Transportation Research A, 21A, 287–302.

Kitamura, R., and T. van der Hoorn (1987) Regularity and irreversibility of weekly travel behavior. Transportation, 14, 227–51.

Mannering, F., and C. Winston (1985) Dynamic models of household vehicle ownership and utilization: an empirical analysis. Rand Journal of Economics, 16, 215–36.

Meurs, H., P. van de Mede, J. Visser, and L. van Wissen (1987a) Analysis of Panel Data. Bureau Goudappel Coffeng, Deventer, The Netherlands.

Meurs, H., F. Gloerich, P. van de Mede, J. Visser, and M. Klok (1987b) The analysis of change in mobility in the Netherlands. Paper presented at the 5th International Conference on Travel Behavior, Aix-en-Provence, France.

Mogridge, M.J.H. (1983) The Car·Market. Pion, London.

Onnen, M.E.K., C.W.F. van Knippenberg, and T. van der Hoorn (1987) Why do people buy cars? Paper presented at the 5th International Conference on Travel Behavior, Aix-en-Provence, France.

Recker, W.W., T.F. Golob, M.G. McNally, and J.D. Leonard (1987) Dynamic tests of the time-space model of complex travel behavior. Paper presented at the 5th International Conference on Travel Behavior, Aix-en-Provence, France.

Smith, N.C., D.A. Hensher, and N. Wrigley (1986) Modelling Discrete Choice Outcome Sequences with Panel Data: An Application to Automobile Transactions. Dimensions of Automobile Demand Project Working Paper No. 18, School of Economic and Financial Studies, Macquarie University, North Ryde, Australia.

Theil, H. (1971) Principles of Econometrics. John Wiley and Sons, New York.

Town, S.W. (1983) Reasons for car acquisition. In Research for Transport Policies in a Changing World: Proceedings of World Conference on Transport Research, Hamburg, Vol. I, pp. 321–330 (SNV Studiengesellschaft Nahverkehr: Hamburg).

Train, K., and M. Lohrer (1983) Vehicle ownership and usage: an integrated system of disaggregate demand models. Paper presented at the 62nd Annual Meeting of the Transportation Research Board, Washington, DC

van der Hoorn, T., and R. Kitamura (1987) Evaluation of the predictive accuracy of cross-sectional and dynamic trip generation models using panel data. Paper presented at the 66th Annual Meeting of the Transportation Research Board, Washington, DC

Author Biography

Ryuichi Kitamura Upon receiving his Ph.D. degree from the University of Michigan, Dr. Kitamura joined the Civil Engineering Department of the University of California at Davis, where he is currently an associate professor. His research interest has concentrated in the area of travel behavior analysis, in particular, trip chaining analysis, activity-based travel analysis, and panel analysis of travel behavior. He recently completed co-editing a special issue of *Transportation* on activity-based travel analysis, and is currently organizing a special session on dynamic analysis of travel behavior for the forthcoming World Conference on Transport Research in Yokohama, Japan. Dr. Kitamura chairs the Subcommittee on Activity and Travel Pattern Analysis of the Transportation Research Board.