

ANALYSIS OF ATTRITION BIASES AND TRIP REPORTING ERRORS FOR PANEL DATA

RYUICHI KITAMURA

Department of Civil Engineering, University of California at Davis, Davis, California 95616,
U.S.A.

and

PIET H. L. BOVY

Institute for Town Planning Research, Delft University of Technology, Berlageweg 1,
2600GA Delft, The Netherlands

Abstract—The objective of the study is to assess the nature and magnitude of reporting errors and attrition biases in the first two waves of a household panel survey data set that consists of weekly trip diaries. This paper investigates the relationship between reporting errors in a first-wave diary and the subsequent decision to participate in a second-wave survey and the relationship between the same first-wave errors and the errors in the second-wave survey for those households which chose to participate. An econometric procedure is developed which accounts for attrition bias and relates reporting errors from the two waves. The model system offers a mechanism by which relationships among mobility, trip reporting errors, and attrition behavior in a panel survey can be statistically evaluated. It consists of three elements: wave-one trip equation, attrition probability model, and wave-two trip equation. The conditional distribution of the error terms of these three components is evaluated assuming chronological dependency among the errors. The result is used to develop a correction term for attrition bias and a term which accounts for the dependency between the wave-one error and wave-two error. Results from empirical application of the model system are reported together with the use of the attrition model to develop an efficient sample weighting factor that fully utilizes the information in the survey results.

INTRODUCTION

Multi-day travel data offer the opportunity to examine day-to-day variations in travel patterns and thus enable the researchers to develop inferences on how activities are organized and trips are scheduled over the survey period.† At the same time, collecting multi-day data may create added problems if respondents' ability and intention to keep accurate travel diaries decline over time (Brög *et al.* 1982; Brög, Meyburg, and Wermuth, 1983; Hendrikx, 1982; Golob and Meurs, 1986). In addition to this problem of reporting errors, there arises the problem of attrition bias (Maddala, 1978; Hausman and Wise, 1979; Hensher, 1985; Hensher and Wrigley, 1986) when multi-day data are collected from the same respondents at two or more points over time. This is the case for the Dutch National Mobility Panel data.

The Dutch Panel data consist of seven-day diaries collected in waves (Golob, Schreurs, and Smit, 1986). The first wave of diaries was collected during a three-week period of March, 1984, the second wave in September, 1984, and the third wave in March, 1985. An analysis of the first-wave diaries indicates that reporting bias does exist in the data set (Golob and Meurs, 1986); it is noted that the bias increases over time during the diary period due to an increase in the percentage of respondents reporting no travel at all for an entire day, and also due to increased under-reporting of walking trips. Detailed observations of reporting bias over diary sequence days are given in Golob and Meurs (1986).

The attrition rate, i.e. the rate at which households dropped out of the panel, is 32% between the first and second wave, a disturbingly high rate. Golob *et al.* (1985) report that households of certain demographic and socio-economic characteristics tended to leave the panel. Such tendencies in attrition result in a substantially different sample composition in the second wave and lead to the first type of attrition bias problem: Naive

†Examples of multi-day travel behavior analysis include: Hanson and Huff (1983), Koppelman and Pas (1984), Kitamura (1986), and Kitamura and Van der Hoorn (1987).

comparison of sample means of mobility indicators between the two waves may be meaningless or even misleading.

In addition, there exists the second type of attrition bias, i.e. bias in estimators of model coefficients. It is quite probable that unobserved factors influencing a household's decision to leave the panel are correlated with unobserved factors influencing its members' propensities to make and report trips. If this is the case, ordinary estimators of model coefficients are biased (e.g. Maddala, 1978; Heckman, 1979). The nature and magnitude of these attrition biases, and the relations in reporting biases between the waves of the Dutch Panel data are not known yet.

The objective of the study is to assess the nature and magnitude of reporting and attrition biases in the first two waves of the Dutch Panel data. The effort is motivated by the recognition that causal relationships between travel behavior and the travel environment cannot be evaluated correctly unless these biases are properly accounted for (a transit fare increase and the weather are major differences in the travel environment between the first two waves of the Dutch Panel data). Specifically the following two questions are addressed in this paper: How are a household's reporting errors in the first-wave diary related to its decision to participate in or to refuse the second-wave survey; and how are the first-wave reporting errors related to second-wave errors, given that the household chose to participate in the second-wave survey?

In order to examine these issues, an econometric model system is developed which accounts for attrition biases and relates reporting errors from the two waves. This model system makes it possible to quantitatively evaluate the relationships among reporting errors, mobility, and attrition behavior of participants of a panel survey. Although previous empirical investigations have found the tendencies that those who under-reported their trips tended to drop out of a survey and that less mobile respondents tended to decline cooperation, no tools have existed for their statistical treatment. The model system developed in this study offers a mechanism through which such tendencies can be quantified.

The household is used as the unit of analysis. This reflects the proposition that the decision to participate in or decline the second-wave survey is a household decision reflecting collectively its members' experiences and reactions to the diary survey of the first wave. It is considered that interactions among household members are important factors contributing to the household's response to a survey. Household-based analysis is preferred here also because the preparation of the original data file considered a household not responding and eliminated the trip diaries of its members altogether if trip diaries were not available and usable for a substantial number of its members (Instituut voor Longitudinaal Beleidsonderzoek, 1985).

In the next section, we present the structure of the model, which consists of three components: wave-one trip equation, attrition probability model, and wave-two trip equation. The conditional distribution of the error terms of these three components is evaluated assuming chronological conditionality among the errors. The result is used to develop a correction term for attrition bias and a term relating the wave-one error and wave-two error. This model is applied to the Dutch Panel data to evaluate the magnitude of attrition bias and to determine the relation between wave-one trip reporting and wave-two trip reporting.

Model structure

Let the total number of reported trips aggregated over the one-week survey period be the dependent variable of the analysis and let

$$y_i = \alpha' \mathbf{V}_i + \xi_i \quad (1)$$

where y_i is the number of trips for household i , α is a vector of coefficients, \mathbf{V}_i is a vector of exogenous variables, and ξ_i a random error term. Since y_i is the number of reported trips, ξ_i reflects not only unexplained propensity to travel but also reporting errors that are uncorrelated with the variables in \mathbf{V}_i .

Evaluating the exact magnitude of reporting errors from travel diaries is an impossible task simply because the number of trips actually made by a household cannot be determined from the survey results and because nonsystematic random disturbances and random reporting errors cannot be separated. Similarly, the coefficient vector α represents the effects of the exogenous variables on reporting errors as well as those on mobility. The two sets of effects again cannot be separated. Accordingly, when the reported number of trips is used as the dependent variable, a standard estimator of α would yield a biased estimator of the exogenous variables' effects on mobility. In this study, ξ_i is interpreted to represent random disturbances in mobility and random reporting errors uncorrelated with the variables in \mathbf{V}_i ; and α is interpreted to reflect the systematic effects of \mathbf{V}_i on both mobility and reporting errors.† In the rest of this paper the term reporting errors refers to nonsystematic reporting errors which are not correlated to \mathbf{V}_i .

It is quite likely that a household's reporting error and unexplained propensity to travel are correlated between the two waves; those households which under-reported their trips in the first wave perhaps tended to under-report trips in the second wave also. Similarly, those which made more trips than expected in the first wave may have also done so in the second wave. Let the error term from the first wave be ξ_{1i} and that from the second wave be ξ_{2i} . Since these two error terms represent reporting error and unexplained mobility from the respective waves, it is expected that they are positively correlated for a given household.

Furthermore, it is also expected that the reporting error from the first wave is correlated with the likelihood that the household will participate in the second wave survey; members of households which under-reported their trip making in the first wave may have had less enthusiasm about the survey and had been discouraged to fill in the diary, therefore, may be more likely to drop out from the second survey. This probable relationship between the first-wave reporting error and attrition can be accounted for by assuming correlation between the first wave error term, ξ_{1i} , and the error term associated with attrition.‡ A model system that incorporates these probable correlations is presented below.

Travel behavior in the two waves of diary data and the attrition between them are expressed as:

Wave 1

$$y_{1i} = \alpha_1' \mathbf{V}_{1i} + \xi_{1i} \quad (2a)$$

Attrition

$$A_i = \beta' \mathbf{X}_i + \epsilon_i$$

(2b)

If $A_i \geq 0$, $a_i = 1$ (participation)
 If $A_i < 0$, $a_i = 0$ (nonparticipation)

Wave 2

$$y_{2i} = \alpha_2' \mathbf{V}_{2i} + \xi_{2i} \quad \text{if } a_i = 1 \quad (2c)$$

where y_{si} is the number of trips reported by household i aggregated over the diary period of wave s ; α_s and β are coefficient vectors; \mathbf{V}_{si} and \mathbf{X}_i are vectors of exogenous variables; and A_i is a latent variable and $a_i = 1$ if household i participates in the second-wave

†This is not to say that accounting for possible biases in coefficient estimates due to the use of the reported number of trips is not important. The issue to the contrary has important implications to travel behavior modeling where the current practice is to use the reported number of trips as an accurate measure of mobility. The subject, however, lies beyond the scope of this study.

‡Another possible factor that may be at work is the effect of participating in the panel survey and filling out a diary on a household member's travel behavior itself. Possible presence of such effects is suggested by Brog *et al.* (1983). No attempt is made in this study to account for this bias in behavior itself.

survey. Note that y_{2i} is not observable if $a_i = 0$, in which case household i rejects the second-wave survey.

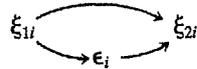
Let

$$(\xi_{1i}, \xi_{2i}, \epsilon_i) \approx MVN(0, \Sigma)$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{1\epsilon}\sigma_1 \\ & \sigma_2^2 & \rho_{2\epsilon}\sigma_2 \\ & & 1 \end{pmatrix} \quad (3)$$

where $\text{Var}(\epsilon_i) = 1$ for normalization. The normality assumption implies that attrition is described by a binary probit model.

In applying the model of eqns (2a-c), we assume the following chronological conditionality among the error terms:



The error term of the attrition equation is conditioned on a realization value of the error term of the first-wave trip equation, but not vice versa. Similarly, the error term of the first-wave trip equation and that of attrition equation jointly influence the error term of the second wave trip equation, but no reverse influence is assumed. This is reasonable to assume since reporting errors and other unobservable elements in the first-wave diary are not likely to have been conditioned on those of the second-wave survey, which was in the distant future at the time of the first-wave survey (the separation between the two waves in the Dutch Panel is 6 months). The one-way conditionality between ϵ_i and ξ_{2i} is a standard treatment in selectivity correction procedures (e.g. Heckman, 1979; Maddala, 1983).

CONDITIONAL DISTRIBUTION OF ERROR TERMS

Consider standard bivariate normal random variates, U_1 and U_2 . The conditional distribution of U_2 , given U_1 , is normal with

$$E(U_2|U_1) = \rho_{12}U_1$$

$$\text{Var}(U_2|U_1) = 1 - \rho_{12}^2 \quad (4)$$

where ρ_{12} is the correlation coefficient between U_1 and U_2 (Johnson and Kotz, 1972, pp. 86-87). Note that the variance of U_2 given U_1 does not depend on the value that U_1 takes on. Applying these relations to ξ_{1i} and ϵ_i

$$E(\epsilon_i|\xi_{1i}) = \frac{\rho_{1\epsilon}}{\sigma_1} \xi_{1i} \quad (5)$$

Accordingly, ϵ_i given ξ_{1i} can be expressed as

$$\epsilon_i = \frac{\rho_{1\epsilon}}{\sigma_1} \xi_{1i} + u_i$$

$$u_i \sim N(0, 1 - \rho_{1\epsilon}^2) \quad (6)$$

and u_i is i.i.d.

Next, consider the conditional distribution of ξ_{2i} given ξ_{1i} and $a_i = 1$, i.e. household i chose to participate in the second-wave survey. The distribution of ϵ_i in this case is

truncated as $\epsilon_i \geq -\beta \mathbf{X}_i$. The desired conditional distribution of ξ_{2i} is normal with mean $E(\xi_{2i}|\xi_{1i}, \epsilon_i \geq -\beta \mathbf{X}_i)$.

In order to evaluate this expectation, first consider the conditional distribution of U_2 and U_3 , given U_1 , where U_1 through U_3 are from a standard trivariate normal distribution. This conditional distribution is bivariate normal with parameters (Johnson and Kotz, 1972, p. 115)

$$\begin{aligned} \boldsymbol{\mu}_{23|1} &= (\rho_{12}U_1, \rho_{13}U_1) \\ \boldsymbol{\Sigma}_{23|1} &= \begin{pmatrix} 1 - \rho_{12}^2 & \rho_{23} - \rho_{12}\rho_{13} \\ \rho_{23} - \rho_{12}\rho_{13} & 1 - \rho_{13}^2 \end{pmatrix} \end{aligned} \quad (7)$$

Note that the off-diagonal element divided by the square-root of the diagonal elements

$$\rho_{23|1} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{(1 - \rho_{12}^2)^{1/2}(1 - \rho_{13}^2)^{1/2}} \quad (8)$$

defines the partial correlation between U_2 and U_3 given U_1 . Applying this result, the conditional distribution of ξ_{2i} and ϵ_i , given ξ_{1i} , is bivariate normal with parameters

$$\boldsymbol{\mu}_{2\epsilon|1} = \begin{pmatrix} \frac{\rho_{12}\sigma_2}{\sigma_1} \xi_{1i}, \frac{\rho_{1\epsilon}}{\sigma_1} \xi_{1i} \end{pmatrix} \quad (9a)$$

$$\boldsymbol{\Sigma}_{2\epsilon|1} = \begin{pmatrix} (1 - \rho_{12}^2)\sigma_2^2 & (\rho_{2\epsilon} - \rho_{12}\rho_{1\epsilon})\sigma_2 \\ (\rho_{2\epsilon} - \rho_{12}\rho_{1\epsilon})\sigma_2 & 1 - \rho_{1\epsilon}^2 \end{pmatrix} \quad (9b)$$

For standard normal variates, we have the following relationship that is frequently used to develop correction terms for selectivity:

$$E(U_2|U_1 \geq h) = \rho_{12} \frac{\phi(h)}{\Phi(-h)} \quad (10)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution function, respectively (Johnson and Kotz, 1970, p. 81; 1972, pp. 112–13). Using this result with the conditional mean vector and variance-covariance matrix of eqns (9a–b), we obtain

$$E(\xi_{2i}|\xi_{1i}, \epsilon_i \geq -\beta \mathbf{X}_i) = \frac{\rho_{12}\sigma_2}{\sigma_1} \xi_{1i} + (1 - \rho_{12}^2)^{1/2}\sigma_2\rho_{2\epsilon|1} \frac{\phi(W_i)}{\Phi(-W_i)} \quad (11)$$

where

$$W_i = \frac{-\beta \mathbf{X}_i - \frac{\rho_{1\epsilon}}{\sigma_1} \xi_{1i}}{(1 - \rho_{1\epsilon}^2)^{1/2}}$$

and $\rho_{2\epsilon|1}$ is defined similarly as in eqn (8). The conditional distribution of ξ_{2i} is normal with the mean given by eqn (11). The variance of ξ_{2i} is obtained as

$$\text{Var}(\xi_{2i}|\xi_{1i}, \epsilon_i \geq -\beta \mathbf{X}_i) = (1 - \rho_{12}^2)\sigma_2^2\{1 + \rho_{2\epsilon|1}^2(W_i\lambda_i - \lambda_i^2)\} \quad (12)$$

where

$$\lambda_i = \frac{\phi(W_i)}{\Phi(-W_i)}$$

Note that the variance of ξ_{2i} varies across i .

CORRECTION TERMS FOR SAMPLE SELECTIVITY

Suppose α_1 of eqn (2a) is estimated and $\hat{\xi}_{1i}$ is obtained as

$$\hat{\xi}_{1i} = y_{1i} - \hat{\alpha}_1' \mathbf{V}_{1i} \quad (13)$$

Noting eqns (5–6), let the attrition eqn (2b) be rewritten as

$$A_i = \boldsymbol{\beta}' \mathbf{X}_i + \gamma_1 \hat{\xi}_{1i} + u_i \quad (14)$$

Then u_i is i.i.d. normal across observations with mean 0 and variance $1 - \rho_{1\epsilon}^2$. The standard maximum likelihood probit procedure applied to eqn (14) yields a consistent estimator of $\boldsymbol{\beta}$ and γ_1 up to a scale constant. Estimate $\hat{\gamma}_1$ serves as an estimator of $\rho_{1\epsilon}/\sigma_1$.

The mean and variance of ξ_{2i} in the second-wave trip eqn (2c), given ξ_{1i} and $\epsilon_i \geq -\boldsymbol{\beta}' \mathbf{X}_i$, have been obtained as eqns (11–12). Then similar to eqn (6)

$$\xi_{2i} = \frac{\rho_{12}\sigma_2}{\sigma_1} \xi_{1i} + (1 - \rho_{12}^2)^{1/2} \sigma_2 \rho_{2\epsilon|1} \lambda_i + u_{2i} \quad (15)$$

where λ_i is as defined in eqn (12) and u_{2i} is an independent normal random variable with mean 0 and variance of eqn (12). The coefficients of the trip equation can be consistently estimated by estimating, instead of eqn (2c),

$$y_{2i} = \boldsymbol{\alpha}_2' \mathbf{V}_{2i} + \gamma_2 \hat{\xi}_{1i} + \gamma_3 \hat{\lambda}_i + u_{2i} \quad (16)$$

where

$$\hat{\lambda}_i = \frac{\phi(\hat{W}_i)}{\Phi(-\hat{W}_i)}$$

$$\hat{W}_i = -\hat{\boldsymbol{\beta}}^* \mathbf{X}_i - \hat{\gamma}_1^* \hat{\xi}_{1i}$$

and $\hat{\boldsymbol{\beta}}^*$ and $\hat{\gamma}_1^*$ are coefficient estimates obtained by estimating the attrition eqn (14). The comparison of eqns (15) and (16) indicates that $\hat{\gamma}_2$ and $\hat{\gamma}_3$ serve as estimators of

$$\rho_{12}\sigma_2/\sigma_1 \quad \text{and} \quad (1 - \rho_{12}^2)^{1/2} \sigma_2 \rho_{2\epsilon|1}$$

respectively. The ordinary least squares applied to eqn (16) will yield unbiased but inefficient estimators of $\boldsymbol{\alpha}_2$, γ_2 and γ_3 . This is because of heteroscedasticity in u_{2i} of eqn (16) as is evident from eqn (12). The standard formula to estimate standard errors for regression coefficients still applies to test the null hypothesis of no attrition bias, i.e. $\gamma_3 = 0$ (Heckman, 1979).

EMPIRICAL ANALYSIS OF DUTCH PANEL DATA

The objective of this empirical analysis is three-fold. Firstly, it aims to evaluate the correlation among the error terms of the three equation model system. In behavioral terms, the analysis attempts to evaluate the magnitude of the correlation between wave-one and wave-two reporting errors and to determine how a household's decision to participate in or decline the second-wave survey is related to its mobility and reporting errors in the first wave. Secondly, the analysis assesses the magnitude of biases in coefficient estimates due to attrition. Thirdly, the study demonstrates how the predicted probability of attrition evaluated for each household can be used as a weighting factor to correct possible biases due to attrition.

Dutch panel survey

The Dutch National Mobility Panel survey was established in order to evaluate changes over time of mobility levels of the Dutch population and to assess the impacts

upon mobility of transit fare increases. The panel households, spread over 20 municipalities across the Netherlands, were selected by a stratified sampling method using as controlling factors household lifecycle, household income, and type of public transit services provided. Seven-day travel diaries are used to collect information on trips made by household members of 12 years of age or older. Three waves of diaries have been collected as of the writing of this paper with six-month intervals. Data from the first two waves are available for this analysis.

The first wave of survey was conducted in March, 1984, with diary starting dates spread randomly over a two-week period. Immediately after the completion of the survey, a significant increase in transit fares took place throughout the nation. Of the 2,185 households contacted through home interviews 1,764, or 80.7%, provided usable responses. The resulting data file contains travel diaries for 3,863 persons.†

The second-wave survey started in September, 1984. Both survey periods are free from influences of holidays or special events. Diary starting dates for the households which were in the first-wave survey were again spread over a two-week period, with the starting days of the week being identical to those of the first wave. No home interview was conducted for the wave-one participants; these households received the survey materials by mail and were requested to mail them back. This presumably contributed significantly to the high attrition rate discussed below.‡ Further details on the Dutch Panel survey can be found in Bureau Goudappel Coffeng (1983, 1985), Instituut voor Longitudinaal Beleidsonderzoek (1984, 1985), Golob *et al.* (1986), Golob and Meurs (1986), and Kitamura and Bovy (1985).

Sample

The data file analyzed in this study contains 1,764 households in the first wave and 1,655 households in the second wave. Of the 1,764 households in the first wave, 1,165 (66.5%) were included in the second-wave file after eliminating cases with incomplete or unusable diaries. Since the objective of this study is to evaluate the relationship between attrition and the error term of the first-wave trip equation and the relationship between the wave-one error and wave-two error, no attempt is made to examine the consistency of the trip records in the diaries and all relevant diary data are included in the analysis. As noted earlier, the household is used as the unit of analysis and the diary data are aggregated at the household level.

The attrition rate is defined as the ratio of the number of households which left the panel to the original number of households in the first wave. Households are considered to have dropped out of the panel if: they could not be located; they declined the second-wave survey; or they did not provide usable responses. The majority of the drop-out households belong to the second category. Note that the attrition behavior is simplified and treated here as having binary responses. This simplification applies to most cases because rejections to participate in the second survey at all were the predominant cause of attrition.§

Preliminary analysis of attrition

Sample frequencies of participation in the second-wave survey are presented in Fig. 1 against several household attributes and total number of reported trips. As noted by Golob *et al.* (1986), low-income households, smaller households, households of old persons, and households without cars tended to decline the second-wave survey. House-

†No response to the first-wave survey is of course another potential source of bias. Methods of correcting biases due to the sample selectivity in the first-wave are discussed in, e.g., Maddala (1983). Such correction is not applied in this study since focus of the analysis at this stage is on trip reporting, unexplained mobility, and attrition behavior.

‡New households added in the second wave were contacted at home by interviewers and requested to mail back completed questionnaires, which resulted in a response rate of 81.6%.

§There were cases where households participated in the survey partially but not completely. For example, household members filled out diaries only for the first few days of the diary week. Such cases can be incorporated into the analysis by extending the binary attrition model to a polychotomous one (see Terza, 1980, for a trichotomous example). This is a possible future extension of the model presented in this study.

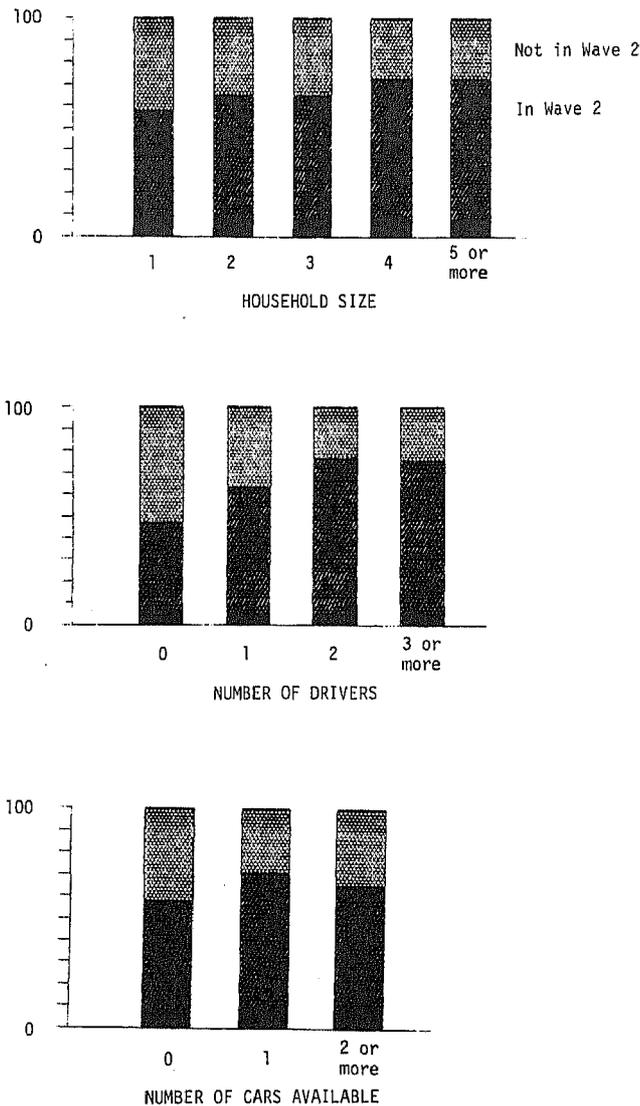


Fig. 1. Household attributes and second-wave panel participation.

holds without drivers tended to leave the panel also. It is known that respondents with less interest in the subject matter of a survey tend not to cooperate with the data collection effort (Brög and Meyburg, 1980; Ampt *et al.*, 1983). Drivers are presumably more mobile and intensive users of transportation systems and accordingly are more interested in the survey itself and tended to participate in the second-wave survey. Similar inferences can be made with respect to car ownership and attrition.

Quite notable is the strong and clear relationship between education and attrition. In this tabulation, the level of education is defined for each household as that of the household member with the highest level of education. As expected and as previous empirical results suggest (Brög *et al.*, 1982), households with lower education tended to decline the second diary survey. The monotonous relationship shown in the figure between education and attrition is quite noteworthy.

Also notable is the association between the number of reported trips and attrition. The rate of second wave participation ranges from the mere 23.5% among the households that reported no trips at all in the wave-one survey to over 80% among those that reported 100 or more trips. The mean household weekly trip rate is 56.5 for the households that remained in the panel and 39.3 for the households that did not. The mean weekly trip rate per diary is, respectively, 24.3 and 19.0 for the two groups. It is clear that attrition is closely related to the reported number of trips.

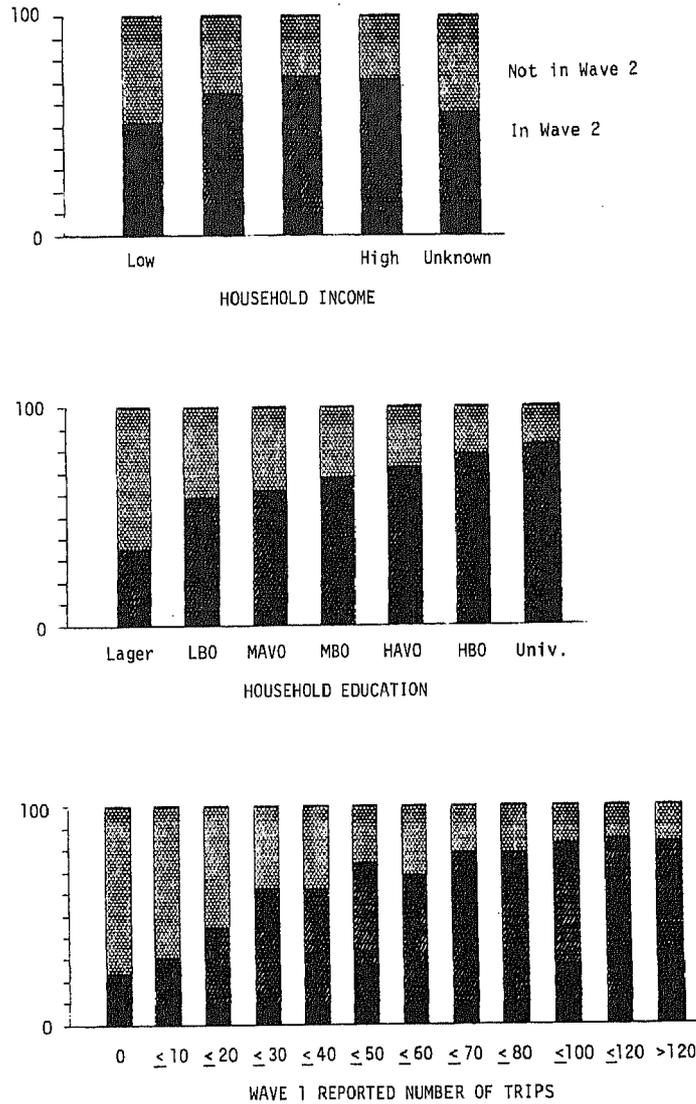


Fig. 1. (Continued).

First wave trip equation

The dependent variable of the trip equations of this study is the total number of trips reported in the diaries by household members who were 12 years or older.† Considered in the model development as explanatory variables are: household size, number of persons filling out the diary, number of adults, number of children by age group, number of workers, number of drivers, lifecycle stage, household car ownership, household income, education, and type of transit services provided. Dummy variables are developed for many of the variables to account for possible nonlinear effects.

Table 1 shows the estimated coefficients of the three-equation model system (the total in the table is reduced to 1973 due to missing data. The definition of the variables used in the model system can be found in Appendix Table). The weighted least squares method is used to estimate the trip equations with the weight being the reciprocal of the square-root of the number of persons who kept the diary (*NDIARIES*). This is based on the assumption that the error term for the total number of trips reported by each household member is homoscedastic, therefore the variance of the error term associated

†It is noted that the members of a household participating in the two diary surveys are not always identical; although infrequent, there are occasions such as some household member reached the age of 12 years between the two waves, therefore were requested to fill out the diary in the second wave.

Table 1. Estimation results: Wave-one, wave-two trip equations and attrition model

	TRIP EQUATION		ATTRITION EQUATION
	Wave 1	Wave 2	
Const.			-1.201 (-3.58)
NDIARIES	22.47 (39.94)	19.97 (30.80)	.595 (-6.47)
ONECAR			-.263 (-2.59)
TWOCARS			-.537 (-3.51)
Lifecycle Stage			
STAGE 1	5.64 (3.85)	3.85 (2.26)	-.236 (-2.18)
STAGE 5			-.321 (-2.00)
STAGE 6	-2.77 (-2.12)	-3.65 (-2.46)	
STAGE 7	-5.62 (-4.51)	-8.10 (-5.03)	
LARGE CITY			-.262 (-2.45)
LOW INCOME			-.196 (-1.89)
Education			
LAGER	-3.99 (-3.02)	-5.50 (-2.76)	-.470 (-2.79)
LBO/MAVO			-.262 (-2.88)
MBO			-.212 (-2.09)
MBO/MAVO	3.43 (3.44)	5.22 (4.30)	
HBO	9.35 (8.95)	10.83 (7.69)	
UNIV	10.05 (8.30)	13.42 (8.40)	
NDRIVERS	3.87 (6.05)	2.10 (2.91)	.197 (3.10)
NADULTS	-4.96 (-6.40)	-.51 (-0.42)	
NCHILD (<6 yr)	1.89 (2.63)	2.46 (3.46)	
NCHILD (<12 yr)	3.53 (4.32)	1.82 (2.06)	
NCHILD (<18 yr)			-.202 (-1.89)
NCHILD (>18 yr)			-.462 (-6.06)
$\hat{y}_{1i}/\sqrt{\text{NDIARIES}}$.049 (4.45)
$\hat{\epsilon}_{1i}$.82 (19.35)	.017 (5.84)
$\phi(\hat{w}_i)/\Phi(\hat{w}_i)$		-1.60 (-0.94)	
N	1743	1127	1743
r ²	.89	.93	
F	1268.0	1046.4	
α	<.00005	<.00005	
L(\hat{C})			-1132.3
L($\hat{\beta}$)			-971.1
χ^2			322.4
df			15

(): t-statistics
L(\hat{C}): Log-likelihood with the constant term alone
L($\hat{\beta}$): Log-likelihood with the estimated coefficients

with the household is proportional to *NDIARIES* (the independence assumption underlying this proposition may not be entirely satisfied, but this is not considered to be crucial).

The number of persons filling out the diary, *NDIARIES*, is the most significant variable in the wave-one trip equation with a regression coefficient of 22.5. This is not surprising since the dependent variable is the number of reported trips totaled over all diaries filled out by members of a household. This variable alone accounts for 86.9% of the total variation in the weighted least square analysis. Although other variables in the model are statistically highly significant, most of the variation is explained by *NDIARIES*. The constant term, which if included would have a value of -0.0235 with a t-statistic of -0.80 , is omitted in the model of Table 1.

Household education, which is represented in this model by 4 dummy variables (*LAGER*, *MBO/MAVO*, *HBO*, *UNIV*), shows a clear and strong positive association with the number of reported trips. Households with individuals with higher education tended to be more mobile and/or reported their trips more accurately.

Other variables with positive contributions are: number of drivers (*NDRIVERS*), number of children of 0 to 6 years old [*NCHILD* (<6 yr)], number of children of 7 to 11 years old [*NCHILD* (<12 yr)], and households of young couples without children (*STAGE1*). Older households without children (*STAGE6*) and pensioners (*STAGE7*) tended to have fewer reported trips. The negative coefficient of number of adults (*NADULTS*) implies that a child of 12 years or older on average had more trips reported in the diary than an adult member of the household.

Household income and car ownership do not significantly influence the total number of reported trips, presumably because walk and bicycle trips are included in the total.

In sum, the number of trips reported in the wave-one diary is mainly a function of household structure and education.

Attrition model

The estimated coefficients of the attrition probit model are also presented in Table 1. A positive coefficient in the table implies a positive contribution of the variable to the probability of participation in the second-wave survey. The number of persons keeping the diary (*NDIARIES*) has a significant positive coefficient. This is consistent with the household size effect presented earlier in a univariate context. Education again plays an important role in the probit attrition model. The coefficient of *LAGER*, the lowest level of education, has a large negative coefficient, and the coefficient value increases as the level of education increases. The estimated coefficient of the dummy variable for low income households (*LOWINCOME*) is also negative, while the coefficient of number of drivers (*NDRIVERS*) shows a positive contribution of this variable. These relations have also been seen earlier in Fig. 1.

This analysis indicates that households with older children tended to drop out of the panel, as indicated by the negative coefficients of the number of children between 12 and 17 years old [*NCHILD* (<18 yr)], number of children older than 18 years [*NCHILD* (\geq 18 yr)], and also life-cycle stage 5 (*STAGES*) that includes households with at least one child of 12 years of age or older.

Contrary to the result of the two-way tabulation shown in Fig. 1, the car ownership dummy variables (*ONECAR*, *TWOCARS*) in this multivariate context exhibit negative coefficients; other factors being equal, a household with more cars tended to drop out of the panel. This apparent contradiction is presumably due to the strong associations car ownership has with many other household characteristics that influence the household's participation decision.

The coefficient of the predicted number of trips from wave-one trip equation divided by the square-root of the number of household members who kept the diary ($\hat{y}_{1i}/\sqrt{NDIARIES}$) is positive and significant. Those households with a larger expected number of trips per person in the first wave tended to remain in the panel. The result offers strong empirical evidence that positive correlation exists between household mobility and panel participation. While such correlations were suggested or observed by heuristic methods in earlier studies (Baanders *et al.*, 1982), this study is the first to offer a statistical mechanism that allows rigorous quantification of the correlation.

The coefficient of $\hat{\xi}_{1i}$ is also positive and significant with a *t*-value of 5.84. Those households which either (or both) made more trips than expected or reported trips accurately than average, tended to participate in the second-wave survey and those which were less mobile than expected or under-reported trips had negative $\hat{\xi}_{1i}$ and tended to reject the second diary. This is an expected result and similar relationships have been suggested between reporting and participation (Brög *et al.*, 1982). Although it is not possible for us to determine which of the two, unexplained mobility or under-reporting of trips, is the major contributor, the result strongly supports the conjecture that those who under-reported their trips in the first wave tended to drop out.

Second-wave trip equation

The second-wave trip equation includes as exogenous variables the residual from the wave-one equation and a term to account for attrition bias in addition to the set of explanatory variables used in the wave-one equation. The objective of the analysis here is to evaluate the error term correlation and also assess the stability of weekly trip generation between the two waves.

The estimation result (Table 1) shows a strong effect of the first-wave residual term, whose coefficient is 0.822 with a *t*-statistic of 19.35. There exists strong positive correlation between the first-wave residual and the second-wave residual; a household's unexplained mobility and reporting errors are strongly correlated between the two waves.

The attrition bias correction term, defined as $\phi(\hat{W}_i)/\Phi(-\hat{W}_i)$, on the other hand, does not have a significant coefficient. We may conclude for this sample that, given that

a household chose to participate in the wave-two survey, its unobserved propensity to do so (i.e. the error term of the attrition model) does not influence the number of trips reported in the second-wave. This is a plausible conclusion. Since the result indicates that ϵ_i is not significantly correlated with ξ_{2i} , the latter is approximately homoscedastic and the t -statistics presented in the table serve as useful approximations.

The estimated coefficients are in general quite comparable between the two waves in terms of their signs, magnitudes, and estimated t -statistics. Significant differences however can be found in the coefficients of the number of members who kept the diary (*NDIARIES*), number of drivers (*NDRIVERS*), and number of adults (*NADULTS*) (the difference is significant at $\alpha = 0.1$ for *NDIARIES* and at $\alpha = 0.05$ for the other two variables). Based on these three coefficients, the contribution that an additional adult holding a driver's license has to the total number of trips is evaluated as 21.38 trips in the first wave and 21.56 trips in the second wave. The stability shown by drivers is noteworthy. The contribution of an adult without a driver's license is 17.51 and 19.46, and a minor diary keeper shows the contribution of 25.24 and 23.62, for the respective waves. The increased contribution of nondrivers in the second wave when the weather is more favorable is intuitively agreeable. Overall, the comparison of the two trip equations suggest relative stability in household trip making between the waves.

Nature and magnitude of biases when correction terms are omitted

The above results have clearly indicated that the error term of the wave-one trip equation significantly affects the error terms of the attrition model and wave-two trip equation. A practical and important question is the significance of the biases in coefficient estimates that may result when the correction terms are omitted. Table 2 shows the estimation results of alternative wave-two trip equation which were estimated to examine this. Model a is the same model as shown in Table 1 and contains the correction term for attrition selectivity and the wave-one residual. The correction term for attrition bias is omitted in model b, and the correction term and wave-one residual are both omitted in model c (note that the estimated t -statistics of model c are not consistent since the model is mis-specified with the significant wave-one residual omitted).

The differences in the estimated coefficient vectors between model a and model b are small and the regression coefficient (r^2) shows no difference. This result is not

Table 2. Comparison of wave-two trip equations estimated with and without correction terms

	MODELS		
	a	b	c
NDIARIES	19.97 (30.80)	19.88 (31.00)	20.75 (27.26)
Lifecycle Stage			
STAGE 1	3.85 (2.26)	4.13 (2.47)	4.50 (2.26)
STAGE 6	-3.65 (-2.46)	-3.41 (-2.33)	-1.72 (-0.99)
STAGE 7	-8.10 (-5.03)	-7.93 (-4.96)	-6.32 (-3.33)
Education			
LAGER	-5.50 (-2.76)	-5.49 (-2.75)	-4.87 (-2.05)
MBO/HAVO	5.22 (4.30)	4.75 (4.30)	4.81 (3.66)
HBO	10.83 (7.69)	10.00 (9.13)	10.04 (7.71)
UNIV	13.42 (8.40)	12.48 (10.05)	12.06 (8.17)
NDRIVERS	2.10 (2.91)	1.94 (2.77)	1.09 (1.31)
NADULTS	-.51 (-0.42)	-1.15 (-1.15)	-1.73 (-1.45)
NCHILD (<6 yr)	2.46 (3.46)	2.43 (3.43)	3.55 (4.21)
NCHILD (<12 yr)	1.82 (2.06)	1.75 (1.98)	2.89 (2.76)
$\hat{\xi}_{1i}$.82 (19.35)	.80 (21.55)	
$\hat{\phi}(\hat{W}_i)/\hat{\sigma}(-\hat{W}_i)$	-1.60 (-0.94)		
N_2	1127	1127	1127
r^2	.93	.93	.90
F	1046.4	1126.9	835.2
α	<.00005	<.00005	<.00005
ρ^2	.72	.72	.60

Note: r^2 is obtained from the weighted least squares procedure. The correlation coefficient, ρ^2 , is evaluated for the unweighted number of trips predicted by the model and the unweighted observed number of trips.

Table 3. Comparison of attrition models estimated with and without correction terms

	MODELS			
	a	b	c	d
Const.	-1.201 (-3.58)	-1.097 (-3.29)	.148 (1.05)	.147 (1.05)
NDIARIES	.595 (6.47)	.564 (6.21)	.344 (4.85)	.334 (4.75)
ONECAR	-.263 (-2.59)	-.251 (-2.50)	-.264 (-2.63)	-.252 (-2.52)
TWOCARS	-.537 (-3.51)	-.551 (-3.63)	-.566 (-3.72)	-.576 (-3.82)
Lifecycle Stage				
STAGE 1	-.236 (-2.18)	-.225 (-2.09)	-.190 (-1.77)	-.183 (-1.71)
STAGE 5	-.321 (-2.00)	-.325 (-2.03)	-.272 (-1.70)	-.278 (-1.75)
LARGE CITY	-.262 (-2.45)	-.266 (-2.50)	-.276 (-2.60)	-.280 (-2.65)
LOW INCOME	-.196 (-1.89)	-.194 (-1.87)	-.139 (-1.37)	-.140 (-1.38)
Education				
LAGER	-.470 (-2.79)	-.497 (-2.97)	-.959 (-7.49)	-.948 (-7.46)
LBO/MAVO	-.262 (-2.88)	-.284 (-3.15)	-.480 (-6.24)	-.484 (-6.35)
MBO	-.212 (-2.09)	-.220 (-2.18)	-.326 (-3.30)	-.326 (-3.32)
NDRIVERS	.197 (3.10)	.202 (3.21)	.261 (4.22)	.261 (4.25)
NCHILD (<18 yr)	-.202 (-1.89)	-.180 (-1.71)	-.137 (-1.31)	-.122 (-1.18)
NCHILD (≥18 yr)	-.462 (-6.06)	-.440 (-5.84)	-.404 (-5.44)	-.389 (-5.29)
$\hat{y}_{1i}/\sqrt{NDIARIES}$.049 (4.45)	.045 (4.13)		
$\hat{\epsilon}_{1i}$.017 (5.84)		.016 (5.60)	
$L(\hat{\beta})$	-971.1	-988.6	-981.3	-997.3
$-2(L(0)-L(\hat{\beta}))$	473.9	439.1	453.6	421.6
df	16	15	15	14
$-2(L(\hat{c})-L(\hat{\beta}))$	322.4	287.6	302.6	270.1
df	15	14	14	13

$L(0) = -1208.1$, and $L(\hat{c}) = -1132.3$, where $L(0)$ is the log-likelihood with no coefficients. The differences, $-2(L(0) - L(\hat{\beta}))$ and $-2(L(\hat{c}) - L(\hat{\beta}))$ are chi-square distributed with the degrees of freedom shown below them.

surprising because the correction term for the attrition bias has an insignificant coefficient. When the wave-one residual is also omitted (model c), differences in the coefficient estimates become clear for *STAGE6*, *NDRIVERS*, *NADULTS*, *NCHILD* (<6 yr), and *NCHILD* (<12 yr). More importantly, the estimated *t*-statistics for these variables are substantially different between model a and model c. The result indicates that model specification effort using panel data may be misguided by inconsistent *t*-statistics that result from omitting residuals from previous waves and points to the importance of accounting for biases due to correlated error structures.

Similar comparison is made in Table 3 for the attrition equation which contains the residual from the wave-one trip equation to correct the bias due to the correlation between the wave-one error term and the attrition error term. The result presented in Table 1 is repeated in Table 3 as model a.

While the estimated coefficient of the correction term is highly significant (model a), omitting this term does not appear to result in substantially different coefficient estimates (model b). Unexplained mobility and reporting errors from wave one do affect attrition behavior, but this effect is only through the error term of the attrition model. It is worthy to note that, in spite of this, exclusion of the correction term does substantially influence the predicted probability of attrition.

Model c of Table 3 contains the correction term, but omits the predicted number of trips from the wave-one trip equation ($\hat{y}_{1i}/\sqrt{NDIARIES}$). Comparison of the coefficient vectors of model a and model c clearly shows the large impact of this variable on the coefficient values, especially on the coefficients of the education dummy variables.

Inclusion of $\hat{y}_{1i}/\sqrt{NDIARIES}$ together with the wave-one residual can be interpreted to represent the hypothesis that attrition behavior is dependent on the behavior during the first-wave diary period, a case of history dependence. The log-likelihood values of model a and model d yield a chi-square value of 52.3 ($df = 2$). Clearly these two variables are highly significant and attrition is in fact dependent on the behavior during the first wave.

Use of the reciprocal of attrition probability as weight

One of the practical applications of the model system of this study is the development of weight for the wave-two sample in order to account for the effects of attrition; since

households of certain characteristics tended to leave the panel, the wave-one and wave-two samples are no longer comparable in their composition unless the wave-two households are properly weighted.

The reciprocal of the probability to participate in the second-wave survey obtained from the attrition model can be used as such a weight. The weight thus developed fully utilizes the information that the survey results offer and is based on attrition behavior observed at a disaggregate level. Its use can be preferred over the more typical approach of weighting observations based on sample distributions across cells defined by a few socioeconomic or demographic variables. Specifically noted is the fact that this weight reflects unexplained mobility and reporting errors from the first wave. Also note that information from all households, including those dropped out, is utilized in estimating the coefficients of the attrition model, and therefore is used in defining the weight.

Table 4 presents the sample mean of the number of trips per household by mode, and the mean number of mode users per household. All statistics are totals for the one-week diary period. The wave-one mean weekly trip rate is 50.70 trips, or 7.24 trips per day per household. About 45% of these trips are made by car, 7.1% by public transit, 31% by bicycle, and 15% by walking. The average number of persons in a household who made car trips during the diary period is 1.87, while the corresponding number of public transit users is 0.49 for bus, tram, and metro (BTM) and 0.25 for train.

The second column of Table 4 shows unweighted sample means from the second wave. Only those 1,165 households who participated in both the wave-one and wave-two surveys are included in this column. The number of trips in general increased, except for BTM and walk trips. The increase in bicycle trips is quite notable.

The third column shows sample averages obtained by applying the weight described above. Note that the number of observations, N , which is inflated by the weight, is quite close to the number of observations in the first wave. Contrary to the unweighted averages, the weighted sample means now indicate that the number of trips slightly decreased in the same period, except for the slight increase in the number of train trips and bicycle trips. This result is not surprising; those wave-one participants who remained in the panel are the ones who tended to have positive residuals in the wave-one trip equation, i.e. tended to make a more-than-expected number of trips and/or reported their trips more accurately. Consequently the unweighted means of the second wave are larger than the wave-one means. When the households are weighted and this attrition bias is accounted for, the numbers of trips and mode users are relatively stable between the two waves.

CONCLUSION

An econometric model of travel behavior was developed in this study and applied to the first two waves of the Dutch National Mobility Panel data set. The model system

Table 4. Weighted and unweighted sample means of mobility indicators from wave one and wave two

	WAVE 1		WAVE 2	
	All Hhlds Unweighted	Wave 1 Households * Unweighted	Wave 1 Households * Unweighted	Weighted
N	1764	1165	1165	1729**
Total No. of Trips	50.70	53.44	53.44	48.96
No. of Car Trips	22.69	23.62	23.62	21.81
No. of BTM Trips	2.29	2.00	2.00	2.03
No. of Train Trips	1.30	1.49	1.49	1.34
No. of Bicycle Trips	15.82	18.44	18.44	16.42
No. of Walk Trips	8.59	7.89	7.89	7.36
No. of Car Users	1.87	1.96	1.96	1.86
No. of BTM Users	.49	.46	.46	.47
No. of Train Users	.25	.28	.28	.27
No. of Bicycle Users	1.39	1.56	1.56	1.44
No. of Walk-Trip Makers	1.32	1.32	1.32	1.25

* Replacement households added in wave-two survey excluded.

** Sample size expanded by the weight.

consists of trip equations for the respective waves and a probit attrition probability model. The development of the estimation procedure assumed chronological dependencies among the error terms of the model components. The model system offers a mechanism by which relationships among mobility, trip reporting errors, and attrition behavior in a panel survey can be statistically evaluated.

Strong correlation was found between the residuals of the trip equations for the two waves and also between those of the first-wave trip equation and attrition probability. This led to the conclusion that those who were less mobile and/or under reported their trips in wave one tended to drop out of the panel in the second-wave survey, and also that those who accurately reported their trips in wave one tended to do so in wave two as well. The study also showed the magnitudes of biases in coefficient estimates that may arise when these correlations are ignored.

As a practical application of the model system, this study has demonstrated how a sample weighting factor can be developed from the attrition probability model. The weight developed using the attrition model fully utilizes the information that the survey results offer and can be preferred over the typical method where observations are weighted according to sample distributions of a few socioeconomic or demographic variables. In addition, this weight reflects unexplained mobility and reporting errors from the first wave. The application of the weight has shown that unweighted comparison of sample means between the two waves may be misleading. The econometric model system developed here is simple to use with panel data, yet versatile enough to be applied to various aspects of longitudinal behavior.

Acknowledgement—The authors are grateful to anonymous referees for their comments and suggestions. This research was performed while the first author was on sabbatical leave at the Institute for Town Planning Research, Delft University of Technology. The funding provided by Rijkswaterstaat, Dutch Ministry of Transport, is gratefully acknowledged.

Appendix Table. Definition of the variables in the model system

VARIABLE	DEFINITION
NDIARIES	Number of household members who kept the diary
ONECAR	1 if one car is available to household
TWOCARS	1 if two or more cars are available to household
Lifecycle Stage	
STAGE 1	1 if household of a couple, male less than 35 years old, with no children at home
STAGE 5	1 if household with at least one child older than 12 years
STAGE 6	1 if household of a couple, male between 35 and 64 years old, with no children at home
STAGE 7	1 if household of a couple, male more than 64 years old, with no children at home
LARGECITY	1 if large metropolitan area
LOWINCOME	1 if household belongs to the lowest income category
Education	
LAGER	1 if household education is at elementary school level
LBO/MAVO	1 if household education is at lower vocational school level or intermediate general preparatory school level
MBO	1 if household education is at intermediate vocational school level
MBO/HAVO	1 if household education is at intermediate vocational school or higher general preparatory school level
HBO	1 if household education is at higher vocational school level
UNIV	1 if household education is at university level
NDRIVERS	Number of drivers in household
NADULTS	Number of adults in household
NCHILD (<6 yr)	Number of children less than 6 years old
NCHILD (<12 yr)	Number of children between 6 and 11 years old
NCHILD (<18 yr)	Number of children between 12 and 17 years old
NCHILD (>18 yr)	Number of children 18 years of age or older

REFERENCES

- Ampt E., Bechinger W., Brög W. and Messelink, A. (1983) Understanding travel survey data validation. Paper presented at the 1983 Australian Transport Research Forum, Canberra, Australia.
- Baanders A., Garden J. M. and Brög W. (1982) Panels: attractions and pitfalls. Paper presented at PTRC Summer Annual Meeting, Warwick, England.
- Brög W. and Meyburg A. H. (1980) Nonresponse problem in travel surveys: an empirical investigation. *Transp. Res. Rec.* 775, 34–38.

- Brög W., Erl E., Meyburg A. H. and Wermuth M. J. (1982) Problems of non-reported trips in surveys of nonhome activity patterns. *Transp. Res. Rec.* **891**, 7-15.
- Brög W., Meyburg A. H. and Wermuth M. J. (1983) Development of survey instruments suitable for determining nonhome activity patterns. *Transp. Res. Rec.* **944**, 7-13.
- Bureau Goudappel Coffeng (1983) Een Panel voor Longitudinaal Verplaatsingsonderzoek: Beschrijving en Analyse van de Mobiliteitsontwikkeling. Deventer, The Netherlands.
- Bureau Goudappel Coffeng (1985) Panel Steekproefopzet en-Verwerking. Deventer, The Netherlands.
- Golob T. F. and Meurs H. (1986) Biases in response over time in a seven-day travel diary. *Transpn*, **13**, 163-81.
- Golob J. M., Schreurs L. J. M. and Smit J. G. (1986) The design and policy applications of a panel for studying changes in mobility over time. In: *Behavioural Research for Transport Policy*, VNU Science Press, Utrecht, The Netherlands, pp. 81-95.
- Hanson S. and Huff J. O. (1983) Repetition and variability in longitudinal records of individual travel behavior. Paper presented at the 30th North American Meetings of the Regional Science Association, Chicago, IL.
- Hausman J. A. and Wise D. A. (1979) Attrition bias in experimental and panel data: the Gary income maintenance experiment. *Econometrica* **47**, 455-73.
- Heckman J. J. (1979) Sample selection bias as a specification error. *Econometrica* **47**, 153-161.
- Hendriks F. W. M. (1982) Data-verzameling in verkeersonderzoek middels huysenquetes. Centraal Bureau voor de Statistiek, Heerlen, The Netherlands.
- Hensher D. A. (1985) Issues in the Pre-analysis of Panel Data. School of Economic and Financial Studies, Macquarie University, North Ryde, New South Wales, Australia.
- Hensher D. A. and Wrigley N. (1986) Statistical modelling of discrete choices in discrete time with panel data. In: *Behavioural Research for Transport Policy*, VNU Science Press, Utrecht, The Netherlands, pp. 97-116.
- Instituut voor Longitudinaal Beleidsonderzoek (1984) *Onderzoek Tarievenplan: Verslag van de Screening en de Erste Meting*. Amsterdam, The Netherlands (prepared for Ministerie van Verkeer en Waterstaat, Projectbureau Integrale Verkeers—en Vervoerstudie).
- Instituut voor Longitudinaal Beleidsonderzoek (1985) *Verslag van de Tweede Meting Longitudinaal Verplaatsings-Onderzoek—Najaar 1984*. Amsterdam, The Netherlands (prepared for Ministerie van Verkeer en Waterstaat, Projectbureau Integrale Verkeers—en Vervoerstudie).
- Johnson N. L. and Kotz S. (1970) *Distribution in Statistics: Continuous Univariate Distributions—1*. John Wiley and Sons, New York, NY.
- Johnson N. L. and Kotz S. (1972) *Distribution in Statistics: Continuous Multivariate Distributions*. John Wiley and Sons, New York, NY.
- Kitamura R. (1986) An analysis of weekly activity patterns and travel expenditure. In: *Behavioural Modelling in Geography and Planning* (Edited by R. G. Golledge and H. Timmermans), Croom Helm, London (forthcoming).
- Kitamura R. and Bovy P. H. L. (1985) Attrition and Trip reporting in the First Two Waves of Dutch Mobility Panel Data. Institute for Town Planning Research, Delft University of Technology, Delft, The Netherlands.
- Kitamura R. and Van der Hoorn T. (1987) Regularity and irreversibility of weekly travel behavior. *Transpn*, **15** (forthcoming).
- Koppelman F. S. and Pas E. I. (1984) Estimation of disaggregate regression models of person trip generation with multiday data. In: *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory* (Edited by J. Volmuller and R. Hamerslag), VNU Science Press, Utrecht, The Netherlands, pp. 513-31.
- Maddala G. S. (1978) Selectivity problems in longitudinal data. *Annales de l'INSEE*, No. 30-31, 423-50.
- Maddala G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge, England.
- Terza J. (1980) Heckman's method extended to polychotomous choice models. *Proceedings of the American Statistical Association: Business and Economics Section*, 466-470.