Title

Travel Trends Using the Puget Sound Panel Survey: A Generalized Estimating Equations Approach

Authors
Yee, Julie
Niemeier, Debbie

Publication Date
1998

Pergamon

Transpn Res.-B, Vol. 32, No. 1, pp. 21–34, 1998

0191-2615/98 $19.00+0.00

PII: S0191-2615(97)00010-6

TRAVEL TRENDS USING THE PUGET SOUND PANEL SURVEY: A GENERALIZED ESTIMATING EQUATIONS APPROACH

JULIE L. YEE
Division of Statistics, University of California, Davis, CA 95616, U.S.A.

and

DEBBIE A. NIEMEIER*
Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, U.S.A.

(Received 17 June 1996; in revised form 21 April 1997)

Abstract—This paper examines longitudinal mode use trends using four waves of the Puget Sound Transportation Panel. The analysis is conducted using generalized estimating equations for model estimation. In addition to examining mode use frequencies over time, we also consider mode use trends conditioning on household income and lifecycle stage. As expected, results indicate an overall increase in the number of worktrips made between 1989 and 1993 and these trips were marked by increasing use of single occupancy vehicles. The full parameters of the model were also to estimate the rate of increase in terms of percentage increase and their confidence intervals. Results indicate that the mean number of worktrips made by driving alone significantly increased from wave 1 to wave 4; with a 95% C.I. the rate of percent increase was estimated between 8.2 and 24.5%. The ranges for rates of change in high occupancy vehicle modes and non-motorized worktrip frequencies overlap with the range for single occupancy vehicle rate of change, and thus, it cannot be said that rate of change for the high occupancy modes was significantly different from the single occupancy rates of change. The rate of change in the mean frequency for the high occupancy–transit mode is not only below the range for single occupancy vehicle trips but also suggests, with 95% confidence, the rate of percent decrease was between 2.88 and 44.0%. © 1998 Elsevier Science Ltd

Keywords: quasi-likelihood, generalized linear model, generalized estimating equations.

1. INTRODUCTION

Recent summary statistics, based on the Nationwide Personal Transportation Study (NPTS), suggest that single occupancy vehicle trips in the U.S. have increased in the past decade while public transport trips have declined. In 1983, single occupancy vehicles accounted for 71.1% of all journey to work trips, increasing to 83.0% in 1990. Concomitantly, between 1983 and 1990, public transport journey to work trips decreased from 4.5 to 4.0% (Hu and Young, 1993). Likewise, U.S. census data suggests that private vehicle use for the journey to work trip increased by approx. 27% between 1960 and 1990 while public transit use declined by as much as 60% during the same period (Rossetti and Eversole, 1993).

Much of the mode use literature associated with trends analysis relies on cross-section data to describe the various longitudinal travel patterns. The potential problems associated with travel behavior analysis using cross-section data have been well documented (e.g. Kitamura, 1990) and include, for example, lack of temporal insight and omitted and confounding variables (Golob, 1990). More recently, researchers have explored travel behavior with panel data, thus overcoming some of the difficulties associated with cross-section data. However, much of this research has been used to examine the effects of specific transportation policies on travel behavior.

For example, Kitamura *et al.* (1990) examined how travel patterns change when telecommuting is considered using panel data. Others have examined the effects of staggered work hours on travel

*Author for correspondence.

behavior (Golob and Guiliano, 1989) and changes in mobility (van Wissen and Meurs, 1989) using panel data. Perhaps the most notable panel data are the Dutch National Mobility Panel which includes weekly travel diaries and household and personal demographic data. Golob (1990) used the Dutch Panel to examine the relationship between travel time expenditures and car ownership while Meurs (1990) examined the characteristics of trip generation.

Absent from the travel behavior literature is a U.S. based analysis of mode use trends using panel survey data. The purpose of this paper is to examine mode use trends using four waves of the Puget Sound Transportation Panel (PSTP). The analysis is conducted using generalized estimating equations (GEE) for model estimation. In addition to examining mode use, we also consider mode use trends conditioning on exogenous variables such as household income and lifecycle stage.

The paper begins with a description of the PSTP data and the major estimation issues and research questions. In Section 3, a theoretical model is presented; Section 4 describes the model estimation. In Section 5, the results are described and finally, conclusions are presented.

2. THE PSTP

The PSTP is a longitudinal data that consists of four waves of travel data collected during the years 1989–1993 (see Murakami and Watterson, 1990). Each wave is organized into three data files: one each containing household, person, and trip diary information for every household member of driving age. Each household is represented by a single record in the household files indexed by a household identification number and carrying attribute information such as household income and lifecycle. Similarly, each person is represented by a single record in the person file and is indexed by a household and person identification number. The person files contain profiles of the individual participants with information such as age, sex and occupation. Finally, the trip diary file includes trip attributes for every trip taken during the two day travel period. Each individual trip is described in terms of the trip purpose, mode, and other related attributes; this file is indexed by the household and person identification and trip number.

Using the trip diary files for each wave, trip modes can be categorized into four mutually exclusive categories:

(a) Single occupant vehicle;
(b) HOV-pool (high occupancy vehicles that include carpool, vanpool, and taxi);
(c) HOV-transit (high occupancy modes that include bus and paratransit); and
(d) non-motor (walk and bike).

Eight additional modes represented in the data were excluded from this analysis. These include motorcycle, school bus, ferry/car, ferry/foot, monorail, boat, train, and airplane and constitute only a small portion of the sample.

We are particularly interested in analyzing mode use over time on the subset of respondents participating in all four waves. Additionally, we restrict the data to work-related trips for those individuals with work trip information in all four waves and only those individuals with reported income and household information in all four waves. The first two waves contained a few households whose incomes were categorized under an alternate scheme indicating only whether they made less than or greater than \$30,000 as opposed to the \$35,000 cutoff point used in our categorizing scheme. These subjects represented a small proportion of the sample and were also omitted from the analysis.

Under these conditions, we have a by-wave sample size of 519 subjects, each with 16 observations (four modes and four waves), for a total of 8304 observations. There are 222 records in which the subject is associated with the same household as another subject. We will assume that individuals from the same household behave independently of one another. Table 1 presents basic trip summary statistics for each wave.

Table 2 presents the basic demographic features associated with continuing participants in each wave. As others have noted (Goulias and Ma, 1996), the proportion of higher income groups increased over time. Approximately 65.9% of the respondents made greater than \$35,000 in wave 1, compared to 82.5% in wave 4. The proportion of lifecycle types also changed over time. Lifecycle groups 1, 3, and 6 increased over time while groups 4 and 7 declined in number.

Table 1.

	Wave 1	Wave 2	Wave 3	Wave 4
Mean no. of trips (SE)				
SOV	1.92 (0.05)	2.34 (0.10)	2.22 (0.08)	2.33 (0.08)
HOV-pool	0.34 (0.03)	0.37 (0.04)	0.37 (0.04)	0.41 (0.05)
HOV-transit	0.23 (0.03)	0.22 (0.03)	0.16 (0.02)	0.18 (0.03)
Non-motor	0.12 (0.02)	0.17 (0.03)	0.12 (0.02)	0.15 (0.03)
Percent of total trips				
SOV	73.8	75.5	77.2	75.8
HOV-pool	12.9	12.0	12.9	13.5
HOV-transit	8.6	7.0	5.7	5.8
Non-motor	4.7	5.5	4.2	4.9

The analysis of longitudinal count data such as we have defined requires specific statistical techniques (e.g. Kitamura, 1990; Maddala, 1987). If, for each of k subjects, the kth set of observed counts, Y_k, is defined to have expected value μ_k, then the standard linear model $\mu_k = \mathbf{X}_k'\beta$ is limited in two basic respects. First, the range of μ_k is not restricted; this hinders practical use in modeling, for example, count data or proportions. Second, the standard linear model assumes independent normally distributed errors with error variances independent of μ_k. Each individual k contributes correlated observations to the full likelihood of the panel data. Without understanding the nature of the correlation between these observations, the contribution to the likelihood by each subject cannot be known, much less used to computationally fit a model.

3. GENERALIZED ESTIMATING EQUATIONS

Generalized linear models (GLM) addresses the limitations of the traditional modeling approach by specifying a monotone differential link function $g(\mu_k)$ that is equated to $\mathbf{X}_k'\beta$, where $\mathbf{X}_k$ is a vector of covariate measures and β is a vector of coefficients relating $\mathbf{X}_k$ to $g(\mu_k)$. A general discussion of the model foundation may be found in McCullagh and Nelder (1989). $\mathbf{Y}_k$ is assumed to have a known distribution in the exponential family with a variance that is a known function of μ_k. For example, proportions may be modeled as a logit model with link function, $g(\mu_k) = \log(\frac{\mu_k}{1-\mu_k}) = \mathbf{X}_k'\beta$, thus restricting each component of μ_k to the range [0,1].

Suppose the interest lies in evaluating how Y_{ij}, the total frequency of trips made during a specified wave j and of mode type i, changes over time, where $i = 1, ..., I$ and $j = 1, ..., J$. We begin by specifying a distribution for the Y_{ij}s. If we assume a Poisson distribution, the canonical link is the natural log function with mean and variance $\exp(\mathbf{X}_{ij}'\beta)$. A general wave-mode GLM may now be formulated as:

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \text{wave}_j + \sum_{s=2}^{I} \beta_{2,s} 1\{s = i\} + \sum_{s=2}^{I} \beta_{3,s} \text{wave}_j * 1\{s = i\} \tag{1}$$

Table 2.

	Wave 1	Wave 2	Wave 3	Wave 4
Income (%)				
< \$35,000	34.1	23.7	18.7	17.5
> \$35,000	65.9	76.3	81.3	82.5
Lifecycle type (%)				
(1) Any child, < 6 yr old	18.1	17.1	13.9	12.7
(2) All children, 6–17 yr	29.3	31.6	30.6	29.7
(3) 1 adult, < 35 yr	2.3	2.1	1.7	1.3
(4) 1 adult, 35–64 yr	7.1	7.7	8.9	10.2
(5) 1 adult, ⩾65 yr	0.4	0.6	0.8	0.8
(6) 2+ adults, < 35 yr	3.9	3.4	1.3	1.2
(7) 2+ adults, 35–64 yr	37.2	35.8	39.9	40.8
(8) 2+ adults, ⩾65 yr	1.7	1.5	2.9	3.3

where μ_{ij} is the mean of Y_{ij} and $1\{\cdot\}$ is the mode indicator function. The summation in the latter two terms does not include $s = 1$ since the baseline model is already set at mode type 1. The mean μ_{ij} measures the propensity of subjects in the population to take a trip in wave j by mode i. It is assumed that $\mu_{ij} > 0$, that is, there is always some propensity among the population to travel by mode i and in wave j, no matter how small. A small value of μ_{ij} would be associated with many observations of Y_{ij} equaling zero. The GLM estimate for the vector β can be obtained by using the estimating equation,

$$\mathbf{D}'\mathbf{H}^{-1}(\mathbf{Y} - \mu) = 0 \tag{2}$$

where $\mu = g^{-1}(\mathbf{X}'\beta)$, $\mathbf{D}$ represents the matrix $\frac{\partial\mu}{\partial\beta}$ and, assuming independence between observations, $\mathbf{H}$ is a diagonal covariance matrix. With a panel data model specified as in eqn (1), and assuming I modes and J waves, there are IJ correlated observations and, without modifications, the standard GLM should not be applied.

Liang and Zeger (1986) and Zeger and Liang (1986) extend the GLM estimating equations to account for correlated measurements using quasi-likelihood. The approach, known as Generalized Estimating Equations (GEE), allows the distribution of $\mathbf{Y}_k$ to be known only to the extent that the mean can be reasonably expressed by $\mu_k = g^{-1}(\mathbf{X}_k'\beta)$ with the variance expressed as a function of the mean, $\mathbf{H}(\mu_k)/\phi$, where ϕ is a scale parameter. Further, in the GEE approach, an empirical variance function is estimated so that model parameters and their variances are consistently estimated even if the correlation structure is misspecified.

The basic form of the quasi-likelihood estimating functions for a correlated response vector is defined for the IJ-vector $\mathbf{Y}$ measured from a single individual. The coefficients may be estimated by implementing the score function given by,

$$\mathrm{U}(\beta) = \mathbf{D}'\phi\mathbf{H}^{-1}(\mathbf{Y} - \mu) \tag{3}$$

where $\mathbf{D}$ is a $IJ \times p$ matrix where the (s, t)th element is given by $\partial\mu_s/\partial\beta_t$, i.e. the partial derivative of the sth component of the IJ-dimensional vector μ with respect to the tth component of the p-dimensional vector β. The function U has properties which are similar to the derivative of the log-likelihood function. For this reason, the estimating equations,

$$\mathrm{U}(\beta) = 0$$

are known as quasi-likelihood estimating equations. The covariance matrix for U is

$$i_\beta = -\mathrm{E}\left(\frac{\partial}{\partial\beta}\mathrm{U}\right)\Big|_\beta = \mathbf{D}'\mathbf{H}^{-1}\mathbf{D}\phi$$

which is analogous to the Fisher information derived for the usual ordinary likelihood functions. The covariance matrix for the estimates $\hat{\beta}$ is given by the inverse

$$\mathrm{Cov}(\hat{\beta}) = i_\beta^{-1} = (\mathbf{D}'\mathbf{H}^{-1}\mathbf{D})^{-1}/\phi.$$

The quasi-likelihood can be extended to a IJK dimensional model vector encompassing the measurements from K individuals. Assuming that respondents are independent, then eqn (3) represents the contribution of each individual to the quasi-likelihood. The estimating equations are then,

$$U(\beta) = \sum_{k=1}^{K} \mathbf{D}_k'\phi\mathbf{H}_k^{-1}(\mathbf{Y}_k - \mu_k) = 0 \tag{4}$$

with solution denoted by $\hat{\beta}$. The corresponding covariance matrix for $\hat{\beta}$ is

$$\left[-E\left(\frac{\partial}{\partial\beta}\mathrm{U}\right)\right]^{-1}\Bigg|_\beta = \left(\sum_{k=1}^{K}\mathbf{D}_k'\mathbf{H}_k^{-1}\mathbf{D}_k\right)^{-1}\Big/\phi.$$

The computational solution of the GEE, eqn (4), requires iteration using an initial estimate of β^0 and the recursive assignment,

$$\hat{\beta}^{t+1} = \hat{\beta} + \left[\sum_{k=1}^{K} \hat{\mathbf{D}}_k' \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{D}}_k\right]^{-1} \left[\sum_{k=1}^{K} \hat{\mathbf{D}}_k' \hat{\mathbf{H}}_k^{-1} (\mathbf{Y}_k - \hat{\mu}_k)\right],$$

where $\hat{\mathbf{D}}_k$, $\hat{\mathbf{H}}_k$, and $\hat{\mu}_k$ are functions of $\hat{\beta}$ and updated at every iteration using the current estimator $\hat{\beta}^t$. If $\hat{\beta}^0$ is sufficiently close to $\hat{\beta}$, then the sequence of $\hat{\beta}^t$ converges to $\hat{\beta}$, and is a sensible starting value for $\hat{\beta}$. The estimated covariance matrix for $\hat{\beta}$ is given by

$$\hat{i}_{\beta}^{-1} = \left[-E\left(\frac{\partial}{\partial\beta}\mathbf{U}\right)\right]^{-1}\Big|_{\hat{\beta}=\beta} = \left(\sum_{k=1}^{K} \hat{\mathbf{D}}_k' \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{D}}_k\right)^{-1} \hat{\phi}, \tag{5}$$

where $\hat{\phi}$ is the estimate for the scale parameter and defined by

$$\hat{\phi}^{-1} = \frac{1}{IJK - p} \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(Y_{ijk} - \hat{\mu}_{ijk})^2}{\hat{\mu}_{ijk}} = \frac{\mathrm{X}^2}{IJK - p}$$

where X^2 is the generalized Pearson statistic (McCullagh and Nelder, 1989).

In specifying the covariance matrix functional **H**, the generalized estimating equations (GEE) can be used to treat longitudinal data where **H** is not only a function of μ but also of an additional set of parameters. Specifying **H** in the form suggested by Liang and Zeger (1986) leads to

$$\mathbf{H}_k = \mathbf{A}_k^{\frac{1}{2}} \mathbf{R} \mathbf{A}_k^{\frac{1}{2}} / \phi$$

where **R** is a suitable correlation matrix for the outcomes, $\mathbf{Y}_k$, and $\mathbf{A}_k$ is the $IJ \times IJ$ diagonal matrix with diagonal elements $\phi\mathrm{Var}(Y_{ijk})$, $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$. Using eqn (5) to estimate the variances of the parameters implies that the correlation matrix **R** has been correctly specified. In practice, it is difficult to ascertain that the true correlation has in fact been specified. Liang and Zeger propose a 'working' robust empirical estimate of the correlation structure to protect against misspecification. The alternate covariance matrix estimator is of the form:

$$\mathbf{V} = \left(\sum_{k=1}^{K} \mathbf{D}_k' \mathbf{H}_k^{-1} \mathbf{D}_k\right)^{-1} \left(\sum_{k=1}^{K} \mathbf{D}_k' \mathbf{H}_k^{-1} \mathrm{Cov}(\mathbf{Y}_k) \mathbf{H}_k^{-1} \mathbf{D}_k\right) \left(\sum_{k=1}^{K} \mathbf{D}_k' \mathbf{H}_k^{-1} \mathbf{D}_k\right)^{-1} \phi. \tag{6}$$

Liang and Zeger also noted that, under mild regularity conditions, β and **V** provide consistent estimates even when **H** is misspecified. Furthermore, it can be shown that

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \xrightarrow{L} W \sim \mathrm{Normal}\left(0, \lim_{n\to\infty} n\mathbf{V}\right).$$

This justifies the construction of t-ratios, which for large sample sizes, may be referred to a normal probability table. With data from a finite sample, $\mathrm{Cov}(\mathbf{Y}_k)$ need only be replaced by the IJ-dimensional square matrix $(\mathbf{Y}_k - \hat{\mu}_k)(\mathbf{Y}_k - \hat{\mu}_k)'$. In this framework, $\mathbf{H}_k(\mu_k)/\phi$ is known as the 'working' covariance matrix. Although $\mathbf{Y}_k$ is believed to contain correlated elements, $\mathbf{H}_k$ can still be specified as if the structure was entirely independent. This would correspond to an independent working correlation matrix and is similar to the solution by ordinary quasi-likelihood with the exception that **V** differs from the covariance estimator in eqn (5) by a correction factor,

$$\left(\sum_{k=1}^{K} \hat{\mathbf{D}}_k' \hat{\mathbf{H}}_k^{-1} \mathrm{Cov}(\mathbf{Y}_k) \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{D}}_k\right) \left(\sum_{k=1}^{K} \hat{\mathbf{D}}_k' \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{D}}_k\right)^{-1},$$

that makes it robust to departures from the assumption that **H** is correctly specified. With the PSTP data, this is advantageous in that it relaxes the necessity of understanding the nature of correlation of any two frequencies observed for the same person.

4. MODEL SPECIFICATION

The model given by eqn (1) may be specified for the PSRC panel data in the following form:

$$\begin{aligned}\log(\mu_{ij}) = \beta_0 + \beta_1 j + \beta_{22}1\{i=2\} + \beta_{23}1\{i=3\} + \beta_{24}1\{i=4\} \\ + \beta_{32}j1\{i=2\} + \beta_{33}j1\{i=3\} + \beta_{34}j1\{i=4\}\end{aligned} \tag{7}$$

This model reflects the four modes and four waves represented in the panel data and results in a $16K \times 9$ data matrix, where the rows are made up of the 16 responses from each of the K individuals. The columns contain the response **Y**, seven covariates and one identifying variable used to match rows of information corresponding to the same individual. The number of parameters is $P = 8$, but only seven covariates need to be specified since the usual intercept term, β_0, in the design matrix is automatically included.

Using the mode-wave model specified in (1), the response vector is multivariate $\mathbf{Y} = (\mathbf{Y}_1, ..., \mathbf{Y}_K)'$, where $\mathbf{Y}_k = (Y_{11k}, Y_{12k}, Y_{13k}, Y_{14k}, Y_{21k}, ..., Y_{24k}, ... Y_{44k})'$, with mean $\mu_k = E(\mathbf{Y}_k) = \exp(\mathbf{X}_k'\beta)$ where β is a 8×1 vector of coefficients to be estimated and $\mathbf{X}_k$ is a 16×8 matrix and each row of $\mathbf{X}_k$ is given as

$$X_{ijk} = (1,\ j,\ 1\{i=2\},\ 1\{i=3\},\ 1\{i=4\},\ j1\{i=2\},\ j1\{i=3\},\ j1\{i=4\})', \tag{8}$$

for $i = 1, ..., 4$ and $j = 1, ..., 4$ with $1\{\cdot\}$ denoting the indicator function. To avoid over-parameterizing the model, the coefficients β_{21} and β_{31}, corresponding to the effects associated with mode 1, the reference, are not included. The variance–covariance matrix of **Y** is given by $\mathrm{Cov}(\mathbf{Y}) = \mathbf{H}(\mu)/\phi$, where, for the PSTP model, $\mathbf{H}(\mu) = \mathrm{Diag}(\mu)$.

The analyses can be performed using Splus® (Statsci, 1995) with programming extensions by Carey and McDermott (1995) based on the GEE method of Liang and Zeger and Zeger and Liang. To solve GEE problems approx. 50 megabytes of temporary virtual memory is required. The output consists of estimates of the coefficients, and the robust estimate of the variance of the estimated coefficients used by Liang and Zeger. Since the method is semi-parametric, there are no likelihood functions and goodness-of-fit tests are not available. However, this is not expected to be a problem since the model constraints are relaxed by typical standards.

A continuous wave variable is defined to indicate each wave of travel data. Although the wave variable may be treated as a categorical variable with four levels, it is better used as a continuous variable representing the time factor between the four periods. In this way, trends with time can be described by the effects of the wave variable in the statistical model. The model coefficients β_0 and β_1 can be interpreted as the intercept and slope for the linear relation between wave and $\log(\mu_{ij})$ when $i = 1$, i.e. the mode of travel is by car. In particular, the slope β_1 quantifies the effect on $\log(\mu_{1j})$ for increasing wave numbers. Since the effect of wave on $\log(\mu_{1j})$ is additive, then the effect of wave on μ_{1j} is multiplicative. Consider the relationship between the two models corresponding to two consecutive wave numbers j and $j+1$. Using eqn (7),

$$\begin{aligned}\mu_{1,j+1} &= \exp\{\beta_0 + \beta_1(j+1)\} \\ &= \exp\{\beta_0 + \beta_1 j\}\exp\{\beta_1\} \\ &= \mu_{1j}\exp\{\beta_1\}.\end{aligned}$$

In other words, the average frequency of SOV worktrips changed between any two consecutive waves by a factor of $\exp\{\beta_1\}$. When $i = 2$, 3, or 4, the intercept and slope for the relation between wave and $\log(\mu_{ij})$ are adjusted by the coefficients β_{2i} and β_{3i}, respectively. So, for modes $i = 2$, 3, and 4, the intercept and slope for the linear relation between wave and $\log(\mu_{ij})$ are $\beta_0 + \beta_{2i}$ and $\beta_1 + \beta_{3i}$, respectively; the average frequency of trips made by mode i, $(i \neq 1)$, changed between any two consecutive waves by a factor of $\exp \beta_1 + \beta_{3i}$.

As coefficients for wave-related terms, β_1 and $\beta_1 + \beta_{3i}$ are the primary parameters describing the rates of changes occurring in mode frequencies over wave time. The coefficient β_1 is negative if the mean frequency of mode 1 (SOV) worktrips is decreasing, positive if increasing, and zero if there were no changes. Similarly, $\beta_1 + \beta_{3i}$ is negative if the mean frequency of worktrips for mode i, ($i = 2$, 3, or 4), is decreasing, positive if increasing and zero if there were no changes. Each of the single parameters β_{3i} may be thought of as a measure for the rate of change over time of mode i frequency relative to that of driving alone.

Since the vector of estimates $\hat{\beta}$ is approximately normal with an approximate covariance matrix given by **V**, then confidence intervals may be calculated using the covariances supplied by **V**. For $\text{se}(\hat{\beta}_1) = \sqrt{\hat{\text{var}}(\hat{\beta}_1)}$, where $\hat{\text{var}}(\hat{\beta}_1)$ is the diagonal element of **V** which estimates $\text{var}(\hat{\beta}_1)$, then a $(1-\alpha)$ 100% confidence interval for β_1 is $\left(\hat{\beta}_1 - z_{\alpha/2}\text{se}(\hat{\beta}_1),\ \hat{\beta}_1 + z_{\alpha/2}\text{se}(\hat{\beta}_1)\right)$. Thus, a $(1-\alpha)$ 100% confidence interval for $\exp\beta_1$, the factor by which μ_{1j} changes between consecutive waves, is

$$\left(\exp\left\{\hat{\beta}_1 - z_{\alpha/2}\text{se}(\hat{\beta}_1)\right\}, \exp\left\{\hat{\beta}_1 + z_{\alpha/2}\text{se}(\hat{\beta}_1)\right\}\right).$$

Similar confidence intervals for $\exp\{\beta_1 + \beta_{3i}\}$ are

$$\left(\exp\left\{\hat{\beta}_1 + \hat{\beta}_{3i} - z_{\alpha/2}\text{se}(\hat{\beta}_1 + \hat{\beta}_{3i})\right\}, \exp\left\{\hat{\beta}_1 + \hat{\beta}_{3i} + z_{\alpha/2}\text{se}(\hat{\beta}_1 + \hat{\beta}_{3i})\right\}\right),$$

where,

$$\text{se}(\hat{\beta}_1 + \hat{\beta}_{3i}) = \sqrt{\hat{\text{var}}(\hat{\beta}_1) + \hat{\text{var}}(\hat{\beta}_{3i}) + 2_{\text{cov}}(\hat{\beta}_1, \hat{\beta}_{3i})}.$$

The change in SOV mean worktrip frequency between the fourth and first waves may be expressed as $\exp\{(4-1)\beta_1\}$, and $\exp\{(4-1)(\beta_1 + \beta_{3i})\}$ for modes 2, 3, or 4. The respective confidence intervals are

$$\left(\exp\left\{3\hat{\beta}_1 - z_{\alpha/2}\text{se}(3\hat{\beta}_1)\right\}, \exp\left\{3\hat{\beta}_1 + z_{\alpha/2}\text{se}(3\hat{\beta}_1)\right\}\right)$$

and

$$\left(\exp\left\{3(\hat{\beta}_1 + \hat{\beta}_{3i}) - z_{\alpha/2}\text{se}(3(\hat{\beta}_1 + \hat{\beta}_{3i}))\right\}, \exp\left\{3(\hat{\beta}_1 + \hat{\beta}_{3i}) + z_{\alpha/2}\text{se}(3(\hat{\beta}_1 + \hat{\beta}_{3i}))\right\}\right)$$

where

$$\text{se}\left(3\hat{\beta}_1\right) = \sqrt{3}\text{se}\left(\hat{\beta}_1\right)$$

and

$$\text{se}\left(3\left(\hat{\beta}_1 + \hat{\beta}_{3i}\right)\right) = \sqrt{3}\text{se}\left(\hat{\beta}_1 + \hat{\beta}_{3i}\right).$$

Finally, it is important to also note the limitations of this study. First, survey participants were randomly selected using a stratified sampling protocol based on mode use proportions derived from previous research (for additional details see Murakami and Watterson, 1990). Respondents were recruited using three methods: random telephone digit dialing, contacting prior participants in the Seattle Metro transit surveys, and solicitation of volunteers on randomly selected bus routes. The random telephone digit dialing method was the primary way of collecting participants who drive alone or carpool. The latter two methods target transit users. The sample groups were obtained separately and controlling for the proportion of transit users, therefore, it is not appropriate to use this data to compute regionwide mode proportions. Moreover, to conclude that the results of the analysis applies to the general population, the probability of returning travel diaries in all four waves must be assumed independent of travel behavior.

5. RESULTS

Table 3 presents the results for the model specified in eqn (7). Notice that the robust standard errors from **V** are generally smaller than those obtained from naively assuming an independence correlation structure, allowing detection of more significant effects. Both the wave and the wave by mode3 terms (corresponding to the coefficients β_1 and β_{33}) are significant while the wave by mode2 and wave by mode4 terms (corresponding to the coefficients β_{32} and β_{34}), are not. These results imply that: (1) the mean frequency of single occupancy vehicle (SOV) worktrips is increasing

Table 3.

Variable	Coefficient	Naive SE	Robust SE	t-ratio
Intercept	0.67	0.05	0.04	18.75*
Wave	0.05	0.02	0.01	4.16*
Mode2	−1.81	0.14	0.13	−13.61*
Mode3	−2.05	0.17	0.16	−13.15*
Mode4	−2.67	0.21	0.20	−13.29*
Wave×mode2	0.01	0.05	0.05	0.27
Wave×mode3	−0.15	0.07	0.05	−3.03*
Wave×mode4	−0.03	0.08	0.07	−0.42

*$P \leqslant 0.05$, $n = 8304$.

significantly between waves; (2) the rates at which the mean frequencies of HOV-pools and non-motor worktrips change over time are not statistically different from the rate of increase in the mean frequency of SOV worktrips; and (3) conversely, the rate at which the mean frequency of the HOV-transit worktrips changed over time is statistically different than the rate of change in the frequency of SOV worktrips ($t = -3.0$, $P \leqslant 0.05$). Moreover, this rate of change $\beta_1 + \beta_{33}$ is estimated as -0.1, indicating an overall decrease in the mean frequency over time.

Generally, wave effects are reflected by the t-ratios for the wave-related coefficients, β_1, β_{32}, β_{33} and β_{34}, however, examination of these coefficients alone gives only a partial indication of the overall travel trends. More usefully, the full parameters of the model can be used to estimate mean trip frequencies. Table 4 shows the estimated mean frequencies for the four modes for waves 1 and 4 along with estimates of the rate of increase in terms of percentage increase (PI) and their confidence intervals.

The results are consistent with the coefficients and their t-ratios given in Table 3, but provide additional insight. With 95% confidence, the mean number of worktrips made by driving alone significantly increased from wave 1 to wave 4 (1989–93); the 95% C.I. indicates that the rate of percent increase was between 8.2 and 24.5%. The ranges for rates of change in HOV-pool and non-motor worktrip frequencies overlap with the range for SOV rate of change, corresponding to the insignificant t-ratios for the wave by mode coefficients for these modes; thus, it cannot be said that rate of change for these modes was significantly different from the SOV rates of change. Moreover, the confidence intervals for percent increase in frequencies for HOV-pool and non-motor trips overlap zero, indicating that no significant change in frequencies of trips by these modes was found.

Although these results are inconclusive in determining whether the trip frequencies for these modes changed, with 95% confidence we can say that the mean frequency of the HOV-pool trips did not decrease by more than 7.72% nor increase by more than 58.10%. Likewise, with 95% confidence, the mean frequency of non-motor trips did not decrease by more than 27.11% nor increase by more than 56.35%. Conversely, the rate of change in the mean frequency for HOV-transit is not only below the range for SOV trips (as is consistent with Table 3) but also suggests an estimated overall negative rate of change; with 95% confidence, the rate of percent decrease was between 2.88 and 44.0%.

The generalized linear model used above for the GEE naturally extends to include additional covariates. The design matrix **X** and coefficient vector are both augmented to include the effects represented by the additional terms. Analogous to ordinary regression modeling, when the interest

Table 4.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% (lower)	PI: 95% (upper)
SOV	2.04	2.37	1.16	16.07	8.20	24.51
HOV-pools	0.34	0.41	1.21	20.79	−7.72	58.10
HOV-transit	0.23	0.17	0.74	−26.26	−44.00	−2.88
Non-motor	0.14	0.15	1.07	6.75	−27.11	56.35

The fitted mean frequencies of worktrips per two consecutive weekdays for wave 1 and wave 4, and the percent increase between the two waves.

Table 5.

Variable	Coefficient	Naive SE	Robust SE	t-ratio
Intercept	0.50	0.10	0.07	7.20*
Wave	0.09	0.04	0.03	3.39*
Mode2	−1.77	0.28	0.27	−6.56*
Mode3	−1.63	0.27	0.24	−6.94*
Mode4	−2.66	0.40	0.45	−5.92*
Income	0.24	0.12	0.08	2.85*
Wave×mode2	−0.07	0.11	0.10	−0.67
Wave×mode3	−0.08	0.11	0.08	−1.03
Wave×mode4	−0.02	0.16	0.19	−0.08
Wave×income	−0.06	0.04	0.03	−1.87
Mode2×income	−0.03	0.32	0.31	−0.09
Mode3×income	−0.71	0.35	0.33	−2.18*
Mode4×income	−0.02	0.47	0.51	−0.04
Wave×mode2×income	0.09	0.13	0.11	0.86
Wave×mode3×income	−0.04	0.14	0.11	−0.37
Wave×mode4×income	−0.01	0.18	0.20	−0.07

*$P \leqslant 0.05$, $n = 8304$.

lies in determining whether or not the values of other covariates are associated with changes in mode frequencies by wave, interaction terms are constructed and included in **X** and β.

5.1. Income effects

To test the effects of income, the subjects were identified as having household incomes either greater than or less than $35,000 (the median King County income) and a categorical covariate, *Income*, was added to the model to assess mode frequencies between the two income groups. This model was fitted with all possible interactions between the three factors: *Wave*, *Mode* and *Income*. Table 5 presents the coefficients and their t-ratios.

One striking result is that the two-way interaction between *Wave* and *Mode3* is no longer significant; however, the *Wave* term is still significant. With all other covariates held fixed, the rate at which the mean frequency of HOV-transit worktrips changed over time is not statistically different than the rate of change in the mean frequency of SOV worktrips. In the earlier analysis, without the income variable, the interaction effect reveals an overall trend in the PSTP sample for a tendency to make proportionally more SOV worktrips and fewer HOV-transit worktrips. In this analysis, the effect of the two-way interaction between wave and mode is interpreted in the presence of income. For households with unchanging income status, there is an increasing number of SOV worktrips over time (the *Wave* term is positive and significant) while the rates of change in worktrip frequencies over time observed for the remaining modes are not significantly different from the increase observed for SOV. This can be observed in Tables 6 and 7, where the incomes

Table 6.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% (lower)	PI: 95% (upper)
SOV	1.79	2.34	1.30	30.44	11.87	52.10
HOV-pools	0.28	0.30	1.06	5.68	−41.23	90.06
HOV−transit	0.32	0.33	1.01	1.47	−33.10	53.91
Non-motor	0.12	0.15	1.25	24.63	−57.07	261.79

Table 7.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% (lower)	PI: 95% (upper)
SOV	2.16	2.37	1.10	9.80	0.89	19.51
HOV-pools	0.36	0.43	1.18	18.04	−12.00	58.35
HOV-transit	0.18	0.14	0.75	−24.81	−49.86	12.76
Non-motor	0.14	0.14	1.01	0.67	−33.17	51.64

Table 8.

Lifecycle group	Sample size
(1) Any child, less than 6 yr	1284
(2) All children, between 6–17 yr	2516
(3) 1 adult, less than 35 yr	156
(4) 1 adult, 35–64 yr	704
(5) 1 adult, 65 yr or greater	52
(6) 2 or more adults, less than 35 yr	204
(7) 2 or more adults, 35–64 yr	3192
(8) 2 or more adults, 65 yr or greater	196

are examined separately,* by noticing that the confidence intervals for percentage increase for mode 1 are greater than zero but also overlap with each of the confidence intervals of the other modes, which are not statistically different from zero. The tables also suggest that respondents with higher incomes generally make more worktrips by all modes except for HOV-transit, in which they make considerably fewer trips.

The statistical significance of the two-way interaction between mode3 and income hints at the difference between the two analyses with regard to the changes in HOV-transit use over time. The analysis indicates that HOV-transit mode use frequencies differ between the income groups; as might be expected, respondents with household incomes greater than $35,000 make fewer worktrips by HOV-transit. Since, by definition, we are looking at the same households over each of the four waves, this implies a possible increase in the number of households earning greater than $35,000 vice versa. Goulias and Ma (1996) have shown that PSTP household incomes increased between 1989–93. As respondents moved from one income group to the next, their use of both HOV-transit and SOV modes for worktrips changed. For respondents remaining in the same income group for each of the four waves, only SOV mode use increased significantly.

Including income and all possible interactions in one model effectively results in a separate model for each of the two groups (i.e. each of the *Intercept*, *Wave*, *Mode*, and *Wave by Mode* terms are all adjusted by the indicator variable, *Income*). Alternatively, the analysis was performed using separate income groups, with two models estimated. Both analyses are similar in that they estimate the parameters of the same pairs of models. However, in the former (aggregated) analysis, the covariance is assumed homogeneous over the entire dataset. In the latter (separated) analysis, this assumption is not necessary since it fits separate covariances for each income group.

5.2. Lifecycle type effects

The analysis was repeated with the covariate household type (lifecycle group). The eight groups lifecycle types are listed in Table 8.

To study the effect of *Wave* for each group in an aggregated analysis involves a model with 64 coefficients and creates computational difficulties. As an alternative, the analysis was run separately for the eight household lifecycle groups. The total sample was divided into eight subsamples by lifecycle group using the same analysis technique as noted for income. Results for household types 5 and 8, consisting of adults older than 65 years, failed to converge. These households tend to make fewer worktrips and accordingly, have estimated means close to zero; Poisson data close to zero have variances approaching zero and consequently, not infrequently a near-singular covariance matrix.

The remaining lifecycle groups reveal heterogeneous travel patterns. For lifecycle 1 (households with a child less than 6 years old), there was a significant increase only in the mean HOV-pool use (Table 10), increasing at a rate between 5.19 and 225.93%. The large variances in this subgroup might be attributed to within group heterogeneity. There appears to have been an increase in use of HOV-transit as well (~53%), but this increase is not statistically significant at the type I error level of 0.05.

*To examine incomes separately, the sample is divided according to income classifications for analyses. For those subjects whose income classification changed during the four waves, the portion of their data vector corresponding to the period in which they made less than $35,000 was included in the analysis for incomes less than $35,000, and the remaining portion was included in the over $35,000.

Table 9.

Variables	Coefficients	Naive SE	Robust SE	t-ratio
Intercept	0.75	0.12	0.09	8.76*
Wave	0.05	0.04	0.04	1.10
Mode2	−2.25	0.33	0.34	−6.58*
Mode3	−2.88	0.47	0.46	−6.28*
Mode4	−2.57	0.48	0.49	−5.21*
Wave×mode2	0.16	0.12	0.11	1.41
Wave×mode3	0.09	0.17	0.13	0.70
Wave×mode4	−0.11	0.19	0.20	−0.55

*$P \leqslant 0.05$.

Table 10.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% lower bound	PI: 95% upper bound
SOV	2.22	2.57	1.15	15.47	−10.70	49.32
HOV-pools	0.27	0.51	1.85	85.16	5.19	225.93
HOV-transit	0.14	0.21	1.53	53.27	−25.14	213.79
Non-motor	0.15	0.13	0.83	−17.15	−72.76	152.03

For lifecycle 2, households with all children between the ages of 6 and 17, the only significant change was an increase in SOV worktrips of a rate between 6.23% and 39.26% (Table 12). Mean worktrip frequencies for the other modes did not significantly change.

Households with one adult less than 35 years old, lifecycle type 3, had a tendency towards reduced use of non-motor vehicles for worktrips with a decrease of 87.97–95.41%. There were no significant changes with the other mode frequencies although there was a marginally significant reduction in HOV-pools (a 90% CI for the percentage of decrease of worktrips by HOV-pools would be 17.10–88.8%).

For lifecycle group 4, households with 1 adult between 35 and 64 years old, no significant wave effects were found. This was a moderately sized sample at 704, much larger than the groups of 156 and 204 represented by two other groups in this study, both of which had a sufficiently large

Table 11.

Variables	Coefficients	Naive SE	Robust SE	t-ratio
Intercept	0.62	0.01	0.07	8.68*
Wave	0.07	0.03	0.02	2.84*
Mode2	−1.58	0.25	0.25	−6.41*
Mode3	−2.20	0.33	0.33	−6.67*
Mode4	−2.91	0.41	0.54	−5.36*
Wave×mode2	−0.09	0.09	0.09	−0.99
Wave×mode3	−0.12	0.12	0.11	−1.01
wave×mode4	0.04	0.14	0.19	0.20

*$P \leqslant 0.05$.

Table 12.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% lower bound	PI: 95% upper bound
Car (alone)	1.98	2.41	1.22	21.63	6.23	39.26
HOV-pools	0.37	0.35	0.93	−7.50	−44.75	54.86
HOV-transit	0.20	0.17	0.86	−14.16	−54.27	61.11
Non-motor	0.11	0.15	1.36	−36.23	−53.01	295.97

Table 13.

Variables	Coefficients	Naive SE	Robust SE	*t*-ratio
Intercept	0.67	0.29	0.17	3.94*
Wave	0.00	0.11	0.06	0.05
Mode2	−1.57	0.94	0.76	−2.06*
Mode3	−2.19	1.11	1.65	−1.33
Mode4	−2.04	1.72	0.99	−2.07*
Wave×mode2	−0.40	0.44	0.23	−1.71
Wave×mode3	−0.26	0.49	0.53	−0.48
Wave×mode4	−0.87	1.01	0.09	−9.38*

*$P \leqslant 0.05$.

Table 14.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% lower bound	PI: 95% upper bound
Car (alone)	1.97	1.98	1.01	0.94	−29.48	44.48
HOV-pools	0.27	0.08	0.30	−69.53	−90.73	0.15
HOV-transit	0.17	0.08	0.47	−53.09	−97.68	850.12
Non-motor	0.11	0.01	0.07	−92.57	−95.41	−87.97

Table 15.

Variables	Coefficients	Naive SE	Robust SE	*t*-ratio
Intercept	0.59	0.18	0.13	4.45*
Wave	0.05	0.06	0.04	1.36
Mode2	−2.18	0.58	0.45	−4.87*
Mode3	−1.11	0.40	0.44	−2.51*
Mode4	−3.03	0.70	0.69	−4.39*
Wave×mode2	−0.03	0.20	0.16	−0.18
Wave×mode3	−0.17	0.15	0.14	−1.21
Wave×mode4	0.24	0.22	0.20	1.16

*$P \leqslant 0.05$.

Table 16.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% l. bound	PI: 95% u. bound
Car (alone)	1.90	2.23	1.18	17.84	−7.05	49.40
HOV-pools	0.21	0.22	1.08	7.70	−56.34	165.70
HOV-transit	0.53	0.38	0.72	−28.45	−62.93	38.11
Non-motor	0.12	0.28	2.41	140.78	−26.73	691.20

Table 17.

Variables	Coefficients	Naive SE	Robust SE	*t*-ratio
Intercept	0.42	0.24	0.19	2.18*
Wave	0.22	0.10	0.07	3.31*
Mode2	−1.56	0.60	0.58	−2.71*
Mode3	−2.37	1.15	0.88	−2.69*
Mode4	−2.06	0.92	0.77	−2.66*
Wave×mode2	−0.07	0.25	0.26	−0.27
Wave×mode3	−0.41	0.56	0.18	−2.32*
Wave×mode4	−0.33	0.43	0.19	−1.68

*$P \leqslant 0.05$.

Table 18.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% lower bound	PI: 95% upper bound
Car (alone)	1.89	3.61	1.92	91.50	30.34	181.36
HOV-pools	0.37	0.58	1.56	56.99	−56.07	453.93
HOV-transit	0.12	0.06	0.56	−44.30	−78.11	41.69
Non-motor	0.17	0.12	0.72	−28.12	−75.60	111.77

Table 19.

Variables	Coefficients	Naive SE	Robust SE	*t*-ratio
Intercept	0.67	0.09	0.07	9.93*
Wave	0.03	0.03	0.02	1.42
Mode2	−1.70	0.22	0.20	−8.41*
Mode3	−1.74	0.28	0.26	−6.72*
Mode4	−2.42	0.33	0.30	−8.10*
Wave×mode2	0.03	0.08	0.08	0.43
Wave×mode3	−0.30	0.11	0.09	−3.29*
Wave×mode4	−0.09	0.12	0.11	−0.82

*$P \leq 0.05$.

Table 20.

	Wave 1 frequency	Wave 4 frequency	Increase factor	Percent increase	PI: 95% lower bound	PI: 95% upper bound
Car (alone)	2.03	2.24	1.11	10.75	−3.80	27.51
HOV−pools	0.38	0.47	1.23	22.66	−19.65	87.26
HOV−transit	0.26	0.12	0.45	−55.25	−72.76	−26.48
Non-motor	0.16	0.14	0.85	−15.21	−53.82	55.67

enough sample size to detect changes between waves. Since the standard errors are relatively large, this suggests the group was very heterogeneous in their travel behavior (Tables 15 and 16).

For lifecycle group 6, households with more than two or more adults less than 35 yr, there was an increased use of worktrips made by driving alone with the percentage increase estimated to be between 30.34 and 181.36% (Table 18). The other modes did not have significantly different worktrip frequencies over the waves.

For lifecycle group 7, households with more than one adult between 35 and 64 yr old, there was a significant decrease, between 26.48 and 72.76%, in use of HOV-transit. The other mode frequencies show no statistically significant change (Table 20).

6. CONCLUSIONS

In conclusion, there was an increase in the number of worktrips made between 1989 and 1993 and these were marked by increasing use of single occupancy vehicles. HOV-transit mode use simultaneously declined, an effect apparently associated with the increasing income of the population. The two income groups defined by those making greater or less than $35,000 exhibited comparable mode use behavior with the exception of HOV-transit. Particularly, people in both income brackets tended to make increasingly more SOV trips as wave increased, however, people in households with the higher income bracket made fewer worktrips overall by HOV-transit than individuals in household's in the lower income bracket. When considering the factor by lifecycle group, it was found that different household types varied. The following summarizes the significant changes with wave: households having all children between 6 and 17 or households having two or more adults less than age 35 made increasingly more worktrips by car over time; households with a child less than 6 years old made increasingly more trips by HOV-pool over time;

households with more than one adult between 35 and 64 yr made increasingly fewer worktrips by HOV-transit over time; and households with one adult less than 35 years made increasingly fewer non-motor worktrips over time.

The full parameters of the model were also used to estimate the rate of increase in terms of PI and their confidence intervals. These results suggest that the mean number of SOV worktrips significantly increased from wave 1 to wave 4 with a 95% C.I. for the rate of percent increase of between 8.2 and 24.5%. The rates of change in HOV-pool and non-motor worktrip frequencies overlap with the range for SOV rate of change, and thus, it cannot be said that rate of change for these modes was significantly different from the SOV rates of change. The rate of change in the mean frequency for HOV-transit is not only below the range for SOV trips but also suggests, with 95% confidence, the rate of percent decrease was between 2.88 and 44.0%.

REFERENCES

Carey, V. and McDermott, A. (1995) Generalized linear models for clustered data via generalized estimating equations (GEE), [computer code for Splus software (v3.1)], Channing Laboratory, Harvard Medical School.

Golob, T. and Guiliano, G. (1989) A simultaneous equation model of employee attitudes to a staggered work hours demonstration project. Working paper UCI-ITS-WP-89-1, Institute of Transportation Studies, University of California, Irvine, CA.

Golob, T. (1990) The dynamics of household travel time expenditures and car ownership decisions. *Transportation Research A* **24**, 443–464.

Goulias, K. G. and Ma, J. (1996) Analysis of longitudinal data from the Puget Sound Transportation Panel. Task report, Pennsylvania Transportation Institute, Pennsylvania State University.

Hu, P. and Young, J. (1993) 1990 *NPTS Databook, Nationwide Personal Transportation Survey, FHWA-PL-94-010A*. Office of Highway Information Management, Federal Highway Administration, HPM-40, Washington, DC.

Kitamura, R. (1990) Panel analysis in transportation planning: an overview. *Transportation Research A* **24**, 401–515.

Kitamura, R., Goulias, K., and Pendyala, R. (1990) Telecommuting and travel demand: An impact assessment for the State of California Telecommute Pilot Project Participants. Research report UCD-TRG-RR-90-8, Institute of Transportation Studies, University of California, Davis, CA.

Liang, K. Y. and Zeger, S. L. (1986) Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Maddala, G. (1987) Recent developments in the econometrics of panel data analysis. *Transportation Research A* **21**(4/5)), 303–326.

McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*, 2nd edn., Chapman and Hall, New York.

Meurs, H. (1990) Dynamic analysis of trip generation. *Transportation Research A* **24**, 427–442.

Murakami, E. and Watterson, T. (1990) Developing a household travel panel survey for the Puget Sound Region. *Transportation Research Record*, 1285, 40–46.

Rossetti, M. and Eversole, B. (1993) *Journey to Work Trends in the United States and its Major Metropolitan Areas, HW362*. Office of Highway Information Management, Federal Highway Administration, HPM-40, Washington, DC.

Statsci (1995) *S-PLUS Programmers Manual*. Mathsoft, Inc., Seattle, WA.

Van Wissen, L. and Meurs, H. (1989) The Dutch mobility panel: experiences and evaluation. *Transportation* **16**, 99–119.

Zeger, S. L. and Liang, K. Y. (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121–1302.