NBER WORKING PAPER SERIES

HEALTH, CHILDREN, AND ELDERLY LIVING ARRANGEMENTS: A MULTIPERIOD-MULTINOMIAL PROBIT MODEL WITH UNOBSERVED HETEROGENEITY AND AUTOCORRELATED ERRORS

Axel Boersch-Supan

Vassilis Hajivassiliou

Laurence J. Kotlikoff

John N. Morris

Working Paper No. 3343

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1990

This research was supported by the National Institute on Aging, grant no. 3 P01 AG 05842. Dan Nash and Gerald Schehl provided valuable research assistance. This paper is part of NBER's research program in Aging. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3343
April 1990

HEALTH, CHILDREN, AND ELDERLY LIVING ARRANGEMENTS: A MULTIPERIOD-MULTINOMIAL PROBIT MODEL WITH UNOBSERVED HETEROGENEITY AND AUTOCORRELATED ERRORS

ABSTRACT

This paper develops a general multiperiod-multinomial probit model for panel data to estimate the living arrangements of the elderly. The model has the following features:

(a) In each period choices do not necessarily obey the assumption of independence of irrelevant alternatives.

(b) Unobserved person-specific attributes are treated as random effects. These random effects may also be correlated across alternatives.

(c) In addition, unobserved choice-specific utility components may persist over some time, creating an autoregressive and/or heteroscedastic error structure.

The model is estimated by simulating the choice probabilities in the likelihood function. We examine several variants of the specification of the correlation structure and investigate the extent the biases created by ignoring intertemporal correlations.

Axel Boersch-Supan
University of Mannheim
Postfach 10 34 62
Gebaude A5,6
6800 Mannheim 1
GERMANY

Vassilis Hajivassiliou
Cowles Foundation, Yale University
Department of Economics
Box 2125 Yale Station
New Haven, CT 06520

Laurence J. Kotlikoff
Boston University and
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138

John N. Morris
Hebrew Rehabilitation Center for the
Aged
1200 Centre Street
Boston, MA 02131

1. INTRODUCTION

Decisions by the elderly with respect to their living arrangements (e.g., living alone, living with children, or living in a nursing home) seem best modelled as a discrete choice problem in which the elderly view certain choices as closer substitutes than others. For example, living with children may more closely substitute for living independently than living in an institution. Unobserved determinants of living arrangements at a point in time are, therefore, quite likely to be correlated. In the parlance of discrete choice models this means, that the assumption of independence of irrelevant alternatives (IIA) will be violated. Indeed, a number of recent studies of living arrangements of the elderly document the violation of IIA.[1]

In addition to relaxing the IIA assumption of no intratemporal correlation between unobserved determinants of competing living arrangements, one should also relax the assumption of no intertemporal correlation of such determinants. The assumption of no intertemporal correlation underlies most studies of living arrangements, particularly those estimated with cross-section data. While cross section variation in household characteristics can provide important insights into the determinants of living arrangements[2], the living arrangement decision is clearly an intertemporal choice and a potentially complicated one at that. Because of moving and associated transactions costs, elderly households may stay longer in inappropriate living arrangements than they would absent such costs. In turn, households may prospectively move into an institution "before it is too late to cope with this change". That is, households may be substantially out of long-run equilibrium if a cross-section survey interviews them shortly before or after a move. Moreover, persons may acquire a taste for certain types of living arrangements. Such habit formation introduces state dependence. Ideally, therefore, living arrangement choices should be estimated with panel data, with an appropriate econometric specification of intertemporal linkages.

These intertemporal linkages include two components. The first component is the linkage through unobserved person-specific attributes, i.e., unobserved heterogeneity through time-invariant error components. An important example is health status, information on which is often missing or unsatisfactory in household surveys. Health status varies over time, but has an important person-specific, time-invariant component that influences housing and living arrangement choices of the elderly. Panel data discrete choice models that capture unobserved heterogeneity include Chamberlain's (1984) conditional fixed effects estimator and one-factor random effect models, such as those proposed by McFadden (1984, p.1434).

However, not all intertemporal correlation patterns in unobservables can be captured by time-invariant error components. A second error component should, therefore, be included to control for time-varying disturbances, e.g., an autoregressive error structure. Examples of the source of error components that taper off over time are the above-mentioned cases of prospective moves and habit formation. Similar effects on the error structure arise when an elderly person, due to unmeasured transactions costs, stays longer in a dwelling than he/she would in the absence of such costs.

Ellwood and Kane (1988) and Börsch-Supan (1988) apply simple models to capture dynamic features of the observed data. Ellwood and Kane (1988) employ an exponential hazard model, while Börsch-Supan (1988) uses a variety of simple Markov transition models. Neither approach captures both unobserved heterogeneity and autoregressive errors. In addition, living arrangement choices are multinomial by nature, ruling out univariate hazard models. Börsch-Supan, Kotlikoff and Morris (1989) also fail to deal fully with heterogeneous and autoregressive unobservables. Their study attempts to finesse these concerns by describing the multinomial-multiperiod choice process as one large discrete choice among all possible outcomes. By invoking the IIA-assumption, a small subset of choices is sufficient to identify the relevant parameters. This approach, that converts the problem of repeated intertemporal choices to a static problem of choosing, ex-ante, the time path of living arrangements is easily criticized both because of the IIA-assumption and the presumption that individuals decide in advance their future living arrangements.

While researchers have recognized the need to estimate choice models with unobserved determinants that are correlated across alternatives and over time, they have been daunted by the high dimensional integration of the associated likelihood functions. This paper uses a new simulation method developed in Börsch-Supan and Hajivassiliou (1990) to estimate the likelihood functions of living arrangement choice models that range, in their error structure, from the very simple to the highly complex. Compared with previous simulation estimators derived by McFadden (1989) and Pakes and Pollard (1989), the new method is capable of dealing with complex error structures with substantially less computation. The Börsch-Supan-Hajivassiliou method builds on recent progress in Monte-Carlo integration techniques by Geweke (1989) and Hajivassiliou and McFadden (1990). It represents a revival of the Lerman and Manski (1981) procedure of approximating the likelihood function by simulated choice probabilities overcoming its computational disadvantages.

The following section, 2, develops the general structure of the choice probability integrals and spells out alternative correlation structures. Section 3 presents the estimation procedure, termed simulated maximum likelihood (SML). Section 4 describes our data, and Section 5 reports results. Section 6 concludes with a summary of major findings.

2. ECONOMETRIC SPECIFICATIONS OF ALTERNATIVE ERROR PROCESSES

Let I be the number of discrete choices in each time period, and T be the number of waves in the panel data. The space of possible outcomes is the set of I^T different choice sequences $\{i_t\}$, $t=1,\ldots,T$. To structure this discrete choice problem, we assume that in each period choices are made according to the random utility maximization hypothesis, i.e.,[2]

$$i_t \text{ is chosen} \Leftrightarrow u_{it} \text{ is maximal element in } \{u_{jt} \mid j=1,\ldots,t\}, \tag{1}$$

where the utility of choice i in period t is the sum of a deterministic utility component $v_{it} = v(X_{it},\beta)$, which depends on the vector of observable variables X_{it} and a parameter vector β to be estimated, and a random utility component ϵ_{it}:

$$u_{it} = v(X_{it},\beta) + \epsilon_{it}. \tag{2}$$

We model the deterministic utility component, $v(X_{it},\beta)$, as simply the linear combination $X_{it}\beta$.[3]

Since the optimal choice delivers maximum utility, it is not the utility level of the maximal choice, but rather the differences in utility levels between the best choice and any other choice that is relevant for the elderly's decision. The probability of a choice sequence $\{i_t\}$ can, therefore, be expressed as integrals over the differences of the unobserved utility components relative to the chosen alternative. Define

$$w_{jt} = \epsilon_{jt}-\epsilon_{it} \qquad \text{for } i=i_t,\ j\neq i_t. \tag{3}$$

These $D = (I-1) \times T$ error differences are stacked in the vector w and have a joint cumulative distribution function F.

For alternative i to be chosen, the error differences can be at most as large as the differences in the deterministic utility components. The areas of integration are therefore

$$A_j(i)=\{ w_{jt} \mid -\infty \le w_{jt} \le X_{it}\beta - X_{jt}\beta \} \quad \text{for } j\neq i \tag{4}$$

and the probability of choice sequence $\{i_t\}$ is

$$P(\{i_t\}|\{X_{it}\};\beta,F) = \int_{\{w_{j1} \in A_j(i_1)|j=1..I,j\neq i_1\}} \times \ldots \times \int_{\{w_{jT} \in A_j(i_T)|j=1..I,j\neq i_T\}} dF(w). \tag{5}$$

Unless the joint cumulative distribution function F and the area of integration $A_j = A_j(i_1) \times \ldots \times A_j(i_T)$ are particularly benign, the integral in (5) will not have a closed form. Closed-form solutions exist if F is a member of the family of generalized extreme-value (GEV) distributions, e.g., the cross-sectional multinomial logit (MNL) or nested multinomial logit (NMNL) models, contributing to the popularity of these specifications. Closed-form solutions also exist if these models are combined with a one-factor random effect that is again extreme-value distributed (e.g., McFadden, 1984).

GEV-type models have the disadvantage of relatively rigid correlation structures. They cannot embed the more general intertemporal correlation patterns exposited in the introduction. Concentrating on the first two moments, we assume a multivariate normal distribution of the w_{jt} in (3), characterized by a covariance matrix M which has $(D+1) \times D/2 - 1$ significant elements: the correlations among the w_{jt} and the variances except one in order to scale the parameter vector β in the deterministic utility components $v(X,\beta)$. This count represents many more covariance parameters than GEV-type models can handle. Moreover, our specification of M is not constrained by hierarchical structures as is the case in the class of nested multinomial logit (NMNL) models.

We estimate this multiperiod-multinomial probit model with different specifications of the covariance matrix M.

(A) The simplest specification M=I yields a pooled cross-sectional probit model that is subject to the independence of irrelevant alternatives (IIA) restriction and ignores all intertemporal linkages. The D=(I-1) x T dimensional integral of the choice probabilities factors into D one-dimensional integrals.

There are several ways to introduce intertemporal linkages:

(B) A random-effect structure is imposed by specifying

$$\epsilon_{i,t} = \alpha_i + \nu_{i,t},\ \nu_{i,t} \text{ i.i.d., } i=1,..,I-1.$$

This yields a blockdiagonal equi-correlation structure of M with (I-1) parameters $\sigma(\alpha)$ in M which need to be estimated. This structure allows for a factorization of the integral in (5) in I-1 T-dimensional blocks which, in turn, can be reduced to one dimension because of the one-factor structure.

(C) An autoregressive error structure can be incorporated by specifying

$$\epsilon_{i,t} = \rho_i \cdot \epsilon_{i,t-1} + \nu_{i,t},\ \nu_{i,t} \text{ i.i.d., } i=1,..,I-1.$$

Again, this yields a blockdiagonal structure of M where each block has the familiar structure of an AR(1) process. I-1 parameters ρ_i in M have to be estimated.

(D) The last two error structures can also be combined by specifying

$$\epsilon_{i,t} = \alpha_i + \eta_{i,t},\ \eta_{i,t} = \rho_i \cdot \eta_{i,t-1} + \nu_{i,t},\ \nu_{i,t} \text{ i.i.d., } i=1,..,I-1.$$

This amounts to overlaying the equi-corrrelation structure with the AR(1) structure. It should be noted that $\sigma(\alpha_i)$ and ρ_i are separately identified only if $\rho_i<1$.

We now drop the IIA assumption. There are several distinct possibilities, depending on the intertemporal error specification.

(E) Starting again with (A) and ignoring any intertemporal structure, the simplest possibility is to assume that the $\epsilon_{i,t}$ are uncorrelated across t but have correlations across i which are constant over time. With the proper reordering of the elements in the stacked vector w, a simple blockdiagonal structure of M emerges with T*(I-1)-dimensional blocks. In this case (I-2) variances and (I-1)*(I-2)/2 covariances can be identified.

(F) This specification can be overlayed with the random effects specification. This destroys the blockdiagonality although the one-factor structure allows a reduction of the dimensionality of the integral in (5). (I-1) variances of the random effects $\sigma(\alpha_i)$ can be identified in addition to the parameters in specification (E). Rather than allowing inter-alternative correlation in the $\nu_{i,t}$ (specification F1), it is also possible to make the random effects α_i correlated (specification F2).

(G) Alternatively, specification (E) can be overlayed with an autoregressive error structure by specifying

$$\epsilon_{i,t} = \rho_i \cdot \epsilon_{i,t-1} + \nu_{i,t}, \quad \mathrm{corr}(\nu_{i,t},\nu_{j,s}) = \omega_{ij} \quad \text{if } s=t, \text{ else } 0.$$

The $\nu_{i,t}$ are correlated across alternatives, but uncorrelated across periods. The familiar structure of an AR(1) process is additively overlayed with the blockdiagonal structure of specification (E). I-1 additional parameters ρ_i in M have to be estimated.

(H) Finally, all three features -- interalternative-correlation, random effects, and autoregressive errors -- can be combined. The resulting error process is

$$\epsilon_{i,t} = \alpha_i + \eta_{i,t}, \qquad \eta_{i,t} = \rho_i \cdot \eta_{i,t-1} + \nu_{i,t}, \qquad i=1,..,I-1,$$

$$\text{with } \mathrm{corr}(\nu_{i,t},\nu_{j,s}) = \begin{cases} 0 & \text{if } t \neq s \\ \omega_{ij} & \text{if } t=s \end{cases} \quad \text{and} \quad \mathrm{cov}(\alpha_i,\alpha_j) = \sigma_{ij}$$

which implies

$$\mathrm{cov}(\epsilon_{i,t},\epsilon_{j,s}) = \sigma_{ij} + \rho_i^{(t-s)} \frac{\sqrt{(1-\rho_i^2)} \cdot \sqrt{(1-\rho_j^2)}}{1-\rho_i\rho_j} \omega_{ij} .$$

This model encompasses all preceding specifications as special cases. Again, all parameters are identified if $\rho_i<1$, $i=1,..I-1$, although, in practice, the identification of this general specification may become shaky when there is only a small number of sufficiently long spells in different choices.

3. ESTIMATION PROCEDURE: SIMULATED MAXIMUM LIKELIHOOD

The likelihood function corresponding to the general multiperiod-multinomial choice problem is the product of the choice probabilities (5):

$$\mathcal{L}(\beta,M) = \prod_{n=1}^{N} P(\{i_{t,n}\} | \{X_{it,n}\};\beta,M) \tag{6}$$

where the index n denotes an observation in a sample of N individuals, and the c.d.f. F in (5) is assumed to be multivariate normal and characterized by the covariance matrix M. Estimating the parameters in (6) is a formidable task because it requires, in the most general case, for each observation and each iteration in the maximization process an evaluation of the $D=(I-1)*T$ dimensional integral in (5).

One may be tempted to accept the efficiency losses due to an incorrect specification of the error structure and simply ignore the correlations that make the integral in (5) so hard to solve. However, unlike the linear model, an incorrect specification of the covariance matrix of the errors M biases not only the standard errors of the estimated coefficients, but also the structural coefficients β themselves. The linear case is very special in isolating specification errors away from β.

Numerical integration of the integral in (5) is computationally infeasible since the number of operations increases with the power of D, the dimension of M. Approximation methods, such as the Clark-approximation (Daganzo, 1981) or its variant proposed by Langdon (1984), are tractable -- their number of operations increases quadratically with D -- but they remain unsatisfactory since their relatively large bias cannot be controlled by increasing the number of observations. Rather, we simulate the choice probabilities $P(\{i_{t,n}\}|\{X_{it,n}\};\beta,M)$ by drawing pseudo-random realizations from the underlying error process.

The most straightforward simulation method is to simulate the choice probabilities $P(\{i_{t,n}\}|\{X_{it,n}\};\beta,M)$ by observed frequencies (Lerman and Manski, 1981):

$$\tilde{P}(i_{tn}) = N_{tn}(i)/N_{tn}, \qquad (7)$$

where N_{tn} denotes the number of draws or replications for individual n at period t, and

$$N_{tn}(i) = \text{count}(\ u_{itn} \text{ is maximal in } \{u_{jtn} \mid j=1,..,t\}\ ). \qquad (8)$$

One then maximizes the simulated likelihood function

$$\tilde{\mathcal{L}}(\beta,M) = \prod_{n=1}^{N} \prod_{t=1}^{T} N_{tn}(i)/N_{tn}. \tag{9}$$

However, in order to obtain reasonably accurate estimates (7) of small choice probabilities, a very large number of draws is required. That results in unacceptably long computer runs.

We exploit instead an algorithm proposed by Geweke (1989) which was originally designed to compute random variates from a multivariate truncated normal distribution. This algorithm is very quick and depends continuously on the parameters β and M. One concern is that it fails to deliver unbiased multivariate truncated normal variates.[4] However, as Börsch-Supan and Hajivassiliou (1990) show, the algorithm can be used to derive unbiased estimates of the choice probabilities. We sketch this method in the remainder of this section.

Univariate truncated normal variates can be drawn according to a straightforward application of the integral transform theorem: Let u be a draw from a univariate standard uniform distribution, $u \in [0,1]$. Then

$$e = G^{-1}(u) = \Phi^{-1}[\ (\Phi(a)-\Phi(a))\cdot u + \Phi(a)\ ] \tag{10}$$

is distributed N(0,1) s.t. $a \le e \le b$, since the c.d.f. of a univariate truncated normal distribution is

$$G(z) = \frac{\Phi(z) - \Phi(a)}{\Phi(a) - \Phi(b)}, \tag{11}$$

where Φ denotes the univariate normal cumulative distribution function. Note that e is a continuously differentiable function of the truncation parameters a and b. This continuity is essential for computational efficiency.

In the multivariate case, let L be the lower diagonal Cholesky factor of the covariance matrix M of the unobserved utility differences w in (3),

$$L \cdot L' = M \ . \qquad (12)$$

Then draw sequentially a vector of D=(I-1)*T univariate truncated normal variates

$$e = N(0,I) \quad \text{s.t.} \quad a \leq L \cdot e \leq \infty, \qquad (13)$$

where the D-dimensional vector a is defined by equation (4):

$$a_{jt} = X_{it}\beta - X_{jt}\beta \qquad \text{for } i=i_t,\ j\neq i_t. \qquad (14)$$

Because L is triangular, the restrictions in (13) are recursive. (For notational simplicity, e and a are in the sequel simply indexed by i=1,...,D):

$$\begin{aligned} & e_1 = N(0,1) \\ & \text{s.t.} \quad a_1 \leq l_{11} \cdot e_1 \leq \infty \quad \Leftrightarrow \quad a_1/l_{11} \leq e_1 \leq \infty \\ & e_2 = N(0,1) \\ & \text{s.t.} \quad a_2 \leq l_{21} \cdot e_1 + l_{22} \cdot e_2 \leq \infty \quad \Leftrightarrow \quad (a_2 - l_{21} \cdot e_1)/l_{22} \leq e_2 \leq \infty \\ & \text{etc.} \end{aligned} \qquad (15)$$

Hence, each e_i, i=1,..,D, can be drawn using the univariate formula (10). Finally, define

$$w = Le \ . \qquad (16)$$

Then (12) implies that w has covariance matrix M and is subject to

$$a \leq Le \leq \infty \quad \Leftrightarrow \quad a \leq w \leq \infty \qquad (17)$$

as required.

The probability for a choice sequence $\{i_{tn}\}$ of observation n is the probability that w falls in the interval given by (4), which is the probability that e falls in the interval given by (13), i.e.,

$$P(\{i_{tn}\}) = \mathrm{Prob}(\, a_1/l_{11} \le e_1 \le \infty) \cdot \mathrm{Prob}((a_2 - l_{21} \cdot e_1)/l_{22} \le e_2 \le \infty \mid e_1) \cdot \ldots \tag{18}$$

For a draw of a D-dimensional vector of truncated normal variates $e_r=(e_{r1},\ldots,e_{rD})$ according to (15), this probability is simulated by

$$\tilde{P}_r(\{i_{tn}\}) = (\, 1 - \Phi(a_1/l_{11})\,) \cdot (\, 1 - \Phi(a_2 - l_{21} \cdot e_{r1})/l_{22})\,) \cdot \ldots \tag{19}$$

and the choice probability is approximated by the average over R replications of (19):

$$\tilde{P}(\{i_{tn}\}) = \frac{1}{R} \sum_{r=1}^{R} \tilde{P}_r(\{i_{tn}\}) \,. \tag{20}$$

Börsch-Supan and Hajivassiliou (1990) prove that $\tilde{P}$ is an unbiased estimator of P in spite of the failure of the Geweke algorithm to provide unbiased expected values of e and w.

Like the univariate case, the generated draws as well as the resulting simulated probability of a choice sequence depend continuously and differentiably on the parameters β in the truncation vector a and the covariance matrix M. Hence, conventional numerical methods such as one of the conjugate gradient methods or quadratic hillclimbing can be used to solve the first order conditions for maximizing the simulated likelihood function

$$\tilde{\mathcal{L}}(\beta,M) = \prod_{n=1}^{N} \sum_{r=1}^{R} \tilde{P}_r(\{i_{tn}\}) \,. \tag{21}$$

This differs from the frequency simulator (7) that generates a discontinuous objective function with the associated numerical problems.

Moreover, as described by Börsch-Supan and Hajivassiliou (1990), the choice probabilities are well approximated by (20) even for a small number of replications, independent of the true choice probabilities. This is in remarkable contrast to the Lerman/Manski frequency simulator that requires that the number of replications be inversely related to the true choice probabilities. The Lerman/Manski simulator thus requires a very large number of replications for small choice probabilities.

Finally, it should be noted that the computational effort in the simulation increases nearly linearly with the dimensionality of the integral in (5), D=(I-1)*T, since most computer time is involved in generating the univariate truncated normal draws.[5] For reliable results it is crucial to compute the cumulative normal distribution function and its inverse with high accuracy. The near-linearity permits applications to large choice sets with a large number of panel waves.

4. DATA, VARIABLE DEFINITIONS, AND BASIC SAMPLE CHARACTERISTICS

The paper employs data from the Survey of the Elderly collected by the Hebrew Rehabilitation Center for the Aged (HRCA). This survey is part of an ongoing panel survey of elderly in Massachusetts that began in 1982. Initially, the sample consisted of 4040 elderly, aged 60 and above. In addition to the baseline interview in 1982, reinterviews were conducted in 1984, 1985, 1986, and 1987. The sample is stratified and consists of two populations. The first population represents about 70 percent of the sample and was drawn from a random selection of communities in Massachusetts. This first subsample is in itself highly stratified to produce an overrepresentation of the very old. The second population that comprises the remaining 30 percent is drawn from elderly participants of the 27 Massachusetts home health care corporations. In the second population the older old are also overrepresented. The sample selection criteria, sampling procedures and exposure rates are described in more detail in Morris et.al. (1987) and Kotlikoff and Morris (1987).

In addition to basic demographic information collected in the baseline interview, each wave of the HRCA panel contains questions about the elderly's current marital status, living arrangements, income, and number and proximity of their children. The surveys pay particular attention to health status, recording the presence and severity of diagnosed conditions and determining an array of functional (dis-)abilities.

Table 1 presents the age distribution of the elderly at baseline in 1982. The average age is 78.5, 78 percent are age 75 or more, and 20 percent are age 85 or more. Among the U.S. noninstitutionalized population aged 60 and over, 27.9 percent are age 75 or more, while only 5.5 percent are over age 85. The overrepresentation of the oldest old in our sample is indicated by the impressive number of 8 centenarians in our sample! Because the sample overrepresents the very old, it is also characterized by a very large proportion of women. In 1982, 68.7 percent of the interviewed elderly were female; by 1986, this percentage had risen to 70.7 percent.

The lower part of Table 1 provides information about family relationships and the isolation of some of the elderly. In 1982, 32.9 percent of the elderly in the HRCA baseline sample were married and 55.0 percent were widowed. Four years later, 26.7 percent of the surviving elderly were married and 61.4 percent were widowed. As of 1986, 41.4 percent of the elderly report no children, 15.2 one, 17.8 two, 12.7 three, and 12.8 percent four or more children. Because the elderly in the sample are quite old, some of their children are elderly themselves, and some children may even have died earlier than their parents. A total 47.0 percent of the elderly have siblings who are still alive; 25.5 percent of all elderly report that they have no relatives alive at all; and 39.3 percent report that they have no friends.

Average yearly income of the elderly rises between 1984 and 1986 from

$8,750 to $10,500. This 20 percent increase is larger than the concomitant growth in average income for the general population which was only 13.2 percent. It is interesting to note that elderly without children have a significantly lower income ($7,500) than elderly with at least one child ($9,500) in 1984, although in 1986, this difference becomes smaller ($9,700 as opposed to $10,750).

One of the major strengths of the HRCA survey is its detailed information on the health status of the elderly. Three kinds of health measures are reported: a subjective health index, an array of diagnosed conditions, and an array of functional ability measures. The subjective health index (SUBJ) is coded "excellent" (1), "good" (2), "fair" (3), or "poor" (4). The presence and severity of seven chronic illnesses are reported: cancer, mental illness, diabetes, stroke, heart disease, hypertension, and arthritis. Each of these illnesses are scored as: "not present" (0), "present but does not cause limitation" (1), and "present and causes limitation" (2). We condense this information in a summary measure, ILLSUM, the (unweighted) sum of all seven scores. Five measures of functional ability are used: the distance an elderly person can walk or wheel, whether an elderly person can take medication, can attend to his/her personal care, can prepare his/her own meals, and can do normal housework. The first measure is scored from 0 to 5, representing mobility from "can walk more than half mile" down to "confined to bed". The other measures can attain five values, representing "could do on own", "needs some help sometimes", "needs some help often", "needs considerable help", and "cannot do at all" with associated scores from 0 to 4. Similar to the chronic illnesses, we condense these indicators in a simple summary measures of functional ability, ADLSUM, the (unweighted) sum of all five scores.

Börsch-Supan, Kotlikoff and Morris (1989) discuss more sophisticated measures, the correlation among the several measures of health status, and their relative performance in predicting living arrangements. While the subjective health rating performs poorly and is barely correlated with the measures of functional ability and diagnosed conditions, ILLSUM and ADLSUM are as good in predicting living arrangement choices as more sophisticted summary measures of health status.

Although the 1982 sample did not include institutionalized elderly, subsequent surveys have followed the elderly as they moved, including moves into and out of nursing homes. The type of institution was carefully recorded in the survey instrument. In addition, in each wave the non-institutionalized elderly were asked who else was living in their home. This provides the opportunity to estimate a quite general model of living arrangement choice including the process of institutionalization, having conditioned on not being institutionalized at the time of the first interview. In the longitudinal analysis we distinguish three categories of living arrangements:

(1) Independent living arrangements: the elderly's household does not contain any other person besides the elderly individual and his/her spouse, if the elderly individual is married and his/her spouse lives with him/her.

(2) Shared living arrangements: the elderly's household contains at least one other adult person beside the elderly individual and his/her spouse. In most cases, the household contains only the elderly, his/her spouse, and the immediate family of one of his or her children, including a child-in-law. Less frequently, the household also contains other related or unrelated persons.

(3) Institutional living arrangements: This category includes elderly who are living on a permanent basis in a health-care facility.

The institutional living arrangements category comprises the entire spectrum ranging from hospitals and nursing homes to congregate housing and boarding houses. Living arrangements are reported as of the day of the interview -- therefore, temporary nursing home stays are not recorded unless they happen to be at the time of interview. Rather, most nursing home stays in our data set represent permanent living arrangements.[6] It is important to keep this in mind when comparing the frequency and risk of institutionalization in this paper with numbers in studies that focus on short-term nursing home stays.

Table 2 presents the distributions of living arrangements in the five waves of the HRCA panel. The frequencies in this table are strictly cross-sectional and are based on all elderly who were living at the time of each cross-section and for whom living arrangements were known.

Most remarkable is the decreasing, but very high proportion of elderly living independently in spite of the very old age of most of the elderly in the sample. Approximately one out of every six elderly shares a household with his or her own children, whereas very few elderly share a household with distantly-related or unrelated persons. The dramatic increase over time in the proportion of institutionalized living arrangements reflects two effects that must be carefully distinguished. Institutionalization increases because the sample ages and health deteriorates as is obvious from Table 2. This effect is confounded by the way the sample was drawn. In 1982, the sample is non-institutionalized by design. Only a few elderly happened to become institutionalized between the time of the sample design and the actual interview. Four years later, more than one fifth of the surviving elderly live in an institution, almost all in a nursing home. As of 1986, very few elderly live in the "new" forms of elderly housing, such as congregate housing or continuing care retirement communities.

Table 3 examines the temporal evolution of living arrangements. It enumerates all living arrangement sequences that are observed among the 1196 elderly whose living arrangements could be ascertained from 1982 through 1986. A little less than half (47.8 percent) of the elderly maintain the same living arrangement from 1982 through 1986. Another 21.0 percent died before 1986 without an observed living arrangement transition. This stability confirms the results by Börsch-Supan (1988) and Ellwood and Kane (1988). About 40 percent of the sampled elderly lived independently from 1982 through 1986. Another 15.6 percent remained independent until they died prior to 1986. Another 24.6 percent lived for at least some time with their children, and 21.1 percent experienced at least one stay in an institution. The most frequently observed transition is from living independently to being institutionalized. These sequences are observed for 42.4 percent of all elderly who change their living arrangement at least once. Only 13.7 percent change from living independently to living with their children. Most other sequences are very rare.

5. ESTIMATION RESULTS

For the longitudinal econometric analysis, we extract a small working sample of 314 elderly who were interviewed in all five waves, whose living arrangements could be ascertained in all five waves, and who have reliable data on all covariates in all five waves. This results in a sample biased toward the more healthy elderly. While we have not done so here, the econometric model can easily be extended to accommodate sample truncation due to exogenous factors, most importantly death and health-related inability to conduct an interview. Table 4 presents a description of the variables employed and the usual sample statistics of this subsample.

The presentation of results is organized according to four intertemporal specifications (pooled cross-sections, random effects, autoregressive errors, and random effects plus autoregressive errors) and two or three specifications of correlation pattern across alternatives (the IIA assumption; correlation between random effects, if applicable; and the full

MNP model). Three replications (draws) were used to simulate the choice probabilities entering the loglikelihood function. Using fewer replications produces less reliable results, but increasing the number of replication up to nine as we did for the final estimate does not change results in any substantive way.

The goodness-of-fit in the various specifications is examined in Table 5. This table reports the value of the simulated loglikelihood function at estimated parameter values and the pseudo-R^2 associated with this loglikelihood value.[7] The cross-sectional estimates yield a pseudo-R^2 of more than 40 percent, a satisfactory fit for this kind of data. However, introducing random effects in order to account for unobserved time-invariant characteristics dramatically increases the fit. If shocks are allowed to taper off in a first order autoregressive process rather than to persist in form of a random effect, the fit is even better. Finally, the combination of random effects and the AR(1) structure yields significantly better results than if either specification is employed separately.[8] Clearly, the unobserved utilities of this model include both time-invariant and time-varying components.

Correlation across alternatives is also present. The full multinomial probit specifications (rightmost numbers in Table 5, denoted by "MNP") fare everywhere significantly better than the models that obey the IIA assumption (leftmost numbers in Table 5, denoted by "IIA"). Inter-alternative correlation appears to work through the contemporary error components rather than through the random effects, as can be seen by comparing the numbers below "RE-Corr" with those below "MNP".

Detailed estimation results follow in Tables 6 through 9. These four tables correspond to the four intertemporal specifications (pooled cross-sections, random effects, autoregressive errors, and random effects plus autoregressive errors). The two or three panels in each table pertain to the correlation pattern across alternatives: the leftmost panel relate to the IIA assumption, the rightmost panel to a full MNP model. In the models with random effects, the middle panel reports on the estimation with correlated random effects. For each variable we measure (1) the relative influence on the likelihood of living alone relative to the likelihood of becoming institutionalized (e.g., AGE1), and (2) the relative influence on the likelihood of living with others relative to the likelihood of becoming institutionalized (e.g., AGE2).

We first comment on the cross-sectional results, Table 6. Four variables describe the influence of demographic characteristics on the living arrangement choices of the elderly person. Age per se decreases both the likelihood of living alone and of living with others relative to the likelihood of becoming institutionalized, holding all other variables constant, particularly health. Female elderly are more likely to live alone. The number of children considerably increases the likelihood of a shared living arrangement. These results are as expected. A surprising result, however, is the insignificance of the indicator variable for being married.

Three variables measure health. While neither the subjective health rating (denoted by "SUBJ") nor the score of diagnosed conditions (denoted by "ILLSUM") predict living arrangement choices very well, the score of functional ability (denoted by "ADLSUM") is by far the most significant variable. The performance of the functional ability index confirms the results of most health-oriented studies of institutionalization.[9] The poor performance of subjective health ratings in predicting living arrangement choices is perhaps not so surprising given that this variable exhibits, on average, very little change over time, in spite of distinct changes over time in average functional ability scores (see Table 4).

The results reveal a significant income effect. The higher the income of the elderly person, the less likely they are to be institutionalized. The direction of the income effect is in line with most previous studies, although many studies fail to measure this income effect with much precision.[10] It is quite difficult to construct a variable measuring the relative costs of ambulatory and institutional care for the included Massachusetts communities. Hence, there are no prices included in our estimation.

In the right panel of Table 6, two contemporary covariance terms are estimated. The IIA-assumption of the left panel is clearly rejected as can be seen by the large difference in the loglikelihood values. The unobserved component in the utility of living independently exhibits significantly less variation than in the utility of the other two choices. Note that the t-statics are measured around the null hypotheses $\sigma(\nu_i)=1$, $corr(\nu_i,\nu_j)=0$ for $i \neq j$, and relate to the following reparametrized values: the t-statistic of $\sigma(\nu_i)$ refers to $\exp(\sigma(\nu_i))$, and the t-statistic of $corr(\nu_i,\nu_j)$ refers to $(\exp(corr(\nu_i,\nu_j))-1)/(\exp(corr(\nu_i,\nu_j))+1)$. This parametrization implicitly imposes the inequalities $\sigma(\nu_i)\geq 0$ and $|corr(\nu_i,\nu_j)|\leq 1$.

The coefficient estimates remain qualitatively unchanged when the IIA-assumption is dropped in favor of a cross-sectional, multinomial probit analysis. However, some coefficients change their relative numerical magnitudes. The income effect, to take just one example, is strengthened relative to the influence of the measure of functional ability.

We now put the panel structure into place. Introduction of random effects (see Table 7) dramatically raises the pseudo-R^2 to almost 60 percent. Some of the time-invariant characteristics become less significant, while the time-varying variables come out much stronger. Such an effect might be expected because the time-varying variables have falsely captured some effects in each cross-section that are now attributed to the random effects. Note that time-invariant characteristics are identified in the random effects model as opposed to a fixed effects specification.

In Table 8, autoregressive error components link the different waves instead of random effects. Finally, Table 9 reports on the full model, where the random effects are augmented by two autoregressive error components. The autocorrelation coefficients ρ_i are highly significant, and they drastically reduce the significance of the random effect terms in the combined specification, Table 9. However, they do not replace the random effects. While they are close to one, the large t-statistics imply that they are significantly different from one. In addition, the likelihood ratio statistic shows a significant difference between the specification in Table 9 with those in Tables 7 and 8. We conclude that the unobserved utilities determining living arrangements of the elderly include both time-invariant and time-varying components. The panel is too short, however, to precisely separate the two error structures as is evident by the high standard errors of the random effect terms at the bottom of Table 9.

The demographic, health, and income variables are remarkably stable across the different specifications of the covariance matrix, in spite of their different fits in terms of achieved likelihood values and quite different numerical values of covariance elements. This stability pertains both to alternative intertemporal and inter-alternative correlation patterns. The likelihood of living independently decreases dramatically with age, even after correcting for the decline in health and functional ability, as measured by the variables "ADLSUM" and "ILLSUM". The gender gap -- elderly men are more likely to live in institutions, elderly women are more likely to live independently -- is evident across all specifications. As opposed to other studies, elderly women are also more likely to live with children.[11] The larger the number of living children, the more probable is living together with one of them.

Among the health variables, the simple functional ability index employed in this paper performs best. It is the most significant variable in the model. In the presence of this variable, subjective health ratings have no predictive power whatsoever. The simple index of diagnosed conditions is weakly significant, but a more detailed analysis of the included illnesses may produce better results.

Finally, economics does matter. The income effect is measured precisely and robustly across all specifications. It is slightly underestimated in cross-sectional analysis, and slightly overestimated in the pure random effects model.[12] Those elderly with higher incomes choose institutions less frequently. Gauged by this willingness to spend income not to enter an institution, institutions appear to be an inferior living arrangement. The elderly's income may be spent on ambulatory care, thereby making living independently feasible in spite of declining functional ability. The ability to buy ambulatory services out of the elderly's income may also increase the likelihood of living with children rather than becoming institutionalized because these services substitute some of the burden that otherwise rests solely on the children. In addition, income may be spent on avoiding institutionalization by making transfer payments to children so that the children are more willing to take in their parents.[13] The results also suggest that increasing income of the elderly does not raise their probability of living alone relative to the probability of living with their children.

6. CONCLUDING REMARKS

The simulated likelihood method works well and requires a very small number of replications. It easily accommodates highly complex error structures and can handle different error structures without major programming effort.

Two main conclusions follow from the estimation results. First, a careful specification of the temporal error process dramatically improves the fit. It also appears that ignoring intertemporal linkages does bias some estimation results numerically, although the different specifications produce qualitatively similar coefficients of the substantive parameters.

Second, living arrangements choices are predominantly governed by functional ability, and, to a lesser degree (but still statistically and numerically significant) by age. The income effect is measured precisely and robustly. Institutions are an inferior living arrangement as measured by the willingness to spend income not to enter an institution. A somewhat surprising result is that changes in marital status do not appear to matter a great deal. The only supply factor that is included in our analysis, the number of living children, is, as can be expected, a significant factor for choosing shared living arrangements.

There are several weak points in the statistical analysis. The autoregressive specification "solves" the initial value problem by invoking a stationarity assumption. This is unsatisfactory, particularly with a short panel, such as in this application. It is possible to estimate a simple non-parametric specification of the initial value distribution, although in practice, the random effects should capture a great deal of these effects.

The sample is selective because it includes only survivors. Whether this sample selection is innocent in the sense of not biasing the estimated coefficients remain to be studied. There is no problem if the choice of a living arrangement leaves mortality and morbidity probabilities unaffected.

If, however, mortality and morbidity are, ceteris paribus, higher in nursing homes (e.g., because of inferior treatment), there is a serious sample selection problem.

Our panel of five waves is short. The identification difficulties apparent in Table 9 are indicative of this short panel length. However, the dramatic differences in goodness-of-fit indicate that, even in a short panel, the rewards for controlling for intertemporal linkages are quite sizeable.

FOOTNOTES

1. Examples are quoted in Börsch-Supan, 1986.

2. Including some rule to break ties.

3. X_{it} is a row vector, while β is a column vector.

4. This was pointed out first by Paul Ruud.

5. The matrix multiplications and the Cholesky-decomposition in (12) require operations that are of higher order. However, the generation of random numbers takes more computing time than these matrix operations even for reasonably large dimensions.

6. Garber (1988) presents evidence on the distribution of lengths of stay in a nursing home.

7. The pseudo-R^2 is defined as 1-(actual likelihood)/(likelihood at zero coefficients and identity covariance matrix).

8. Significance as measured by the likelihood ratio statistic.

9. See Garber, 1988, for a survey of health-oriented studies of institutionalization.

10. See Börsch-Supan, Kotlikoff, and Morris, 1989, for a survey.

11. Börsch-Supan, Kotlikoff, and Morris, 1989, report the opposite for the same basic data set, but a much less selected sample.

12. These differences are not statistically significant.

13. See Kotlikoff and Morris (1988) on this "bribery" hypothesis.

REFERENCES

Börsch-Supan, A., 1986, Household Formation, Housing Prices, and Public Policy Impacts, Journal of Public Economics 25.

Börsch-Supan, A., 1988, Elderly Americans: A Dynamic Analysis of Household Dissolution and Living Arrangement Transitions, Paper presented at The Conference on Issues in the Economics of Aging, Phoenix, AZ.

Börsch-Supan, A., Kotlikoff, L., and Morris, J., 1989, "The Dynamics of Living Arrangements of the Elderly: Health and Family Support." National Bureau of Economic Research Working Paper

Börsch-Supan, A., and Hajivassiliou, V., 1990, "Smooth Unbiased Multivariate Probability Simulators for Limited Dependent Variable Models" Cowles Foundation Working Paper, Yale University.

Chamberlain, G., 1984, Panel Data, in: Zvi Griliches and Michael D. Intriligator (eds.), Handbook of Econometrics, vol. 2, Amsterdam: North Holland.

Daganzo, C., 1981, Multinomial Probit, New York: Academic Press.

Ellwood, D.T., and Kane, T.J., 1988, The American Way of Aging: An Event History Analysis, Paper presented at The Conference on Issues in the Economics of Aging, Phoenix, AZ.

Garber, A.M., 1988, Predicting Nursing Home Utilization Among the High-Risk Elderly, Paper presented at The Conference on Issues in the Economics of Aging, Phoenix, AZ.

Geweke, J., 1989, Efficient Simulation from the Mutlivariate Normal Distribution Subject to Linear Inequality Constraints and the Evaluation of Constraint Probabilities, Duke University Working Paper.

Hajivassiliou, V., and McFadden, D., 1990, The Method of Simulated Scores for the Estimation of LDV Models: With An Application to the Problem of External Debt Crises, Discussion Paper, Yale University.

Kotlikoff, L.J., and Morris,J., 1987, How Much Care Do the Aged Receive from their Children? A Bimodal Picture of Contact and Assistance, NBER Working Paper No. 2391.

Kotlikoff, L.J., and Morris,J., 1988, Why Don't the Elderly Live With their Children? A New Look, Paper presented at The Conference on Issues in the Economics of Aging, Phoenix, AZ.

Langdon, M.G., 1984, Methods of Determining Choice Probability in Utility Maximising Multiple Alternative Models, Transportation Research B 18.

Lerman, S., and Manski, C., 1981, On the Use of Simulated Frequencies to Approximate Choice Probabilities, in: Manski, C., and McFadden, D., Structural Analysis of Discrete Data with Econometric Applications, Cambridge, Mass.: MIT-Press.

McFadden, D., 1984, Qualitative Response Models, in: Z. Griliches and M. Intriligator, Handbook of Econometrics, Amsterdam: North Holland.

McFadden, D., 1989, A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration, Econometrica 57.

Morris, John M., Claire E. Gutkin, Clarence C. Sherwood, and Ellen Bernstein. 1987. "Interest in Long Term Care Insurance." Final Report in connection with HCFA Cooperative Agreement no. 18-C-98375/1, June.

Pakes, A., and Pollard, D., 1989, Simulation and the Asymptotics of Optimization Estimators, Econometrica 57.

Table 1: DEMOGRAPHIC CHARACTERISTICS

(1) AGE DISTRIBUTION AT BASELINE 1982

	60+	65+	70+	75+	80+	85+	90+	95+	100+
Count	212	233	231	985	826	400	150	32	8
Percent	6.9	7.6	7.5	32.0	26.8	13.0	4.9	1.0	0.3

(2) MARITAL STATUS

	1982	1984	1985	1986	1987
Married	32.9	29.3	28.6	26.7	
Widowed	55.0	58.8	59.4	61.4	
Never Married	8.2	8.1	8.2	8.3	
Divorced/Sep.	3.9	3.7	3.7	3.6	

(3) NUMBER OF CHILDREN IN 1986

NUMBER OF CHILDREN:

	0	1	2	3	4	5	6	7	8+
Count	1275	468	549	392	189	87	51	31	35
Percent	41.4	15.2	17.8	12.7	6.1	2.8	1.7	1.0	1.1

(4) ISOLATED ELDERLY

Percentage of Elderly in 1986 without ...

Children	Siblings	Children or Siblings	Any Relatives	Friends	Any Relatives or Friends
41.4	53.0	31.2	25.5	39.3	24.5

Source: HRCA Survey of the Elderly, Working Sample of 3077 Elderly

Table 2: LIVING ARRANGEMENTS OF THE ELDERLY
(Percentages)

	1982	1984	1985	1986	1987
Independent Living Arrangements:					
Alone:	56.8	51.2	50.5	46.4	
With Spouse:	18.5	14.0	11.9	10.8	
Total:	75.3	65.2	62.4	57.2	
Shared Living Arrangements:					
Alone With Kids:	16.6	17.4	15.7	13.7	
With Spouse and Kids:	1.4	1.7	1.8	1.8	
Other Relatives or Nonrelatives Present:	5.9	5.9	5.7	5.1	
Total:	23.9	25.0	23.2	20.6	
Institutional Living Arrangements:					
Convent, Rectory, CCRC, Congregate Housing, or Retirement Home:	0.0	0.2	0.7	0.6	
Foster Home, Community or Domestic Care:	0.0	0.2	0.2	0.3	
Nursing Home (ICF):	0.2	5.4	8.0	11.6	
Nursing Home (SNF):	0.0	2.9	3.5	7.0	
Rest Home (level IV):	0.0	0.4	0.7	1.3	
Hotel, Boarding or Rooming House: •	0.6	0.3	0.3	0.2	
Hospital:	0.0	0.4	1.1	1.2	
Total:	0.8	9.8	14.5	22.2	
Number of Observations:	3070	2965	1130	2331	

Source: HRCA Survey of the Elderly (cross-sectional subsamples of elderly with completed interviews)

Table 3: LIVING ARRANGEMENT SEQUENCES 1982, 1984, 1985, 1986

Sequence	IIII	IIIC	IIIO	IIIN	IIID	IICI	IICC	IICN	IIOI
Count	474	17	6	40	3	1	8	2	2
Percent	39.63	1.42	0.50	3.34	0.25	0.08	0.67	0.17	0.17
Sequence	IIOO	IION	IINI	IINN	IIND	IIDD	ICII	ICIN	ICCC
Count	1	3	1	42	1	110	1	1	20
Percent	0.08	0.25	0.08	3.51	0.08	9.20	0.08	0.08	1.67
Sequence	ICCN	ICOO	ICNN	ICDD	IOII	IOIO	IOCN	IOOI	IOOO
Count	2	1	4	6	1	1	1	3	6
Percent	0.17	0.08	0.33	0.50	0.08	0.08	0.08	0.25	0.50
Sequence	IONN	IODD	INCC	INNO	INNN	INND	INDD	IDDD	CIII
Count	2	4	1	1	47	2	26	74	3
Percent	0.17	0.33	0.08	0.08	3.93	0.17	2.17	6.19	0.25
Sequence	CIIC	CIIO	CIDD	CCII	CCCI	CCCC	CCCO	CCCN	CCCD
Count	1	1	1	6	6	87	4	18	1
Percent	0.08	0.08	0.08	0.50	0.50	7.27	0.33	1.51	0.08
Sequence	CCNN	CCDD	CODD	CNII	CNNN	CNDD	CDDD	OIII	OINN
Count	8	36	1	1	12	7	11	6	1
Percent	0.67	3.01	0.08	0.08	1.00	0.59	0.92	0.50	0.08
Sequence	OCCC	OCCN	OCNN	OCDD	OOIN	OOCI	OOCC	OOCO	OOCN
Count	2	1	2	1	1	2	1	11	2
Percent	0.17	0.08	0.17	0.08	0.08	0.17	0.08	0.92	0.17
Sequence	OOOO	OOON	OONI	OONN	OODD	ONNN	ONDD	ODDD	NIII
Count	7	1	1	6	9	4	3	7	1
Percent	0.59	0.08	0.08	0.50	0.75	0.33	0.25	0.59	0.08
Sequence	NICC	NICN	NIDD	NCNN	NNNN				
Count	1	1	1	1	4				
Percent	0.08	0.08	0.08	0.08	0.33				

Source: HRCA Survey of the Elderly (1196 Elderly, excludes elderly not interviewe or without ascertained living arrangement in at least one wave)

Notes: Living Arrangements are denoted by:
I=Lives Independently, C=Lives with Children, O=Lives with Other Relatives or Nonrelatives, N=Lives in Nursing Home, D=Dead.

Table 4: VARIABLE DEFINITIONS AND STATISTICS IN LONGITUDINAL SUBSAMPLE

DEPENDENT VARIABLE:

Choice	Definition	Sample Frequency 1982	1984	1985	1986	1987
1	Independent Living Arrangements	.790	.742	.732	.697	.643
2	Shared Living Arrangements	.210	.229	.220	.236	.223
3	Institutional Living Arrangements	.000	.029	.048	.067	.134
Number of Observations:		314	314	314	314	314

EXPLANATORY VARIABLES:

Variable	Definition	Sample Average 1982	1984	1985	1986	1987
AGE	age of elderly person	78.2	80.2	81.2	82.2	83.2
FEMALE	1 if female, 0 if male	0.85	0.85	0.85	0.85	0.85
KIDS	number of living children	2.31	2.31	2.31	2.31	2.31
MARRIED	1 if married 0 if widowed or not married	.178	.134	.121	.115	.105
SUBJ	subjective health rating	2.74	2.65	2.60	2.64	2.65
ADLSUM	score of functional disability	5.25	5.75	5.82	6.27	7.38
ILLSUM	score of diagnosed conditions	3.47	3.40	3.70	3.98	4.12
INCOME	real annual income (in $1000 1987)	6.10	6.18	6.27	6.85	7.19

Note: Each explanatory variable is interacted with choice 1 (=living independently) and choice 2 (=living with children or others) while choice 3 (=living in an institution) is the base category.

Table 5: ESTIMATION RESULTS: GOODNESS OF FIT
(Loglikelihood values, pseudo-R^2 in parentheses)

Pooled Cross-Sections ($\epsilon_{i,t} = \nu_{i,t}$):

IIA	MNP
-996.46	-957.94
(0.422)	(0.445)

Random Effects Included ($\epsilon_{i,t} = \alpha_i + \nu_{i,t}$):

IIA	RE-Corr	MNP
-715.70	-711.79	-671.93
(0.585)	(0.587)	(0.610)

First Order Autoregressive Errors Included ($\epsilon_{i,t} = \rho_i \cdot \epsilon_{i,t-1} + \nu_{i,t}$):

IIA	MNP
-673.72	-652.74
(0.609)	(0.622)

Random Effects and First Order Autoregressive Errors Included:

($\epsilon_{i,t} = \alpha_i + \eta_{i,t}, \quad \eta_{i,t} = \rho_i \cdot \eta_{i,t-1} + \nu_{i,t}$)

IIA	RE-Corr	MNP
-648.07	-647.60	-632.45
(0.624)	(0.625)	(0.633)

Note: Three different specifications of correlations across alternatives are employed, denoted by:

IIA: Independence of irrelevant alternatives imposed i.e., $\sigma(\nu_i,\nu_j)=\sigma(\alpha_i,\alpha_j)=0$.

RE-Corr: Random effects correlated i.e., $\sigma(\alpha_i,\alpha_j)\neq 0$, $\sigma(\nu_i,\nu_j)=0$.

MNP: Unobserved time-specific utility components correlated i.e., $\sigma(\nu_i,\nu_j)\neq 0$, $\sigma(\alpha_i,\alpha_j)=0$.

Table 6: POOLED CROSS-SECTIONAL PROBIT ESTIMATES

Error structure: $\epsilon_{i,t} = \nu_{i,t}$

Variable	IIA (Specification A) Estimate	IIA (Specification A) t-stat	MNP (Specification E) Estimate	MNP (Specification E) t-stat
AGE1	-.0319	-2.64	-.0234	-2.87
AGE2	-.0169	-1.39	-.0159	-1.87
FEMALE1	.4490	1.81	.3687	1.72
FEMALE2	.4163	1.56	.3102	1.38
KIDS1	.0447	.99	.0624	1.54
KIDS2	.1325	2.86	.1258	2.86
MARRIED1	.4243	1.21	.1870	.66
MARRIED2	-.3468	-.92	-.3640	-1.20
SUBJ1	.1263	1.08	.0843	.81
SUBJ2	-.0658	-.54	-.0333	-.29
ADLSUM1	-.2343	12.58	-.1769	10.08
ADLSUM2	-.1239	-6.61	-.1132	-5.22
ILLSUM1	-.0256	-.66	-.0242	-.68
ILLSUM2	-.0195	-.48	-.0139	-.36
INCOME1	.0788	2.45	.0809	2.61
INCOME2	.0922	2.86	.0905	2.92
CONSTANT1	5.5292	4.92	4.1058	5.65
CONSTANT2	2.7875	2.45	2.5686	3.26
stddev(ν_1)	1.0000	(fix)	.2834	-2.36
corr(ν_1,ν_2)	.0000	(fix)	.4465	1.72
loglik	-996.46		-957.88	
loglik at zero	-1724.82		-1724.82	
pseudo-R^2	42.23%		44.46%	
nobs	1570		1570	

<u>Note:</u> In this and the following tables, the t-statistics of the elements of the covariance matrix refer to the re-parametrized estimated values. They are evaluated around zero for correlations and around one for standard deviations.

Table 7: RANDOM EFFECTS PROBIT MODEL

Error Structure: $\epsilon_{i,t} = \alpha_i + \nu_{i,t}$

	IIA (Spec. B)		RE-Corr (Spec. F1)		MNP (Spec. F2)	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
AGE1	-.0570	-2.64	-.0604	-3.05	-.0643	-3.50
AGE2	-.0307	-1.22	-.0311	-1.40	-.0360	-1.79
FEMALE1	.5597	1.38	.4370	1.11	.7641	2.21
FEMALE2	1.0004	1.82	1.2543	2.37	.8631	2.16
KIDS1	.0329	.38	.0094	.12	.0586	.78
KIDS2	.2235	2.16	.2036	2.24	.1398	1.73
MARRIED1	.6279	1.29	.5589	1.20	.3121	.73
MARRIED2	.2165	.38	.1706	.31	-.1039	-.22
SUBJ1	.0889	.50	.1023	.60	.0521	.33
SUBJ2	-.1938	-1.00	-.2192	-1.18	-.0756	-.46
ADLSUM1	-.2985	-11.28	-.2850	-11.05	-.2472	-10.12
ADLSUM2	-.1824	-6.24	-.1716	-6.04	-.1981	-7.17
ILLSUM1	-.0905	-1.53	-.0977	-1.73	-.0900	-1.66
ILLSUM2	-.0743	-1.10	-.0741	-1.16	-.0704	-1.23
INCOME1	.1190	2.28	.1149	2.30	.0988	2.29
INCOME2	.1361	2.59	.1328	2.64	.1074	2.47
CONSTANT1	9.2564	4.71	9.3513	5.12	8.9092	5.21
CONSTANT2	3.9987	1.75	3.4848	1.68	5.2459	2.78
$stddev(\nu_1)$	1.0000	(fix)	1.0000	(fix)	.5833	-2.79
$corr(\nu_1,\nu_2)$	.0000	(fix)	.0000	(fix)	.7485	4.81
$stddev(\alpha_1)$	1.1305	1.03	.9650	-.29	.7386	-2.21
$stddev(\alpha_2)$	1.9847	7.93	1.7468	5.23	1.1366	.71
$corr(\alpha_1,\alpha_2)$	.0000	(fix)	-.5495	-3.18	.0000	(fix)
loglik		-717.79		-711.79		-671.93
loglik at zero		-1724.82		-1724.82		-1724.82
pseudo-R^2		58.38%		58.73%		61.04%
nobs		1570		1570		1570

Table 8: PROBIT MODEL WITH AUTOREGRESSIVE ERRORS

Error Structure: $\epsilon_{i,t} = \rho_i \cdot \epsilon_{i,t-1} + \nu_{i,t}$

Variable	IIA (Specification C) Estimate	t-	MNP (Specification G) Estimate	t-stat
AGE1	-.0458	-3.23	-.0368	-2.51
AGE2	-.0237	-1.63	-.0033	-.16
FEMALE1	.2286	.91	.4414	1.79
FEMALE2	.6579	2.27	.6295	1.56
KIDS1	.0176	.34	.0541	.97
KIDS2	.1351	2.50	.1801	2.50
MARRIED1	.1352	.44	.2048	.66
MARRIED2	-.1184	-.35	-.3845	-.93
SUBJ1	-.0146	-.12	.0100	.08
SUBJ2	-.1266	-1.03	-.1055	-.72
ADLSUM1	-.1972	-11.06	-.1953	-8.15
ADLSUM2	-.1419	-7.83	-.1286	-4.92
ILLSUM1	-.0464	-1.18	-.0300	-.70
ILLSUM2	-.0511	-1.24	-.0285	-.55
INCOME1	.0635	2.06	.0910	2.36
INCOME2	.0694	2.25	.1007	2.58
CONSTANT1	7.2253	5.66	5.6732	4.08
CONSTANT2	3.6772	2.79	.8886	.45
$stddev(\nu_1)$	1.0000	(fix)	.2678	-3.27
$corr(\nu_1,\nu_2)$	.0000	(fix)	.0137	.08
ρ_1	.9278	10.40	.9065	7.53
ρ_2	.8059	15.56	.8648	19.13
loglik		-673.73		-652.74
loglik at zero		-1724.82		-1724.82
pseudo-R^2		60.94%		62.16%
nobs		1570		1570

Table 9: RANDOM EFFECTS PROBIT MODEL WITH AUTOREGRESSIVE ERRORS

Error Structure: $\epsilon_{i,t} = \alpha_i + \eta_{i,t}$, $\eta_{i,t} = \rho_i \cdot \eta_{i,t-1} + \nu_{i,t}$

	IIA (Spec. D)		RE-Corr (Spec. H1)		MNP (Spec. H2)	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
AGE1	-.0646	-3.96	-.0644	-3.74	-.0513	-3.60
AGE2	-.0421	-2.32	-.0424	-2.25	-.0279	-1.43
FEMALE1	.6071	1.80	.6237	1.84	.5791	1.90
FEMALE2	.9769	2.41	.9257	2.24	.7492	1.62
KIDS1	.0469	.66	.0500	.71	.0465	.79
KIDS2	.1554	1.96	.1534	1.94	.1666	1.99
MARRIED1	.1969	.50	.1960	.49	.2004	.57
MARRIED2	-.1502	-.34	-.1549	-.35	-.3729	-.83
SUBJ1	.0461	.32	.0421	.29	.1059	.79
SUBJ2	-.0724	-.47	-.0683	-.44	-.0450	-.28
ADLSUM1	-.2358	-10.01	-.2356	-10.09	-.2201	-10.50
ADLSUM2	-.1811	-7.27	-.1826	-7.29	-.1612	-6.35
ILLSUM1	-.0848	-1.67	-.0843	-1.67	-.0864	-1.89
ILLSUM2	-.0694	-1.26	-.0703	-1.28	-.0718	-1.28
INCOME1	.0866	2.11	.0869	2.06	.0892	2.23
INCOME2	.0943	2.29	.0942	2.22	.0987	2.44
CONSTANT1	8.9868	6.30	8.9608	5.88	7.2120	5.59
CONSTANT2	5.2089	3.25	5.3660	3.21	3.3559	1.92
stddev(ν_1)	1.0000	(fix)	1.0000	(fix)	.0278	-3.77
corr(ν_1,ν_2)	.0000	(fix)	.0000	(fix)	-.3898	-2.59
stddev(α_1)	.0027	-.14	.1288	-1.98	.0022	-.16
stddev(α_2)	1.0582	.34	1.0239	.13	.0054	-.16
corr(α_1,α_2)	.0000	(fix)	1.0000	.05	.0000	(fix)
ρ_1	.9499	7.87	.9571	6.87	.9865	2.75
ρ_2	.6692	7.67	.6946	7.08	.8719	20.54
loglik		-648.07		-647.60		-632.45
loglik at zero		-1724.82		-1724.82		-1724.82
pseudo-R^2		62.43%		62.46%		63.33%
nobs		1570		1570		1570

Table 10: COVARIANCE MATRIX OF RANDOM UTILITY TERM IN SPECIFICATION H

Error Structure: $\epsilon_{i,t} = \alpha_i + \eta_{i,t}$,

$$\eta_{i,t} = \rho_i \cdot \eta_{i,t-1} + \nu_{i,t}, \quad i=1,\ldots,I-1,$$

where $\mathrm{corr}(\nu_{i,t},\nu_{j,s}) = \begin{cases} 0 & \text{if } t \neq s \\ \omega_{ij} & \text{if } t=s \end{cases}$ and $\mathrm{cov}(\alpha_i,\alpha_j) = \delta_{ij}$

implying

$$\mathrm{cov}(\epsilon_{i,t},\epsilon_{j,s}) = \delta_{ij} + \rho_i^{(t-s)} \frac{\sqrt{(1-\rho_i^2)}\cdot\sqrt{(1-\rho_j^2)}}{1-\rho_i\rho_j} \omega_{ij}, \quad i,j=1,\ldots,I-1.$$

	t=	1			2			3			4			5		
s	j	i=1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
	1	.03	-.08	.0	.03	-.07	.0	.03	-.06	.0	.03	-.05	.0	.03	-.04	.0
1	2	-.08	4.17	.0	-.08	3.64	.0	-.08	3.17	.0	-.07	2.76	.0	-.07	2.41	.0
	3	.0	.0	2.0	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0
	1	.03	-.08	.0	.03	-.08	.0	.03	-.07	.0	.03	-.06	.0	.03	-.05	.0
2	2	-.07	3.64	.0	-.08	4.17	.0	-.08	3.64	.0	-.08	3.17	.0	-.07	2.76	.0
	3	.0	.0	1.0	.0	.0	2.0	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0
	1	.03	-.08	.0	.03	-.08	.0	.03	-.08	.0	.03	-.07	.0	.03	-.06	.0
3	2	-.06	3.17	.0	-.07	3.64	.0	-.08	4.17	.0	-.08	3.64	.0	-.08	3.17	.0
	3	.0	.0	1.0	.0	.0	1.0	.0	.0	2.0	.0	.0	1.0	.0	.0	1.0
	1	.03	-.07	.0	.03	-.08	.0	.03	-.08	.0	.03	-.08	.0	.03	-.07	.0
4	2	-.05	2.76	.0	-.06	3.17	.0	-.07	3.64	.0	-.08	4.17	.0	-.08	3.64	.0
	3	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0	.0	.0	2.0	.0	.0	1.0
	1	.03	-.07	.0	.03	-.07	.0	.03	-.08	.0	.03	-.08	.0	.03	-.08	.0
5	2	-.04	2.41	.0	-.05	2.76	.0	-.06	3.17	.0	-.07	3.64	.0	-.08	4.17	.0
	3	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0	.0	.0	1.0	.0	.0	2.0