Individualised life tables: investigating dynamics of health, work and cohabitation in the UK

Working Paper 108

Martin Karlsson

Oxford Institute of Ageing

Les Mayhew

Cass Business School

Ben Rickayzen

Cass Business School

February 2008

Oxford Institute of Ageing Working Papers

Editor: Kenneth Howse

www.ageing.ox.ac.uk

Individualised Life Tables

Investigating Dynamics of Health, Work and Cohabitation in the UK

Martin Karlsson*
Oxford Institute of Ageing

Les Mayhew
Cass Business School

Ben Rickayzen
Cass Business School

February 28, 2008

Abstract

A life table is a table which shows, at each age, the probability that a person in a given population will die before their next birthday. It can be used to calculate life expectancy and healthy life expectancy for people of different ages. In this work, using longitudinal datasets and panel data methods, we produce life tables for different subgroups of the population, defined according to cohabitation status, employment and other factors.

As a first step, we estimate the dynamics of factors which are of particular importance in people's lives: health, labour market participation, cohabitation and mortality. The significance of these variables is twofold: they determine the well-being of individuals, but the variables also determine the resources available to the individuals in times of ill health. Using the British Household Panel Survey, we analyse the extent to which these variables are influenced by one another, and by exogenous factors such as education and ethnicity. Estimating a system of probit models using simulation techniques, we are able to distinguish the effects of the exogenous and endogenous variables from state dependence and unobserved heterogeneity. We also correct for attrition and the initial conditions problem.

We estimate time trends in mortality, health and other dependent variables to investigate whether a compression of morbidity has occurred in the recent past. Finally, the parameter estimates are used to simulate life tables for various sub-groups in the population and compare measures of life expectancy and healthy life expectancy for different groups.

JEL Classification C5, I12

Keywords disability, cohabitation, mortality, labour supply, maximum simulated likelihood, attrition

*Corresponding author: martin.karlsson@ageing.ox.ac.uk.

1 Introduction

Most developed countries' populations are ageing rapidly with consequent implications for public spending on long-term care (LTC), pensions and health care. The UK dependency ratio (the number of retired people per 100 people of working age) is projected to increase from 24 today to 40 in 2040. Although substantial, the increase is lower than in many other countries. In Japan, for instance, the ratio is projected to increase from 30 today to 65 in 2040 (United Nations, 2006).

Such demographic changes are expected to have a significant impact on the demand for LTC. Most consumers of LTC are over age 80; for example, in England, almost 80 per cent of care home inhabitants belong to this age group (Bajekal, 2002). Since increasing life expectancy causes this group to grow at a faster rate than the general retired population, there is concern that the demographic burden could make the current system of financing LTC unsustainable. Indeed, in the UK, there is already a trend towards concentrating resources only on individuals with severe disability (Karlsson et al, 2004).

However, the impacts of an ageing population do not stop there. UK government policy is to increase state retirement age over the next decades. Between 2010 and 2020 female state pension age will increase in steps to 65 by 2020 at which point it will be equal to male state pension age. Thereafter, it will rise to 66 between 2024 and 2026 and then eventually to 68 between 2044 and 2046. Currently the average age of the working population (i.e. people between age 16 and state pension age) is 38.7 years but by 2025 it will have increased to 40.6 years, partly as a result of changes in state pension age and partly due to an ageing population. The result is that the working age population will increase from 36.8 million to 40.9 million by 2025 after taking changes in state pension age into account. According to ONS population projections, of the total, 12.4m in 2025 will be aged 50+ as compared with 9.4 million in 2007, an increase of 3m.

Although life expectancy is undoubtedly increasing, the question arises as to whether older workers will be healthy enough to work or whether they will become an increasing burden on the economy by swelling the ranks of people on long term sickness and disability benefits. Furthermore, if they do continue to work, will they be as productive as younger workers? The direct and indirect benefit costs, for example, are put at £30 billion pounds a year and the annual loss of output at three times that figure. This is before taking possible changes in age-related productivity change into account for those that remain in work (Blake and Mayhew, 2006).

To prepare for all possible eventualities policy makers need to be able calibrate social policy over the next few decades accordingly but this requires a much greater understanding of health life trajectories and disablement processes. Currently there is no mechanism for quantifying the percentages of people who will be unable to work at higher retirement ages, whether healthy life expectancy as well as life expectancy is increasing, which sub-groups are the most vulnerable to sickness and disability, the extent to which risk is socially rather than biologically determined, and if risk in these cases can be manipulated through the policy process.

1.1 Competing hypotheses

Relatively little is known about long-term trends and the determinants of the disablement process. One important issue that has not yet been resolved is the long-term trends in healthy life expectancy and disabled life expectancy. Three competing hypotheses have been proposed. The most optimistic one, suggesting a compression of morbidity, was proposed by Fries (1980).

According to this perspective, adult life expectancy is approaching its biological limit so that if disability spells can be postponed to higher ages the result will be an overall reduction in the time spent disabled. By contrast, Gruenberg (1977) suggested an expansion of morbidity based on the argument that the observed decline in mortality was mainly due to falling accident rates. The third hypothesis was proposed by Manton (1987) according to whom the development in mortality and morbidity is a combination of the two, which could lead to an expansion of the time spent in good health as well as the time spent in disability.

Official statistics, however, are surprisingly inconclusive as to which of the three hypotheses prevails in reality (Bone et al, 1995, Bebbington & Darton, 1996, Bebbington and Comas-Herrera, 2000). In general, results seem to be sensitive to the definition of disability (activities of daily living (ADLs) or limiting long-standing illness) as well as to the severity of disability taken into account.

Despite this ambiguity in the statistics, the long-term trends have very strong implications for the future funding of long-term care. In a long-term projection model, Karlsson et al (2006) find that a pessimistic scenario ('expansion of morbidity') implies some 2 million disabled older people greater than the most optimistic scenario ('compression of morbidity'). The implications for public finances are similar: in the pessimistic scenario, the element of the tax rate necessary to finance formal long-term care would have to increase by around 80 per cent of its present level, whereas virtually no increase would be necessary in the optimistic scenario. Similar differences arise in the supply of demand for informal care (i.e. unpaid care provided by spouses, children or other members of the local community): with an optimistic scenario, there is virtually no shortfall of informal carers in the next few decades, whereas the pessimistic scenario leads to a serious deficit of informal care that will eventually strain public finances.

In research related to that reported in this paper, we found, for example, that working lives could be extended through already well established public health measures such as smoking cessation (Cass, 2007)[1]. Using a very simple model we found, for example, that a female non-smoker has a HLE (healthy life expectancy) of 28.1 years at age 50 and only 19.6 years if a smoker. At a state retirement age of 60, about 74% of female non-smokers would be expected to be

[1] Official Cass press release to coincide with smoking ban in Britain.

healthy and 62% of female smokers. At a state retirement age of 66, this changes to 67% and 51% respectively which suggests that there is an ever widening gap.

This paper focuses on two particular aspects of the disablement process, namely the effects of cohabitation and the effects of labour market participation. It blends the classical techniques of life tables to measure survival risk at different ages with the social dimension based on individual social characteristics that have been shown in previous studies to influence longevity. However, the other new elements to our research are that we develop a methodology that is able to estimate healthy life expectancy as well as life expectancy . In this way we are able to model the life trajectories of up to 64 different categories or sub-groups of population.

The rest of this paper sets out our hypotheses in more detail and also the data sources and assumptions used. We then describe our modelling approach in detail and how we test the relationships in the data before incorporating them into a life table from which we are able to estimate key aggregate characteristics such as healthy life expectancy and life expectancy at different ages. Since these considerations primarily affect older workers we focus our attention on the 50+ age group, but our parameter estimates could be used to calculate life expectancy at any adult age.

1.2 The Relationship between health and cohabitation

Cohabitation is of particular importance for several reasons. Firstly, it is strongly correlated with health (a relationship which seems to be stronger for higher ages; Lillard and Panis, 1996) and it is of great interest to know whether this correlation reflects a causal effect – so that changing cohabitation patterns would have implications for health – or merely reflects self-selection into and out of cohabitation (i.e. people who cohabit are healthier at the outset). Separating causation and correlation leads to a host of methodological challenges that will be considered below.

Secondly, knowing the relationship between cohabitation and disability is important for analysing the implications of ageing for long-term care. Informal care comprises a substantial part of total long-term care resources and around 75 per cent of all LTC recipients in the UK receive informal care according to Karlsson et al (2006). It is a common concern that there may be a shortage of informal carers if certain discernible trends carry on in the future. These trends are, inter alia, the increase in single person households, the rising number of childless older people and the increase in the proportion of females in paid employment. It should be noted, however, that there are some trends that could be expected to countervail these threats to informal care provision. These could include, for instance, a decreasing age at which people retire, together with an improvement in health among younger retirees. This scenario implies that there will be a larger pool of able retirees available in the future to provide informal care; however, the opposite might be more likely because of planned increases in state pension age.

There is little previous empirical research in the field. Brown (2000) performs a simple empirical analysis of the National Survey of Families and Households (waves 87-88), estimating the effects of cohabitation and relationship characteristics, allowing for self-selection into cohabitation, on psychological well being. Brown found no evidence of self-selection, but observed that simple cohabitation is less beneficial to psychological health than marriage. The main explanation seems to be poorer relationship quality in cohabitation relationships.

Cheung (2000) looked at cohabitation and mortality amongst British women. Analysing the Health and Lifestyle Survey, Cox regressions were used in order to allow for self-selection into cohabitation. This is one of few studies allowing for reverse causality from health to marital status (i.e. people being healthy having a higher propensity to be married). Having adjusted for age and marital selection factors, being single was significantly associated with higher mortality, but being divorced or widowed was not. Another study that tries to compensate for reverse causation is Goldman et al (1995). They analyse marital status, health and mortality amongst older people, controlling for baseline health (i.e. before a change in marital status), socioeconomic status and social networks. The main finding is that marriage affects mortality only for men, and that the effect is modest. Widowed men are more likely to be disabled, whereas single women are actually healthier than married counterparts.

Finally, Lillard and Panis (1996) use a simultaneous equations model to estimate the relationship between health, marital status and mortality, with instrumental variables to account for the reverse causality problem. One of their hypotheses is that the selection effect has a 'demand side' (i.e. healthy people are more attractive) and a supply side (i.e. unhealthy people have more to gain from marriage), and they find indications of both: the explained part of health status tends to be negatively correlated with marriage, whereas the unexplained part is positively correlated.

Hence, if the good health status is attributable to personal characteristics, it tends to reduce the propensity to get married, whereas the propensity goes up for a person whose good health is not attributable to personal characteristics. For example, this result would imply that adverse health effects from unemployment (an observable characteristic) are connected with a reduced chance of being married, whereas the opposite holds for individual variations in health that cannot be explained by such personal characteristics. The paper by Lillard and Panis represents the most rigorous attempt to take the reverse causality issue into account; however, the models estimated do not allow for random changes over time in the dependent variable, or autocorrelation (i.e. that these random changes are persistent once they occur).

A good overview of the empirical research to date is provided by Wilson & Oswald (2005). After reviewing a large number of articles on the relationship between cohabitation and health - psychological, physical and in connection with mortality - they identify the following general conclusions:

- Marriage reduces the risk of psychological illness
- Marriage tends to increase life expectancy

- Marriage makes people healthier & happier
- Men tend to gain more from the advantageous effects of marriage.
- There is not only a guardian effect (i.e. changes in risk behaviours) - marriage seems to have other positive effects on health as well.
- The quality of the relationship is important

1.3 The relationship between health and employment

Concerning the relationship between health and work, there is an emerging literature within economics which aims at taking a person's health into account when modelling their labour market participation. As pointed out by Disney et al (2006), health can influence the decision to work in many ways, such as:

- People in poorer health might experience higher disutility from work than healthy people
- Poor health might reduce the income from work, to the extent that the individual's reduced productivity is reflected in their wage
- Poor health might entitle the individual to non-wage income, which is conditional on not working (such as incapacity benefits).

All these factors suggest that there is a positive relationship between health and work. There is also potentially an effect of poor health which works in the opposite direction: an individual in poor health might need a higher income to cover medical expenses – which would imply a higher propensity to work.

One of the most rigorous analyses of health and retirement was made by French (2005). French estimates a life cycle model of consumption and labour supply, taking health into account. Health, mortality and wages are all assumed to be exogenous; hence, the individual maximises lifetime utility with respect to a given stochastic process for health and mortality. The model is estimated using the method of simulated moments. The author finds that at any point in the life cycle, the effect of health on hours worked is significant. Nevertheless, health only explains a small amount of the total variation in work-hours over the life cycle. Furthermore, health seems to affect participation as such more than the number of hours worked, but also here the explanatory power of health is limited. Hence, declining health only explains around 10 per cent of the reduction in participation between 55 and 70.

One limitation with French's analysis, however – and this drawback applies to Heyma's (2004) approach and the model by Domeij and Johannesson (2006) as well – is that health is treated as exogenous. There are, however, several reasons to believe that the analysis of the effect of health on retirement is marred with endogeneity problems. As Disney et al (2006) point out, reported health status may be problematic for a number of reasons:

- Inactive individuals have an incentive to over-report poor health, to justify their inactiveness
- Individual heterogeneity is of great importance
- Individuals with permanent and very poor health may never have worked – and can hence not be observed 'retiring'
- Ill health may impact on other labour market attributes of the worker (productivity etc).
- The health stock may be endogenous to the labour market state of the individual.

To remedy this problem, it has been suggested that a time-dependent 'health stock' should be estimated, based on self-assessed health and a number of 'objective' health criteria. This approach is believed to overcome some of the endogeneity problems mentioned above. Using this approach, Disney et al (2006) estimate a fixed effects panel data model and a survival model for the retirement decision. They find that, in both specifications, current health shocks and lagged health has the expected effects on labour market participation. Furthermore, they find no evidence that health shocks are 'asymmetric' in the sense that a deterioration in health has a different effect on labour supply than an improvement in health. The authors also find that individuals living in couples have significantly higher probabilities of both being in paid employment.

It is questionable, however, to what extent the 'health stock' approach solves the endogeneity problem. Disney et al (2006) find that the alternative of simply adding up various impairments does worse in the estimation. However, it is likely the variables used to estimate the health stock suffer from the same problems as self-assessed health. Furthermore, it might be the case that some of the 'objective' health conditions considered are completely irrelevant for the occupation in which the individual is or could be active. In this respect, self-assessed health contains some information which is relevant for the analysis which is not contained in the various 'objective' measures, and that is the extent to which health represents an obstacle to the labour market participation of a certain individual as perceived by that individual.

In summary, previous empirical research has found a strong link between health and labour supply, but the importance of this link remains an open issue. For instance, Bound et al (1999) find that the health effects on labour market behaviour are very sensitive to assumptions concerning the parameters of the individual's utility function. One problem in many studies is that they fail to account for the potential two-way interplay between health and labour market participation. Furthermore, many authors find that marital status is an important factor in retirement behaviour – and to some extent this variable suffers from the same endogeneity problems as health. In this paper, we seek to remedy these problems.

1.4 Aim and structure of the paper

In this paper, we make use of all available waves of the British Household Panel Survey (BHPS) data in order to study the determinants of disability, cohabitation and employment of individuals over time. Our research has three main aims. Firstly, we analyse the interdependencies between mortality, employment, cohabitation and disability. Secondly, we study time trends in the various dependent variables jointly to explore relationships between them. Thirdly, we examine socioeconomic differences – as captured by educational attainment – in the four dependent variables. Educational attainment is particularly important in this context. Firstly, it is very convenient as a socioeconomic indicator as it normally remains constant over most of the life course. Secondly, it is a well established result in health economics and epidemiology that education is an important factor in explaining socioeconomic differences in health (cf Fuchs, 2004). Thirdly, empirical studies of marital matching indicate that education is an important aspect of a person's 'marriageability' (Wong, 2003). Hence, excluding education might lead to an overestimation of the importance of cohabitation status for health. Similar concerns arise for the other dependent variables.

There are two main methodological challenges. The first is that we seek to estimate a dynamic model, where previous realisations of the dependent variables influence current outcomes. Secondly, we seek to distinguish causation from correlation in the relationship between the various dependent variables. This requires using simulation techniques that allow for systematic differences between individuals which are not discernible in the data. For instance, it might be that healthier people are considered more attractive for marriage or cohabitation, and our method is one way to correct for this type of reverse causality.

The paper is organised as follows. In the next section, our methodological approach is outlined and the dataset presented. After that, in Section 3, we present our results. The last section concludes.

2 Methodological approach

Our econometric model is a system of equations; one for health, one for cohabitation, one for work and one for survival. We follow the estimation technique suggested by Börsch-Supan et al (1993) and adapt it to our problem. Firstly, we allow for unobserved heterogeneity in all dependent variables. In other words, we do not assume that all differences in trajectories of health, cohabitation and so forth are attributable to observable characteristics (age, gender, education, ethnicity) but we exploit the longitudinal character of the dataset to allow for systematic differences between individuals which emerge from the analysis.

Disability status varies over time, but it also has an important, time-invariant component reflecting the fact that some people are "structurally" healthier than others (due to genetic predisposition or preferences towards risk factors, for example). The same goes for the other dependent variables, where it can be assumed that people are likely to be structurally different from each other. Furthermore, there are reasons to believe that these person-specific attributes are correlated across equations. For example, it is plausible that unobservable characteristics for health and survival are correlated. Not accounting for these unobserved characteristics could lead to inconsistent estimates of the causal effects of the various variables.

However, not all unobserved differences can be captured by components which do not vary over time. Hence, we also allow for time-varying disturbances, which are potentially correlated across the equations and potentially exhibiting autoregression, meaning that shocks are potentially persistent. A natural example of this would be a positive shock which leads to an individual finding work in another labour market area – an event which could potentially be correlated with a change in cohabitation status if the partner is unable or unwilling to relocate.

2.1 Estimating Equations

We now define estimating equations and then investigate the error structure more closely. We estimate four different equations, one with the hazard rate for mortality and three for the different states which survivors can be in (concerning work, cohabitation and retirement). For the mortality hazard rate, as for the other equations of the model, we use a probit model. The probit model suggests survival depends on a latent, unbounded and continuous survival variable A^*_{it}:

$$A^*_{it} = \alpha^a_i + X_{it}c_a + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} \end{bmatrix} d_a + \varepsilon^a_{it} \tag{1}$$

where α^a_i is an individual constant, capturing structural differences between individuals, X_{it} is a set of time-varying exogenous variables (age, squared age and year), $W_{i,t-1}$ is the individual's labour market status in the previous year, $C_{i,t-1}$ is the individual's cohabitation status in the previous year, and $H_{i,t-1}$ is the individual's health status in the previous year. Finally, ε^a_{it} is a random disturbance.

As in the standard probit model, realised survival is defined according to a switching function:

$$A_{it} = \left\{ \begin{array}{lll} 1 & if & A_{i,t-1} \cdot A^*_{it} \geq 0 \\ 0 & \text{otherwise} & \end{array} \right. .$$

Similarly, the estimating equations for the three other dependent variables are:

$$W^*_{it} = \alpha^w_i + X_{it}c_w + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} \end{bmatrix} d_w + \varepsilon^w_{it} \tag{2}$$

$$C^*_{it} = \alpha^c_i + X_{it}c_c + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} & W_{it} \end{bmatrix} d_c + \varepsilon^c_{it} \tag{3}$$

$$H^*_{it} = \alpha^h_i + X_{it}c_h + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} & W_{it} & C_{it} \end{bmatrix} d_h + \varepsilon^h_{it} \tag{4}$$

where W_{it}^{*} is a latent continuous variable related to the work decision, C_{it}^{*} is the corresponding latent variable for cohabitation and H_{it}^{*} is the latent variable for health. Furthermore, we allow for the current labour market status to affect cohabitation and health (in addition to the values from the previous year) and for the current cohabitation status to influence health. Identification of the system requires that we leave the contemporaneous variables out of one of the equations. Since our main interest is in the health variable, we will keep both endogenous variables in that equation.

Just as for the survival variable above, the actual realisation of the labour market status is determined according to

$$W_{it} = \begin{cases} 1 & if \quad A_{i,t} \cdot W_{it}^{*} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and similar conditions apply to the health and cohabitation variables, respectively.

Hence, we allow for state dependence in all the dependent variables. One problem in models with state dependence is that the initial conditions are not random draws from an unconditional distribution, but to some extent reflect pre-existing differences between individuals. Ignoring this fact would lead to biased estimates of the parameters of the model.

There are three possible solutions to this problem. The simplest solution is to treat the initial conditions as non-random constants in each cross-section. However, this approach entails the untenable assumption that the initial outcomes of the dependent variables are independent of unobserved heterogeneity and any independent variables. Alternatively, Heckman (1981) suggested approximating the conditional distribution of the initial condition. However, this method can be computationally burdensome and is not suitable for our needs. Finally, Wooldridge (2000) has suggested an alternative approach. It entails conditioning the distribution of the unobserved effect on the initial values and any exogenous explanatory variables. This approach enables us to choose a flexible and convenient auxiliary distribution, and hence we do not need to assume that the initial observations are drawn from a steady-state distribution. If we assume normality of the auxiliary distribution, the correction for the initial conditions problem becomes particularly straightforward.

Hence, we assume that

$$\boldsymbol{\alpha}_i | \mathbf{Y}_{i0}, \mathbf{Z}_i \backsim N\left(\boldsymbol{\theta}_0 + f\left(\mathbf{Y}_{i0}\right)\boldsymbol{\theta}_1 + \mathbf{Z}_i\boldsymbol{\theta}_2, \Omega\right) \tag{5}$$

where $\boldsymbol{\alpha}_i$ is the vector of individual effects for individual i (one for each regression), $\mathbf{Y}_{i0} = \begin{bmatrix} W_{i0} & C_{i0} & H_{i0} \end{bmatrix}$ is the vector of initial observations of the dependent variables, $\mathbf{Z}_i$ is a vector of time-independent exogenous variables and Ω is the covariance matrix of the individual effects (to be outlined in more detail below). The function $f\left(\mathbf{Y}_{i0}\right)$ includes all initial conditions and all possible linear interactions between them. Due to collinearity problems, we left time-dependent exogenous variables out of the specification.

Finally, we look at the correlation of the error terms more closely. Firstly, consider the fixed individualised effects $\boldsymbol{\alpha}_i$. We parameterise the covariance matrix of $\boldsymbol{\alpha}_i$ as

$$\Omega = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \omega_{ww} & \omega_{wc} & \omega_{wh} \\ 0 & \omega_{cw} & \omega_{cc} & \omega_{ch} \\ 0 & \omega_{hw} & \omega_{hc} & \omega_{hh} \end{pmatrix}$$

thus assuming that there is no random component to the individual effect in the survival equation (reflecting the lower informativeness of that variable) whereas the individual effects of the other three equations are to vary and be correlated with each other.

Concerning the vector of transitory shocks $\boldsymbol{\varepsilon}_{it}$, we assume a similar covariance structure:

$$Cov\left(\boldsymbol{\varepsilon}_{it}\right) = \begin{pmatrix} 1 & \sigma_{aw} & \sigma_{ac} & \sigma_{ah} \\ \sigma_{wa} & 1 & \sigma_{wc} & \sigma_{wh} \\ \sigma_{ca} & \sigma_{cw} & 1 & \sigma_{ch} \\ \sigma_{ha} & \sigma_{hw} & \sigma_{hc} & 1 \end{pmatrix}.$$

Thus, the off-diagonal parameters, denoted σ_{jk}, capture how random shocks (represented by the variable $\boldsymbol{\varepsilon}_{it}$) to the various dependent variables are correlated. If this variable is significant, there are random factors which systematically affect two dependent variables simultaneously. For example, a positive shock which leads to an individual finding a job in a distant region might also represent a negative shock to his/her cohabitation status. Again, disregarding this effect would lead to an overestimation of the causal relationship between the variables.

Finally, we also allow for autocorrelation in the error terms. This effect, represented by parameters ρ_j below, shows to what extent shocks to the dependent variables are persistent over time. These do not have an obvious interpretation but are necessary once we allow for persistence (i.e. state dependence) in the dependent variables since otherwise the coefficients would be biased.

Using all this information, we can build the $4T \times 4T$ covariance matrix Σ for the combined error terms $\boldsymbol{\alpha}_i+\boldsymbol{\varepsilon}_{it}$. A typical element in this matrix is determined as

$$\Sigma_{ij} = \sigma_{ab} + \rho_a^{|t-s|} \frac{\sqrt{1-\rho_a^2}\sqrt{1-\rho_b^2}}{1-\rho_a\rho_b} \omega_{ab}$$

where[2] $a = \left\lceil \frac{i}{4} \right\rceil$ and $b = \left\lceil \frac{j}{4} \right\rceil$ identify the corresponding estimating equation,

$$t = i - 4\left(a-1\right)$$

[2] $\lceil x \rceil$ denotes the value of x rounded upwards to the next integer. Hence, we order the columns and rows in the covariance matrix first according to year and then according to equation.

and

$$s = j - 4\,(b - 1)$$

identify the corresponding years. Hence, the element Σ_{ij} tells us how the error term at time t in equation a is correlated with the error term at time s in equation b. The parameter σ_{ab} is the covariance of the individual effect in equation a with the corresponding individual effect in equation b. Likewise, ω_{ab} is the covariance of shocks in equation a with shocks in equation b. Whenever $a = b$, the corresponding variance is included.

2.2 Estimating Procedure: Maximum Simulated Likelihood

We estimate the model outlined in equations (1) to (4) using maximum likelihood. However, given that the four dependent variables are limited to taking on discrete values only, estimating a dynamic model with the type of error structures we have outlined above poses some challenges. The main problem is that the likelihood function attains so many dimensions that it becomes intractable.

However, maximum simulated likelihood offers a solution to this problem. The idea of this estimator is to generate several series of error terms which are consistent with the data actually observed. We employ an algorithm proposed by Geweke (1989). In short, we draw a series of numbers from a uniform distribution and then transform them (in a straightforward application of the integral transform theorem) into a truncated normal variable that fits the observed data. The Geweke algorithm produces unbiased estimates of the parameters, and once it has been implemented, standard maximum likelihood techniques can be used to estimate the model.

In general, the simulation estimator produces consistent estimates of the parameters of the econometric model. Furthermore, Börsch-Supan and Hajivassiliou (1993) find that for 20 simulations per observation, the simulation bias is negligible. Hence, the estimator seems to be appropriate for our purposes.

2.3 A note on identification

It should be noted that identification is based on the distributional assumption that the system is jointly normally distributed. Exclusion restrictions could have been used in order to further improve the robustness of the estimates (cf. Wilde, 2000). We decided not to follow this route, however, since we are looking for a simple model which can be used for population simulations, and besides it is unclear to what extent there are suitable variables available for such an exercise. Instead, we seek to assure that endogeneity problems are kept to a minimum by using the correlated effects model and allowing the transitory shocks to be correlated across equations.

Furthermore, it should be noted that the assumption of a normal hazard function is non-standard, but it is required so that the simulation estimator can be used. Future research should analyse the sensitivity of our results with respect to this assumption.

2.4 Hypothesis Testing

We use a sample consisting of all permanent members of the BHPS. A consideration is the extent to which these individuals can be pooled in one single regression model. If different subsets of the population are systematically different, separate estimations for each subset might produce better results. This is of particular relevance for the sex dimension (i.e. risk factors and selection might work differently on men and women). To test for whether a pooled regression was more appropriate we performed a likelihood ratio test. The test statistic is

$$LR = 2\left(\ln L_1 - \ln L_2\right) \tag{6}$$

where L_1 is the model estimated with separate coefficient estimates for males and L_2 is the baseline model where sex-specific coefficients are constrained to be equal to zero. Obviously, if the parameters are the same for the two subgroups, forcing them to be equal will not change the likelihood very much, and the statistic in (10) will take on a low value. If, on the other hand, the assumption of equal coefficients is very restrictive, the unconstrained model will have a much better fit, reflected in $L_1 >> L_2$, and the test statistic will be significantly different from zero. Under weak regularity conditions the Likelihood Ratio test statistic is approximately chi-square distributed with degrees of freedom equal to the difference between the dimensions of the unrestricted and restricted models (i.e., the number of sex-specific parameters).

2.5 The Dataset

For the estimation, we use the first fourteen waves of the British Household Panel Survey. In this subsection, we define the variables used, explain how we have treated missing values and provide some summary statistics.

2.5.1 Variables

The variables used for estimation are presented in *Table 1.* The definitions are mostly obvious.

Missing variable values are a particularly large problem in this work, since excluding individuals with missing observations is not an option as it would bias the mortality rates. In general, some 2-3 per cent of observations were missing. Some of these were quite easy to impute from earlier or later observations: for instance, somebody who has a university degree in a certain year will have a university degree in any subsequent year. However, attrition problems in long data panels are well known.

By means of illustration, we fitted a simple probit model of the binary variable *NOMISS* (taking the value 1 if an individual does not attrite in any of the waves, and the value 0 otherwise). Parameter estimates are provided in *Table*

Table 1: Definition of Variables.

Variable	*Definition*	*BHPS Variables*
A	Alive	-
W	Has work	wJBHAS, wJBOFF
C	Married or cohabiting	wMASTAT
H	Health limits daily activities	wHLLT
c	Constant	-
Sex	Male or Female	xSEX
Age	Calendar year minus birth year	wDOBY
Age^2	Age squared	wDOBY
E1	University Degree	wQFACHI
E2	A levels or equivalent	wQFACHI
E3	O levels of equivalent	wQFACHI
E4	None of these	wQFACHI
Year	Calendar year (base: 1992)	-
$Year^2$	Calendar year squared	-
Non-white	Race is non-white	xRACE

2. According to the estimates, people who left the sample were on average more likely to be single, male, and non-white; as these variables turn out to be significant. In addition, the probability of remaining in the sample is monotonically increasing with the level of education: the parameter estimate for university graduate is more than twice the corresponding estimate for individuals with O levels. In the specification of *Table 2*, we treated recorded deaths as non-missing in all subsequent waves. We also tried an alternative specification where deaths are treated as missing observations after the first occurence, with similar results. Hence, the characteristics leading to attrition seem to be relatively independent of the exact specification.

Whenever attrition is based on factors which are systematically related to the response variable, even after conditioning on explanatory variables, a sample selection problem can result. This bias would make the parameter estimates useless for projection purposes, since it would be based on the relatively advantaged group which remains in the dataset, and not the entire population.

Table 2: Probit regression of NOMISS on model variables (1991).

Variable	*Coefficient*	*Std Error*	*T Stat*	*P Value*
Work	0.0393	0.0324	1.2100	0.2250
Cohabitation	0.0880	0.0303	2.9000	0.0040
Health	-0.0528	0.0401	-1.3200	0.1870
Sex	0.1070	0.0262	4.0800	0.0000
Age	0.0325	0.0043	7.5200	0.0000
Age^2	-0.0002	0.0000	-5.1300	0.0000
Edu1	0.3880	0.0542	7.1500	0.0000
Edu2	0.3161	0.0389	8.1200	0.0000
Edu3	0.1891	0.0345	5.4800	0.0000
Non-white	-0.4970	0.0680	-7.3100	0.0000
Constant	-0.9702	0.1003	-9.6800	0.0000
Pseudo R^2	0.0274			
N	9,865			

Hence, we need to correct for this potential bias.

In principle, there are two alternative ways to correct for attrition bias. One is to estimate a separate selection equation and allow for correlation in unobservables between this selection equation and the estimating equations. This type of correction is computationally burdensome, however, and is more suitable for linear panel data models (Wooldridge, 2002). The other alternative is inverse probability weighting (IPW). The assumption underlying IPW is that, conditional on certain variables in the first time period, $\mathbf{Z}_{i0}$, observed variables at time t ($\mathbf{Y}_{it}$ and $\mathbf{X}_{it}$) are independent of participation in the panel (s_{it}):

$$\Pr\left(s_{it}=1|\mathbf{Y}_{it},\mathbf{X}_{it},\mathbf{Z}_{i0}\right)=\Pr\left(s_{it}=1|\mathbf{Z}_{i0}\right) \tag{7}$$

This assumption has been called *selection on observables.* In our estimation, we make use of all dependent and independent variables from the first round; hence, $\mathbf{Z}_{i0} = (\mathbf{Y}_{i0}, \mathbf{X}_{i0})$. As shown by Wooldridge (2002), it is possible to include more covariates in the estimation of selection probabilities over time. However, such an approach relies on a stronger ignorability condition than the one in equation (7) and besides it requires attrition to be an absorbing state (which it is not). Hence, we decided to make use of first period variables only. We estimated attrition probabilities by means of an ordinary probit model and then weighted observations by the inverse of their predicted probability:

$$\ln L = \sum_{t=1}^{T} \sum_{i=1}^{N} \ln L_{it}^{1/\widehat{s}_{it}}$$

2.5.2 Descriptive Statistics

In what follows, we will provide some simple cross-tabulations of the raw data which we use in the estimates. We include all 9,865 permanent members of the panel. We start out by showing how disability evolves with age in *Table 3*. We have defined disability relatively widely as failing one or more ADLs.

Table 3: Disability Status (ADLs) by age, 1991. (Number of individuals, percentage in italics).

Age	Healthy	Disabled	Total
0-40	4,312	275	4,587
	94.00%	*6.00%*	*100.00%*
41-60	2,594	418	3,012
	86.12%	*13.88%*	*100.00%*
61-80	1,481	496	1,977
	74.91%	*25.09%*	*100.00%*
81+	168	121	289
	58.13%	*41.87%*	*100.00%*
Total	8,555	1,310	9,865
	86.72%	*13.28%*	*100.00%*

Table 3 shows the well documented relationship between health and age. For instance, among people in their fifties, less than 15 per cent have any physical impairment, whereas at the highest ages, more than 40 per cent of people have at least one impairment.

Next, we look at the role of cohabitation. In *Table 4*, we cross-tabulate the initial wave by health status and cohabitation status; again, the disabled status is assumed to be when one or more ADLs are failed. The two seem to be correlated; a person not cohabiting is thirty per cent more likely to be disabled than a person who is cohabiting.

Table 4: Health Status (ADLs) by cohabitation status, 1991.

Cohabit	*Healthy*	*Disabled*	*Total*
No	2,874	532	3,406
	84.38%	*15.62%*	*100.00%*
Yes	5,681	778	6,459
	87.95%	*12.05%*	*100.00%*
Total	8,555	1,310	9,865
	86.72%	*13.28%*	*100.00%*

Furthermore, the cohabitation status in the initial year seems to be quite a good predictor of the health status (including death) in subsequent years. In *Table 5* we cross-tabulate the cohabitation status in 1991 with the health status in 1996. Clearly, people who were cohabiting in 1991 had a higher chance of being alive and healthy in 1996. The mortality rate, in particular, seems to be high for non-cohabiting people when compared to cohabiting people.

Table 5: Health Status 1996 (ADLs) by cohabitation status 1991.

Cohabit	*Healthy*	*Disabled*	*Dead*	*Total*
No	1,914	432	235	2,581
	74.16%	*16.74%*	*9.10%*	*100.00%*
Yes	4,157	747	250	5,154
	80.66%	*14.49%*	*4.85%*	*100.00%*
Total	6,071	1,179	485	7,735
	78.49%	*15.24%*	*6.27%*	*100.00%*

In *Table 6*, we present the health status in 1996 by employment status in

1991. Not very surprisingly, individuals who were working in 1991 were much more likely to be alive and healthy five years later, than their non-working counterparts.

Table 6: Health Status 1996 (ADLs) by working status 1991 (individuals aged 26-60).

Work	*Healthy*	*Disabled*	*Dead*	*Total*
No	742	295	41	1,078
	68.83%	*27.37%*	*3.80%*	*100.00%*
Yes	3,207	360	47	3,614
	88.74%	*9.96%*	*1.30%*	*100.00%*
Total	3,949	655	88	4,692
	84.16%	*13.96%*	*1.88%*	*100.00%*

Finally, we look at the relationship between health and education. Figures are presented in *Table 7.* The education variable reflects the self reported educational attainment. The health variable is disability in 1996. As expected, a higher educational attainment is correlated with better health. The effect of education on health seems to be present at all levels of education.

2.5.3 Transition Rates

Finally, we provide some estimates of transition probabilities within the sample. As we have three dependent variables which are not mutually exclusive (Work, Cohabit, Health) and the two exclusive states Attrited and Dead, we have ten different states in total for each gender. The transition probabilities for females are given in *Table 8.* Each row corresponds to a certain initial state, and each column corresponds to a certain state in the following year. So for instance, the third row shows the distribution over different states at time t, conditional on individuals having reported being working, cohabiting and healthy (WCH) at time $t-1$. Accordingly, such an individual faces a 4.3 per cent risk of working and cohabiting but not being healthy (WC) in the next period.

The strong degree of persistence in the states shows up in the high probabilities on the diagonals. There are some exceptions to this pattern, however. For example, a person who is working and cohabiting but not healthy at time $t-1$, has a forty per cent chance of recovering to full health in the next period.

The corresponding transition matrix for males is given in *Table 9.* The results are very similar, even though males seem to be slightly more mobile between states than females. On the other hand, several transitions seem to occur with negligible probability, such as, for example, a transition from being working, cohabiting and healthy (category 'WCH') to being working only (category 'W').

Table 7: Health Status 1996 (ADLs) by educational attainment 1991.

Education	*Healthy*	*Disabled*	*Dead*	*Total*
None	2,132	754	373	3,259
	65.42%	*23.14%*	*11.45%*	*100.00%*
O Levels	2,103	231	61	2,395
	87.81%	*9.65%*	*2.55%*	*100.00%*
A Levels	1,301	137	44	1,482
	87.79%	*9.24%*	*2.97%*	*100.00%*
University	535	57	7	599
	89.32%	*9.52%*	*1.17%*	*100.00%*
Total	6,071	1,179	485	7,735
	78.49%	*15.24%*	*6.27%*	*100.00%*

As a final remark, it is worth noticing that amongst individuals who have left the sample, there is a six per cent chance of reappearing in the sample in the next year (roughly equal for males and females). Most of the individuals who re-appear in this way are working, cohabiting and healthy. Hence, treating attrition as an absorbing state would lead to a loss of several observations (1,659 to be exact).

3 Results

3.1 Hypothesis Testing

Becauise the dataset we are using is large, we decided to base specification on smaller subsets of the data consisting of 5 per cent of the sample or 493 individuals in total. The results of the Likelihood ratio test are presented in *Table 10.*

Clearly, the parameter estimates for males and females are significantly different. Hence, we proceeded to estimate the model for the full dataset separately for males and females.

Table 8: Transition Rates, females.

	Att.	Dead	WCH	WC	WH	W	CH	C	H	None
Att.	0.939	0.011	0.015	0.001	0.009	0.001	0.009	0.004	0.007	0.005
Dead	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
WCH	0.043	0.001	0.823	0.043	0.020	0.001	0.058	0.007	0.003	0.000
WC	0.038	0.005	0.396	0.394	0.017	0.007	0.052	0.085	0.003	0.003
WH	0.065	0.001	0.086	0.005	0.729	0.038	0.007	0.001	0.060	0.008
W	0.070	0.002	0.050	0.026	0.426	0.273	0.008	0.010	0.052	0.084
CH	0.049	0.005	0.104	0.005	0.004	0.000	0.711	0.091	0.026	0.004
C	0.051	0.027	0.018	0.019	0.001	0.001	0.204	0.644	0.010	0.025
H	0.054	0.016	0.007	0.001	0.061	0.004	0.018	0.002	0.680	0.157
None	0.066	0.058	0.001	0.001	0.009	0.008	0.005	0.009	0.199	0.645

Table 9: Transition Rates, males.

	Att.	Dead	WCH	WC	WH	W	CH	C	H	None
Att.	0.941	0.007	0.023	0.002	0.011	0.000	0.005	0.004	0.005	0.002
Dead	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
WCH	0.052	0.002	0.859	0.030	0.019	0.000	0.033	0.005	0.002	0.000
WC	0.072	0.015	0.402	0.385	0.008	0.011	0.031	0.072	0.002	0.002
WH	0.085	0.002	0.094	0.003	0.724	0.023	0.005	0.000	0.058	0.005
W	0.062	0.012	0.068	0.015	0.406	0.316	0.006	0.000	0.034	0.080
CH	0.044	0.014	0.080	0.003	0.003	0.000	0.714	0.122	0.016	0.003
C	0.051	0.050	0.014	0.011	0.000	0.000	0.174	0.678	0.006	0.015
H	0.072	0.022	0.013	0.001	0.135	0.003	0.017	0.003	0.610	0.124
None	0.059	0.082	0.001	0.000	0.021	0.013	0.003	0.017	0.207	0.595

3.2 Parameter Estimates

Estimation results for the **mortality hazard rate** are presented in *Table 11*. In the table, parameter estimates are grouped according to the definitions in section 2. The first three rows contain parameter estimates for time-varying exogenous variables; the following three rows contain parameter estimates for time-varying endogenous variables; the next vector contains coefficients for variables which remain constant over time; and finally we present estimates related to the correlation structure of the error terms. Furthermore, the left part of the table reports parameter estimates for males and the right part contains the corresponding estimates for females. We report pseudo R^2 as a crude measure of goodness of fit – which applies to the whole system of equations. This statistic suggests our model has a reasonably good fit to the data, implying that around two thirds of the variation is explained by the model for males as well

Table 10: Test Results. LR Test for pooling of males and females.

	Total	Males	Females
Likelihood, full model		-1840.9609	-1555.4329
Likelihood, restricted model	-3613.5421		
LR Chi2(df)	217.1483(95)		
P	0.0000		

as females.

Concerning the **time-varying exogenous variables** (age, age squared and the linear time trend) we notice that they are all significant and with the expected signs. Mortality increases with age, at an increasing rate. Also the year effect is positive and statistically significant, suggesting that the survival probability has improved over the sample period. However, this trend seems to be more pronounced for males than for females, which suggests that the gap in life expectancy between the sexes might indeed be closing.

Concerning **time-varying endogenous variables** (W_{t-1}, C_{t-1}, and H_{t-1}) we notice that cohabitation status and health in the previous period are important determinants of current-period male survival probability, whereas the previous employment status is insignificant. This is also to be expected since the bulk of mortality occurs post-retirement. For females, however, only previous health is an important determinant of mortality.

Concerning the **constant explanatory variables**, denoted θ, we notice that education has the expected sign, but that the education effect on mortality seems to be non-monotonic for males: individuals with only O-levels seem to have better survival prospects than those who also have A-level qualifications. For females, the education coefficients are generally smaller and statistically less significant. Also, the initial year employment status is significant for males – thus suggesting that there is a long-term relationship between an individual's propensity to work and their mortality rate. For females, on the other hand, initial period cohabitation and health status are significant and positive, but not employment. This finding suggests that the initial employment status is of less importance to females – possibly because there is less selection into employment based on health (and more based on family characteristics) taking place.

Finally, concerning the parameters of the correlation structure of the error terms, we notice that there is a strong and positive relationship between the unobserved heterogeneity in the survival and in the work dimension (ω_{AW}) for both sexes. This finding suggests that a substantial part of the correlation between survival and employment is due to background factors affecting both. Similar, but weaker, correlations are observed between survival and cohabitation (ω_{AC}), as well as survival and health (ω_{AH}). Interestingly, the correlation in

Table 11: Estimation results, Survival

Parameter vector	Variable	Males		Females	
		Coefficient	*p Value*	*Coefficient*	*p Value*
c	*Age*	-1.856	0.00	-1.604	0.06
	Age²	-0.817	0.08	-1.138	0.07
	Year	0.147	0.00	0.099	0.01
d	W_{t-1}	0.080	0.11	0.087	0.17
	C_{t-1}	0.134	0.01	-0.048	0.32
	H_{t-1}	0.585	0.00	0.431	0.00
θ	*Constant*	2.931	0.00	3.225	0.00
	University	0.201	0.01	0.132	0.12
	A Levels	0.071	0.12	0.090	0.11
	O Levels	0.105	0.01	0.079	0.08
	Non-white	0.226	0.01	0.782	0.00
	W_0	0.415	0.06	0.937	0.88
	C_0	-0.014	0.86	0.151	0.03
	H_0	-0.009	0.89	0.144	0.00
	$W_0 C_0$	-0.194	0.43	-0.713	0.91
	$W_0 H_0$	-0.464	0.04	-1.090	0.86
	$C_0 H_0$	0.021	0.79	-0.013	0.86
	$W_0 C_0 H_0$	0.256	0.32	0.790	0.90
ϱ	ϱ_A	-0.645	0.00	-0.566	0.00
ω	ω_{AW}	0.763	0.00	0.799	0.00
	ω_{AC}	0.702	0.00	-0.601	0.00
	ω_{AH}	0.410	0.00	0.842	0.00
	Pseudo R^2	0.646		0.627	
	N	4,577		5,288	

unobserved heterogeneity between survival and cohabitation is of opposite signs for males and females.

Next, results for the **work** equation are presented in *Table 12*. The table follows the same structure as *Table 11*. Concerning the **time-varying exogenous variables** (age, age squared and the linear time trend), we find that the non-linear relationship between age and employment is strongly statistically significant. Also the time trend is positive, implying that employment opportunities, especially for males, have improved strongly over the time period.

Concerning **time-varying endogenous variables** (W_{t-1}, C_{t-1}, and H_{t-1}) it is notable that we find very strong state dependence in the work variable. Also the cohabitation status and health in the previous year are important determinants of current-year employment status. As expected, health is a more important factor than cohabitation. However, whereas previous cohabitation tends to increase the male employment rate, the effect is negative for females.

Concerning the **constant explanatory** variables (θ), we get the reasonable

Table 12: Estimation results, Employment

Parameter vector	Variable	Males Coefficient	Males p Value	Females Coefficient	Females p Value
c	*Age*	7.412	0.000	7.123	0.000
	Age^2	-11.120	0.000	-10.276	0.000
	Year	0.140	0.000	0.084	0.000
d	W_{t-1}	2.111	0.000	2.117	0.000
	C_{t-1}	0.193	0.000	-0.084	0.000
	H_{t-1}	0.425	0.000	0.338	0.000
θ	*Constant*	-2.668	0.000	-2.737	0.000
	University	0.365	0.000	0.316	0.000
	A Levels	0.152	0.000	0.164	0.000
	O Levels	0.172	0.000	0.132	0.000
	Non-white	-0.075	0.055	-0.105	0.004
	W_0	0.159	0.286	0.437	0.000
	C_0	-0.387	0.000	0.059	0.452
	H_0	0.174	0.050	0.172	0.016
	$W_0 C_0$	0.712	0.000	0.070	0.536
	$W_0 H_0$	0.161	0.292	-0.070	0.485
	$C_0 H_0$	0.282	0.008	-0.003	0.972
	$W_0 C_0 H_0$	-0.542	0.002	0.005	0.969
ϱ	ϱ_W	-0.340	0.000	-0.347	0.000
ω	ω_{AW}	0.763	0.000	0.799	0.000
	ω_{WC}	0.643	0.000	-0.352	0.000
	ω_{WH}	0.259	0.000	0.621	0.000
σ	σ_{WW}	0.122	0.000	0.092	0.000
	σ_{WC}	0.023	0.000	0.034	0.000
	σ_{WH}	0.046	0.000	0.002	0.770
	N	4,577		5,288	

result that people with higher education tend to work longer, although again, the employment prospects are not increasing monotonically with higher male education. Nevertheless, parameter estimates are very similar for males and females. Furthermore, the initial cohabitation status seems to have a strong predictive power for future employment status for males, whereas the other initial conditions seem to be of less importance. For females, it is rather initial employment status and health that are important, whereas cohabitation seems to be largely irrelevant.

Finally, concerning the parameters of the correlation structure of the error terms, we notice the somewhat unexpected finding that although shocks to employment are strongly correlated with shocks to the other dependent variables, for males this corrrelation seems to be weaker concerning shocks to health (ω_{WH}). However, the individual effect in employment for males exhibits a stronger correlation with health (σ_{WH}) than with cohabitation (σ_{WC}), suggesting that unobserved individual differences are more important in explaining the correlation between the two. For females, one interesting finding is that there is positive correlation in unobserved characteristics between employment and cohabitation (σ_{WC}). Hence, this correlation in background factors tends to offset the negative direct effect which cohabitation has on female employment rates.

Next, we present parameter estimates for the cohabitation variable in *Table 13*. Again, the **time-varying exogenous variables** are largely as expected, although it interesting to notice that the cohabitation probability actually has a positive trend – for both sexes. This is good news for policy-makers who have been concerned about an expected breakdown of family structures.

Table 13: Estimation results, Cohabitation

		Males		Females	
Parameter vector	**Variable**	*Coefficient*	*p Value*	*Coefficient*	*p Value*
c	*Age*	2.552	0.000	1.692	0.000
	Age^2	-3.383	0.000	-3.138	0.000
	Year	0.086	0.000	0.040	0.115
d	W_{t-1}	0.286	0.000	0.386	0.000
	C_{t-1}	2.926	0.000	3.137	0.000
	H_{t-1}	0.144	0.000	0.028	0.354
	W_t	-0.334	0.000	-0.461	0.000
θ	*Constant*	-2.325	0.000	-1.933	0.000
	University	0.158	0.000	0.105	0.004
	A Levels	0.079	0.000	0.029	0.320
	O Levels	0.096	0.000	-0.047	0.070
	Non-white	0.101	0.004	-0.084	0.059
	W_0	0.303	0.007	0.190	0.098
	C_0	0.993	0.000	0.789	0.000
	H_0	0.243	0.000	0.107	0.191
	$W_0 C_0$	-0.467	0.000	-0.101	0.452
	$W_0 H_0$	-0.207	0.069	0.039	0.748
	$C_0 H_0$	-0.430	0.000	-0.146	0.112
	$W_0 C_0 H_0$	0.352	0.008	0.011	0.940
ϱ	ϱ_C	-0.181	0.000	-0.228	0.000
ω	ω_{AC}	0.702	0.000	-0.601	0.000
	ω_{WC}	0.643	0.000	-0.352	0.000
	ω_{CH}	-0.168	0.000	-0.538	0.000
σ	σ_{CC}	0.020	0.000	0.138	0.000
	σ_{WC}	0.023	0.000	0.034	0.000
	σ_{CH}	0.006	0.382	0.026	0.001
	N	4,577		5,288	

For the **time-varying endogenous variables** (W_{t-1}, C_{t-1}, and H_{t-1}) we find that previous realisations of the dependent variables are strongly significant, and with expected signs. The state dependence in the cohabitation variable is very strong, but for males also previous health is important and positively related to current cohabitation status. Employment also is strongly significant, and our results suggest that a transition into employment actually decrease the probability of cohabitation for both sexes.

Concerning the **constant explanatory variables** (θ), one interesting finding is that cohabitation also is strongly linked with education, so that people with a university degree are much more likely to be cohabiting than others. However, the education gradient in cohabitation is much more pronounced for males than for females. Also, non-white males are more likely to be cohabiting than the white majority, whereas the ethnicity coefficient is negative for females. Furthermore, it is noticeable that the initial conditions are very important in determining subsequent cohabitation status for males. For females, on the other hand, the initial state seems to be of less importance - with the notable exception of initial cohabitation. Finally, it can be noted that the variance in fixed unobservable factors (σ_{CC}) is relatively low for males, suggesting that most of the persistence in cohabitaion is related to observable factors and hence captured by the initial conditions. For females, on the other hand, such unobserved heterogeneity seems to be of greater importance. Interestingly, shocks to cohabitation and health are negatively correlated for both sexes (as captured by σ_{CH}), a finding which might be capturing the effect on health of bereavement.

Finally, we report results for the health equation in *Table 14*. Again, we get the predictable result that age is an important factor in determining an individual's health. What is more unexpected , however, is the negative time trend, which is strongly significant for both sexes. This finding is seemingly consistent with official statistics - which suggest that the long-term trends in disability depend on the definition of disability used. However, it should be noted that the ambiguity concerning trends in healthy life expectancy is related to the proportion of additional life years spent in ill health. Our results, on the other hand, suggest a negative trend in the health variable itself, which would actually suggest that *all* years added to life are spent in poor health unless the trend is offset by changes in the other independent variables. Hence, it could be speculated that this estimated negative trend reflects increased prevalence of obesity or other adverse trends which are assumed to have negative implications for public health, but that this negative trend is neutralised by concurrent improvements in employment or educational attainment.

Next, concerning the **time-varying endogenous variables**, we find that all but health itself are strongly significant for males. Hence, previous and past employment both have positive effects on health. Concerning cohabitation, the overall effect is positive, and particularly a transition into cohabitation is found to have a positive effect on health. For females, employment is of great importance whereas the other variables seem to be less relevant. Quite surprisingly, the effect of previous health on current health is estimated to be negative and significant for females, thus suggesting that all persistence in bad health is due

Table 14: Estimation results, Health

Parameter vector	Variable	Males Coefficient	Males p Value	Females Coefficient	Females p Value
c	*Age*	0.588	0.156	4.111	0.000
	Age^2	-1.121	0.004	-5.046	0.000
	Year	-0.348	0.000	-0.430	0.000
d	W_{t-1}	0.129	0.000	0.009	0.656
	C_{t-1}	-0.416	0.000	-0.015	0.548
	H_{t-1}	0.015	0.397	-0.031	0.031
	W_t	0.633	0.000	0.377	0.000
	C_t	1.171	0.000	0.073	0.012
θ	*Constant*	-0.555	0.000	-1.017	0.000
	University	0.362	0.000	0.211	0.000
	A Levels	0.292	0.000	0.176	0.000
	O Levels	0.126	0.000	0.170	0.000
	Non-white	-0.174	0.000	-0.274	0.000
	W_0	0.414	0.018	1.201	0.000
	C_0	-0.832	0.000	0.028	0.684
	H_0	1.722	0.000	1.664	0.000
	$W_0 C_0$	-0.085	0.660	-1.158	0.000
	$W_0 H_0$	-0.523	0.004	-1.190	0.000
	$C_0 H_0$	0.045	0.677	0.035	0.646
	$W_0 C_0 H_0$	0.296	0.145	1.104	0.000
ϱ	ϱ_H	0.627	0.000	0.590	0.000
ω	ω_{AH}	0.410	0.000	0.842	0.000
	ω_{WH}	0.259	0.000	0.621	0.000
	ω_{CH}	-0.168	0.000	-0.538	0.000
σ	σ_{HH}	0.612	0.000	0.513	0.000
	σ_{WH}	0.046	0.000	0.002	0.770
	σ_{CH}	0.006	0.382	0.026	0.001
	N	4,577		5,288	

to permanent factors and not state dependence.

Next, concerning the constant variables, we observe a clear education gradient in health, which is somewhat stronger for males than for females. The higher an individual's educational attainment is, the higher is their probability of being healthy. Ethnicity is significant: non-whites are estimated to be in poorer health than the white majority.

Finally, there is a high degree of unobserved heterogeneity as captured by the parameter σ_{HH}. Interestingly, for males, this individual effect is strongly correlated with the unobserved heterogeneity in employment status, but not with cohabitation - whereas the opposite holds for females.

4 Simulations

In this section, we provide some examples of how our results can be used for population simulations. We do this in two steps. First, we show how the parameter estimates can be used to simulate life trajectories for an entire population. In a second step, we analyse how moving an individual from one state to another – but keeping all other personal characteristics constant – changes the life trajectory of that individual. These changes are then compared with the overall differences between the two sub-populations.

4.1 Population Simulations

In this part, we simulated a population of 100,000 individuals for all different combinations of initial states. Since we have four levels of education and three different states (W, C and H), we have 32 ($= 4 \cdot 2^3$) different combinations to consider for each sex. Given the richness of the material we only report the expected time spent in different states for different types of individuals. We focus on 50 year old individuals, but the simulation could of course be made for any age. We did not produce any simulations for the non-white population since there are indications that this subsample is not very representative.

In *Table 15*, we set out the simulation results for males. The first column shows the number of individuals in the respective groups, for a general population of 100,000. We partitioned the 2004 wave of the BHPS to obtain these values. Columns 2-5 contain the educational attainment and initial states of the different individuals. Column 6 reports healthy life expectancy, column 7 overall life expectancy and column 8 reports the time spent working. The bottom row provides weighted averages for the entire population. The groups are ranked according to their healthy life expectancy (column 6).

According to our estimates, individuals who have a university degree and are healthy, working and cohabiting at the outset have the highest chances of living a long healthy life. Their healthy life expectancy is 25 years higher than the healthy life expectancy of an individual with no formal education who is unhealthy and not working at the outset. We also see that the population as a whole can expect to spend 8.5 years (ie 33.4 − 24.9 = 8.5 years) of their remaining lives in disability.

In *Table 16*, we report corresponding results for females. Whilst females have 3.5 years more of overall life expectancy than males, their healthy life expectancy is actually lower. Hence, females spend 13.5 years in disability on average. They also retire earlier than males and can expect to cohabit for a shorter period than males.

4.2 Comparative Analysis

In the next step, we analyse what our model predicts would happen if an individual were moved from one starting state to another, without changing the

Table 15: Life trajectories for 50-year old individuals. Base year: 2004. Males.

1	2	3	4	5	6	7	8
N	*EDU*	*W*	*C*	*H*	*HLE*	*LE*	*WLE*
12,969	1	1	1	1	31.6	38.2	18.0
1,120	1	0	1	1	28.8	36.9	7.8
24,158	2	1	1	1	28.4	34.7	15.3
18,938	3	1	1	1	27.7	34.9	15.4
930	1	1	0	1	27.1	38.0	17.8
1,310	2	0	1	1	25.7	33.3	5.6
1,490	3	0	1	1	25.1	33.6	5.7
11,659	4	1	1	1	24.5	31.6	13.3
650	2	1	0	1	24.1	34.1	15.1
2,050	3	1	0	1	23.4	34.6	15.3
1,870	4	0	1	1	22.0	30.3	4.1
190	1	0	0	1	20.5	33.8	5.3
1,030	4	1	0	1	20.3	31.3	13.0
2,890	1	1	1	0	18.4	32.2	14.3
1,210	2	0	0	1	18.0	30.0	3.5
1,210	3	0	0	1	17.3	30.4	3.6
370	1	0	1	0	16.8	32.4	5.9
190	1	1	0	0	16.3	34.4	13.0
3,360	2	1	1	0	16.0	28.8	11.8
4,200	3	1	1	0	15.2	29.0	11.9
1,680	4	0	0	1	14.9	27.4	2.4
750	2	0	1	0	14.4	28.9	3.9
370	2	1	0	0	13.6	30.7	10.4
470	3	0	1	0	13.6	29.1	4.1
90	3	1	0	0	13.0	31.1	10.5
2,150	4	1	1	0	12.8	26.1	9.8
560	4	0	1	0	11.3	26.0	2.7
0	1	0	0	0	11.0	29.5	4.5
90	4	1	0	0	10.5	27.9	8.4
470	2	0	0	0	9.2	25.8	2.9
370	3	0	0	0	8.4	26.1	2.9
1,210	4	0	0	0	6.5	23.1	1.8
100,000					**24.9**	**33.4**	**13.4**

other characteristics. This analysis is possible since we have assumed correlated individual effects in equation (5) and thus taken the starting position into account. As a point of reference, we compare these estimates with the corresponding differences between the relevant subgroups according to *Table 15* and *Table 16* above.

In *Table 17*, we show the effects of a change in the initial working state from non-working to working for males. To understand this table, we need to explain what we mean by the terms "gap" and "gain".

The term "gap" is the difference between the two average periods for the relevant sub-groups of *Table 15*, where one average spell is calculated with the characteristic "working" and the other is calculated with the characteristic "not working". For example, the "gap" in healthy life expectancy is 1.6 years when

Table 16: Life trajectories for 50-year old individuals. Base year: 2004. Females.

1	2	3	4	5	6	7	8
N	*EDU*	*W*	*C*	*H*	*HLE*	*LE*	*WLE*
1,550	1	0	1	1	29.4	41.1	5.5
8,360	1	1	1	1	29.1	39.9	14.8
1,470	2	0	1	1	28.4	39.8	4.1
12,690	2	1	1	1	28.2	38.8	12.9
4,260	3	0	1	1	28.2	39.6	3.8
22,060	3	1	1	1	28.0	38.6	12.6
4,020	4	0	1	1	25.8	37.2	2.8
9,830	4	1	1	1	25.4	36.2	10.9
1,010	1	1	0	1	21.8	39.7	13.6
1,320	2	1	0	1	21.0	38.6	11.8
1,630	3	1	0	1	20.9	38.5	11.5
1,930	4	1	0	1	18.6	35.9	9.8
310	1	0	0	1	17.5	36.5	3.9
2,940	1	1	1	0	17.0	34.2	12.8
850	2	0	0	1	16.8	35.5	2.8
150	1	0	1	0	16.8	37.4	4.3
2,010	3	0	0	1	16.6	35.2	2.6
620	2	0	1	0	16.2	36.4	3.2
3,560	2	1	1	0	16.2	33.2	10.9
6,040	3	1	1	0	16.0	32.9	10.6
850	3	0	1	0	16.0	36.2	2.9
3,410	4	0	0	1	15.1	32.9	1.9
460	1	1	0	0	14.1	39.1	12.1
1,320	4	0	1	0	14.1	33.8	2.1
3,480	4	1	1	0	14.0	30.5	8.9
310	2	1	0	0	13.4	38.1	10.3
620	3	1	0	0	13.3	37.7	10.0
540	4	1	0	0	11.5	35.5	8.4
230	1	0	0	0	8.4	32.3	3.0
310	2	0	0	0	8.1	31.3	2.1
540	3	0	0	0	7.9	30.9	2.0
1,320	4	0	0	0	6.8	28.9	1.3
100,000					**23.4**	**36.9**	**10.0**

comparing employed males who are cohabting, unhealthy and educated to university degree level (18.4 years) with that of their unemployed counterparts (16.8 years). Overall, we see that the difference in healthy life expectancy between working and non-working individuals is 3.7 years at the age of 50, as reported in the bottom row of the table. The difference in overall life expectancy is 2.5 years, and the difference in working life expectancy is, as expected, much larger: 9.5 years.

The "gain" is based on a counterfactual experiment. Imagine that we take a male who is in state A (for instance, unemployed, healthy and single) and we move him into state B (for instance, working, healthy and single) without changing the unobservable characteristics - as captured by the α_i terms defined in equation (5). The "gain" is how this new, artificial, group - having observable

Table 17: Effects of changes in the initial work state. Males. Base year: 2004

				HLE		LE		WLE	
N	EDU	C	H	*Gap*	*Gain*	*Gap*	*Gain*	*Gap*	*Gain*
190	1	0	1	6.6	0.6	4.2	0.1	12.5	3.3
1,210	3	0	1	6.1	0.4	4.2	0.1	11.7	3.1
1,210	2	0	1	6.1	0.6	4.1	0.2	11.6	3.2
1,680	4	0	1	5.5	0.6	4.0	0.2	10.6	3.0
1,120	1	1	1	2.8	0.3	1.4	0.1	10.2	3.0
1,310	2	1	1	2.7	0.5	1.4	0.1	9.7	3.2
1,490	3	1	1	2.6	0.5	1.3	0.2	9.7	3.2
1,870	4	1	1	2.5	0.5	1.3	0.1	9.2	3.3
0	1	0	0	5.3	6.8	4.9	0.4	8.5	2.2
370	1	1	0	1.6	0.9	-0.2	0.3	8.4	3.4
750	2	1	0	1.5	0.9	-0.1	0.3	7.8	3.3
470	3	1	0	1.6	0.9	-0.1	0.4	7.8	3.3
370	3	0	0	4.6	0.7	5.0	0.3	7.6	3.1
470	2	0	0	4.5	0.6	5.0	0.3	7.5	3.1
560	4	1	0	1.5	0.8	0.0	0.4	7.1	3.2
1,210	4	0	0	4.0	0.6	4.8	0.3	6.6	2.9
14,279				**3.7**	**0.5**	**2.5**	**0.2**	**9.5**	**3.1**

characteristics of group B but unobservable characteristics of group A - fares in comparison with their counterparts (in terms of unobservable characteristics) of group A.

If one considers the entire "gap", the "gain" component is the change one gets when the observable characteristics are changed and the residual (i.e. "gap" minus "gain") is the effect of changing the unobservable characteristics. The unobservable characteristics would be anything which is not captured by the independent variables – e.g. intelligence, preferences, social networks etc.

Turning back to *Table 17*, it should be noted from the bottom row that moving a non-working male of age 50 into employment has the potential to increase their healthy life expectancy by 0.5 years (reported in the 'gain' column). This is much less than the actually observed difference between working and non-working individuals (reported in the 'gap' column), which is 3.7 years. Similarly, even though currently working individuals can expect to work for 9.5 years more than non-working individuals; those who are initially not working gain only 3.1 years of expected working time by being moved into employment. In the table, we have ranked individuals according to their potential gains in working life expectancy from a transition into employment. According to this ranking, university educated individuals who are initially cohabiting and unhealthy stand to gain the most from a transition into employment (3.4 years). At the other end of the scale, single individuals who are initially unhealthy and hold a university degree stand to gain the least (2.2 years) from such a transition. The obvious policy implication is that targeted measures to bring people back to work will have quite different long-term effects depending on which sub-group they affect.

The corresponding results for females are presented in *Table 18*. For females, we notice that the gap in healthy life expectancy between employed and unemployed individuals is much smaller than for males (the average difference is 1.5 years compared to 3.7 years for males). The differences in working life expectancy, however, are similar. On the other hand, the estimated gain in working years from a transition into employment is estimated to be much larger for a female than for a male. Accordingly, a previously unemployed female can on average expect to gain 4.4 working years from a transition into employment. The spread in estimated gains is larger than for males. Hence, a healthy, cohabiting female with a university degree would gain 5.4 working years from a transition into employment (which is a substantial share of the gap to their working counterparts, estimated to be 9.4 years). On the other hand, an unhealthy single female without educational qualification can expect to gain only 3.3 years from such a transition. This is far below the observed discrepancy between the two groups, which is estimated to be 7.1 years. Hence, differences between employed and unemployed individuals might be a very poor indicator of the potential gains for an individual from a transition into employment.

Table 18: Effects of changes in the initial work state. Females. Base year: 2004

				HLE		LE		WLE	
N	EDU	C	H	*Gap*	*Gain*	*Gap*	*Gain*	*Gap*	*Gain*
310	1	0	1	4.3	3.6	3.2	-2.7	9.7	4.7
1,550	1	1	1	-0.3	-0.3	-1.2	-4.1	9.4	5.4
230	1	0	0	5.7	3.5	6.8	-2.5	9.1	5.0
850	2	0	1	4.2	1.7	3.1	-5.3	8.9	3.8
2,010	3	0	1	4.2	1.1	3.3	-4.7	8.9	4.1
1,470	2	1	1	-0.2	-2.2	-1.1	-6.5	8.8	4.7
4,260	3	1	1	-0.2	-2.6	-1.0	-5.8	8.8	5.1
150	1	1	0	0.2	0.9	-3.2	-4.7	8.5	5.0
310	2	0	0	5.3	1.6	6.8	-5.2	8.2	3.9
4,020	4	1	1	-0.5	-3.4	-1.0	-6.8	8.1	4.6
540	3	0	0	5.4	1.3	6.8	-4.4	8.0	4.1
3,410	4	0	1	3.6	0.4	3.0	-5.4	8.0	3.6
620	2	1	0	-0.1	-0.9	-3.2	-7.2	7.8	4.1
850	3	1	0	0.0	-1.4	-3.3	-6.6	7.7	4.5
1,320	4	0	0	4.8	0.4	6.6	-5.5	7.1	3.3
1,320	4	1	0	-0.1	-2.0	-3.3	-7.3	6.8	3.8
23,220				**1.5**	**-1.0**	**0.6**	**-5.7**	**8.3**	**4.4**

If we compare the potential gains from a transition into employment in the two tables, it becomes evident that two traits seem to be of greater importance than the others: gender and cohabitation. Females stand to gain much more from a transition into employment than males, and cohabiting individuals typically get better gains than single individuals. This pattern is quite different from the pattern for the overall differences between working and non-working individuals, where instead educational attainment (for females) and health status (for males) seem to be the most important factors.

5 Conclusion

In this paper, we have developed an econometric model to estimate the population dynamics of survival, employment, cohabitation and disability in the UK population. Using panel data techniques for limited dependent variables and simulation techniques, we are able to estimate a model which aims at explaining how these variables are affected by exogenous factors, and also by each other, in a dynamic setting. We use the entire BHPS and are able to fit a model which seems to have very high explanatory power. By using inverse probability weights to correct for attrition, we try to ensure that the sample remains representative over time. This is a particularly difficult task as far as health is concerned, since individuals who move into institutional care are excluded from the sample.

Our econometric model has reasonable goodness of fit and delivers a host of interesting findings. Firstly, it was found that the parameters for males and females are significantly different, so that separate models should be estimated for each gender. Concerning survival probabilities, we found that cohabitation status and health status both have a strong impact on subsequent survival probabilities. Furthermore, initial conditions also have strong explanatory power for survival. There are, however, some indications that the recording of deaths is incomplete in the sample and that this problem might not have been entirely solved by correcting for attrition.

Concerning employment status, we observed a very strong education gradient, which seems to be very similar for both sexes. Accordingly, people with university degrees are much more likely to be working than those with only secondary education, and individuals with secondary education are much more likely to be working than those with no education. This finding is consistent with Butt et al (2008). Furthermore, we find that cohabitation significantly increases the labour market participation rate for males, but reduces it for females – the latter probably due to having children.

Regarding cohabitation status, we found that education seems to be an important explanatory factor for males but not necessarily for females. For females, on the other hand, there seems to be a somewhat stronger state dependence in the endogenous variables, so that previous cohabitation and employment explain a large proportion of the variation. As regards health status, we find that education is important for both sexes, but much more so for males. Furthermore, unobserved heterogeneity seems to be of particular importance for health. One interesting finding concerns the time trends in health: the time trend for survival is positive for both sexes, whereas the corresponding time trend for health is negative. This implies that our model gives support for the extension of morbidity hypothesis mentioned in the introduction.

We also used the parameter estimates to perform some simple population projections for 50-year olds, using 2004 (the latest available wave of the BHPS) as our initial year. Parameters from our regression analysis could be used to produce life tables for different subgroups of the population. We reported this information in the form of 'risk ladders', which provide simple summary statistics such as life expectancy, healthy life expectancy and working life expectancy for all the subgroups considered.

We found that the estimated population averages were largely consistent with recent ONS statistics, which we see as a further indication that our model is good (Government Actuary's Department, 2006). Since we defined disability rather widely, we find that the expected time spent in disability is 8.5 years for males and 13.5 years for females, whereas the healthy life expectancy is 24.9 years for males and 23.4 years for females. So females become disabled before males but end up living longer on average. Although not reported above, our model can also be used to predict the availability of a spouse carer during the time spent in disability. According to our estimates, males spend only three out of their 8.5 years in disability on average without a cohabiting partner (i.e. as widowers, divorcees, or never-married/cohabited). Due to typical differences in age at marriage and in longevity, females are projected to spend around 6.3 our of their total 13.5 years of disability without a partner. This is the time period during which the need for formal long-term care services would be particularly strong. According to our estimates, there is considerable variation in how long this period is – ranging from two years to more than 20 years, depending on various characteristics at 50.

In the last section, we illustrated the importance of unobserved heterogeneity by analysing the effects of moving individuals from one starting state to another without changing their other characteristics. It transpired that the resulting life expectancies can be quite different from the averages of the 'destination state'. For example, the difference in working life expectancy between working and non-working females is around 8 years on average. However, moving an individual from non-employment to employment would only result in a gain of 4.4 years on average. Hence, policy analysis should not be based on simple comparisons of individuals in different states.

One immediate application of our model would be to estimate the impact of changes in state pension age. Assuming that the state retirement is exogenous to individual health, we can use our model to estimate how large a proportion of a certain cohort remains healthy until certain ages. For example we found that males who are working and healthy at 50, have an 81.9 per cent chance of still being healthy at age 60. The corresponding figure for the entire male population (i.e., including also those who are initially not working and/or not healthy) is 75.5 per cent and the corresponding figure for initially working and healthy females is 77.0 per cent (69.4 per cent for the entire female population). However, the state pension age of around 65 coincides with a time of relatively quick deterioration in health for males and females. Hence, only 72 per cent of the initially working and healthy males can be expected to be healthy at 67 (females: 68.6 per cent) and the corresponding figure for the age of 70 is 66.3 per cent (females: 64.0 per cent). In conclusion, increasing the state pension ages might not be a very effective solution to counteract the impact of ageing on the economy. On the other hand, our estimates clearly show that working in itself has a positive effect on health for both sexes, so this issue would clearly require some further investigation.

6 References

Bajekal, M. (2002), *Care Homes and their Residents*, London: The Stationery Office.

Bebbington, A. C. & Darton, R. A. (1996). *Healthy life expectancy in England and Wales: recent evidence*. PSSRU Discussion Paper 1205.

Bebbington, A & Comas-Herrera, A. (2000). *Healthy Life Expectancy: Trends to 1998, and the Implications for Long Term Care Costs*, PSSRU Discussion Paper 1695 (London School of Economics).

Blake, David P. and Mayhew, Leslie D. (2006), "On the Sustainability of the UK State Pension System in the Light of Population Ageing and Declining Fertility" . *Economic Journal*, Vol. 116, No. 512, pp. F286-F305

Bone, M., Bebbington, A. C., Jagger, C., Morgan, K. & Nicolaas, G. (1995). *Health expectancy and its uses*. London: HMSO.

Börsch-Supan, A. & V. A. Hajivassiliou (1993), "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variables models", *Journal of Econometrics* 58, pp 347-68.

Bound, J. M. Schoenbaum, T. R. Stinebrickner and T. Waidmann (1999), "The dynamic effects of health on the labor force transitions of older workers", *Labour Economics*, 6:2, pp. 179-202.

Brown, S.L. (2000), "The Effect of Union Type on Psychological Wellbeing: Depression Cohabitants Versus Marrieds", *Journal of Health and Social Behaviour*, 41, 241-255.

Butt, Z., Haberman, S., Verrall, R. & Wass, V (2008), Calculating compensation for loss of future earnings: estimating and using work life expectancy, *J.R. Statist. Soc.* A 171, Part 4, pp 1-37

Cass (2007), *Giving up smoking adds seven years to good health, says new research*; mimeo, Cass Business School.

Cheung, Y.B. (2000), "Marital Status and Mortality in British Women: A Longitudinal Study", *International Journal of Epidemiology*, 29, 93-99.

Disney R, C. Emmerson and M. Wakefield (2006), “Ill health and retirement in Britain: A panel data-based analysis.” *Journal of Health Economics* 25: 621-649.

Domeij, D. and Johannesson, M. (2006), ”Consumption and health”, *Contributions to Macroeconomics*, 6(1): pp. 1314–1314.

French, E. (2005), “The effects of health, wealth, and wages on labour supply and retirement behaviour”, *Review of Economic Studies*, 72, 397–427.

Fries J (1980), "Aging, Natural Death and the Compression of Morbidity." *New England Journal of Medicine* 303:130-5.

Fuchs, V. R. (2004), "Reflections on the socio-economic correlates of health", *Journal of Health Economics*, 23:4, pp. 629-36.

Geweke, J. (1989) *Efficient Simulation from the Multivariate Normal Distribution Subject to Linear Inequality Constraints and the Evaluation of Constraint Probabilities*. Duke University, Durham, N.C: Mimeo.

Goldman, N., S. Korenman & R. Weinstein (1995), "Marital Status and Health Among the Elderly", *Social Science and Medicine*, 40, 1717-1730.

Government Actuary's Department (2006), *Period and cohort expectation of life tables. 2006-based projections*; dataset available at www.gad.gov.uk.

Gruenberg EM (1977). "The failures of success". *Milbank Memorial Foundation Quarterly/Health and Society*, vol.55, pp.3-24.

Heckman, J.J. (1981), "The incidentical parameters problem and the problem of initial conditions in estimating a discrete-time data stochastic process", Chapter 4, pp 179-95, MIT Press: Cambridge, MA in *Structural Analysis of discrete data with econometric applications*, edited by C.F. Manski and D. McFadden.

Heyma, A.. (2004), "A structural dynamic analysis of retirement behavior in the Netherlands", *Journal of Applied Econometrics*, Vol. 19, No. 6, pp. 739-759.

Karlsson, M., L. Mayhew and B. Rickayzen (2007), "Long term care financing in 4 OECD countries: fiscal burden and distributive effects". *Health Policy*, 80, pp. 107-34

Karlsson M., Mayhew, L., Plumb, R and Rickayzen, B. (2006), "Future costs for long-term care. Cost projections for long-term care for older people in the United Kingdom", *Health Policy* 75, 187–213

Lillard, L.A. and C.W.A. Panis (1996), "Marital Status and Mortality: The Role of Health", *Demography*, 33, 313-327.

Manton KG (1987). "Response to 'an introduction to the compression of morbidity"'. *Gerontologica Perspecta*, vol.1, pp.23-30.

Pickard, L. (2002), "The decline of intensive intergenerational care of older people in Great Britain, 1985-1995", *Population Trends*, Office for National Statistics.

United Nations, Population Division of the Department of Economic and Social Affairs (2006), *World Population Prospects: The 2006 Revision*.

Wilde, J. (2000), "Identification of multiple equation probit models with endogenous dummy regressors", Economics Letters 69(1), pp. 309-12.

Wilson, C. M. and Oswald, A. J. (2005) *How does marriage affect physical and psychological health? A survey of the longitudinal evidence*, IZA Discussion Paper No 1619

Wong, Linda (2003), "Structural Estimation of Marriage Models", *Journal of Labor Economics*, Vol 21:3, 699-727.

Wooldridge, J. (2000), *The initial conditions problem in dynamic nonlinear panel data models with unobserved heterogeneity.* Mimeo, University of Michigan.

A Simulating a population based on model estimates

The maximum likelihood procedure has provided us with parameter estimates for the econometric model. For simplicity, we partition these parameters into four groups: one is denoted β, one is denoted d the third one denoted θ, and the fourth one is denoted κ. The parameter vector θ contains all parameters related to the initial conditions problem, i.e. the parameters θ_0, θ_1 and θ_2 from equation (7). Hence, this parameter vector relates to all variables which remain constant over the projection period. The parameter vector d contains all parameters related to state dependence. The parameter vector β contains all parameters related to time-varying exogenous variables such as age and time, and the parameter vector κ contains the parameters of the covariance matrix of the error terms (Σ; we have denoted these parameters ρ, σ and ω in the paper).

The different sets of parameters are outlined in *Table 19* below.

Table 19: Parameters used in simulation.

Parameter Vector	**Contains**
β	Time-varying IV:s
d	State dependence
θ	Initial conditions
κ	Covariance matrix

Obviously, the first three sets of parameter vectors contain parameters for each of the four estimating equations - hence, we can define vectors β_a, d_a and θ_a for the parameters of the survival equation, and similarly for the other three equations.

A.1 Simulating a subpopulation

We determine a sample size N, in this case $10,000$. Assuming a maximum life length of $T = 100$ years from the start year (since we focus on 50-year olds, this is reasonable), we need to simulate a matrix of error terms and then make sure they have the appropriate correlations with each other (determined by Ω).

A.1.1 Simulating Error Terms

First, we simulate a matrix of standard normals:

$$v = F^{-1}(u)$$

where u is a $4T \times N$ matrix with each element $u_{ij} \backsim U(0,1)$ and $F^{-1}(\cdot)$ is the inverse of the cdf of a standard normal distribution. Obviously, v is also $4T \times N$ and the observations are iid with mean zero and variance 1.[3]

Next, we build the $4T \times 4T$ covariance matrix Σ. This matrix is defined as

$$\Sigma_{ij} = \sigma_{ab} + \rho_a^{|t-s|} \frac{\sqrt{1-\rho_a^2}\sqrt{1-\rho_b^2}}{1-\rho_a\rho_b}\omega_{ab}$$

[3] We do not report here how the first period error terms are derived. In order to obtain these, we use a simulation algorithm similar to that of the maximum likelihood procedure detailed in the main text. This way we take into account that the conditional distribution of v depends on the starting value of the dependent variables of the model.

where[4] $a = \left\lceil \frac{i}{4} \right\rceil$ and $b = \left\lceil \frac{j}{4} \right\rceil$ identify the corresponding estimating equation,

$$t = i - 4\,(a - 1)$$

and

$$s = j - 4\,(b - 1)$$

identify the corresponding years. Hence, the element Σ_{ij} tells us how the error term at time t in equation a is correlated with the error term at time s in equation b. The parameter σ_{ab} is the covariance of the individual effect in equation a with the corresponding individual effect in equation b. Likewise, ω_{ab} is the covariance of shocks in equation a with shocks in equation b. Whenever $a = b$, the corresponding variance is included.

Since Σ is positive definite and symmetric, we can carry out Cholesky decomposition. Hence, we define the $4T \times 4T$ matrix L as the lower diagonal Cholesky factor of the covariance matrix Σ:

$$L \cdot L' = \Sigma.$$

Then, premultiplying the matrix of simulated error terms v by the Cholesky factor,

$$e = (Lv)'$$

we get a $N \times 4T$ matrix of error terms, distributed according to $N\,(0, \Sigma)$. Next, we partition the matrix e so that we get one $N \times T$ matrix for each estimating equation. We denote this matrices e^a, e^w, e^c and e^h for convenience. They are derived from the oringinal matrix according to the equation

$$e^a = e \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \otimes I_T$$

and similar for the four other equations.

A.1.2 Determining the individual effect

Next, we need to determine the individual effect. The deterministic part (see equation (7)) is the same for all individuals in a subgroup. A subgroup is characterised by their values of the independent variables which remain constant over time as well as the initial state, represented by the variables W_0, C_0 and H_0. If we denote all other independent variables which remain constant (i.e. education, sex, ethnicity) Z, we can define the vector

$$G = \begin{pmatrix} W_0 & C_0 & H_0 & W_0C_0 & W_0H_0 & C_0H_0 & W_0C_0H_0 & Z \end{pmatrix}.$$

[4] $\lceil x \rceil$ denotes the value of x rounded upwards to the next integer.

Then, the deterministic part of the individual effect in equation j ($j \in \{a, w, c, h\}$) can be defined as

$$\alpha_j = G\theta_j$$

where we have suppressed the random part of α from equation (7) since it is included in the matrix e. Also, since all individuals in a certain subgroup have the same individual effects α_j (again, ignoring the random term), we suppressed the individual index i used in equation (7).

A.1.3 Simulating outcomes

Having defined the individual effect α_j, and constructed the matrix of error terms e, it is straightforward to simulate a population. This is done recursively, starting in year 1 and calculating the current state in all dimensions (A, W, C, H) for all simulated individuals, then moving on to the next year. Hence, we use the following procedure:

$$A_{it} = A_{i,t-1} \cdot 1\left[\alpha_a + X_t\beta_a + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} \end{bmatrix} d_a + e_{it}^a \geq 0\right]$$

where A_{it} takes on the value 1 if simulated individual i survives to period t; X_t represents exogenous variables changing over time (age, time trend) and $1[\cdot]$ is the indicator function, taking on the value 1 whenever the expression in the square brackets is true.

For the other dependent variables, we follow the same procedure; also taking into account that individuals need to be alive to be working, cohabiting and healthy. Hence,

$$W_{it} = A_{it} \cdot 1\left[\alpha_w + X_t\beta_w + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} \end{bmatrix} d_w + e_{it}^w \geq 0\right]$$

and then, for the remaining two, we also add simultaneously determined variables (such as W_{it}); hence:

$$C_{it} = A_{it} \cdot 1\left[\alpha_c + X_t\beta_c + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} & W_{it} \end{bmatrix} d_c + e_{it}^c \geq 0\right]$$

$$H_{it} = A_{it} \cdot 1\left[\alpha_h + X_t\beta_h + \begin{bmatrix} W_{i,t-1} & C_{i,t-1} & H_{i,t-1} & W_{it} & C_{it} \end{bmatrix} d_h + e_{it}^h \geq 0\right].$$

Hence, at the end of this exercise, we have four $N \times T$ matrices of simulated outcomes for the four variables A, W, C and H. This means that we can easily obtain life expectancy measures by simple matrix manipulations. For example:

$$LE = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} A_{it}}{N} + 0.5.$$

Similarly, healthy life expectancy can be calculated as:

$$HLE = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} H_{it}}{N} + 0.5H_0.$$

And the same goes for other combinations of the dependent variables, such as working healthy life expectancy etc.

A.2 Analysing the effects of changing status

Next, we want to analyse the effect of moving an individual from a certain starting state to another one, without changing other characteristics. Obviously, the difference in, say, life expectancy between two groups i and j (which we call 'gap' in the paper) is simply the difference between the two:

$$\Delta LE = LE_i - LE_j.$$

However, when we consider moving an individual from one state to another, we want to take into account the fact that they can be assumed to be different from individuals in the destination category - and this is arguably the reason why they were actually in a different category at the outset. This differences between individuals belonging to different groups are captured by the individual effect α_i in our model. Hence, when we analys the effect of moving an individual, we want to calculate counterfactual outcomes, based on a simulation where we keep α_i constant but change the starting position in accordance with the destination category.

Consider the survival equation above. In this new setting, the 'counterfactual' survival in period 1 would be

$$A'_{i1} = 1\left[\alpha_a + X_t\beta_a + \left[\begin{array}{ccc} W'_0 & C'_0 & H'_0 \end{array}\right] d_a + e^a_{i1} \geq 0\right]$$

where $\alpha_a = G\theta_a$ is determined according the *actual* starting position $\left[\begin{array}{ccc} W_0 & C_0 & H_0 \end{array}\right]$ of the group we are considering, whereas the state dependence vector $\left[\begin{array}{ccc} W'_0 & C'_0 & H'_0 \end{array}\right]$ is determined by the *counterfactual* starting position of the destination group. The same procedure is used for the other four dependent variables.

Now, if we denote the life expectancy calculated according to this counterfactual experiment by LE', we could calculate the expected gain from moving from one starting position to another one as $LE' - LE$. We have called this difference 'gain' in the paper. Obviously, this number can be larger or smaller than the difference ΔLE defined above.