

**A STRUCTURAL AND EMPIRICAL MODEL OF
SUBSISTENCE ACTIVITY BEHAVIOR AND INCOME**

Chandra R. Bhat
Research Assistant Professor
Transportation Center
Northwestern University
Evanston, Illinois 60208

and

Frank S. Koppelman
Professor of Civil Engineering
and Transportation
Northwestern University
Evanston, Illinois 60208

Abstract

This paper develops a structural and empirical model of subsistence activity behavior and income. Subsistence activity decisions (work participation and hours of work decisions) and income have an important bearing on activity and travel behavior of individuals. The proposed structural model represents an effort to analyze subsistence activity behavior and income earnings to support a better understanding, and reliable forecasting, of individual travel behavior. The empirical model formulates and estimates an integrated model of employment, hours of work and income which takes account of interdependencies among these choices and their structural relationships with other relevant variables. Social factors that inhibit an individual's employment and work hours decision and affect an individual's income are incorporated in the model. A sample of households from the Dutch National Mobility Panel is used in the empirical analysis.

1. Introduction

Activity-based analysis has been the focus of many research studies in recent years. Very broadly, activity-based analyses attempt to obtain a better understanding of the behavioral basis for individual decisions regarding participation in activities in certain places at given times. The specific application of the activity-based approach to travel analysis is referred to as activity-based travel analysis (the reader is referred to Kitamura 1988a, Jones et al. 1990 or Bhat 1991 for a recent review of activity-based travel studies). The basis of this approach is that travel is the result of movement of individuals among locations to pursue activities scattered in space (Oi & Shuldiner 1962). Hence, by understanding the need to participate in activities, improved knowledge of travel can be obtained. In contrast to traditional trip-based analyses which directly model travel behavior, activity-based analyses place primary emphasis on modeling activity behavior.

While there has been considerable work on activity-based travel analysis, most of this work has focused on the spatial and temporal linkage of activities; that is, the scheduling of activities (see for example, Clarke 1986; Kitamura & Kermanshah 1983; Recker et al. 1986). The agenda of activities for participation, and one or more associated characteristics of the participation, are considered as predetermined. The few studies which have focused on activity agenda determination (Damm 1980; Van der Hoorn 1983; Hirsh et al. 1986) do not consider the influence of household needs and the complex interactions among household members on individual activity generation.

Bhat & Koppelman (1992) developed a comprehensive framework of individual activity agenda generation. They view an activity agenda as comprising a list of activities (that will be participated-in over a particular time period) along with the attributes of frequency, duration, destination of activity performance, mode to destination and time-window for participation. Activity scheduling is viewed as the appropriate sequencing of activities within the activity agenda and the determinant of the precise temporal dimension of activity participation. Activity agenda generation and activity scheduling are intricately linked. Activity participation and attributes of the participation will depend on the convenience in, and costs of, sequencing activities. The effect of scheduling opportunities on activity agenda generation is represented in the form of composite measures such as accessibility and density of opportunity to pursue activities (which may be viewed as surrogate measures of convenience and opportunity for activity sequencing). Such a structure assumes that detailed activity sequencing and activity chaining issues are not considered in individual decisions regarding participation (and accompanying characteristics of this participation). The resultant

activity agenda is subsequently subjected to detailed sequencing to form a satisfactory travel-activity pattern.

The individual activity agenda generation process is divided into four main modules in Bhat and Koppelman's framework. The first module, the household needs module, involves development of the household subsistence activity patterns (comprising subsistence activity patterns of each individual member and the resulting individual income earnings) and generation of household maintenance needs.¹ The second module is the household auto-ownership model. The subsistence activity block of the household needs module and the auto ownership module influence the third module that pertains to the allocation of automobile(s) and household maintenance activities among household members. After the allocation process, the individual plans on how he or she should fulfill the assigned out-of-home maintenance activities. Simultaneously, decisions on participation (and attributes of the participation) in leisure activities are made. This planning and decision mechanism forms the basis for "constructing" the overall individual activity agenda (which can then be processed using existing activity scheduling models to develop individual travel-activity patterns) and is the focus of the fourth module, the programming module. The current paper focuses on the subsistence activity component of the household needs module.

Subsistence activity behavior refers to two inter-related decisions in this paper -- the work participation or employment decision and the hours of market work decision (or individual labor supply choice).

Subsistence activity and income earnings have a considerable influence on overall activity behavior. It is well established that participation in non-work activities is contingent on time availability after fulfillment of work activity and is scheduled around the more structured and rigid work activity (Kitamura 1984; Clarke 1986). Individuals' participation, and amount of participation, in work affects the allocation of obligatory household activities among household members. The income that an individual brings into the household (relative to the other members) may be viewed as a measure of the "bargaining power" of that individual in the household and also affects

¹ The classification of household needs into subsistence and maintenance needs is based on the activity typology employed by Reichman 1976. According to this classification, subsistence activity refers to the supply of work or work-related business. Maintenance activity pertains to the purchase and consumption of goods and services to satisfy household or personal physiological and biological needs and includes grocery shopping, personal business such as banking, going to the post office etc., and household work.

household activity allocation among members.² Finally, household income determines the potential of a family to consume goods and leisure and consequently determines the financial potential for non-work activity participation.

Subsistence activity and income are clearly important variables in activity analysis. The objective of this paper is to develop a model that facilitates a better understanding of the factors affecting subsistence activity decisions and income earnings. Such a model will constitute an important component of the activity-based forecasting system. The next section of the paper discusses the data source and sample used in the empirical analysis. The third section advances the model system of subsistence activity behavior and income and presents the methodology to estimate model parameters. The fourth section presents the empirical specification and discusses the results of the model. Important conclusions are summarized in section 5.

2. Data Source and Sample Formation

The data source used in the present study is the Dutch National Mobility Panel. This panel was instituted in 1984, and involves weekly travel diaries and household and personal questionnaires collected at biannual and annual intervals. Ten waves (a wave refers to cross-sectional data at one time point) were collected between March 1984 and March 1989. A stratified sampling scheme was adopted to ensure adequate households in policy relevant subpopulations (van Wissen & Meurs 1989). Additional households were included to replace households which dropped out of the study in an intermediate wave. This replacement was determined by appropriate refreshment techniques to preserve the representativeness of the sample. Each wave consists of about 1800 households sampled according to category of municipality, household income, and household composition.

Data for our analysis was obtained from waves 1, 3, 5, 7 and 9 of the panel collected during the spring of each year between (and including) 1984 and 1988. The data was screened to include only couple or nuclear family households³ in which the husband is employed. We removed households in which the husband was unemployed because there were too few of them to undertake any meaningful analysis of husband's employment. Households in which adults are self-employed

² This relationship between resources of an individual and her/his bargaining power originates in the resource theory of sociology (Blood & Wolfe 1969; Heer 1963).

³ A nuclear family household comprises two adults, a male and a female, with one or more children below the age of 18. A couple household comprises two adults of opposite sex.

were excluded because the concept of income is not clearly defined for such individuals. Households with seniors over 60 years and/or disabled persons were removed from the sample due to their low rate of employment. The resulting sample, which includes 2279 observations of nuclear and couple family households, was used in the analysis.

3. Model System

The model system is developed for couple or nuclear family households in which the husband is employed. Since we found that a majority of husbands (> 95%) work on a full-time basis, we focus on the wife's subsistence activity behavior in this paper.

The endogenous variables in the model system are husband's income, wife's employment choice, wife's hours of work, and wife's income. In this section, we develop the simultaneous equation system of the model and also present the econometric procedure used in estimation. The simultaneous equation system accommodates the qualitative and limited-dependent nature of the endogenous variables.⁴ In principle, this simultaneous model system can be estimated by full-information maximum likelihood methods; that is, the likelihood function corresponding to the complete system may be explicitly derived and maximized with respect to all the unknown parameters. However, the joint distribution of the random variables of the system involves a four-dimensional, multivariate and multi-truncated normal distribution, where each of the variables is a function of many unknown parameters. Maximizing this likelihood function is extremely difficult. In addition, it is doubtful that such full-information maximum likelihood estimates, which are asymptotically fully efficient if the model is specified correctly, are sufficiently robust against various misspecifications of the model, such as variable exclusions and nonnormality (Greene 1990; Hanoch 1980). We adopt a limited-information maximum likelihood estimation method in this paper. The limited information method, though not efficient, is computationally simpler and provides consistent estimates for all the model parameters. The procedure will also, by and large,

⁴ Qualitative variables refer to dummy variables which can take only one of two discrete values. Limited dependent variables refer to variables which are limited in their range because of some underlying stochastic choice mechanism (Maddala 1983). In the model, wife's employment is a qualitative variable, while the remaining endogenous variables are limited-dependent variables. Earlier econometric studies of equation systems with qualitative and limited dependent endogenous variables include Lee (1981), Heckman (1978) and Amemiya (1979). Applications of such systems in the transportation field include Golob & Meurs (1988), Kitamura (1988b) and van Wissen (1989).

confine the effect of any specification errors to the particular equation in which it appears and thus is more robust to misspecifications.

In the limited information procedure, the husband's income and wife's employment equations are estimated individually. The husband's income variable occurring on the right hand side of other equations is replaced by an imputed value obtained from the estimation of the husband's income equation (this imputed value is an unbiased estimator of the actual value). The wife's hours of work and income equations are estimated in combination with the wife's employment equation to account for the censored nature of these endogenous variables based on wife's employment.⁵

In the following sections, we discuss the structure and estimation of each equation in the system. We use the subscript i to denote observations (or households).

3.1. Husband's Income

The first equation in the structural system is husband's income. We assume a suitable monotonic transformation of income so that husband's income may be expressed as a linear function of independent variables. Two issues arise at this point. One, the selection of the monotonic transformation, and two, the grouped nature of income (that is, data being recorded in categories rather than on a continuous scale).

The class of monotonic transformations of income is restricted to "power transformations" as suggested by Box & Cox (1964). In practice, the empirical transformation that reasonably suits income data is the natural logarithm (Heckman 1974). An extensive treatment of the appropriateness of the log transformation for income may be found in Mincer (1974).

The grouped nature of income is addressed by defining a continuous index function (also referred to as a latent function) for the logarithm of husband's income, I_{hi}^* . We do not observe I_{hi}^* but observe that I_{hi}^* falls into a certain interval. The first equation of our system is then written as:

$$I_{hi}^* = \pi_h' X_{hi} + v_{hi}$$

$$I_{hi} = j, \text{ if } \frac{a_{j-1}}{p_i} < I_{hi}^* \leq \frac{a_j}{p_i}, j = 1, 2, \dots, J \quad (1)$$

⁵ Wife's hours of work and income are censored variables because they are observed only if the wife is employed.

where v_{hi} is a normal random error term with mean 0 and variance σ_h , X_{hi} is a vector of exogenous variables affecting husband's (log) income and π_h is a corresponding vector of parameters. The a_j 's represent known threshold values for each income category j . These thresholds are normalized by the price index p_i to obtain the equivalent real-income censoring bounds. The J income intervals exhaust the real line and hence we assume $a_0/p_i = -\infty$ and $a_J/p_i = +\infty$. Representing the cumulative standard normal by Φ , the probability that husband's income falls in category j may be written as

$$\text{prob}(I_{hi} = j) = \Phi\left(\frac{A_{j,i} - \pi_h' X_{hi}}{\sigma_h}\right) - \Phi\left(\frac{A_{j-1,i} - \pi_h' X_{hi}}{\sigma_h}\right), \quad (2)$$

where $A_{j,i}$ is the upper real-income censoring bound for category j and individual i .

Defining a set of dummy variables

$$M_{ij} = \begin{cases} 1 & \text{if } I_{hi} \text{ falls in the } j\text{th category} \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

$(i = 1, 2, \dots, N, j = 1, 2, \dots, J)$

the likelihood function for estimation of the parameters π_h and σ_h is

$$L_h = \prod_{i=1}^N \prod_{j=1}^J \left[\Phi\left(\frac{A_{j,i} - \pi_h' X_{hi}}{\sigma_h}\right) - \Phi\left(\frac{A_{j-1,i} - \pi_h' X_{hi}}{\sigma_h}\right) \right]^{M_{ij}}. \quad (4)$$

Initial parameter values for the maximum likelihood search are obtained by assigning to each income observation, its conditional expectation based on the marginal distribution of I_{hi}^* and regressing these conditional expectations on the vector of exogenous variables.

An imputed value for husband's (log) income is computed from equation (1) as $\hat{I}_{hi}^* = \hat{\pi}_h' X_{hi}$ and is used for husband's (log) income in subsequent equations.

3.2. Wife's Employment

The second equation in our model system is the wife's employment decision. Wife's employment choice is a function of exogenous variables and household assets or unearned income. In our model, husband's (log) income is treated as unearned income to the wife; that is, the wife regards her husband as an "income producing asset" which affects her work decision (Cogan 1980).

We define a latent continuous function E_i^* denoting the wife's employment propensity and view the discrete employment decision E_i as a reflection of this underlying propensity. If this propensity exceeds zero, the wife will work. Otherwise, she will not work. We may write the relationship between the latent employment propensity and the discrete employment decision as follows:

$$\begin{aligned} E_i^* &= \pi_e' X_{ei} + \gamma_e \hat{I}_{hi} + v_{ei} \\ E_i &= 1 \quad \text{if } E_i^* > 0 \\ E_i &= 0 \quad \text{if } E_i^* \leq 0 \end{aligned} \tag{5}$$

where the vector X_{ei} represents a vector of exogenous variables affecting wife's employment. We assume a normal distribution for the random error term v_{ei} with mean zero and unit variance. This is a familiar probit model. The parameters π_e and γ_e are estimated using a univariate probit procedure.

3.3. Wife's Hours of Work

The wife's hours of work equation is conditional on the individual being employed. We use a logarithm transformation for wife's hours of work in our model. In equation form, we write the wife's hours of work (or labor supply) equation as:

$$L_i = \pi_l' X_{li} + \gamma_l \hat{I}_{hi} + v_{li}, \text{ observed only if } E_f = 1. \tag{6}$$

where L_i is the wife's (log) hours of work, X_{li} represents a vector of exogenous variables affecting wife's hours of work, π_l (a vector) and γ_l (a scalar) are parameters to be estimated; v_{li} is a normal random error term with mean zero and variance σ_l . Husband's (log) income appears in equation (6) because it is expected to have a negative effect on wife's work hours due to the positive effect of an increase in unearned income on wife's leisure (Killingsworth 1983).

Limiting our attention to employed wives and estimating a simple regression to estimate the parameters in equation (6) is subject to problems of selection bias (Heckman 1976).⁶ Appropriate estimation procedures for obtaining the hours of work parameters will account for the possible correlation between the error terms in the employment equation and the hours of work equation. Using standard results of truncated bivariate normal distributions (Johnson & Kotz 1972), the parameters of the hours of work equation and the correlation term can be estimated by maximum likelihood methods (the reader is referred to Heckman 1979 or Amemiya 1985 for the maximum likelihood expression).⁷ The initial values for the maximum likelihood estimation are obtained by using the Heckman two-step procedure for sample selection models (Heckman 1979).

3.4. Wife's Income

Wife's income is conditional on her employment status. In addition, it is available only in grouped form. Defining the wife's (log) income as I_{wi}^* and the observed categorical wife's income data as I_{wi} , we write

$$\left. \begin{aligned} I_{wi}^* &= \pi_w' X_{wi} + \gamma_w \hat{I}_{hi} + v_{wi} \\ I_{wi} &= k, \text{ if } \frac{d_{k-1}}{p_i} < I_{wi}^* \leq \frac{d_k}{p_i} \end{aligned} \right\} \text{observed only if } E_i^* > 0 \quad (7)$$

where k is an index for categories ($k=1,2,\dots,K$), d_k represents the thresholds of absolute income and p_i is the price index.⁸ The variable vector X_{wi} contains exogenous variables affecting wife's income and v_{wi} is a normal random error term with mean 0 and variance $\sigma_w \cdot \pi_w$ and γ_w are parameters to be estimated.

⁶ The selection bias occurs due to correlation in unmeasured factors that affect employment status and hours of work. Individuals who are highly motivated and active, may be more likely to work and may also work many hours.

⁷ Maximization of the likelihood function provides estimates of both the hours of work equation and the employment equation. The employment equation has been estimated earlier, and hence we have more than one set of estimates of wife's employment. In fact maximum likelihood estimation of the wife's income equation will provide an additional estimate of parameters in the employment equation. Consequently, there is a multiplicity of employment estimates. All these estimates are consistent and were found to be very close empirically. In our estimation, we will use direct employment probit estimates as the parameters in the employment equation.

⁸ Husband's income affects wife's hours of work. Since wife's income is related to her hours of work, husband's income appears in equation (7).

Wife's income (in log form) is a censored grouped variable (the censoring based on employment). Limiting our attention to observations in the uncensored portion and estimating parameters by a grouped data method similar to the one employed for husband's income equation is subject to problems of selection bias. We overcome this by estimating wife's income jointly with wife's employment. Assuming a bivariate normal distribution between the conditional distributions of the underlying latent wife's employment and income functions, and defining the following parameters and variable vectors

$$\tilde{\pi}_e = (\pi'_e, \gamma'_e), \tilde{\pi}_w = (\pi'_w, \gamma'_w)', \tilde{X}_{ei} = (\hat{I}_{hi}^*)', \text{ and } \tilde{X}_{wi} = (X'_{wi}, \hat{I}_{hi}^*)'$$

the appropriate maximum likelihood procedure for estimation of the parameters is as shown in the following equation.⁹

$$\begin{aligned} \langle f \rangle = & \prod_{i=1}^N [-\Phi(\tilde{\pi}'_e \tilde{X}_{ei})]^{-E_i} \times \\ & \left[\prod_{l=1}^L \left\{ \Phi_2 \left(\frac{D_{k,i} - \tilde{\pi}'_w \tilde{X}_{wi}}{\sigma_w}, \tilde{\pi}'_e \tilde{X}_{ei}, -\rho_{ew} \right) - \Phi_2 \left(\frac{D_{k-1,i} - \tilde{\pi}'_w \tilde{X}_{wi}}{\sigma_w}, \tilde{\pi}'_e \tilde{X}_{ei}, -\rho_{ew} \right) \right\}^{t_{ik}} \right]^{E_i} \end{aligned} \quad (8)$$

where ρ_{ew} is the correlation between the error terms v_e and v_w in wife's employment and income equations respectively, $D_{k,i} = d_k / p_i$ represents the real income thresholds associated with each income category k and observation i , Φ_2 is the cumulative standard bivariate normal function, and T_{ik} is defined as follows:

$$T_{ik} = \begin{cases} 1 & \text{if } I_{wi} \text{ falls in the } k\text{th category} \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

$(i = 1, 2, \dots, N, k = 1, 2, \dots, K).$

Initial parameter values are obtained by a modification of the procedure adopted for husband's income estimation. We assign to each observation in the uncensored region, its conditional expectation based on the marginal distribution of the underlying latent continuous variable I_{wi}^* . We

⁹ We are not aware of any application of this variant of sample selection in econometric literature.

treat these values as the actual continuous income values and apply a Heckman's two step method for sample selection models to obtain start values for the parameters.

4. Empirical Specification and Estimation Results

The choice of variables and the specification adopted in the model was guided by conceptual arguments, empirical evidence provided by earlier labor economic studies, and considerations of parsimony in representation. Table 1 provides a list of exogenous variables used in the model and their definitions. The husband's age variables enable non-linear estimation of the age effect on husband's earning. The wife's age variables facilitate estimation of non-linear effects of wife's age on employment, hours of work and income. The education variables indicate the effect of different levels of education (secondary or high) relative to primary education. Work acceptability is the ratio of total female labor force (that is, all females who are employed or not employed but seeking jobs) to total active female population in each municipality.¹⁰ It represents the degree to which wife's employment is considered acceptable or appropriate in each community.¹¹ Wave dummy variables are introduced in the specification to capture temporal variations of the endogenous variables across all households. Such temporal variations may arise from differences in the state of the economy, e.g., changes in costs of living and/or absolute income earnings.

Price levels are assumed to be constant across regions in this analysis. The Netherlands is a small country and it may not be unreasonable to assume constant price levels in such a compact geographic area (Killingsworth 1983). Thus, variations in the price index arise in this study from time series or wave differences in price level.

The estimation results for each equation are presented and discussed in the following sections. The total number of observations is 2279 in all estimations (571 of the 2279 households have an employed wife).

¹⁰ The data used in the computation of this index was obtained from the Central Bureau of Statistics (CBS), Netherlands.

¹¹ We recognize alternative interpretations of the work acceptability index which may represent a combination of location attributes. Viewed from this perspective, the index may be considered as a parsimonious representation for the set of local factors influencing wife's work participation.

4.1. Husband's Income Equation

The unit of measure used for the husband's income is real annual income in guilders per year. The variables included in the husband's income equation are: husband's age (with splined points at the age of 35 and 45 to capture the potential non-linear impact of age on income),¹² husband's education (introduced as two dummy variables corresponding to secondary and high education levels with primary education as the base), an interaction term of secondary/high education and age beyond 45 to reflect variation in age effects between low education and medium to high education levels,¹³ and wave dummy variables (to represent other differences by time period).

The results of the grouped data estimation of husband's income in log form are shown in Table 2a. All but the wave variables are statistically significant. Age has a positive impact on husband's income presumably because it is a proxy for experience; however, there is a decline in the age effect beyond age 35 from +0.025 to +0.010. This decline in the marginal age effect may be attributed to decreasing returns to scale of experience and/or a deterioration in efficiency and productivity (Mincer 1974). The effect of age beyond 45 is complicated by the interaction of secondary/high education (that is, education level greater than primary level) and age above 45. For individuals with a low education, (log) income decreases beyond the age of 45 at a rate of -0.011 ($=0.025-0.015-0.021$). However, for individuals with medium to high education, the net effect is near zero ($0.025-0.015-0.021+0.009$). These results support the anticipated differential effect of age on productivity based on education level. There is a strong positive influence of the education dummy variables on husband's income. Secondary and high education levels increase husband's income (relative to primary education), with high education having a greater positive influence than secondary education. Further, the age-education interaction reflects an increasing difference between primary education level and higher education levels beyond 45 years.

The marginal effects of age and education level on husband's income (computed for mean variables values) shown in Table 2b highlight the differential importance of these variables.

¹² Analysis using four spline points at the ages of 30, 35, 40, and 45 was not found to be significantly better than the specification with two spline points at 35 and 45.

¹³ No significant age-education interaction was found in the 35-45 years range.

4.2. Wife's Employment Equation

The exogenous variables in the wife's employment equation include a dummy variable for husband's high education, wife's age and age over 40, wife's education level, total number of children and number of children less than 12 years in the household, work acceptability, and wave dummy variables. In addition, wife's employment is influenced by the husband's (log) income.

The wife's employment equation is estimated using a probit model. Estimation results for the employment participation index are given in Table 3. All the parameters except those on the constant, wife's secondary education, and the dummy variable for waves 3 and 5 (base is wave 1) are significant. The estimation results show that husband's income decreases wife's work propensity. This is in keeping with our presumption that leisure is a normal good (Killingsworth, 1983). Husband's education increases his wife's propensity to work. Geerken & Gove (1984) find a similar positive effect of husband's education on wife's employment propensity. Husband's education may be viewed as a measure of his ideological outlook on traditional gender roles. High education of the husband would appear to lead to a more egalitarian allocation of household responsibilities and, consequently, a relaxation of household constraints that affect wife's employment decision. It may also be associated with greater respect for the talents and values of the wife, thus decreasing the impact of any traditional inhibiting factors that influence wife's market work choice. Age increases the propensity to work till 40.¹⁴ Beyond this age, there is a decline in employment propensity. This pattern is consistent with female labor supply studies (for example, see Hanoch 1980 and Cogan 1980) which suggest that the value of non-work time increases at later lifecycle stages, reducing employment propensity. Wife's education has a positive effect on her work propensity, though secondary education does not seem to have a significantly different effect from primary education. The total number of children has a negative impact on wife's inclination to work, and the number of children less than 12 has an added negative influence.¹⁵ As we would expect, the regional social acceptability of wife's work role (work acceptability) has a positive

¹⁴ The choice of 40 as the spline point was made primarily from sample size considerations. The small number of employed wives constrained us to the choice of a single breakpoint. Employing 45 years as a spline point (as for the husband's income specification) was not possible due to the very few wives over 45 years in the sample.

¹⁵ This result is consistent with the notion that children increase the value of a wife's time at home and that they are potential sources of child care costs and other psychological costs of work outside the home (Lundberg 1988).

impact on wife's employment decision. Finally, the sign and magnitude of the two wave dummy variables indicate increases in work intensity in later waves, suggesting a positive period effect.

4.3. Wife's Hours of Work Equation

Estimates of the wife's hours of work equation (in log form) after accounting for sample selection are presented in Table 4a.¹⁶ The exogenous variable specification is similar to that used for the wife's employment equation.

The effect of husband's income and wife's age are not significant. Number of kids does not significantly affect wife's choice of work hours if the children are above the age of 12. All other effects are significant and in the same direction as their impact on employment choice.

The correlation in unmeasured factors that influence work propensity and hours of work choice is positive, but not significant. Marginal effects of significant variables on wife's hours of work are computed from the above estimates and shown in Table 4b.

4.4. Wife's Income Equation

The exogenous variables in the wife's income equation include the wife's age and education variables, work acceptability index, variables associated with the number and age distribution of children, and wave dummy variables.

Estimates of the wife's income equation (in logarithmic form) after accounting for sample selection are presented in Table 5a. As anticipated, the wife's education variables and the work acceptability index have a significant positive effect on wife's income. The magnitude of parameters on the two education dummy variables is consistent with our expectations. Unobserved factors that affect wife's employment propensity and income earning potential are significant and positively correlated as indicated by the estimate of ρ_{el} . This positive correlation indicates that, all observed exogenous factors being equal, employed wives have a higher income earning potential than unemployed wives, if the unemployed wives were to work. Marginal effects of the significant variables (computed for mean variable values) are provided in Table 5b.

4.5. Summary of Estimation Results

The estimation results indicate the effect of exogenous variables on husband's income, wife's subsistence activity choices (wife's employment choice and hours of work), and wife's income.

Husband's income is significantly influenced by age and education with a strong interaction effect.

Wife's employment decision, work hours, and income are negatively influenced by husband's income, but only the effect on wife's employment decision is significant. Husband's high education and the wife's education level have a positive impact on wife's subsistence activity choices and income. The wife's subsistence activity choices and income are also positively affected by wife's age till age 40. However, the age effect becomes negative thereafter.

Wife's employment participation (but not wife's hours of work and income) is negatively affected by the number of children in the household with an incremental effect for small children. Finally, work acceptability in the area of residence has a strong positive effect on wife's subsistence activity choices and income.

5. Conclusion

This paper develops an empirical model of subsistence activity behavior and income using data from the Dutch National Mobility Panel. Subsistence activity behavior and income earnings are very important components, and determinants, of the overall activity agenda of individuals. Since travel is a consequence of participation in activities, a better understanding of subsistence activity behavior and income earnings is critical to the development of improved activity-based travel demand forecasting techniques.

The paper employs a simultaneous equation framework to model husband's income, wife's subsistence activity behavior and wife's income. The results indicate that these variables can be well-explained by individual and household socio-demographics and regional socio-economic variables. The model also emphasizes the complexity involved in the activity-based approach. In particular, the results indicate that there are likely to be strong correlations among unobserved factors affecting activity variables. These correlations necessitate the application of simultaneous

¹⁶ The log likelihood values in the table correspond to joint estimation of the employment and hours of work equations to account for sample selection in hours of work. As mentioned earlier, we obtained estimates of the employment equation during this joint estimation (not shown in the table) which were very close to those in Table 2a.

equation techniques to model activity variables. Despite this complexity, the activity-based approach is conceptually appealing and is likely to provide a significant improvement in travel demand model estimation and forecasting techniques.

The model developed here may be extended and refined in a number of ways. This study confines the sample used in empirical analysis to nuclear family households. It will be useful to apply the same model to different household types and interpret the similarity/dissimilarity in empirical results. An extension to other household types is also important from a forecasting viewpoint, since a model for each of the different household type segments is needed.

The subsistence activity decisions modeled in this study are the employment choice and number of work hours. Individuals in a household may have an added dimension in their choice of employment and work hours. They may stagger their work timings. For example, in the presence of young children, there may be pressure for one adult to stay at home at all times either due to psychological reasons or due to child care costs. Rather than one member giving up employment completely, the individuals may make a conscious attempt to stagger their work timings. Incorporating this added dimension of subsistence activity choice will result in a more realistic decision-making framework. Modeling work timing will also be valuable in developing, and interlinking, other components of the activity framework such as allocation of maintenance activity and cars among individuals in a household and non-work activity participation decisions (Bhat & Koppelman 1992).

References

- Amemiya T (1979) The estimation of a simultaneous equation tobit model. *International Economic Review* 20(1): 169-181.
- Amemiya T (1985) *Advanced Econometrics*, Harvard University Press, Cambridge.
- Bhat CR & Koppelman FS (1992) A conceptual framework of individual activity program generation. *Transportation Research*.
- Bhat CR (1991) Toward a model of activity program generation. Unpublished Ph.D. Dissertation, Dept. of Civil Engineering, Northwestern University, Evanston, Illinois, December.
- Blood RO & Wolfe DM (1969) *Husbands and Wives: The Dynamics of Married Living*. Free Press, Glencoe, Illinois.
- Box G & Cox D (1964) An analysis of transformations. *Journal of the Royal Statistical Society* 26B: 211-243.
- Browning EK & Browning JM (1986) *Microeconomic Theory and Applications*. Little, Brown and Company, Boston.
- Chiswick B (1970) An interregional analysis of schooling and the skewness of income. In: Hansen WL (ed) *Education, Income, and Human Capital*. National Bureau of Economic Research, New York.
- Clarke MI (1986) Activity modelling - A research tool or a practical planning technique? In *Behavioral Research for Transport Policy* (pp 3-15). VNU Science Press, Utrecht, The Netherlands.
- Cogan J (1980) Labor supply with costs of labor market entry. In Smith JP (ed) *Female Labor Supply: Theory and Estimation* (pp 25-89). Princeton University Press, Princeton, New Jersey.
- Damm D (1980) Interdependencies in activity behavior. *Transportation Research Record* 750: 33-40.
- Geerken M & Gove WR (1983) *At Home and at Work: The Family's Allocation of Labor*. Sage Publications, Beverly Hills.
- Golob TF & Meurs H (1988) Development of structural equations models of the dynamics of passenger travel demand. *Environment and Planning* 20A: 1197-1218.
- Hanoch G (1980) A multivariate model of labor supply: Methodology and estimation. In Smith JP (ed) *Female Labor Supply: Theory and Estimation* (pp 249-326).

- Heckman JJ (1974) Shadow prices, market wages and labor supply. *Econometrica* 42(4): 679-694.
- Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5/4: 475-492.
- Heckman JJ (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46(6): 931-959.
- Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47: 153-161.
- Heer DM (1963) The measurement and bases of family power: An overview, *Marriage and Family Living*. 25: 133-139.
- Hirsh M, Prashker JN & Ben-Akiva ME (1986) Dynamic model of weekly activity pattern. *Transportation Science* 20(1): 24-36.
- Hoorn T, van der (1983) Development of an activity model using a one-week activity-diary data base. In Carpenter S & Jones P (eds) *Recent Advances in Travel Demand Analysis* (pp 335-349). Gower, Aldershot, England.
- Johnson N & Kotz S (1972) *Distributions in Statistics: Continuous Multivariate Distribution*. John Wiley, New York.
- Jones PM, Koppelman FS & Orfeuil JP (1990) Activity analysis: State of the art and future directions. *Developments in Dynamic and Activity-Based Approaches to Travel Analysis* (pp 34-55). Gower, Aldershot, England.
- Killingsworth MR (1983) *Labor Supply*. Cambridge University Press, Cambridge.
- Kitamura R & Kermanshah M (1983) Identifying time and history dependencies of activity choice. *Transportation Research Record* 944: 22-30.
- Kitamura (1984) A model of daily time allocation to discretionary out-of-home activities and trips. *Transportation Research* 18b(3): 255-266.
- Kitamura R (1988a) A dynamic model system of household car ownership, trip generation, and modal split: Model development and simulation experiments. In: Proceedings of the 14th Australian Road Research Board Conference, Part 3 (pp 96-111). Australian Road Research Board, Vermont South, Victoria, Australia, 1988.
- Kitamura R (1988b) An evaluation of activity-based travel analysis: A review paper. *Transportation* 15(1/2): 9-60.

- Lee LF (1981) Simultaneous equations model with discrete and censored dependent variables. In: Manski CF & McFadden D (eds) *Structural Analysis of Discrete Data with Econometric Applications* (pp 346-364). MIT Press, Cambridge, Massachusetts.
- Lundberg S (1988) Labor supply of husbands and wives: A simultaneous equations approach. *The Review of Economics and Statistics*: 224-235.
- Maddala GS (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University press, Cambridge, New York.
- Mincer J (1974) *Schooling, Experience and Earnings*. National Bureau of Economic Research, New York.
- Oi WY & Shuldiner PW (1962) *An Analysis of Urban Travel Demands*. Northwestern University Press, Evanston, Illinois.
- Recker WW, McNally MG & Root GS (1986) A model of complex travel behavior: Part I. Theoretical development. *Transportation Research* 20A: 307-318.
- Reichman S (1976) Travel adjustments and life styles: A behavioral approach. In: Stopher P & Meyburg A (eds) *Behavioral Travel Demand Models*. Lexington Books, Lexington, Massachusetts.
- Wissen L, van (1989) A dynamic model of household interactions in activity patterns. Presented at the International Conference on Dynamic Travel Behavior Analysis, Kyoto, July.
- Wissen L, van & Meurs HJ (1989) The Dutch national mobility panel: Experiences and evaluation. *Transportation* 16(2).

List of Tables

Table 1. List of Exogenous Variables in Model

Table 2a. (Log) Husband's Income Estimates

Table 2b. Marginal Effects on Real Value of Husband's Income (guilders)

Table 3. Wife's Employment Propensity Estimates

Table 4a. (Log) Wife's Annual Hours of Work Estimates

Table 4b. Marginal Effects on Wife's Annual Hours of Work

Table 5a. (Log) Wife's Income Estimates

Table 5b. Marginal Effects on Wife's Income

Table 1. List of Exogenous Variables in Model

Variable	Definition
Husband's age	age of husband
Husband's age > 35	(husband's age-35) if husband's age greater than 35, 0 otherwise
Husband's age > 45	(husband's age-45) if husband's age greater than 45, 0 otherwise
Husband's secondary education	1 if husband's education is at secondary level, 0 otherwise
Husband's high education	1 if husband's education is high, 0 otherwise
Husband's age > 45 and sec/high education	(husband's age-45) if husband's age greater than 45 and husband's education is secondary or high, 0 otherwise
Wave 1	1 if household is in wave 1, 0 otherwise
Wave 3	1 if household is in wave 3, 0 otherwise
Wave 5	1 if household is in wave 5, 0 otherwise
Wave 7	1 if household is in wave 7, 0 otherwise
Wave 9	1 if household is in wave 9, 0 otherwise
Wife's age	wife's age
Wife's age > 40	(wife's age-40) if wife's age greater than 40, 0 otherwise
Wife's secondary education	1 if wife's education is at secondary level, 0 otherwise
Wife's high education	1 if wife's education is high, 0 otherwise
Work acceptability*	regional social acceptability of wife's work role
Number of kids	number of children less than 12 in household
Total number of kids	total number of children in household

* calculated as the ratio of female labor force to total female population between 15 yrs and 64 yrs for each region.

Table 2a. (Log) Husband's Income Estimates

Variable	Coefficient	Standard error	Coef./Std.error
Constant	9.131	0.098	93.05
Husband's age			
entire range	0.025	0.003	8.50
> 35 years	-0.015	0.004	-3.54
> 45 years	-0.021	0.006	-3.65
Husband's age education	0.009	0.002	4.31
Husband's education			
secondary	0.203	0.012	16.75
high	0.385	0.013	28.76
Wave variables			
one	0.057	0.016	3.64
five	0.018	0.015	0.84
seven	0.021	0.015	1.19
nine	0.012	0.015	1.38
σ_m	0.207	0.005	42.45
Log Likelihood LL	LL(slopes=0) = -3094 LL(convergence) = -2492		

Table 2b. Marginal Effects on Real Value of Husband's Income (guilders)

Variable	Marginal Effect on Husband's Income
Husband's age (per year)	
< 35 years	702
35 - 45 years	288
> 45 years - prim./sec. educ.	-294
> 45 years - high education	-46
Husband's education	
secondary	5,131
high	10,858

Table 3. Wife's Employment Propensity Estimates

Variable	Coefficient	Standard error	Coef./Std.error
Constant	1.744	3.102	0.56
Husband's (log) income	-0.644	0.324	-2.00
Husband's high educ.	0.251	0.114	2.19
Wife's age			
entire range	0.058	0.010	5.59
> 40 years	-0.127	0.022	-5.70
Wife's education			
secondary	0.083	0.072	1.15
high	0.475	0.101	4.72
Children			
total number	-0.244	0.060	-4.06
no. < 12 years	-0.099	0.050	-1.99
Work acceptability	4.853	0.796	6.09
Wave variables			
three/five	0.132	0.086	1.53
seven/nine	0.307	0.085	3.60
Log Likelihood LL	LL(slopes=0) = -1283 LL(convergence) = -1182		

Table 4a. (Log) Wife's Annual Hours of Work Estimates

Variable	Coefficient	Standard error	Coef./Std.error
Constant	7.630	3.072	2.48
Husband's (log) income	-0.260	0.317	-0.82
Husband's high educ.	0.183	0.106	1.72
Wife's age			
entire range	0.005	0.009	0.55
> 40 years	-0.034	0.020	-1.70
Wife's education			
secondary	0.263	0.067	3.91
high	0.425	0.095	4.50
Children			
total number	-0.016	0.054	-0.29
no. < 12 years	-0.128	0.043	-2.95
Work acceptability	2.189	0.593	3.69
Wave variables			
three/five	0.127	0.084	1.51
seven/nine	0.186	0.080	2.32
σ_l	0.586	0.024	24.88
ρ_{el}	0.075	0.071	1.06
Log Likelihood LL	LL(slopes, $\rho_{el}=0$) = -1826 LL(convergence) = -1686		

Table 4b. Marginal Effects on Wife's Annual Hours of Work

Variable	Marginal Effect on Wife's Hours
Husband's high education	104
Wife's education	
secondary	150
high	243
Children (each child)	
< 12 years	-83
Work acceptability index (shift of 0.1)	125

Table 5a. (Log) Wife's Income Estimates

Variable	Coefficient	Standard error	Coef./Std.error
Constant	10.097	3.399	2.97
Husband's (log) income	-0.410	0.355	-1.16
Husband's high educ.	0.350	0.125	2.79
Wife's age			
entire range	0.016	0.012	1.28
> 40 years	-0.034	0.025	-1.34
Wife's education			
secondary	0.458	0.081	5.62
high	0.980	0.124	7.92
Children			
total number	-0.116	0.074	-1.56
no. < 12 years	-0.051	0.056	-0.91
Work acceptability	3.744	0.954	3.93
Wave variables			
three/five	0.105	0.105	1.00
seven/nine	0.133	0.108	1.24
σ_w	0.671	0.064	10.46
ρ_{ew}	0.447	0.173	2.60
Log Likelihood LL	LL(slopes, $\rho_{ew}=0$) = -1887 LL(convergence) = -1701		

Table 5b. Marginal Effects on Wife's Income

Variable	Marginal Effect on Wife's Income
Husband's high education	2,100
Wife's education	
secondary	2,747
high	5,878
Work acceptability index (shift of 0.01)	225