LIVING ARRANGEMENTS: HEALTH AND WEALTH EFFECTS

Axel Börsch-Supan

Daniel McFadden

Reinhold Schnabel

Working Paper No. 4398

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1993

The first author appreciated the hospitality at the University of California at Berkeley. Financial Assistance was provided by NIA, grant 3 PO AG05842-01. We thank Ed Norton who cleaned up the NBER Economic Supplement. 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 #4398
July 1993

LIVING ARRANGEMENTS: HEALTH AND WEALTH EFFECTS

ABSTRACT

This paper investigates the choice of living arrangements among elderly Americans. It has two specific aims. First, because health is not directly measurable and can only be described by indicators such as *ADLs* and *IADLs*, it explores a new econometric approach to model the influence of the latent health status on living arrangements. Second, it exploits the NBER Economic Supplement of the Longitudinal Study on Aging to investigate the role of housing and financial wealth in the choice of living arrangements.

Axel Börsch-Supan
Department of Economics
University of Mannheim
D-68131 Mannheim
GERMANY
TEL: 49/621/292-5183
and Center for Economic
Policy Research, UK
and NBER

Daniel McFadden
University of California
at Berkeley
Evans Hall
Berkeley, CA 94720
TEL: (510) 643-8428
and NBER

Reinhold Schnabel
Department of Economics
University of Mannheim
D-68131 Mannheim
GERMANY
TEL: 49/621/292-3181

1. Introduction

The choice of a living arrangement -- as an independent household, with adult children or other related or unrelated persons, or in an institution -- has many implications for the well-being of an elderly person. Changes in living arrangements are likely to be associated with changes in the level of care and assistance received by the elderly. Living together with other family members eases situations of illness while living alone makes coping with illnesses harder. Thus, the choice of living arrangements has many external effects. Moreover, living arrangements commonly affect the elderly's eligibility for certain types of government assistance, such as food stamps and supplemental social security, and induces demand for social support services such as district nursing, meals-on-wheels etc. Finally, the change of living arrangements frequently involves the sale of the home by the elderly and may therefore dramatically change the liquid wealth of the elderly. On the other hand, if the elderly tend to stay longer living independently, the balance of the housing market changes because housing becomes relatively more scarce due to the increased length of stay in the family home by the older generation. In short, it is important to understand the determinants of the living arrangement choice.

There is a long line of literature investigating the determinants of living arrangements of the aged. Schwartz, Danziger and Smolensky (1984) employ the Retirement History Survey (RHS) to estimate a binary choice model between living independently and dependently. Their empirical results were mixed, and neither health nor income effects are very strong. Börsch-Supan (1989) estimates a multinomial logit model of living arrangements using data from the Annual Housing

Survey (AHS). As in the paper by Schwartz, Danziger and Smolensky, the data preclude an analysis of institutionalization. In contrast, Garber (1990) concentrates on the determinants of institutionalization and its length using the Channeling Demonstration, while Kotlikoff and Morris (1987, 1990) and Börsch-Supan, Gokhale, Kotlikoff and Morris (1991) analyze the importance of family links in forming multigenerational households.

Papers by Ellwood and Kane (1990), Börsch-Supan (1990), and Börsch-Supan, Hajivassiliou, Kotlikoff and Morris (1991) represent more comprehensive analyses of living arrangements that include both institutionalized and non-institutionalized elderly. All three papers find an increasing proportion of elderly living alone and contribute this to the positive income-elasticity of privacy.

These studies leave several questions unanswered. First, most studies of living arrangements suffer from a less than satisfactory description of health. This is partly due to lack of data but the problem is deeper: Even when health is measured by indicators such as *Activities of Daily Living (ADLs)* and *Independent Activities of Daily Living (IADLs),* or by the presence of conditions such as cancer or Alzheimer's disease, or by simply asking the elderly how she feels, we do not really measure health but a concoction of subjective feelings and objective states that are correlated with health. In the language of econometrics, health is a latent, unmeasurable variable, for which we only observe a set of indicators. One goal of this paper is to develop an econometric framework in order to model this errors-in-variables problem in the discrete decision of living arrangements. We relate latent health to *ADLs* and *IADLs* by a nonlinear version of a multiple-indicator, multiple-cause *(MIMIC)* model which explicitly considers the categorical measurement of the health indicators. We estimate this model using data from the Longitudinal Study on Aging (LSOA).

Another important question which has not been answered is the role of wealth. Does housing wealth tie elderly to their home? This question extends the lock-in discussion (Feinstein and McFadden, 1990; Venti and Wise, 1990) to household formation. What is the role of financial wealth in the demand for old-age institutions? Wealth data is rarely available in elderly surveys, and if so, its value may be questionable. We will explore the NBER Economics Supplement of the LSOA in this respect which contains information on income and assets of the LSOA sample persons in 1990.

The paper is set up as follows. Section 2 introduces the data sources and presents descriptive statistics of our working sample. Estimates based on a standard discrete choice model are briefly described in section 3. In section 4 we discuss the econometric model and address the issues of identification and estimation, while section 5 presents the results and section 6 concludes.

2. The Longitudinal Study on Aging and the NBER Economic Supplement

The Longitudinal Study on Aging is a panel survey based on the 1984 Supplement of Aging to the National Health Interview Survey (NHIS). The National Health Interview Surveys are continuing surveys comprising each year about 100,000 non-institutionalized persons of all ages in about 40,000 households.[1] Interviews are held every week

---

1 See Kovar and Poe (1985).

throughout the year. The Supplement on Aging was added to the NHIS during the 1984 interviews. The SOA included questions on

- Family Structure
- Community and Social Support
- Occupation and Retirement
- Conditions and Impairment, *ADLs* and *IADLs*
- Structural Characteristics of Housing
- Regular Medical Care and Nursing Home Stay
- Health Opinions and Behavior

to all NHIS sample persons aged 65 years and over.[2] The questions were similar to those in the 1984 National Nursing Home Survey (NNHS), so that by combining the two data, estimates for the total elderly population would be possible.

The SOA was explicitly designed to be the first wave of the LSOA. In 1986, 1988 and 1990, all persons aged 70 and above in the 1984 SOA were reinterviewed by computer-assisted telephone interviews with mail follow-up.[3]

Records for participants who gave permission were also matched with the National Death Index and the Medicare Files maintained by the Health Care Finance Administration. While the first wave does not include the institutionalized elderly, sample persons were interviewed in the later waves even when they entered a nursing home or another institution.

In 1990, the NBER added an Economics Supplement to the LSOA. This supplement included a detailed account of personal income

2 See Fitti and Kovar (1987). The response rate to the SOA was 96.7 percent.

3 See NCHS (1991).

sources for each sample person, an inventory of assets including financial and real wealth, and questions about structural housing characteristics. Response rates to these questions were smaller than to the standard LSOA questions, and particularly small to the wealth questions.[4]

As a working sample, we selected only single elderly because almost all married elderly are living independently. In the 1990 cross section, this working sample consists of 2193 elderly between age 76 and 102. The average age in 1990 was almost 83 years.

Table 1 presents descriptive statistics of the most important variables. Even in this sample of the very old and non-married, 63 percent live by themselves. 28.7 percent live with their children, other relatives or non-relatives, and 8.2 percent live in institutions.

81.4 percent of the sample persons are female. The non-white population is underrepresented with only 9.2 percent. On average, the sample persons have 2 children still living.

The economic variables comprise income and wealth. Income is very low, the median is below \$ 2,400. 27.6 percent report no income at all. On the other hand, 63.4 percent have their own home, and except for less than 15 percent of the homeowners, this home is free and clear of mortgages. The Median value of the home is \$ 31,000, and the average value is about \$ 50,000. The discrepancy between mean and median is much larger for financial assets. The median financial assets sum up to only \$ 3,500, while the mean is ten times as large. These numbers are approximately in line with results from SIPP and other surveys (Venti and Wise, 1991)

---

4 The response rate to financial assets was 63.5 percent. Missing values were assigned by Edward Norton using a hot-deck method.

Table 1 also reports on a set of health indicators. We restrict our attention to functional health measures such as the ***activities of daily living (ADLs)*** and the ***independent activities of daily living (IADLs)*** which are measured in four categories (no, some, severe problems in doing *xyz*, and can not do *xyz* at all). The variables are coded such that ***higher*** values for *ADLs* and *IADLs* indicate *less* capability. Functional health indicators have been found most appropriate in describing living arrangements, and superior to subjective health ratings or indicators for the presence and severity of diagnose conditions (Börsch-Supan, Kotlikoff and Morris, 1992). Table 1 lists the percentages of sample persons who have no problems in performing a set of ten activities. *IADLs* were asked only for the non-institutionalized, *ADLs* for all sample persons. The pattern is familiar: most problems occur with walking, and the fewest with eating.

3. The Standard Approach: Multinomial Logit Analysis

Table 2 and 3 present results of a simple multinomial logit model, relating the choice of living arrangements to demographic, economic, and -- in table 3 -- also to the health indicators. Both versions of the discrete choice model show that educated persons are less likely to live with others or in nursing homes, and that the probability to live with others (mainly children) increases with the number of daughters but not significantly with the number of sons. Higher wealth increases the likelihood to live with children, while there is no significant wealth effect in institutionalization, except that the ownership of a house reduces the probability of entering a nursing home.

The contribution of the health indicators in table 3 is highly significant -- the log likelihood increases considerably and the likelihood ratio test statistic is 718.2. However, the inclusion of so many indicators results in multicollinearity and low t-statistics among the individual *ADLs*. This is one reason to contemplate using factor analysis in describing the effect of the health indicators. Exploratory factor analysis, taking the health indicators as if they were continuous indicators, shows that more than three quarters of the variance can be explained by only two factors.

The inclusion of the health indicators does not change the other parameters by a lot. The main exception is age, which becomes insignificant once the functional health measures are taken into account. In turn, personal income, which was insignificant when the health indicators were left out, increases in statistical importance with a negative effect on institutionalization and living with others.

These results essentially reproduce the estimates of Börsch-Supan, Kotlikoff and Morris (1992). This is helpful to know because the latter estimates were obtained from a geographically very restricted sample of Massachusetts elderly, the Hebrew Rehabilitation Center for the Aged (HRCA) sample. Knowing that the HRCA sample is representative at least in the respect of choosing living arrangements gives confidence in the other analyses that have been performed on the basis of this rich data set.[5]

---

[5] Kotlikoff and Morris (1987, 1990); Börsch-Supan, Kotlikoff and Morris (1989); Börsch-Supan, Gokhale, Kotlikoff and Morris (1992); Börsch-Supan, Hajivassiliou, Kotlikoff and Morris (1992).

4. An Econometric Model of the Influence of Latent Health

4.1. Model Specification

Obviously, the contribution of the health indicators in table 2 is highly significant. However, one might doubt that these indicators directly affect the choice of living arrangements. Rather, one might argue that it is the underlying but unobservable health status which affects both, the choice of living arrangements and the set of indicators. The problem boils down to the question of causal links among the four following groups of variables:

- the choice among N_a living arrangements, denoted by u,
- the latent health status, denoted by h^*, N_h-dimensional,
- the health indicators (*ADLs* and *IADLs*), denoted by y_k, $k=1,..,N_y$,
- the demographic and economic exogenous variables, denoted by z_j, $j=1,..,N_z$,

Figures 1 and 2 visualize the two approaches, using the above notation for the four variable groups, and distinguishing latent from observable variables by an asterisk. In addition to latent health, we have two more latent variables. First, the choice between the discrete alternatives u depends on the unobserved utility levels u_i^*, $i=1,..,N_a$. In our case N_a equals 3 (living independently, with others, or in an institution). A person chooses the living arrangement which yields the highest utility level u_i^*

$$(1) \qquad u = i \quad \Leftrightarrow \quad u_i^* = \max(u_j^*, j=1,..,N_a) \ .$$

This is the conventional random utility maximization hypothesis underlying discrete choice.

Second, we also do not precisely observe the health indicators because the sample persons are asked to report their performance in each activity using S_k ordinal categories (e.g., no, some, severe problems in walking, and can not walk at all) rather than a continuous scale. The relation between the k-th observed health indicator y_k and the underlying continuous y_k^* is described by thresholds $\mu_{k,j}$ which will be estimated:

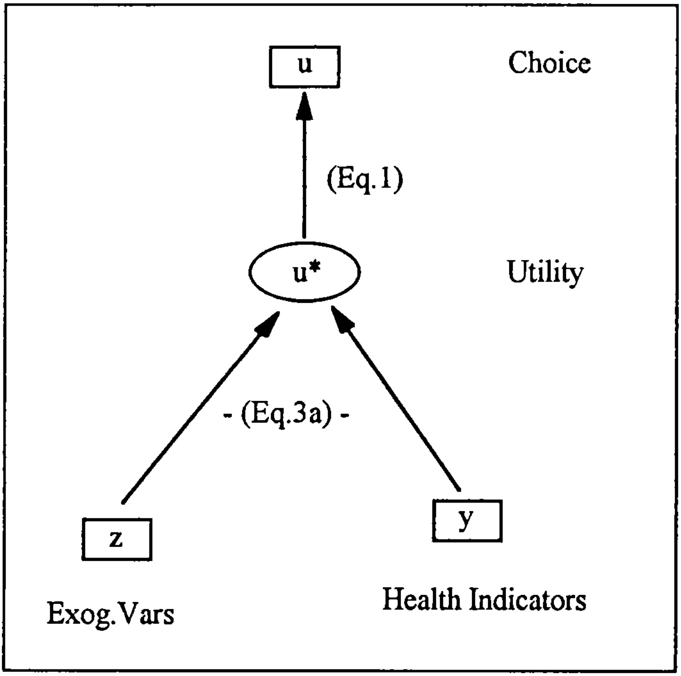
$$\text{(2)} \qquad y_k = j \iff \mu_{k,j-1} < y_k^* < \mu_{k,j}, \quad k = 1,\ldots,N_y,\ j = 1,\ldots,S_k.$$

In the discrete choice model of the preceding section, the choice of living arrangements is directly linked to the ordinal health indicators and to the exogenous variables (figure 1). Moreover, the transmission between ordinal measurement and continuous indicators (equation 2) is ignored. The unobserved utility levels u_i^* are therefore given by:

$$\text{(3a)} \qquad u_i^* = \tilde{\beta}_i' z + \tilde{\gamma}_i' y + \tilde{\varepsilon}_i, \qquad i = 2,\ldots,N_a,$$

where $\tilde{\varepsilon}_i$ denotes an additive error term in the utility of alternative i. Alternative 1 (living independently) is taken as the reference alternative.

Fig. 1: Multinomial Logit Model

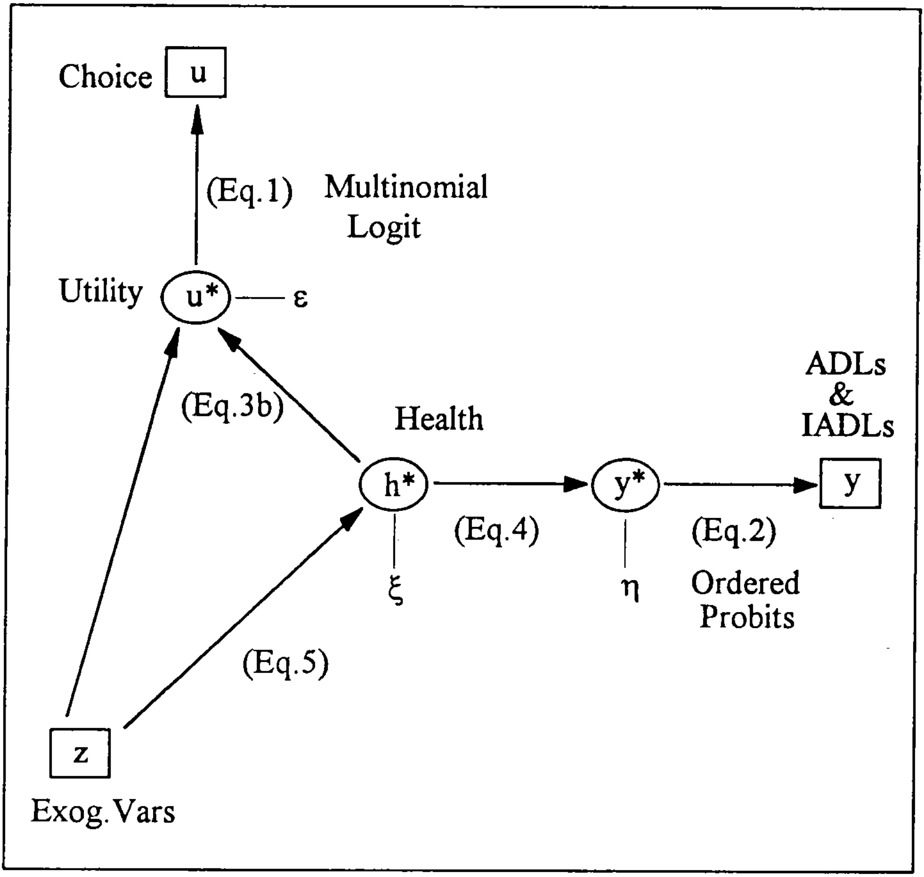


The multiple-indicators, multiple-causes (MIMIC) model endogenizes the indicators y_k. Latent health determines both, indicators and living arrangement choice. The model also takes the categorical measurement of the health indicators and the choice decision into account. Moreover, our MIMIC model distinguishes between the direct influence of the exogenous variables on the living arrangement choice, and the indirect influence via the latent health status.

The unobserved utility levels u_i^* are now determined by:

$$\text{(3b)} \quad u_i^* = \beta_i' z + \gamma_i' h^* + \varepsilon_i, \qquad i = 2,\ldots,N_a .$$

Fig. 2: Non-Linear MIMIC Model



Rather than taking the health indicators as given, they are now determined by the health status in a factor-analytic model:

(4) $$y_k^* = \lambda_k' h^* + \eta_k, \qquad k = 1,..,N_y .$$

Finally, a set of equations expresses the influence of the exogenous variables on the latent health status:

(5) $\quad h_m^* = \delta_m' z + \xi_m, \qquad\qquad m = 1,..,N_h,$

or in stacked form

$$\mathbf{h}^* = \Delta' \mathbf{z} + \xi .$$

One may interprete relation (5) as a production function of health. Due to progress in medical science this function may change over time.

The three sets of equations (3b), (4) and (5) form a nonlinear version of a LISREL model.[6] It is nonlinear in two respects. First, the main dependent variable, the choice of living arrangements, is described by a nonlinear discrete choice model which links the observed choices **u** to the latent utilities **u*** (equation 1). McFadden (1988) introduced this case of factor analysis in the presence of a discrete choice equation, and Morikawa, Ben-Akiva and McFadden (1990) present an application to travel demand.

Our model introduces a second nonlinearity with the additional complication of categorical indicators. The measurement equations (4), which link the indicators y^* with the health status **h*** via the factor loadings λ_k, are described by ordered probit models if we assume the η_k to be normally distributed.

---

6 See Jöreskog and Sörbom (1988).

By inserting (5) into (3b) and (4), we eliminate the health production equation and obtain two sets of reduced form equations on which our estimation will be based:

$$(6) \qquad u_i^* = \beta_i' z + \gamma_i' (\Delta z + \xi) + \varepsilon_i , \qquad i = 2,\ldots,N_a$$

$$= \pi_i' z + \gamma_i' \xi + \varepsilon_i ,$$

and similarly for the factor-analysis equations which determine the health indicators:

$$(7) \qquad y_k^* = \lambda_k' (\Delta z + \xi) + \eta_k , \qquad k = 1,..,N_y ,$$

$$= \psi_k' z + \lambda_k' \xi + \eta_k ,$$

where the reduced form parameters π_i and ψ_k are:

$$(8) \qquad \pi_i' = \beta_i' + \gamma_i' \Delta' \qquad \text{for } i = 2,3 \qquad \text{and}$$

$$\psi_k' = \lambda_k' \Delta' \qquad \text{for } k = 1,\ldots,N_y .$$

4.2. The Likelihood Function

We assume that the three groups of error terms ξ, ε, and η are mutually independent. Moreover, we assume that the ε are extreme-value distributed, resulting in a logit model for the choice equation 6. The η are assumed to be normal, resulting in N_y ordered probit models for the health indicators. The likelihood of an individual who has chosen alternative i and is characterized by the health indicators j_k, k=1...N_y, conditional on ξ_m, the error of the health equation, is therefore a product of the probabilities of a logit model and N_y ordered probit models:[7]

$$(9)\quad L(\beta,\gamma,\Delta,\lambda,\mu|\xi) = \frac{\exp(\beta_i' z + \gamma_i' \Delta' z + \gamma_i' \xi)}{1 + \exp\left(\sum_{j=2}^{N_a} (\beta_j' z + \gamma_j' \Delta' z + \gamma_j' \xi)\right)}$$

$$\times \prod_{k=1}^{N_y} \Big(\Phi(\mu_{k,j_k} - \lambda_k' \Delta' z - \lambda_k' \xi)$$

$$- \Phi(\mu_{k,j_k-1} - \lambda_k' \Delta' z - \lambda_k' \xi) \Big)$$

$$= LOGIT(i, z, \xi) \prod_{k=1}^{N_y} ORDPROBIT(j_k, z, \xi).$$

Finally, we have to eliminate the error terms ξ in equation 9 which represent the latent components of the health status. We accomplish this by integrating over the N_h-dimensional error term,

7 For notational convenience, the index for individual observations is left out.

assuming that the ξ are jointly standard normally distributed, possibly with correlations $\rho(\xi_i,\xi_j)$. The unconditional Likelihoodfunction is:

$$(10) \quad L(\beta,\gamma,\Delta,\lambda,\mu,\rho) = \int_{-\infty}^{\infty} LOGIT(\mathrm{i}, \mathrm{z}, \xi) \prod_{k=1}^{N_y} ORDPROBIT(j_k, \mathrm{z}, \xi) \;\; \varphi(\xi,\rho)\, d\xi \, .$$

We estimate the nonlinear MIMIC model by maximizing the sum of the indivi·dual log-likelihood contributions over the coefficients β, γ, Δ and λ, over the thresholds $\mu_{k,j}$, and (in general) over the correlations $\rho(\xi_i,\xi_j)$ among the latent health components.

4.3. Identification and Estimation

In order to check the identification of the system we start by inspecting the set of equations 7, which make up the ordered probit part of the likelihood function.[8] Maximizing 10 directly identifies the absolute value of the factor loadings λ_k attached to ξ through the term $\lambda_k' \xi$ in the case of orthogonal ξ. The signs are not identified because the thresholds $\mu_{j,k}$ of the ordered probit can be ordered either way. By

---

8 In the sequel, we consider uncorrelated ξ. If the ξ are correlated, also the ρ have to be estimated and additional identification restrictions are required.

counting the elements of δ_m, the coefficients of the exogenous variables in the health equations, and the elements of the reduced form parameters ψ_k in the ordered probit part

$$(11)\quad \begin{pmatrix} \psi_{1k} \\ \psi_{2k} \\ . \\ . \\ \psi_{N_z k} \end{pmatrix} = \lambda_{1k} \begin{pmatrix} \delta_{11} \\ \delta_{21} \\ . \\ . \\ \delta_{N_z 1} \end{pmatrix} + \lambda_{2k} \begin{pmatrix} \delta_{12} \\ \delta_{22} \\ . \\ . \\ \delta_{N_z 2} \end{pmatrix} \qquad \text{for } k = 1,\ldots,N_y$$

$$(Nz \times 1) \qquad (Nz \times 1) \qquad (Nz \times 1)$$

it becomes clear that the structural coefficients δ_m are only identified if $N_y \geq N_h$. Hence, the number of indicators y_k has to be at least as large as the number of latent health dimensions h^*. Since in typical applications the number of indicators tends to be large compared to the number of underlying factors identification of $\boldsymbol{\lambda}$ and δ is easily achieved.

In contrast to the factor loadings $\boldsymbol{\lambda}_k$ in equation 7, the γ_i, the coefficients of health in the choice of living arrangements, are not identifiable through the term $\gamma_i' \xi$ because the scale in the discrete choice model is undetermined. Moreover, β_i, the coefficients of the exogenous variables in the choice of living arrangements, are also not directly identifiable even though δ_m is given:

$$(12)\quad \begin{pmatrix} \pi_{1i} \\ \pi_{2i} \\ . \\ . \\ \pi_{N_z i} \end{pmatrix} = \begin{pmatrix} \beta_{1i} \\ \beta_{2i} \\ . \\ . \\ \beta_{N_z i} \end{pmatrix} + \gamma_{1i} \begin{pmatrix} \delta_{11} \\ \delta_{21} \\ . \\ . \\ \delta_{N_z 1} \end{pmatrix} + \gamma_{2i} \begin{pmatrix} \delta_{12} \\ \delta_{22} \\ . \\ . \\ \delta_{N_z 2} \end{pmatrix}$$

for $i = 2,\ldots N_a$.

The number of elements in $\boldsymbol{\beta}_i$ equals the number of reduced form parameters in $\boldsymbol{\pi}_i$. Since $\boldsymbol{\gamma}_i$ is not identifiable, there is an excess number of structural parameters equal to the number of elements in $\boldsymbol{\gamma}_i$, the number of health dimensions. Hence, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can only be identified by imposing further restrictions.

We explore two possibilities of identifying $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}_i$:[9]

- identification in a cross-section with N_h parameter restrictions on each $\boldsymbol{\beta}_i$, and
- identification in repeated cross-sections exploiting parameter differences in Δ over time.

In the first case, we impose the assumption that at least N_h exogenous variables influence the choice of living arrangements only indirectly via their influence on health, but not directly. This pins down the parameters $\boldsymbol{\gamma}_i$, the impact of health on choice. With $\boldsymbol{\gamma}_i$ given, the remaining $\boldsymbol{\beta}_i$ are just identified.[10]

---

9 Other identification approaches are possible with panel data.

10 Identification of factor analytic models through linear parameter restrictions has been introduced by Jöreskog (1967).

In the second approach, we impose the assumption that the coefficients of the main choice equation (3b) do not change over time, but that the technical progress in medical science changes the health production function (equation 5). With two cross sections t_1 and t_2, we first estimate the reduced form coefficients

$$(13) \qquad \hat{\pi}'_{i(t_1)} = \beta'_i + \gamma'_i \hat{\Delta}'_{(t_1)}$$

$$\hat{\pi}'_{i(t_2)} = \beta'_i + \gamma'_i \hat{\Delta}'_{(t_2)}.$$

Then, γ can be estimated from:

$$(14) \qquad \hat{\pi}'_{i(t_1)} - \hat{\pi}'_{i(t_2)} = \gamma'_i (\hat{\Delta}'_{(t_1)} - \hat{\Delta}'_{(t_2)}),$$

provided that $N_z \geq N_h$.

In either approach to identification, we first estimate the reduced form parameters by maximizing (10) using (8). In a second step, we compute the structural parameters by a minimum distance method (nonlinear generalized least squares) applied to equations 8.

Given the results from the exploratory factor analysis, we assume that two dimensions suffice to describe the latent health status. For simplicity we also impose $\rho(\xi_1,\xi_2) = 0$, although other factor structures can be thought of. Even with $\rho = 0$, the integral in (10) does not factor easily due to the functional form. In order to evaluate the integral we therefore employ two-dimensional Gauss-Hermite integration.

5. Estimation Results

Table 3 presents the reduced form estimates of the nonlinear MIMIC model. The first panel refers to the discrete choice submodel (equation 6) with parameters π, while the second panel represents estimates of the ordered probit submodel (equations 7) with parameters ψ_k, λ_k, and μ_k. In addition to the factor loadings λ_k for the two latent health status variables, and the switch-points μ,[11] some of the structural parameters β_{ik} in the living arrangement choice (equation 3b) can directly be identified because the corresponding δ_{nk} (equation 5) are zero. These are the coefficients of those exogenous variables which appear in the upper but not in the lower panel. The corresponding rows of coefficients are marked by $\beta=\pi$.

The results are encouraging. The large t-values of π and ψ show that the causal links in figure 2 are significant. Moreover, the t-values of μ imply that it makes a difference to account for the categorical nature of the health indicators. We therefore proceed in estimating the structural coefficients. We first pursue identification through parameter restrictions.

In selecting possible restrictions, the main question is which variables are most likely to influence the living arrangement decision only by their indirect impact on health without directly influencing the living arrangement choice. Of the variables included, age per se as well as education certainly do effect the health status but are less likely to directly affect living arrangement choices. This is also clear from the exploratory logit analysis, table 2, where age in both columns and education in the second column become insignificant after including the

---

11 Actually, $\mu2^*=\exp(\mu2)-\mu1$ and $\mu3^*=\exp(\mu3)-\mu1-\mu2$.

health indicators. The estimated coefficient of education, although still significant in the first column, decreases in magnitude and its significance level.

Table 4 presents the estimates. The upper panel displays the living arrangement choice equation. The demographic variables are weaker than in the multinomial logit estimation, table 2, except for the "daughters' effect." Living with children is strongly correlated with the number of daughters who can take care of the elderly. There is no corresponding "son's effect."

Higher financial and housing wealth significantly increases the likelihood to live with children, while the ownership of a house reduces the probability of entering a nursing home. The positive correlation between wealth of the elderly and living together appears to be evidence in favor of the "bribery hypothesis" of Kotlikoff and Morris (1990) -- wealthy elderly who like to be taken care of by their children, are able to bribe the children, who would rather live by themselves, if it weren't for the shared wealth. Unfortunately, we do not know the wealth of the children to shed more light on this issue. Because the wealth of children is commonly highly correlated with the wealth of the parents, the coefficients may also express a supply effect: only wealthy children can take their parents in. As a caveat, the wealth actually reported may rather be household wealth including the younger generations wealth, although the question in the survey instrument was intended to record personal wealth of the elderly sample person.

The strong negative and significant coefficient on the MORTG variable tells us that the few elderly who still have mortgages on their home are unlikely to move to their children (or, though not significantly, into a nursing home). This is easily explained by the fact that almost all elderly with a mortgage are recent movers and unlikely to move again.

Main point of the MIMIC model was capturing the influence of health. Both health variables significantly affect the choice to live with children. While the first factor only affects the probability to live with children, a higher value of the second health factor, indicating a healthier elderly, makes both dependent living arrangements less likely than living independently.

What are the two health factors? If we look at the next block of results -- pertaining to the health measurement equation 4 -- we see that the second factor is strongly associated with the *IADLs*, while the first factor is more related to the first four *ADLs*. Looking at the health production function -- lower panel of table 4, c.f. equation 5 -- we see that the second factor is mainly determined by age, while the most important determinant for the first factor is education. The first health factor works more like a random effect, while the second factor carries the deterministic component associated with the exogenous variables.

The coefficients of the socio-demographic variables in table 4 have a similar pattern as the coefficients in table 2. However, some of the magnitudes change considerably. For example, the coefficients of the RACE and the OWN variable in the first column almost double in magnitude compared to table 2. In general, the changes are largest for those variables that appear in several equations of the system and not only in the choice equation. If we believe in the *a priori* assumptions underlying the MIMIC model, we must conclude that the multinomial logit model yields biased parameter results.

One may also be interested in seeing whether the nonlinear MIMIC model predicts better than the simple multinomial logit model.

This is a weak test of the *a priori* assumptions underlying the MIMIC model. It is weak because the real strength of the structural model is the prediction of the effect of structural changes. However the data do not provide such an experiment.

In order to test the out-of-sample performance, we use the 1986 and 1988 waves of the LSOA. We restricted attention to the unmarried elderly, so the 1986 and 1988 samples are smaller than the 1990 sample due to those elderly who were still married in 1986 or 1988. Table 5 shows the results.

In the *in-sample* prediction, the multinomial logit model fits the sample better than the nonlinear *MIMIC* model. It produces better estimates of the institutionalization probability, and it has an overall higher success rate. This might be expected from an atheoretical model designed to describe the data. The balance changes in the out-of-sample prediction. Now the nonlinear model has a better overall performance, and it is closer in predicting living with others. Again, this reversal is exactly what an econometrician wishes for a model that may mine the sample worse but captures the true structure better. The improvement, however, is rather modest. It would be helpful to have a hold-out sample which consists of different elderly rather than of the same elderly two years prior to the estimation period.

We also pursued the second method to identify the structural parameters in exploiting the variation in the health production function over time, see equation 14. We use the difference between the matrices Δ_{1988} and Δ_{1990}, estimated for the 1988 and 1990 waves, maintaining that the structural coefficients of the choice equation (β, γ) remain constant. It would be preferable to estimate the second set of coefficients from a cross-section as far away from 1990 as possible in order to capture a sufficient change in health technology than using 1988. However, the 1986 wave has only very few institutionalized persons, and the 1984 wave has none. The results are disappointing because $\Delta_{1988} - \Delta_{1990}$ has turned out to be virtually random. The explanation is obvious: two years are too short to induce significant changes in health technology.

6. Conclusions

This paper is a classical exercise in what econometrics is supposed to do -- and where the problems of sophisticated econometrics are. It uses *a priori* knowledge drawn from economics (and from common sense) in order to structure the inference we draw from the data. The multinomial logit model is atheoretical in the sense that it makes no usage of the causal links depicted in figure 1. In turn, the MIMIC model in figure 2 employs a rather involved super structure to guide our inference.

The main problem with this model is of course identification. After all, the MIMIC model uses the same information as the simple multinomial logit model, and only introduces a potentially large number of latent variable constructs. In addition to postulating causal links as in figure 2, additional parameter restrictions were required. In the first case, identification via exclusion restrictions is pretty much in the spirit of conventional simultaneous equation models. In the second case, we assumed that some parameters change over time while others stay constant. Both identifying assumptions can easily be criticized. If the identifying assumptions are false, our estimates are inconsistent. If they are true, we have gained efficiency and have learned more about those structural coefficients that we have estimated.

Our panel data could identify the coefficients of the latent health variables much better than the cross-section models of section 5. The latent health variables have a function very similar to random effects. This is the reason they are so hard to identify in cross sections. By exploiting the panel structure, we could identify, say, two latent health status in 1988 as well as in 1990, possibly correlated over time. The likelihood function would be similar to (10) but would require higher dimensional integration. In further research, we will estimate this model by employing simulation methods for the computation of these integrals, such as the smooth simulated maximum likelihood approach of Börsch-Supan and Hajivassiliou (1993).

Turning back to substance, the coefficients of our main interest were health and wealth. While wealth is an important economic variable in the choice of living arrangements, income has proven to be of little relevance once wealth is included. Health is one of the main predictors of living arrangement choices. This is to be expected. Health is well captured by two factors, one associated with independent activities and strongly related to age, while the other, more person-specific factor is associated with more basic capabilities. Living together with others, mainly children, is positively affected by financial and housing wealth, while homeowners are less likely to become institutionalized.

REFERENCES

Börsch-Supan, A., 1989, Household Dissolution and the Choice of Alternative Living Arrangements Among Elderly Americans, in: David Wise (ed.), The Economics of Aging, Chicago: University of Chicago Press.

——————— 1990, Elderly Americans: A Dynamic Analysis of Household Dissolution and Living Arrangement Transitions, in: David Wise (ed.), Issues in the Economics of Aging, Chicago: University of Chicago Press.

Börsch-Supan, A., L. Kotlikoff, and J. Morris, 1989, The Dynamics of Living Arrangements of the Elderly: Health and Family Support, in: J. Pacolet and C. Wilderom (eds.) The Economics of Care of the Elderly, London, 1991.

Börsch-Supan, A., J. Gokhale, L. Kotlikoff and J. Morris, 1992, The Provision of Time to the Elderly by Their Children, in: David Wise (ed.) Topics in the Economics of Aging. University of Chicago Press.

Börsch-Supan, A., V. Hajivassiliou, L. Kotlikoff and J. Morris, 1992, Health, Children, and Elderly Living Arrangements: A Multiperiod-Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors, in: David Wise (ed.) Topics in the Economics of Aging. University of Chicago Press.

Börsch-Supan, A., and V. Hajivassiliou, 1993, Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models, Journal of Econometrics, forthcoming.

Ellwood, D.T., and Kane, T.J., 1990, The American Way of Aging: An Event History Analysis, in: David Wise (ed.), Issues in the Economics of Aging, Chicago: University of Chicago Press.

Feinstein, J. and D. McFadden, 1988, The Dynamics of Housing Demand by the Elderly, in: David Wise (ed.) The Economics of Aging, University of Chicago Press, Chicago.

Fitti, J.E., and M.G. Kovar, 1987, The Supplement on Aging to the 1984 National Health Interview Survey, U.S.GPO: Washington, D.C.

Garber, A.M., 1990, Predicting Nursing Home Utilization Among the High-Risk Elderly, in: David Wise (ed.), Issues in the Economics of Aging, Chicago: University of Chicago Press.

Jöreskog, K.G., 1967, Some Contributions to Maximum Likelihood Factor Analysis, Psychometrica 32, 443-482.

Jöreskog, K.G., and D. Sörbom, 1988, LISREL 7 - A Guide to the Program and Applications, Chicago.

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.

——————— 1990, Why Don't the Elderly Live With their Children? A New Look, in: David Wise (ed.), Issues in the Economics of Aging, Chicago: University of Chicago Press.

Kovar, M.G., and G.S. Poe, 1985, The National Health Interview Survey Design, 1973-84, and Procedures, 1975-83, U.S.GPO: Washington, D.C.

McFadden, D., 1988, Discrete Response to Unobserved Variables for which there are Multiple Indicators, MIT, mimeo.

Morikawa, T., M. Ben-Akiva, and D. McFadden, 1990, Incorporating Psychometric Data in Econometric Travel Demand Models, MIT, mimeo.

National Center for Health Statistics, 1991, The Longitudinal Study on Aging, Version 4, mimeo, Washington, D.C.

Schwartz, S., Danziger, S., and Smolensky, E., 1984, The Choice of Living Arrangements by the Elderly, in: H.J. Aaron and G. Burtless (eds.), Retirement and Economic Behavior, The Brookings Institution, Washington, D.C.

Venti, Steven F., and David A. Wise, 1990, But They Don't Want to Reduce Equity, in: David Wise (ed.) Issues in the Economics of Aging, University of Chicago Press, Chicago.

——————— 1991, "The Saving Effect of Tax-Deferred Retirement Accounts: Evidence from SIPP", in: B.D. Bernheim and J.B. Shoven (eds.), *National Saving and Economic Performance*, University of Chicago Press.

Table 1: Description of Variables - LSOA 1990

Dependent Variable (Living Arrangements):			
LIVARG	Living independently	63.0%	
	Living with others	28.7%	
	Living in an Institution	8.2%	
Demographic Exogenous Variables:			
AGE90	Age in 1990 (years)	82.6	
SEX	Gender: Female = 1	81.4%	
RACE	Black and Hispanic = 1	9.2%	
EDUC	Highest grade completed in years	10.1	
SONS	Number of living sons	1.01	
DAUGHTERS	Number of living daughthers	1.07	
Economic Exogenous Variables:			
OWN	1 = Homeownership	63.4%	
MORTG	1 = Home Free and Clear	85.7% (of owners)	
			median
INCPERS	Annual Personal Income (in $)	7,748	2,394
HOUSAS	House Value (in $)	38,113 (all)	20,000
		49,684 (owner)	31,000
FINAS	Financial Assets (in $)	36,012	3,500
	- Stocks, Bonds, Mutual Funds	16,517	
	- Savings, other Bank Accounts	19,495	
Health Indicators:			
Activities of Daily Living: Sample person without difficulties			
BATH	Bathing	74.4%	
DRESS	Dressing	82.5%	
EAT	Eating	92.1%	
GETUP	Getting Up from Bed/Chair	76.6%	
WALK	Walking	58.8%	
OUTSD	Getting Outside	75.9%	
TOIL	Toiletting	86.4%	
Independent Activities of Daily Living: Sample person without difficulties			
MEALS	Preparing Meals	75.0% (asked only	
SHOP	Shopping	68.9% for elderly in	
HOUSEW	Doing Light Housework	77.9% households)	

Means and Medians computed on the Working Sample, 2193 elderly. Source: LSOA 1990

Table 2: Multinomial Logit Model - Estimation Results

	Probability to ... rather than to live independently			
	live with children or others		*live in an institution*	
CONSTANT	-3.51220	-3.98	-11.35623	-7.90
AGE90	0.04896	4.82	0.13259	8.24
EDUC	-0.09395	-6.07	-0.08393	-3.22
RACE	0.59230	3.61	-1.09008	-2.24
SEX	-0.17913	-1.32	-0.09591	-0.39
DAUGHTERS	0.15276	3.74	0.06029	0.80
SONS	0.03730	0.88	-0.05156	-0.63
OWN	0.33773	3.02	-1.82145	-8.94
MORTG	-0.95982	-5.85	-0.25593	-0.75
FINASS	0.00124	2.41	0.00100	1.04
HOMEASS	0.00157	1.91	0.00249	1.69
INCPERS	0.00037	0.14	-0.00282	-0.53
Log-Likelihood = -1657.1			*Source: LSOA 1990, 2193 elderly*	

Table 3: Multinomial Logit - Estimation Results with ADLs and IADLs

	Probability to ... rather than to live independently			
	live with children or others		*live in an institution*	
CONSTANT	-1.97921	-2.10	-10.39988	-4.15
AGE90	0.01790	1.61	0.00993	0.39
EDUC	-0.07186	-4.42	-0.02114	-0.52
RACE	0.50494	2.91	-1.49199	-2.56
SEX	-0.24538	-1.74	-0.33389	-0.85
DAUGHTERS	0.14204	3.34	0.07087	0.65
SONS	0.04433	1.00	-0.03871	-0.33
OWN	0.31200	2.68	-1.87982	-6.79
MORTG	-0.90636	-5.30	-0.18384	-0.38
FINASS	0.00150	2.77	0.00249	1.87
HOMEASS	0.00155	1.83	-0.00113	-0.50
INCPERS	-0.00176	-0.60	-0.01341	-2.06
BATH	-0.02732	-0.29	0.04103	0.24
DRESS	-0.09357	-0.74	-0.04878	-0.27
GETUP	-0.04733	-0.42	0.02812	0.14
WALK	-0.02884	-0.34	-0.09203	-0.49
OUTSD	0.00642	0.06	0.01751	0.09
HOUSW	0.22450	2.52	1.63224	5.49
MEAL	0.38729	4.52	0.95070	3.87
SHOP	0.13567	1.98	0.57994	1.76

Log-Likelihood = -1298.0 *Source: LSOA 1990, 2193 elderly*

Table 3: Multiple-Indicator, Multiple-Cause Model: Reduced Form 1990

Living Arrangement Choice (Equation 6)

	Probability to ... rather than to live independently				
	live with children or others		*live in an institution*		
CONSTANT	-5.07076	-4.40	-12.87920	-6.68	
AGE90	0.07081	5.37	0.15511	7.28	
EDUC	-0.12157	-6.12	-0.12186	-3.65	
RACE	0.87069	3.74	-0.76198	-1.45	
SEX	-0.28649	-1.60	-0.15910	-0.49	
DAUGHTERS	0.21032	3.82	0.12253	1.35	($\beta=\pi$)
SONS	0.02979	0.55	-0.04329	-0.42	
OWN	0.27072	1.91	-1.85104	-7.63	
MORTG	-1.18610	-5.13	-0.56230	-1.24	($\beta=\pi$)
FINASS	0.00173	2.28	0.00119	1.11	($\beta=\pi$)
HOMEASS	0.00197	2.00	0.00243	1.20	($\beta=\pi$)

Health Measurement (Equations 7)

	Bathing		*Dressing*		*Getting Up*		*Walking*	
AGE90	0.144	13.00	0.129	11.26	0.104	9.32	0.103	11.73
EDUC	-0.034	-2.55	-0.040	-2.74	-0.043	-2.92	-0.050	-4.26
RACE	0.968	5.49	1.193	6.21	0.799	4.65	0.652	4.90
SEX	0.059	0.41	-0.208	-1.48	0.189	1.35	-0.007	-0.06
OWN	-0.324	-3.02	-0.517	-4.55	-0.479	-4.53	-0.406	-4.79
HEALTH1	-1.084	-13.10	-1.210	-13.74	-1.322	-13.67	-1.123	-14.27
HEALTH2	-1.717	-18.30	-1.702	-18.45	-1.493	-18.49	-1.260	-20.63
MU1	13.254	13.36	12.339	12.28	9.870	9.91	8.367	11.26
MU2*	-0.138	1.89	-0.050	-0.63	0.162	2.27	0.231	4.62
MU3*	-0.785	-6.63	-0.478	-3.58	-0.167	-1.43	-0.118	-1.68

	Going Outside		*Light Housework*		*Preparing Meals*		*Shopping*	
AGE90	0.186	13.92	0.199	11.28	0.208	12.40	0.223	14.97
EDUC	-0.077	-4.89	-0.102	-5.43	-0.135	-6.94	-0.127	-7.50
RACE	1.118	5.82	1.153	5.27	1.264	6.01	1.154	6.15
SEX	0.163	1.04	-0.225	-1.39	-0.029	-0.17	0.420	2.51
OWN	-0.574	-4.69	-0.852	-5.95	-0.832	-5.84	-0.726	-5.85
HEALTH1	-1.376	-13.33	-0.761	-8.31	-0.625	-7.21	-0.718	-9.62
HEALTH2	-2.035	-17.06	-2.621	-14.88	-2.689	-15.55	-2.308	-16.68
MU1	16.645	14.16	17.485	11.51	17.852	12.44	18.791	14.80
MU2*	-0.272	-3.19	-0.749	-6.03	-0.496	-4.47	-0.867	-7.97
MU3*	-0.422	-3.94	-1.078	-6.48	-0.884	-5.99	-1.098	-8.49

Log-Likelihood = -8122. *Source: LSOA 1990, 2193 elderly.*

Table 4: Structural Parameters of the MIMIC Model

Living Arrangement Choice (Equations 3b)

	Probability to ... rather than to live independently			
	live with children or others		*live in an institution*	
CONSTANT	-5.21285	-4.60	-12.49175	-6.59
RACE	1.06272	2.38	-1.09591	-1.69
SEX	-0.01280	-0.04	0.18139	0.47
DAUGHTERS	0.22783	4.18	0.14136	1.58
SONS	0.03278	0.62	-0.03544	-0.34
OWN	0.56104	2.33	-1.34210	-4.55
MORTG	-1.31190	-5.74	-0.52219	-1.17
FINASS	0.00150	1.99	0.00094	0.88
HOMEASS	0.00189	1.94	0.00190	0.95
HEALTH1	-2.72910	-2.16	-1.28417	-0.86
HEALTH2	-0.76748	-2.93	-1.67740	-5.43

Health Measurement (Equations 4)

	Bathing		*Dressing*		*Getting Up*		*Walking*	
HEALTH1	-1.057	-13.15	-1.179	-13.78	-1.333	-14.18	-1.154	-15.17
HEALTH2	-1.635	-19.07	-1.616	-19.30	-1.440	-18.52	-1.273	-21.38
MU1	12.390	16.28	13.005	16.85	11.284	14.75	8.976	14.86
MU2*	-0.164	-2.28	-0.048	-0.63	0.169	2.41	0.270	5.54
MU3*	-0.805	-6.94	-0.506	-3.87	-0.146	-1.27	-0.091	-1.31
	Going Outside		*Light Housework*		*Preparing Meals*		*Shopping*	
HEALTH1	-1.370	-13.66	-0.772	-8.55	-0.638	-7.54	-0.711	-9.66
HEALTH2	-2.039	-17.53	-2.490	-14.97	-2.654	-16.01	-2.369	-17.46
MU1	15.540	15.32	18.795	14.14	19.631	14.84	16.976	14.99
MU2*	-0.242	-2.89	-0.763	-6.35	-0.479	-4.41	-0.895	-8.39
MU3*	-0.461	-4.39	-1.118	-6.84	-0.899	-6.21	-1.160	-9.16

Health Production (Equations 5)

	HEALTH 1		*HEALTH 2*	
AGE90	-0.00216	-0.31	-0.08708	-20.15
EDUC	0.02881	2.97	0.05623	9.56
RACE	0.20709	1.85	-0.42710	-6.83
SEX	0.04999	0.55	0.01938	0.34
OWN	0.02259	0.31	0.32748	7.24

Identification by parameter restrictions *Source: LSOA 1990, 2193 elderly.*

Table 5: Prediction Performance of the Two Alternative Models

	Observed	Multi-nomial Logit	Non-Linear MIMIC Red. Form	Non-Linear MIMIC Str. Form
In-Sample 1990:				
Alone	64.02	79.30	82.58	73.83
With others	28.45	13.41	15.18	23..16
Institution	7.52	7.30	2.33	3.01
Percent Correct		74.15	69.08	63.61
Out-of-Sample 1988:				
Alone	67.50	87.32	86.70	76.97
With others	28.63	9.44	12.07	21.65
Institution	3.87	3.24	1.22	1.38
Percent Correct		68.19	70.96	64.85
Out-of-Sample 1986:				
Alone	70.96	90.30	89.66	78.40
With others	27.62	8.50	10.12	21.18
Institution	1.42	1.20	0.2	0.42
Percent Correct		74.86	72.8	67.07

Source: LSOA 1988 and 1990, 1,963 and 2,193 elderly, respectively