14 July 2004

Ill health and retirement in Britain: A panel data-based analysis

Richard Disney*, Carl Emmerson† and Matthew Wakefield†

Abstract

We examine the effect of ill health on retirement decisions in Britain, using the first eight waves of the British Household Panel Survey (1991-98). As self-reported health status is likely to be endogenous to the retirement decision, we instrument self-reported health by a constructed 'health stock' measure using a set of health indicator variables and personal characteristics, as suggested by Bound *et al* (1999). Using a range of econometric techniques, we show that adverse individual shocks to health stocks are a significant predictor of individual retirement behaviour among workers aged 50 and over. We compare the response of economic activity to our constructed health measure to that arising when direct indicators of functional limitations and specific health problems are used instead. We also test whether adverse and positive health shocks have symmetric effects on transitions in and out of economic activity among this age group.

Key Words Ill health Retirement Response bias
JEL classification H55 I12 J26

Acknowledgements

We are grateful to the Nuffield Foundation for funding the project: 'A study of health and the labour market behaviour of older workers'. Co-funding was provided by the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP). We thank two referees of this journal, Costas Meghir and Frank Windmeijer for valuable suggestions, and also participants at seminars at the Institute for Fiscal Studies, at the NBER Summer Institute 2003, at workshops on ill health and labour supply at the University of Tilburg and the RTN Network on Ageing at Naples, and at the Universities of Essex, Newcastle, St Andrews and Strathclyde for many useful comments. The British Household Panel Survey data used in this paper were made available through the UK Data Archive, funded by the Economic and Social Research Council. The usual disclaimers apply.

* Institute for Fiscal Studies, London and University of Nottingham. *Corresponding author*: School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD
Tel: +44 115 951 5619 Fax: +44 115 951 4159 email: richard.disney@nottingham.ac.uk

† Institute for Fiscal Studies, London

Ill health and retirement in Britain: A panel data-based analysis

1. Introduction

The number of people on disability benefits more than doubled between the late 1970s and the end of the 1990s in the UK. Throughout this period around half of claimants were aged between 50 and the state pension age. The substantial growth in claimants aged 50 and over is shown in figure 1. 'Ill health' is a major reason for retirement among British men, especially for men without access to an occupational pension (Tanner, 1998, Table 7). Indeed 'own ill health' is the most commonly cited reason for retirement among both men and women in the early 1990s (Disney, Grundy and Johnson, 1997, Table 2.19). Similar trends have been observed in other countries such as the Netherlands and the United States (Bound and Burkhauser, 1999).

Figure 1 here

At any point in time there is a strong correlation between observing a person not working and their self-reported overall poor health status, but this may give a misleading impression of the impact of health state on *retirement*. First, individuals who are inactive often have an incentive, for self-esteem if nothing else, to report worse-than-actual health. Second, differences in reported self-assessed health are large, even for individuals in identical labour market states – individual heterogeneity is important. Third, individuals with permanent and very poor health may *never* have worked, so they cannot be observed 'retiring'. Fourth, ill-health may impact on other labour market attributes of the worker (for example, the wage they earn – see Meghir and Whitehouse, 1997) which implies that there are both income and substitution effects on labour supply arising from shocks to the worker's health status. Finally, the health stock may be endogenous to the labour market state of the individual (Kerkhofs, Lindeboom and Theeuwes, 1999).

The potential measurement error and endogeneity of self-reported health status has led some economists to reject the use of such general measures in retirement models completely (such as Myers, 1982) even though they have been, and continue to be, commonly used in this field for want of better measures.[1] A further problem, in the UK at least, is that for those individuals with no private pension rights, disability benefits are the only 'route' into early retirement through the social security programme since the social security pension cannot be received before the state pension age (currently 65 for men and 60 for women). Consequently, there is an inducement for early retirees to utilise the ill-health route and their self-assessed health status will correlate with preferences for early retirement (Blundell and Johnson, 1998).[2]

For the researcher interested in the link between ill health and retirement, one obvious strategy is to substitute more objective measures of ill health (if available) for self-reported health status in the model 'explaining' retirement.[3] Some studies have argued for the intrinsic superiority of this approach, since it eliminates the errors-in-variables and biases arising from the subjective health measure (many such studies are cited in Quinn, Burkhauser and Myers, 1990). But as Bound (1991) points out, we cannot be sure that such proxies are any better predictors of (in)activity than self-reported health status, as the researcher thereby assumes some link between work status and these other health measures.[4] Such a strategy does not eliminate the errors-in-variables problem but replaces it with a similar problem on a proxy variable, and may thereby lose any additional information on the 'true' association between health and behaviour that might be intrinsic to the self-reported 'subjective' measure.[5]

Another pertinent suggestion, explored by Anderson, Burkhauser and Quinn (1986) and Bound, Schoenbaum, Stinebrickner and Waidmann (1999) (hereafter Bound *et al*) is that *changes* in labour market status e.g. 'retirement' (whether permanent or temporary) should be associated with 'shocks' to the individual's underlying 'health stock'. Bound *et al*'s strategy is to construct a latent health stock or index of health for each individual as a function of personal characteristics and health indicators. This constructed variable is used to instrument self-reported health in a panel data model of economic activity in order to explore the relationship between time variations in health and changes in work status (see also Stern, 1989).[6] Modelling health 'shocks', it can be argued, eliminates any person-specific association between characteristics and labour market outcomes (such as fixed preferences for work, or longstanding disability), whilst proxying self-reported health status by time-varying health and personal characteristics should ameliorate any reporting bias in the former.

This paper follows the general strategy suggested by these authors. It exploits the panel element of the data set to construct individual 'health stocks', and uses time variation in these 'stocks' as an explanatory variable in reduced form models of labour market (in)activity amongst a sample of older people in Britain.

Two econometric approaches are used. In the first, linear and non-linear fixed effects estimators are used to examine the impacts of individual health 'shocks' on the economic activity rates of individuals aged between 50 and state pension age. The subjective health question invites the respondent to compare their health to that of people of a similar age, so this approach pins down the association between transitions in and out of paid work and current variations in relative health and allows for respondent heterogeneity. The second approach uses the same data and constructed variables but estimates a hazard function with non-parametric duration dependence. The hazard incorporates lagged and current health, and parameterised individual heterogeneity. Relative to the fixed effects model, such a specification has a more intuitive interpretation as a retirement model and permits greater flexibility both in examining the dynamic impact of health status on retirement and in examining whether the impact of health on movements in and out of the labour market differs according to current work status. These advantages come at the cost of imposing restrictions on the structure of individual heterogeneity, which could be important in this context (see section 3.2).

Whichever approach is used, we find robust evidence that individual health deteriorations lead to a greater likelihood of transitions into economic inactivity later in the working life. Two questions arise from this finding. First, does this two-step procedure do better than, for example, simply augmenting the reduced form retirement equation with individual indicators of 'objective' health limitations and difficulties? We therefore also show that the predictive power of (linear) combinations of objective health limitations in explaining ill health retirement is limited relative to our chosen method of constructing an individual time-varying health stock.

Second, are relative individual health *improvements* associated with a greater likelihood of reverse transitions out of economic inactivity? The fixed effect models do not differentiate between types of state transitions, although symmetry of behavioural responses to positive and negative 'shocks' to health can be tested. In contrast, in the baseline hazard specifications, we utilise the last reported exit from economic activity (if observed) during the period as the indicator of 'retirement'. To examine symmetry of activity responses, we then utilise *all* transitions between economic activity and inactivity as in the fixed effect model. We show some evidence of symmetric response: that is, the probabilities of individual transitions from economic activity to inactivity are strongly associated with deteriorations in relative health status and that the reverse transition is observed with relative improvements in health status. We do not, in this paper, explore the 'feedback' of labour market activity or inactivity on the evolving health stock.

In summary, our range of methods and tests suggests that this method of modelling ill health and economic (in)activity is reasonably robust. We also attempted to test whether the UK's 1995 disability benefit reform, which tightened some eligibility conditions and cut real disability benefits, had any impact on the link between work-related disability and economic inactivity. We find no evidence of any impact in the data. This may of course be the 'true' answer[7], but it may simply be that our 'test' is inappropriate in part because we only observe receipt and not applications for disability benefits. These results are not discussed further here (see Disney, Emmerson and Wakefield, 2003 for further discussion of this issue).

The structure of the paper is as follows. Section 2 describes the construction of our health stock variable. Section 3 uses this variable as an instrument for health status in reduced form labour market models. Section 4 describes our sensitivity analysis to asymmetries in the health-economic activity relationship and to alternative health measures. Section 5 concludes.

2. Modelling the individual's health stock

2.1 Data and modelling issues

To construct an individual's underlying 'health stock', we follow Bound *et al* (1999) in assuming that the i[th] individual's health at time t is determined by a linear combination of exogenous personal characteristics (such as age and education) x_{it}, a vector of detailed personal health indicators (such as functional limitations) z_{it} and unobservables υ_{it} uncorrelated with x_{it} and z_{it}. We allow the impact of these characteristics to vary over time. Denote this (unobserved) health state as η_{it}. So:

$$\eta_{it} = x_{it}'\beta_t + z_{it}'\gamma_t + \upsilon_{it} \quad (1)$$

Although this health state is not observed, a self-reported health status is observed in our data, as a categorical variable with five 'states': ranging from 'excellent' to 'very poor' (the exact form of the question is discussed below). Denote this categorical variable as h_{it}. The latent counterpart to h_{it}, which is denoted by h^*_{it}, is a simple function of η_{it} and a term reflecting reporting error:

$$h^*_{it} = \eta_{it} + \varepsilon_{it} \tag{2}$$

Crucially, we assume that ε_{it} is uncorrelated with υ_{it}. Reporting error may well be correlated with the state in which the individual is located. By using this instrumental variable-type procedure, we are assuming that the errors are uncorrelated with those arising when reporting specific health limitations. Thus we write:

$$\begin{aligned} h^*_{it} &= x'_{it}\beta_t + z'_{it}\gamma_t + [\upsilon_{it} + \varepsilon_{it}] \\ h^*_{it} &= x'_{it}\beta_t + z'_{it}\gamma_t + u_{it} \end{aligned} \tag{3}$$

Assuming that u_{it} is normally distributed, equation (3) can be estimated as an ordered probit.

To reiterate, this time varying individual 'health stock' is constructed to strip the health term in the labour force participation equation of possible subjectivity and endogeneity in individual response to general health-related questions. Using self-reported health status, h_{it}, as a proxy for η_{it} directly will be biased if the reporting error term in equation (2) is correlated with terms in the labour force participation equation that we estimate in the next section. This assumption cannot be tested directly but we assume that it is likely to be the case. In addition, simply entering the z_{it} vector in equation (1) directly into a labour force participation equation will likely induce errors-in-variables biases, because more specific health factors, even if accurately reported, may not perfectly predict current capacity to work.[8] The explanatory power of entering these specific health measures directly is tested directly, at least for fairly simple (linear) specifications. Using the latent variable model in equation (3), Bound *et al* (1999) argue, is a standard measure of dealing with these problems, by using a proxy with error to instrument an endogenous and error-ridden variable such as h^* (see also Griliches, 1974).

The data used in the analysis are drawn from the first eight waves of the British Household Panel Survey (BHPS), 1991-98.[9] This survey is a representative sample of the population of England, Wales and Scotland (south of the Caledonian Canal). Since we are interested in retirement behaviour we use a subsample of people aged 50 to 64 in 1991 who are removed from our panel when they reach the state pension age (65 for men and 60 for women).[10] We use transitions between observed economic activity and inactivity as our measurement of 'retirement' with, for the present, the last observed transition into inactivity (if any) defined as retirement.[11] This selection by age, coupled with the requirement that we observe certain variables (particularly work status and health) leaves us with a sample of 1,440 individuals in 1991, reduced (largely by the upper age cut-off) to 478 by 1998.

One advantage of using this data is that it is a panel that allows us to track individuals over a relatively long period of time: our eight year panel is significantly longer than that available to Bound *et al* (1999), for example. The BHPS also records a rich set of characteristics for individuals in the sample. In what follows we use data on educational achievement, family composition, region of residence and a derived variable on housing wealth,[12] as the components of the x_{it} vector, and the many measures of individual health that are contained in the survey as the components of the z_{it} vector.

These measures of individual health status in the z_{it} vector in equation (1) come in two coded sets of indicators: the first reporting whether or not individuals say that they have certain health problems and difficulties and the second recording whether or not individuals feel that their health limits their ability to perform certain daily activities. The exact questions are described in Appendix 2. The health problems and difficulties that individuals are asked about are: problems with arms, legs, hands, feet, back or neck; difficulty seeing; difficulty hearing; skin conditions and allergies; chest or breathing problems including asthma and bronchitis; heart problems and blood pressure or circulation problems; stomach, liver, kidney or digestive problems; diabetes; anxiety, depression or bad nerves; alcohol or drug related problems; epilepsy; migraine or frequent headaches; other health problems. After being asked "does your health limit your daily activities compared to most people of your age?" the specific activities that BHPS respondents are asked about are: doing the housework; climbing stairs; dressing oneself; walking for at least ten minutes.[13]

These specific indicators of health are used to cleanse the more subjective general assessment of health of response patterns that, we argued, might be state dependent and subject to other biases. This self-reported general health measure h_{it} is a response to the question: "Please think back over the last 12 months about how your health has been. *Compared to other people of your own age*, would you say that your health on the whole has been: excellent; good; fair; poor; very poor; don't know" (our italics).[14]

It is noteworthy that this self-reported health status derives from a question that specifically asks respondents to compare their own health to that of other people of their own age. One interpretation of responses to this question is that they should *not* pick up the likely average decline in health status as the panel of respondents ages: the cumulative effect on retirement decisions of the general deterioration of the health of the cohort with age should be identified elsewhere in the model, primarily by the age terms. Notwithstanding the nature of the health question, however, there does seem to be a general decline in self-assessed health relative to the cohort. This is seen from year-on-year comparisons of the data. The overall effect can be seen by comparing the first and last years of data, presented in Table 1, where we differentiate between the whole sample in each year, and those who were present in both years:[15]

Table 1 here

This, of course, provides another reason for being cautious about simply using the self-reported overall health measure in an analysis of economic (in)activity. The average decline in self-reported health relative to the cohort may arise from a change in self-perception, or a change in the comparison group implicitly used by the respondent (who may be, for example, assessing only those people of similar ages who are still economically active – although our sample includes both the active and the inactive). By using our IV-type approach in conjunction with statistical procedures that allow for sample heterogeneity and individual state dependence, we would argue that this potential problem of interpretation is considerably reduced.

2.2 Estimation

We now estimate the model for the latent 'health stock'. Using as the dependent variable the categorical variable described above, Table 2 depicts the ordered probit underlying equation (3) for 1991.

Table 2 here

Looking at the sample characteristics, the first column of the table reveals that there are more men than women in the sample. This is because we impose an upper age limit of the state pension age, which is lower for women. Over four fifths of respondents are in a couple rather than single. Most of the sample are owner occupiers, 37% own outright and mean housing equity across the whole sample is just over £54,000. Almost half of the sample has no educational qualifications, although there is wide variation in achieved qualifications. Significant proportions of the sample report having difficulties or health problems, notably with arms/legs/hands, lung or heart problems.

Examining the parameter estimates, few personal characteristics, other than those related to health, are significant in explaining self-assessed health. Individuals in a couple and with higher housing equity are likely to report 'better' categorical health status. The education variables are jointly significant with more highly educated people in general reporting better health. The regional dummies are jointly insignificant in explaining self-reported health status, although previous studies of the incidence of disability benefit receipt (such as Disney and Webb, 1991) show strong regional disparities, suggesting that regional labour market differences might be important in explaining economic inactivity. However there is regional variation in health, which it turns out is being picked up by the inclusion of specific health variables as the regional dummies are jointly statistically significant if the health variables are excluded from the set of regressors.

In contrast to personal characteristics, and as expected, many of the health measures are individually significant at the one or five percent levels, and tests of joint significance show that the measures of functional limitations are very significant when considered together, as are the variables recording health conditions and problems. Among these variables only 'climbing stairs' and 'hearing difficulties' have an insignificant impact on self-reported health status.

It is would be desirable to provide some interpretation of the coefficients as 'marginal effects'. But it is well known that it is difficult to interpret the coefficients in an ordered discrete choice model like this ordered probit in this way.[16] A positive coefficient unambiguously means that an increase in the variable concerned will decrease the probability with which an individual is predicted to be in the lowest health category (very poor) and increase the probability with which they are predicted to be in excellent health, and *vice versa*. So the negative sign on coefficients on all of the health variables (with the one exception of 'hearing difficulties') are as we would expect. To get a feel for parameter estimates: if we take a representative individual who has the average (mean) values for characteristics measured by continuous variables and is assigned values of the dummy variable characteristics that are the most common in the data, then the predicted probability of being in excellent health if the person has no chest or breathing problems is 0.35. If they have chest problems it is only 0.10.

The predicted values of the index for each individual from this equation constitute the basis of the individual 'health stocks' in 1991. We normalise these to give the 'health stock' as an absolute deviation of the individual's health index from the cohort mean for each year. Normalising in this way avoids the need to make any further assumptions in order to identify the constant in each separate ordered probit. The predicted individual 'health stocks' are re-assessed each year in the light of new information on self-reported health and the z_{it} and x_{it} vectors contained in the data. We update by running a new ordered probit for each year on same set of independent variables that are used in the first year. We predict the health index and normalise to give the health stock by subtracting the year specific mean, in the same way as described for the 1991 data. In this fashion, we construct an evolving health stock for each individual (relative to the year-on-year average for the sample) in the data set over the period 1991-98.[17] These are the latent health stock measures that will be introduced into the reduced form equations describing the evolution of each individual's economic (in)activity. The descriptive statistics for these constructed health stocks, which are useful for interpreting coefficients in the models that follow, appear in Appendix 1.

3. The impact of health on retirement

3.1 Theory

The economic theory underlying the relationship between ill health and retirement is standard (for example, Lazear, 1986). Agents have preferences over current and future leisure with the values of current and future leisure depending, *inter alia*, on current and expected states of health. Agents form expectations over future states of health. They maximise utility subject to the lifetime budget constraint, and standard dynamic optimisation conditions determine choice of retirement date, R*. Shocks to factors such as income, preferences, or expectations could affect R*. There are various models of how people solve this optimisation problem.[18] There may be some institutional constraints that limit choice of R* but, in the United Kingdom, these are relatively few – participants can annuitise a private pension from age 50 and can both receive pensions and continue to work for another employer. Individuals can also take their state pension and work beyond the state pension age without a retirement test, at least since the abolition of the Earnings Rule in 1989 (Disney and Smith, 2002).

Poorer health status, *ceteris paribus*, will reduce the probability of continued work for several reasons. First, poorer health may raise the current disutility of work. Second, poorer health reduces the return from work if there is a relationship between poor health and low wages. Third, poor health may entitle the individual to non-wage income, such as disability benefits, which is contingent on not being in work. The only counteracting principle is if poor health raises consumption requirements (for example, the cost of health treatment) and requires greater income than can be provided through the disability insurance programme. On the other hand, if poorer health is associated with lower life expectancy, the annualised consumption available from existing wealth is raised which might induce earlier retirement.

As mentioned in the introduction we apply two different econometric techniques to estimate a reduced form of such a retirement model. The first approach exploits variation over time in each individual's health stock by using a fixed effects estimator to look at the impact of *changes* in health on *changes* in labour market status. This has the advantage of controlling for any form of relationship between work status and fixed, unobserved individual characteristics. The second approach also exploits variation in each individual's health by estimating a hazard function that incorporates both lagged and current health. The relative advantage of this approach is its greater flexibility in modelling dynamics. In addition it allows us to examine whether the impact of health on movements in and out of the labour market differs according to current work status.

3.2 Estimation I: fixed effect models

We now examine the role of the constructed health stock in reduced form models of retirement. Here, we estimate a simplified version of the underlying model as described in equation (4) of the previous section, using a fixed effect approach to allow for individual heterogeneity in, for example, preferences for work. We focus on self-reported economic activity and inactivity rather than self-reported 'retirements' (see note 11). Note that in this framework there may be several transitions between economic activity and inactivity by the same individual. Since only the outcome is observed, a discrete choice model that exploits the 'panel' structure of our data and takes account of individual (fixed) effects is appropriate. The fixed effects discrete choice model is:

$$\begin{aligned} lf^*_{it} &= \alpha_i + \eta_{it}\lambda + x'_{it}\beta + \varepsilon_{it}, \quad i = 1....n, \ t = 1...T \\ lf_{it} &= 1 \ \text{ if } lf^*_{it} > 0, \text{ and 0 otherwise} \end{aligned} \tag{4}$$

where lf^*_{it} is the latent variable that indexes the probability of participation of individual i at time t, here defined as whether the individual reports that they are currently working (including self-employment), η_{it} is the unobserved health state, x'_{it} is a vector of other characteristics and α_i is the individual fixed effect. The fixed effects (conditional) logit model is written in general as:

$$\Pr(lf_{it} = 1 | (x_{it}, \eta_{it})) = \frac{e^{a_i + \eta_{it}\lambda + x'_{it}\beta}}{1 + e^{a_i + \eta_{it}\lambda + x'_{it}\beta}} \tag{5}$$

Chamberlain (1980) shows that conditioning the model on the number of individual transitions results in the *conditional likelihood function* being free of the incidental ('fixed effect') parameters, α_i. This also means that a contribution to the likelihood only arises from those groups of observations (of a given individual over time, say) that are not always zero or one – in this case, those who transit 'states' between economic activity and inactivity.

The vector of explanatory variables for economic activity status comprises $\hat{h}_{it}$, which is the predicted value of the underlying health 'stock' η_{it} for the individual, relative to the mean for each year, obtained from estimating equation (3) as described previously, and a vector of time-varying individual characteristics. Because we use the constructed health stock variable, standard errors are bootstrapped. Variable definitions are as follows:

> Dependent variable: 1 if self-reported economically active (employed or self-employed), 0 otherwise.
>
> Age: a quadratic in age (÷10) ≥ 50 and <60 (women) <65 (men)
>
> Couple: if respondent is in a couple = 1 (default = single, widowed, divorced)
>
> No. of children = number resident in household
>
> Regional unemployment rate: the regional unemployment rate at *t*
>
> Housing equity = Value of housing equity in £000. (This may be >=<0)
>
> Health stock = Deviations of individual health stock measure from mean at *t* as defined in Section 2

Table 3 illustrates the results of this exercise for the fixed effect (conditional) logit, and also for a standard logit pooled over all observations (with additional time invariant regressors) and for the standard linear fixed effect model. Since other studies of retirement in Britain suggest that individual fixed effects are important (e.g. Meghir and Whitehouse, 1997 and Blundell, Meghir and Smith, 2002) and the linear model is an inappropriate choice when the variable of interest is dichotomous, we prefer the specification in column 1. However all three are presented because the standard logit is common in the literature on labour market transitions and because coefficients from linear fixed effect models are easier to interpret.

Table 3 here

Irrespective of estimation method, the quadratic in age has a significant impact on transitions out of economic activity (higher order polynomials are rejected by likelihood ratio tests). Individuals in couples are more likely to work (the logit result), and becoming married tends to be associated with an increase in economic activity in the fixed effect specifications, although this latter effect is not statistically significant.

Turning to the other regressors, a higher unemployment rate is associated with a lower probability of economic activity but the level effect is not significant. However the impact of a rising unemployment rate (the fixed effect interpretation) is significantly negative. Moreover, higher (and rising) housing equity, proxying household wealth, also reduces economic activity.

Of most interest in this context is the coefficient on the individual's relative health status, which is strongly positively associated with economic activity (a higher value indicating better health). This is an encouraging result, bearing in mind that we are using a constructed variable proxying an assumed underlying health stock. Moreover, the fixed effect specifications confirm that, for individuals who transit between labour market states, there is a link between changing health stocks and changes in labour market state (notably, for this age group, retirement) rather than simply an underlying association between poor health and inactivity.[19]

3.3 Estimation II: hazard functions

Fixed effect models capture the heterogeneity of response in individual panels but impose strong restrictions on the dynamic structure of the model. To check that the relationship between health and transitions into economic inactivity is robust, we also estimate a hazard rate model. Specifically, we estimate a discrete time proportional hazard model with a gamma mixture distribution to incorporate unobserved individual heterogeneity (see Prentice and Gloeckler, 1978; Meyer, 1990; Jenkins, 1995 and, for estimation in STATA, Jenkins, 1997). We incorporate additional duration dummies to capture duration dependence non-parametrically.

We initially model last observed exits from economic activity (if observed) which we define as 'retirement' since in many cases inactivity will be an absorbing state. Transitions are observed between annual intervals in our data and we do not know in all cases the actual date of exit. Denote these annual intervals $[0 = t_0, t_1)$, $[t_1, t_2)$,… $[t_{k-1}, t_k)$. The probability of exit in the j-th interval for person i is:

$$\begin{aligned} &prob\{T \in [t_{j-1}, t_j)\} = S(t_{j-1}; \eta_{it}, x_{it}) - S(t_j; \eta_{it}, x_{it}) \\ &\text{and} \\ &prob\{T \geq t_{j-1}\} = S(t_{j-1}; \eta_{it}, x_{it}) \end{aligned} \tag{6}$$

where S is the survivor function and other variables are defined as before. Given the proportional hazards assumption, the survivor function in the discrete case is written as:

$$S(t_j; \eta_{it}, x_{it}) = \exp[-\exp(\eta_{it} + x_{it}'\beta + \delta_j)] \ \text{ where } \ \delta_j = \log(H_{it}) \text{ for } j = 1,...k. \tag{7}$$

and where H_t is the integrated baseline hazard at t. The discrete time hazard, h_j, in the j-the interval is:

$$h_j(\eta_{ij}, x_{ij}) = 1 - \exp[-\exp(\eta_{ij} + x_{ij} + \gamma_j)] \text{ with } \gamma_j = \log \int_{t_{j-1}}^{t_j} \lambda_0(\tau) d\tau \tag{8}$$

where γ_j is the baseline hazard in the interval j-1 to j and λ is the instantaneous hazard *rate*.

The specification permits us to include duration dependence non-parametrically using duration dummies. We also incorporate a Gamma distributed random variable ε_i with unit mean and variance $\sigma^2 \equiv \nu$ to describe unobserved individual heterogeneity.[20] We can rewrite (8) including duration dummies, D, and unobserved heterogeneity as:

$$h_j(\eta_{ij}, x_{ij}, D_j) = 1 - \exp\{-\exp[\eta_{ij} + x_{ij} + D_j + \gamma_j + \log(\varepsilon_i)]\} \tag{9}$$

Table 4 provides estimates of the baseline hazard model in (9) using, as the dependent variable, the last observed exit (if any) into economic inactivity. It provides both coefficients and the hazard ratios. As before, the sample comprises individuals aged

50 up to state pension age in 1991, with individual observations removed if the individual reaches state pension age. Also as before, the standard errors are bootstrapped given the inclusion of the estimated health stock variable.

Table 4 here

Although the model structure is rather different from that of the fixed effects models in the previous sub-section, similar results emerge, although we can now separate out age and duration effects. Since duration can only be observed within the sample period for those entering the sample period in employment, we add an additional variable that captures the elapsed spell time in the current state at the start of the period. The results suggest that age can be captured by a linear term, and that the exit rate is increasing with age. Elapsed time in work at the start of the period is not statistically significant. There is a pattern of rising exit probabilities with duration which is however non-monotonic and which may be capturing the non-linearity in the age-exit relationship observed in the fixed effect specifications.

In terms of household characteristics, educational qualifications have no significant impact on the exit rate from economic activity. There are some regional effects, with the North West, Wales and Scotland characterised by higher rates of exit before state pension age. The number of children in the household, as well as gender and living in a couple have only weak effects on the exit rate. Also of weak significance is the regional unemployment rate. It has the opposite sign to that of the fixed effect models but note that the unemployment rate in the hazard measures a 'levels' effect, not a 'difference' effect (the appropriate comparison is with the pooled logit, where the unemployment rate coefficient is insignificant). In contrast to the unemployment term, the value of housing equity has the same effect as in the specifications in Table 3, with higher household wealth increasing the probability of final exit from the workforce.

The health stock measure is now introduced both as a current and lagged level. Both terms are highly significant, and have the 'correct' sign – a better health stock relative to the sample reduces the probability of exit, as does the one period lagged health stock. To help the interpretation of coefficients in this model, in the final column of the table we report 'hazard ratios' which (approximately) measure the proportional effect on the hazard of a one-unit change in the variable in question. A unit increase in either the health stock or its lagged value would (ceteris paribus) decrease the probability of exit from work by approximately 45%. The statistics in *Appendix 1* indicate that while a one unit change in an individual's health stock between any two years is relatively unlikely, around 41% of the people in our sample do in fact experience a one unit change in health stock over the entire period for which they are observed in our data. To see how important changes in health stocks are in the retirement decision, a natural comparison is with the impact of wealth on retirement. The coefficients and associated hazard ratios shown in Table 4 indicate that a similar 45% change (albeit an increase) in the likelihood of exiting work would require a change in housing equity of around £90,000 (from a mean in 1991 of just under £55,000 – see Table 2).[21] These results reinforce those of the previous subsection by again showing the strong link that exists between health status and the timing of retirement and illustrating the robustness of our procedure for modelling health and the transition into economic inactivity.

4. Sensitivity Analysis

4.1 Symmetry of labour market transitions to changes in individual health stocks

This section focuses on the sensitivity of the results to alternative specifications. In particular, this sub-section examines whether the participation response of individuals to health shocks is symmetric – that is, whether improvements in health are associated with transitions into work in the same way that deteriorations are linked to exits from work. The alternative is that there may be a 'ratchet' effect such that after an individual's poor health has caused exit from work, it requires health to recover beyond the threshold that induced exit in order to encourage new efforts to seek work. There are also issues concerning the measure of 'entries' and 'exits' from economic activity and the interpretation of changes in the health stock.

As mentioned in Section 3.2, the hazard models focus on the final observed exit from economic activity – which we termed 'retirement'. However to examine symmetry, we must model *all* exits and entries from economic activity among this age group and test whether the response to deteriorations and improvements in *relative* health are symmetric. It is also interesting is to see whether similar symmetries or asymmetries appear for other conditioning variables in the data set. Separate hazard functions for all exits and entries are depicted in Table 5.

Table 5 here

The results are as follows. First, the hazard for all exits (Table 5) exhibits similar coefficients to that for 'final' exits (Table 4) – the main difference perhaps being the lower coefficient on the lagged health stock. Second, comparing all exits and all entries in Table 5, the hazard model for entries to economic activity rejects the specification of parametric unobserved heterogeneity and so is estimated without heterogeneity, whereas unobserved heterogeneity is relevant to the model of exits from economic activity. In contrast (not shown in Table 5), the duration terms are all highly significant in the hazard for entrants as opposed to the mixed picture for exits (see Table 4). This last re-entry result suggests strong duration dependence in inactivity, as might be expected (a finding that cannot be obtained using the fixed effect models of the previous section). In addition, unlike in the exit model, the unemployment rate term is significant (at the 5% level) for re-entrants and has the expected sign.

Other conditioning variables have a more obvious symmetric impact. The impact of age on exits and entries is almost symmetric in coefficient as well as sign and so, too, is the impact of housing equity, proxying household wealth. Ill health and lagged ill health reduce the probability of economic activity, as in Table 4. For re-entrants, the health terms have the reverse sign and are jointly significant (the likelihood ratio test statistic of 24.86 is well above the critical value for even a 1% significance test) despite the insignificance of lagged health on its own. However while the signs are reversed the impact of improvements in relative health on the probability of re-entry is dampened compared to the effect of health on labour market exits.

The test just described exploits the separation of labour market exits and entries in the hazard specification in order to assess whether responses to changes in our 'health stock' are symmetric across these two kinds of behaviour. Given that an individual's 'health stock' measures their health relative to that of other people of a similar age, it is perhaps misleading to utilise this to evaluate the direction of change in that individual's health status: a measured improvement may simply reflect the fact that the health of the individual has deteriorated by less than the average change across the whole sample. It might therefore be argued that we are not strictly testing symmetry in terms of absolute self-reported health. To examine this further, we also utilised another variable in the data set derived from the response to the question: 'Does your health limit the type of work or amount of work you can do?' [YES/NO]. Since we are interested in how time-varying responses to this variable affect economic activity, we revert in this case to the fixed effects model. This test of symmetry also tests the plausibility of the fixed effects set up which includes exits and entries within a single estimation and therefore implicitly treats them as similar behaviours.

Responses to this variable are of course likely to be strongly endogenous to labour market status, and it is of some interest to see whether our instrumental variable-type technique gives similar results on this variable. We interpret responding negatively to this question as reporting being in 'good health'. We then use the same set of independent variables and a similar modelling structure to that described in Section 2 and Table 2, in order to derive individual year-specific predictions of the probability of responding negatively to this question (results are available on request from the authors). Unlike in the ordered probit used previously, there is just one probability of this kind that can be predicted from what is now a standard probit for a dichotomous variable. Furthermore, in this case an increase in this probability can be interpreted as meaning that health has improved and become less of a constraint on work. The predicted probabilities are inserted in the labour market (in)activity equation in the same way as was our more detailed 'health stock' measure.

Results analogous to those reported in Table 3 but using this new predicted variable **No health limit on work**, are contained in Table 6. Results are similar, but there is one additional variable. Since the probability of health affecting work is now measured as an 'absolute' probability we are able to create a dummy variable indicating that health has improved relative to the previous year.[22] The term **Symmetric health impact** is this dummy multiplied by the predicted health measure. By including this additional regressor, we have a test of a positive 'ratchet effect' of ill health on the retirement probability. A significant negative coefficient on the interacted health term would suggest that an improvement in health has a weaker impact on the probability of transiting from inactivity to activity than the reverse.

Table 6 here

Table 6 provides no robust evidence of asymmetry in the fixed effect logit – that is, the interaction term, although negative as predicted, is insignificant. However the term is significant as well as negative in the logit and linear fixed effect models. The linear fixed effects model gives a measure of this asymmetry that is easier to interpret – an increase of 10% in the likelihood that health adversely affects work increases the probability of exiting economic activity by around 1.6%; whereas the same change in the other direction raises the probability of entering economic activity by just over 1.4%. There is an asymmetry in the impact of health, but it does not have a substantial effect. Taking the results of Table 5 and 6 in conjunction, therefore, both health deteriorations and health improvements affect transitions between economic activity and inactivity among this age group, with some evidence that deteriorations have a slightly larger impact on transition probabilities.

4.2 Alternative measures of individual ill health

In this sub-section, we investigate further the sensitivity of the results to the measure of the health stock. As argued in Section 1, economic activity should be strongly correlated with subjective, work-related, measures of health status in part due to reporting bias. On the other hand, specific indicators of disability may have a weaker correlation with economic activity simply because some of these disabilities may have relatively little impact on capacity to work.

Table 7 examines alternative specifications of the health variable in the hazard rate specifications. Column 1 gives the coefficient and standard error from the generated health stock measure used in Table 4, as a benchmark.

Table 7 here

Column 2 illustrates a specification where, instead of using the constructed health stock two count variables, which count up the number of difficulties in basic physical activities ($j=1$ *to* 4) and the number of health problems ($n=1$ *to* 13), as reported by the individual, both lagged and current, are entered directly into the hazard rate. A higher count in each case should be associated with a greater number of functional limitations or incapacities and therefore with a greater probability of retirement, so these can be regarded as crude self-reported 'disability indices'. Because the counts are directly inserted in the specification, we do not need to bootstrap the standard errors.

It will be noted immediately that higher 'counts' and lagged counts of these variables are positively and significantly associated with the retirement hazard. At first sight, this might suggest that if a simple count of self-assessed 'objective' health difficulties and limitations is significant, there is no need for the two-step procedure as described in Section 2. But simply counting up health factors is equivalent to imposing a common restriction that each health indicator has an identical impact on the individual's overall health 'stock' and therefore on his or her capacity to be economically active. This seems implausible and less informative than the derivation of the individual's health stock by the procedure depicted in Table 2, which implicitly weights each health problem using the additional information contained in the respondent's assessment of their overall health.

If we examine the impact of individual health problems and limitations (Column 3), the problem with the use of self-reported indicators becomes more apparent. Few individual indicators have a statistically significant impact on the probability of exit from the labour force, and this remains the case even if we drop all the lagged terms (not shown). Furthermore, a likelihood ratio test fails to reject the null that these sets of variables do in fact enter with common coefficients (p-stat = 0.203), supporting the notion that aggregating them into a more easily interpretable measure may not be unreasonable although not testing whether this simple additive structure provides the most informative aggregated variable. Equally striking results are obtained if we utilise the fixed effects model (results available on request from the authors) where even the health problems 'count' is an insignificant variable in the economic activity transition model. Within this set up, by a series of sequential restrictions that exclude variables for which we reject significance at the 10% level, we can eliminate all but five of the 'objective' factors. These are 'trouble walking for 10 minutes', 'anxiety and/or depression', 'trouble doing the housework', 'diabetes' and 'migraine or headaches'. The last two are associated with an increased likelihood of working which indicates that not all of the specific measures are (in isolation) appropriate measures of health as it affects the capacity to work.

Overall, therefore, the two step procedure of deriving a measure of the individual's health stock and then using this constructed stock (and its time variation) in a retirement model seems to provide a richer model of the relationship between ill health and economic (in)activity than simply adding linear combinations of self-reported health problems and health limitations. Of course, this does not rule out the possibility that more detailed assessments of the individual's health status, including variables designed explicitly to measure 'capacity for work', may provide a superior understanding of individual work decisions. This is perhaps especially likely if capacity is measured with reference to the demands of an individual's occupation. Moreover, Column (3) of Table 7 still imposes the restriction that the various current and lagged health indicators have a linear impact on the probability of economic activity. Models that allow for specific clusters of factors affecting economic activity based on epidemiological or other-related health assessments might perform better than our current model. However data dimensionality limits the scope for simply including a multiplicity of interaction terms designed to identify specific clusters of factors on an *ad hoc* basis. In the interim, the two-step procedure used here seems to provide plausible and consistent results.

5. Conclusion

This paper represents one of the first attempts to examine the impact of ill-health on retirement in some detail in the United Kingdom, using the British Household Panel Survey. The focus of the paper is on the nature of health measures that are utilised in reduced form retirement models of the type described here. It argues that reporting bias is intrinsic to self-reported measures of general health (especially, in questions that explicitly link health to economic activity status), but that there is a lack of 'fit' between objective measures of disability and functional limitations on the one hand and health as it relates to economic activity on the other. We therefore follow the approach of Bound *et al* (1999) in constructing an underlying 'health stock' of the individual, and in treating temporal variations in this measure as proxying individual-specific 'health shocks' that affect retirement behaviour, using a fixed effect specification for the labour market model. We also input our health stock measure into a hazard specification of the retirement model that has a relatively flexible dynamic structure and captures duration dependence non-parametrically.

The paper shows that a constructed proxy variable of this type has a powerful explanatory effect for transitions between economic activity and inactivity in a reduced form model that incorporates other time-varying covariates. Moreover, the approach seems superior, in terms of explanatory power, to the application of linear reduced form models incorporating disability index-type health measures. The paper tested for symmetry in labour market transitions to changes in health. Little evidence of asymmetry was found. Richer models of the link between retirement and ill health are being developed, but we believe that the approach used here is capable of generating important insights into that link.

6. References

Anderson, K.H., Burkhauser, R. and Quinn, J.F. (1986) 'Do retirement dreams come true? The effect of unanticipated events on retirement plans', *Industrial and Labor Relations Review*, 39, July, 518-526.

Banks, J., Blundell, R., Disney, R. and Emmerson, C. (2002) 'Retirement, pensions and the adequacy of saving: A guide to the debate', Institute for Fiscal Studies *Briefing Note* 29, London.

Blundell, R. and Johnson, P. (1998) 'Pensions and labour market participation in the UK', *American Economic Review*, 88, May, 168-172.

Blundell, R., Meghir, C., and Smith, S. (2002) 'Pension incentives and the pattern of early retirement', *Economic Journal*, 112, March, C153-C170.

Bound, J. (1991) 'Self-reported versus objective measures of health in retirement models', *Journal of Human Resources,* 26, 106-138.

Bound, J. and Burkhauser, R. (1999) 'Economic analysis of transfer programs targeted on people with disabilities', in O. Ashenfelter and D. Card (eds) *Handbook of Labor Economics* Volume 3C, Amsterdam: Elsevier Science, 3417-3528.

Bound, J., Schoenbaum, M., Stinebrickner, T.R. and Waidmann, T. (1999) 'The dynamic effects of health on the labor force transitions of older workers', *Labour Economics*, 6, 179-202.

Chamberlain, G. (1980) 'Analysis of covariance with qualitative data', *Review of Economic Studies*, 47, 225-238.

Contoyannis, P. and Rice, N. (2001) 'The impact of health on wages: Evidence from the British Household Panel Survey', *Empirical Economics*, 26, 599-622.

Disney, R., Emmerson, C., and Wakefield, M. (2003) 'Ill health and retirement in Britain: a panel data-based analysis', *Working Paper* W03/02, Institute for Fiscal Studies, London.

Disney, R., Grundy, E. and Johnson, P. (1997) *The Dynamics of Retirement*, Research Report No. 72, Department of Social Security, HMSO: London.

Disney, R. and Smith, S. (2002) 'The labour supply effect of the abolition of the earnings rule for older workers in the United Kingdom', *Economic Journal*, 112, March, C136-C152.

Disney, R. and Webb, S. (1991) 'Why are there so many long term sick in Britain?' *Economic Journal,* 101, March, 252-262.

Dwyer, D. S. and Mitchell, O. S. (1999), 'Health problems as determinants of retirement: are self-rated measures endogenous?', *Journal of Health Economics,* vol. 18, pages 173–193.

Greene, W.H. (2000) *Econometric Analysis*, 4th edition, Prentice Hall: London.

Griliches, Z. (1974) 'Errors in variables and other unobservables', *Econometrica*, 42, November, 971-998.

Jenkins, S.P. (1995) 'Easy estimation methods for discrete-time duration models', *Oxford Bulletin of Economics and Statistics,* 57 (1), 129-138.

Jenkins, S.P. (1997) 'sbe17: Discrete time proportional hazards regression (pgmhaz)', *STATA Technical Bulletin*, STB-39, September, 22-32.

Jones, A.M., Contoyannis, P. and Rice, N. (2003) 'The dynamics of health in the British Household Panel Study', Department of Economics and Related Studies, University of York: http://www.york.ac.uk/res/herc/papers.html

Kaplan, G. and Camacho, T. (1983) 'Perceived health and mortality: A nine-year follow-up of the human population laboratory cohort', *American Journal of Epidemiology*, 117, 292-304.

Kerkhofs, M., Lindeboom, M. and Theeuwes, J. (1999) 'Retirement, financial incentives and health', *Labour Economics*, 6, 203-227.

Lazear, E. (1986) 'Retirement from the labor force', 305-355 in O. Ashenfelter and R. Layard (eds) *Handbook of Labour Economics, Volume 1*, Elsevier Science Publishers, BV, Amsterdam.

Meghir, C. and Whitehouse, E. (1997) 'Labour market transitions and retirement of men in the UK', *Journal of Econometrics*, 79, 327-354.

Meyer, B.D. (1990) 'Unemployment insurance and unemployment spells', *Econometrica*, 58 (4), 757-782.

Miniaci, R. and Stancanelli, E. (1998) 'Microeconometric analysis of the retirement decision: United Kingdom', OECD Economics Department *Working Papers* No. 206, Paris.

Myers, R. J. (1982) 'Why do people retire early from work?', *Aging and Work*, 5, 83-91.

Prentice, R. and Gloeckler, L. (1978) 'Regression analysis of grouped survival data with application to breast cancer data', *Biometrics*, 34, 57-67.

Quinn, J.F., Burkhauser, R. and Myers, D.A. (1990) *Passing the Torch: The influence of economic incentives on work and retirement*, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Stern, S. (1989) 'Measuring the effect of disability on labor force participation', *Journal of Human Resources*, 24, Summer, 361-395.

Tanner, S. (1998) 'The dynamics of male retirement behaviour', *Fiscal Studies*, 19, May, 175-196.

Wannamethe, L. and Shaper, A.G. (1991) 'Self-assessment of health status and mortality in middle-aged British men', *International Journal of Epidemiology*, 20, 239-245.

Wooldridge, M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, Mass.

Zabalaza, A., Pissarides, C. and Barton, M. (1980) 'Social security and the choice between full-time work, part-time work, and retirement', *Journal of Public Economics*, 14, 245-276.

Table 1:
Self-assessed health, relative to cohort, 1991 and 1998

	Percentage of whole sample reporting relative health status as:		*Percentage of sample present in 1991 and 1998 reporting relative health status as:*	
	1991	1998	1991	1998
Excellent	27.2	17.8	29.5	18.2
Good	43.4	42.9	43.3	42.4
Fair	18.5	25.3	18.9	25.8
Poor	8.3	10.7	6.7	9.9
Very poor	2.5	3.4	1.6	3.7
No. of observations	*1,440*	*478*	*434*	*434*

Source: constructed by the authors from successive waves of the BHPS

Table 2: Results of 1991 Ordered Probit estimating self–reported health status as a function of 'objective' health measures and individual characteristics

Variable	Mean	s. deviation	Coefficient	Standard error
Male	0.572	0.495	0.024	0.068
In a couple	0.819	0.385	0.202 **	0.081
Age/10	5.571	0.400	-52.481	49.703
(Age/10) squared	31.191	4.519	9.047	8.792
(Age/10) cubed	175.560	38.429	-0.517	0.517
Owner-occupier	0.770	0.421	-0.040	0.089
Own-outright	0.368	0.482	0.038	0.072
Housing equity (£,'000)	54.411	64.246	0.001 **	0.001
Regional unemployment rate, %	6.118	1.500	-0.104	0.081
White	0.970	0.170	-0.025	0.180
Number of children in household	0.093	0.380	-0.099	0.081
Regional dummies:				
Live in conurbation	0.308	0.462	0.091	0.103
South West	0.095	0.294	0.162	0.124
East Midlands	0.081	0.272	0.057	0.151
West Midlands	0.097	0.295	0.195	0.162
North West	0.102	0.303	0.376	0.260
Yorkshire/Humberside	0.088	0.283	0.189	0.213
Rest of North	0.069	0.254	0.451	0.351
Wales	0.052	0.222	0.166	0.273
Scotland	0.087	0.282	0.325	0.289
Education dummies:				
Degree	0.076	0.266	0.405 ***	0.122
Other higher qualification	0.128	0.335	0.252 **	0.099
A-level (or equivalent)	0.077	0.267	0.171	0.119
O-level (or equivalent)	0.133	0.339	0.226 **	0.095
Low education	0.105	0.306	0.271 ***	0.103
Health 'difficulties':				
Doing housework	0.055	0.228	-0.675 ***	0.173
Climbing stairs	0.078	0.268	-0.218	0.171
Getting dressed	0.024	0.152	-0.406 *	0.228
Walking 10mins	0.078	0.268	-0.925 ***	0.167
Health 'problems':				
Arms/legs/hands	0.333	0.472	-0.589 ***	0.067
Sight	0.085	0.279	-0.407 ***	0.110
Hearing	0.090	0.287	0.130	0.106
Skin/allergies	0.077	0.267	-0.243 **	0.110
Chest/breathing	0.098	0.297	-0.926 ***	0.103
Heart/blood press.	0.181	0.385	-0.730 ***	0.079
Stomach/digestion	0.066	0.248	-0.538 ***	0.121
Diabetes	0.026	0.160	-1.082 ***	0.186
Anxiety/depression	0.049	0.217	-0.739 ***	0.137
Alcohol/drugs	0.005	0.070	-0.745 *	0.433
Epilepsy	0.008	0.087	-0.721 **	0.335
Migraine	0.070	0.255	-0.377 ***	0.117
Other	0.052	0.222	-0.693 ***	0.134
F-tests:			Chi-Squared	P-Value
Regional dummies			4.92	0.84
Education dummies			19.07 **	0.002
Health 'difficulties'			151.12 ***	0.000
Health 'problems'			352.98 ***	0.000

**: significant at 5% level
***: significant at 1% level

Number of obs = 1,440; Log likelihood = –1,495.48; LR $\chi^2(42)$ = 834.29

A full set of results from the other 7 years is available from the authors on request.

Table 3:
Economic activity equations

	(1) Fixed effects logit			**(2) Logit**			**(3) Linear fixed effects**		
	Coefficient		***s.e.***	***Coefficient***		***s.e.***	***Coefficient***		***s.e.***
Age/10	29.744	**	12.141	8.699	***	3.138	1.557	***	0.553
(Age/10) squared	-3.155	**	1.084	-0.919	***	0.275	-0.172	***	0.049
Couple	0.129		0.807	0.320	**	0.143	0.045		0.043
No. children in household	-0.870		0.630	-0.455	***	0.147	-0.052	**	0.026
Regional unemployment rate	-0.128	**	0.051	-0.014		0.017	-0.007	***	0.003
Housing equity	-0.006	*	0.004	-0.002		0.001	-0.001	**	0.000
Health stock	**0.352**	*****	**0.117**	**0.782**	*****	**0.059**	**0.041**	*****	**0.009**
Constant	n/a		n/a	-19.582	**	8.976	-2.633	*	1.554
No of cases	Obs=2,279		Groups=374	Obs=6,944			Obs=6,944		Groups=1,608
Log likelihood	-628.28			-3,950.16			R^2 within 0.0951 Between 0.0908 Overall 0.0922		

Note: ***= significant at 1% level **=significant at 5% level *= significant at 10% level.
The standard logit model also includes a set of regressors that are time-invariant, such as gender, educational qualifications and regional dummies, which are not included in either the conditional logit or the linear fixed effects model. These comprise all of the variables shown in table 2 except for those relating to health difficulties (with activities of daily living) and to specific health problems (the excluded variables comprise the vector Z_{it} in equation 1 in section 2.1).
Due to the health stock being calculated from the predicted index of eight ordered probits the standard errors are estimated using a bootstrapping technique. 500 bootstraps for each equation have been run. In the conditional logit and the linear fixed effects model we select a random sample of individuals (and then select all observations for that individual), while in the logit model we select randomly across all observations (i.e. observations of the same individual in different years are selected separately). Due to some of the random draws leading to non-convergence of the ordered probits the total number of bootstraps used in the calculation of the standard errors is 482 in the conditional logit and 481 the linear fixed effects model and also (coincidentally) in the pooled logit.

Table 4: Hazard model of final exit from economic activity

	Coefficient		*(s.e)*	Hazard ratio
Sex (male = 1)	-0.229		*(0.225)*	0.796
Living in a couple	-0.425	*	*(0.250)*	0.653
Number of children in house	0.495		*(0.373)*	1.641
Age at start of obs'd spell	0.224	***	*(0.045)*	1.251
Duration dummies:	*(LR-test rejects joint significance, p-stat 0.301)*			
Two years	0.335		*(0.241)*	1.397
Three years	0.591	*	*(0.305)*	1.805
Four years	0.832	**	*(0.373)*	2.297
Five years	0.783	*	*(0.468)*	2.188
Six years	0.876		*(0.552)*	2.401
Seven years	1.199	*	*(0.665)*	3.317
Time in work at start of obs.	-0.006		*(0.010)*	0.994
Housing equity	0.005	**	*(0.002)*	1.005
Regional unemployment rate	-0.132		*(0.082)*	0.876
Education dummies:	*(LR-test rejects joint significance, p-stat 0.968)*			
Degree	0.062		*(0.317)*	1.064
Other high qualification	-0.029		*(0.271)*	0.971
A-level	-0.143		*(0.358)*	0.867
O-level	0.011		*(0.293)*	1.011
Lower qualifications	-0.310		*(0.313)*	0.733
Region dummies:	*(LR-test* cannot *reject joint sig. at 2%, p-stat 0.012)*			
Conurbation	-0.098		*(0.255)*	0.906
South west	0.331		*(0.342)*	1.393
East midlands	0.264		*(0.369)*	1.302
West midlands	0.358		*(0.360)*	1.430
North west	0.688	**	*(0.317)*	1.990
Yorks/Humberside	0.593	*	*(0.340)*	1.809
Rest of north	0.265		*(0.438)*	1.303
Wales	2.112	***	*(0.621)*	8.267
Scotland	1.172	**	*(0.516)*	3.228
Health stock	**-0.583**	*****	***(0.128)***	**0.558**
Lagged health stock	**-0.611**	*****	***(0.147)***	**0.543**
Constant included; Unobserved heterogeneity included using a gamma mixing distribution				
LR test of significance of unobserved heterogeneity fails to reject, P-stat = 0.000012				
No of cases	2,557 observations, 725 spells, 725 individuals.			
Log likelihood	-739.88			

Note: Results are for the Prentice-Gloeckler/Meyer hazard specification with unobserved heterogeneity described by a gamma mixing distribution. Due to the health stock variables being predicted from ordered probits, the standard errors are estimated using a bootstrapping technique. 500 bootstraps for each were run. For each replication of the bootstrap a random sample of individuals was drawn and all observations of each individual were used. Due to 19 of the random draws leading to non-convergence of the ordered probits, the total number of bootstraps used in the calculation of the standard errors was 481.

Table 5: Testing Symmetry:
Hazard models for exit from/entry into economic activity

(a) Exit model

	Coefficient		***(s.e)***	**Hazard ratio**
Sex (male = 1)	-0.137		*(0.170)*	0.872
Living in a couple	-0.208		*(0.226)*	0.812
Number of children in house	0.255		*(0.233)*	1.291
Age at start of obs'd spell	0.133	***	*(0.028)*	1.142
Housing equity	0.003	**	*(0.002)*	1.003
Regional unemployment rate	-0.045		*(0.068)*	0.956
Health stock	**-0.577**	*****	***(0.106)***	**0.562**
Lagged health stock	**-0.360**	*****	***(0.110)***	**0.698**

Dummies included for: duration (No. of years) [a]; education (six levels); region

Constant included

Unobserved heterogeneity included using a gamma mixing distribution

LR test of significance of unobserved heterogeneity fails to reject, P-stat = 0.0003

No of cases	2,732 observations, 799 spells, 747 individuals
Log likelihood	- 924.6

(b) Entry Model

	Coefficient		***(s.e)***	**Hazard ratio**
Sex (male = 1)	0.250		*(0.186)*	1.284
Living in a couple	0.421	*	*(0.249)*	1.524
Number of children in house	-0.336		*(0.292)*	0.714
Age at start of obs'd spell	-0.134	***	*(0.026)*	0.874
Housing equity	-0.004	**	*(0.002)*	0.996
Regional unemployment rate	-0.113	**	*(0.057)*	0.893
Health stock	**0.158**		***(0.122)***	**1.171**
Lagged health stock	**0.404**	*****	***(0.132)***	**1.498**

Dummies included for: duration (No. of years) [a]; education (six levels); region

Constant included

Unobserved heterogeneity not included.

LR test of significance of (gamma) unobserved heterogeneity rejects, P-stat = 0.5

No of cases	1,834 observations, 654 spells, 621 individuals
Log likelihood	- 396.3

[a] Plus a variable indicating the number of years a person had been working/unoccupied when they first entered the sample, if they were working/unoccupied when first observed.

Note: Due to the health stock variables being predicted from ordered probits, the standard errors are estimated using a bootstrapping technique. 500 bootstraps for each were run. In the exit model 482 ran successfully while 488 ran successfully in the entry model.

Table 6:

Symmetry and economic activity equations
(Alternative health question)

	(1) Fixed effects logit			(2) Logit			(3) Linear fixed effects		
	Coefficient		*s.e.*	*Coefficient*		*s.e.*	*Coefficient*		*s.e.*
Age/10	40.184	**	16.147	12.450	***	3.675	1.905	***	0.663
(Age/10) squared	-4.115	***	1.442	-1.220	***	0.319	-0.201	***	0.059
Couple	1.078		1.860	0.271		0.177	0.103	*	0.059
# children in household	-1.372	*	0.778	-0.513	***	0.171	-0.080	***	0.030
Regional unemployment rate	-0.194	**	0.081	-0.012		0.021	-0.009	**	0.004
Housing equity	-0.004		0.004	-0.003	**	0.001	0.000		0.000
Health no limit on work	**1.650**	*********	**0.467**	**2.713**	*********	**0.211**	**0.162**	*********	**0.040**
Symmetric health impact	**-0.180**		**0.125**	**-0.451**	*********	**0.091**	**-0.018**	********	**0.008**
Constant	N/a		N/a	32.920	***	10.588	-3.802	**	1.859
No of cases	Obs=1,495		Groups 271	Obs=5,098			Obs=5,098		Groups=1,239
Log likelihood	-429.31			-2,905.48			R^2 within 0.0777 Between 0.1109 Overall 0.0979		

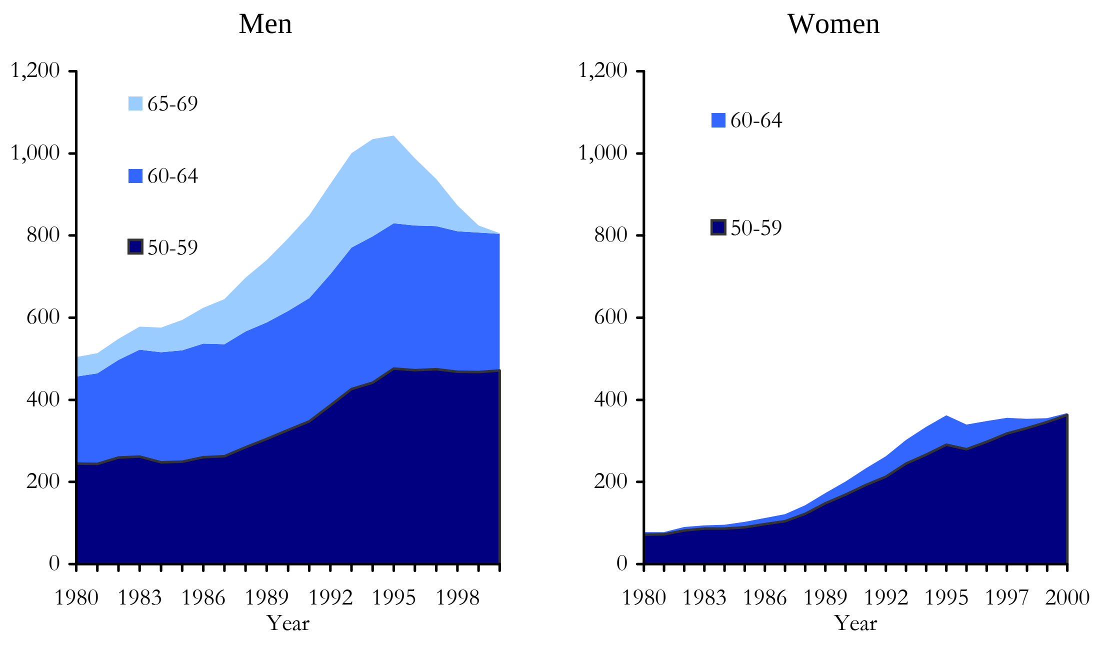
Note: ***= significant at 1% level **=significant at 5% level *= significant at 10% level.
The IV-type technique is applied to the question 'Does your health affect the type or amount of work that you undertake?' [YES/NO].
As in Table 3, the pooled logit model also includes a set of time invariant regressors for which coefficients are not reported.
Also as explained in the note to Table 3, standard errors are estimated using a bootstrapping technique with 500 replications. All random draws converged as the 1st stage estimation is now a probit rather than the ordered probit used in the estimations reported in Table 3.

Table 7: Sensitivity of economic activity to health measure (hazard model)

Specification	*(1)*			*(2)*			*(3)*		
	Coeff	*(se)*	*Haz rat*	Coeff	*(se)*	*Haz rat*	Coeff	*(se)*	*Haz rat*
Predicted health stock	-0.583	*(0.128)*	0.558						
Lag pred. health stock	-0.611	*(0.147)*	0.543						
No of Difficulties with:				1.190	*(0.240)*	3.288			
Doing housework							1.216	*(0.584)*	3.374
Climbing stairs							1.073	*(0.548)*	2.923
Getting dressed							1.042	*(1.013)*	2.835
Walking 10mins							1.281	*(0.526)*	3.602
Lag No. difficulties with:				0.668	*(0.260)*	1.949			
Doing housework							0.531	*(0.743)*	1.701
Climbing stairs							0.865	*(0.676)*	2.374
Getting dressed							1.102	*(1.085)*	3.011
Walking 10mins							0.201	*(0.681)*	1.223
No of Health problems with:				0.184	*(0.099)*	1.203			
Arms/legs/hands							0.180	*(0.203)*	1.197
Sight							0.310	*(0.445)*	1.364
Hearing							0.130	*(0.300)*	1.139
Skin/allergies							-0.356	*(0.364)*	0.700
Chest/breathing							0.175	*(0.354)*	1.192
Heart/blood press.							0.171	*(0.275)*	1.186
Stomach/digestion							0.326	*(0.346)*	1.386
Diabetes							-1.775	*(1.363)*	0.169
Anxiety/depression							1.008	*(0.397)*	2.741
Alcohol/drugs							1.890	*(1.220)*	6.621
Epilepsy							-0.526	*(4.874)*	0.591
Migraine							-0.041	*(0.423)*	0.959
Other							-0.277	*(0.493)*	0.758
Lag No. health probs with:				0.293	*(0.102)*	1.340			
Arms/legs/hands							0.563	*(0.201)*	1.756
Sight							0.556	*(0.375)*	1.744
Hearing							-0.176	*(0.317)*	0.838
Skin/allergies							0.295	*(0.345)*	1.343
Chest/breathing							0.661	*(0.367)*	1.937
Heart/blood press.							0.649	*(0.292)*	1.914
Stomach/digestion							-0.675	*(0.419)*	0.509
Diabetes							0.079	*(1.350)*	1.083
Anxiety/depression							-0.355	*(0.512)*	0.701
Alcohol/drugs							1.621	*(1.094)*	5.056
Epilepsy							-1.506	*(4.899)*	0.222
Migraine							-0.023	*(0.392)*	0.977
Other							0.700	*(0.432)*	2.013
Log likelihood	-739.88			-728.41			-710.33		

Notes to Table 7: Results are for the Prentice-Gloeckler/Meyer hazard specification with unobserved heterogeneity described by a gamma mixing distribution.
The regression reported in column (1) is our baseline model as reported in table 4. The regression in column (2) replaces our generated health stock regressor and its lag with count variables indicating the number of specific health problems and its lag, and the number of difficulties with activities or daily living, and its lag. Instead of entering counts, in column (3) we enter dummies indicating each difficulty or problem. Column (2) therefore amounts to a restriction of column (3) such that within the groups of current and lagged problems and difficulties variables enter with a fixed coefficient.

Figure 1
Numbers of claimants of Invalidity and Incapacity Benefit aged 50 and over, 1980 to 2000



Note: Data in thousands. Source: Banks *et al* (2002)

Appendix 1: Descriptive statistics for constructed health stock variable

Variable	25th percentile	Median	75th percentile	Mean	(s. d.)
Health stock, those in work	-0.055	0.393	0.749	0.275	(0.659)
Health stock, those not in work	-1.044	-0.079	0.503	-0.383	(1.163)
Health stock, all in baseline	-0.381	0.226	0.672	-0.000	(0.961)
Change in health stock	-0.337	-0.003	0.337	-0.006	(0.637)
Difference between max and min health stock (one obs per person)	0.494	0.877	1.308	0.996	(0.694)

In total there are 6,944 person/year observations, (4,043 of people who are in work and 2,901 of people who are out of work) for 1,608 separate individuals. We observe 1,259 people more than once and a total of 5,088 first differences.

These data show that it is uncommon for the measured health stock of an individual to change by as much as one unit between one year and the next: 11 per cent (535 of 5,088 observations) of the year-on-year changes that we observe are of one unit of more in either direction (of these, 273 are declines and 262 are increases).

However, the average difference between maximum and minimum observed health stock (for those observed more than once) is almost exactly one. The median value of this difference is slightly lower than one and 41 per cent (515 of 1,259) of individuals in this sample experience a variation of health stock in excess of one unit.

Appendix 2: Definitions of BHPS health variables

The following is the text of the BHPS health questions from the 1998 questionnaire. The wording is very similar across all waves of the survey. Copies of the BHPS user documentation are available via the website of the University of Essex based Institute of Social and Economic Research (ISER) at http://iserwww.essex.ac.uk/ulsc/bhps/.

Questions on Health and Caring

I would now like to ask you about your health and the use you make of health services.

Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been .....[read out]

Excellent; Good; Fair; Poor; or Very poor; (Don't know)?

Can I check, are you registered as a disabled person, either with Social Services or with a green card?

Yes/No; (Don't know)

Do you have any of the following health problems or disabilities?

[Exclude temporary conditions; code all that apply or code 'none'; read out:]

- None
- *A.* Problems or disability connected with: arms, legs, hands, feet, back, or neck (including arthritis and rheumatism)
- *B.* Difficulty in seeing (other than needing glasses to read normal size print)
- *C.* Difficulty in hearing
- *D.* Skin conditions/allergies
- *E.* Chest/breathing problems, asthma, bronchitis
- *F.* Heart/blood pressure or blood circulation problems
- *G.* Stomach/liver/kidneys or digestive problems
- *H.* Diabetes
- *I.* Anxiety, depression or bad nerves
- *J.* Alcohol or drug related problems
- *K.* Epilepsy
- L. Migraine or frequent headaches
- M. Other health problems (please give details)

Does your health in any way limit your daily activities compared to most people of your age?

Yes/No; (Don't know)

I am going to read you out some activities. Please tell me which, if any, you would normally find difficult to manage on your own?

[Code all that apply or code 'none'; read out:]

a) Doing the housework
b) Climbing stairs
c) Dressing yourself
d) Walking for at least 10 minutes
e) (None of these)

Does your health limit the type of work or the amount of work you can do?

[Include both paid and unpaid work]

Yes/No; (Don't know)

Footnotes

[1] For a survey of US evidence that uses self-reported measures, see Quinn, Burkhauser and Myers (1990). There are relatively few studies of retirement behaviour in the United Kingdom. Zabalza, Pissarides and Barton (1980) use self-reported 'poor health', Meghir and Whitehouse (1997) use changes in self-reported 'health problems' as their proxy for ill health and Miniaci and Stancanelli (1998) use 'ill health as a reason for leaving last job' in modelling the ill health-retirement 'route'. The study by Blundell, Meghir and Smith (2002) is described below.

[2] Of course, receipt of disability benefit is also conditioned on various 'objective' health assessments and work capability tests. These tests have changed over time in the UK and were toughened significantly in 1995 without any apparent change in claimant numbers below state pension age, although new claims were ended once the individual reached state pension age – see Figure 1.

[3] These 'objective' measures may also be self-reported, or be externally assessed by the interviewer or by some form of medical examination. Again in the UK context, Blundell, Meghir and Smith (2002), in their model of the impact of pension incentives on retirement, use a disability 'severity score' calculated from various 'objective' measures of self-reported disability, and also estimate an auxiliary reduced form probit for receipt of Invalidity Benefit.

[4] For example, construction of a disability 'score' such as the well known Disability Living Index may not primarily be motivated by attempts to measure the employment capacity of the individual.

[5] There is some evidence, for example, that self-reported health status is an additional predictor of individual mortality after controlling for observables (Kaplan and Camacho, 1983; Wannamethe and Shaper, 1991).

[6] Dwyer and Mitchell (1999) implement a similar instrumental variables approach to look at the effect of health on the retirement expectations (as opposed to actual retirement) of men using a cross section from the US health and retirement study.

[7] The hypothesis of no effect would not seem inconsistent with the aggregate data (figure 1) which indicate that among the provisions of the reform package it was the exclusion of people aged over the state pension age, rather than the tightening of the health test, which had the impact of reducing claimant numbers.

[8] Bound (1991) illustrates these likely outcomes concerning h^* and the use of 'objective factors', formally: see *ibid* pp.110-114.

[9] The question relating to health status changed in the BHPS in 1999.

[10] Very few people work beyond state pension age in the UK, especially among men. However, we also experimented with imposing no upper limit on age so that the oldest members of our sample had reached age 71 by 1998. This made little difference to our results.

[11] There is a separate question concerning self-reported retirement in the BHPS, but previous analysis of the Retirement Survey (Disney, Grundy and Johnson, 1997) and the BHPS itself suggests evidence of recall bias, and that observed prolonged inactivity is not always defined as 'retirement' whereas self-reported 'retirement' is sometimes associated with economic activity; moreover self-reported retirement has a larger spike at state pension age than is warranted by the pattern of exits into inactivity by age.

[12] Our thanks to Andrew Henley at University College, Aberystwyth for providing the results of his programme that models housing equity in the BHPS.

[13] The BHPS also contains a series of questions that contribute to a constructed index of "subjective well being". These questions ask things like whether or not people feel that recently their concentration has been good, and whether or not they feel that recently they "have been playing a useful part in things". In total there are twelve such questions, but after some experimentation we decided not to use these variables as asking how people feel about themselves may be more subjective than the sets of questions about health problems and limitations on daily activities.

[14] Contoyiannis and Rice (2001) use this exact variable in their analysis of the impact of health on wages. When they use panel IV estimators, the impact of self-reported health status is not significant (*ibid* Tables 3 and 4) suggesting that measurement error (heterogeneity) dominates. Note that we allow for person specific effects at the second stage of our estimation procedure.

[15] The numbers decline because of the nature of the age selection of our sample, and also due to sample attrition over the period of the panel, either through death or non-response. Note that biases due to the association of poor health and mortality should lead to an overestimate of the average health of the ageing cohort since subsequent responses are conditioned on survival. The smaller number in the second set of cells arises because some individuals observed in 1998 failed to respond to all the health questions in 1991.

[16] See Greene (2000, p.876ff.) or Wooldridge (2002, p. 504ff.).

[17] Results of the ordered probits for the years 1992-1998 can be obtained from the authors on request. Qualitatively they are similar to the results for 1991.

[18] For example, individuals may notionally evaluate the utility from retirement now against all future prospective utility streams, including returning to work, as in a dynamic programming problem, or else evaluate retirement now against the highest valued stream from future retirement, assuming retirement is an absorbing state, as in the 'option value' model.

[19] Some panel attrition occurs because individuals die, enter residential care or make untraced moves. A referee made the pertinent point that such transitions may correlate with poor or deteriorating health status, so biasing the relationship between observed economic activity and health. We therefore estimated the equations on the sub-sample who appeared in all waves and only exited pre-1998 due to reaching state pension age. This reduces sample observations – for example in the fixed effect logit to 2062. However parameter estimates are very similar; again by way of illustration the coefficient on the health variable in the fixed effect logit becomes 0.335 (0.352 in Table 3). Results are available from the authors on request. The issue of health-related attrition in the BHPS is further discussed in Jones *et al* (2003).

[20] This is run using Stephen P. Jenkin's *pgmhaz8* command in STATA.

[21] Whilst the period 1997-2003 confirms that very large house price increases can occur over relatively short periods, over the period 1991-98 nominal house prices actually arose by around 30%, suggesting that house prices had a much smaller, albeit potentially significant impact on retirement.

[22] Note that we require at least two observations to observe this change, so that the first observation of each respondent is absent. This is why the sample size is smaller in Table 6 than Table 3.