PAPER PREPARED FOR HESG GLASGOW SUMMER 2004

DRAFT PLEASE DO NOT QUOTE

HEALTH AND RETIREMENT IN THE UK: DURATION ANALYSIS OF THE BRITISH HOUSEHOLD PANEL SURVEY

Jennifer Roberts*,
Sheffield Health Economics Group, ScHARR, University of Sheffield

Nigel Rice
Centre for Health Economics, University of York

Andrew Jones
Dept of Economics and Related Studies, University of York

* Corresponding author
Jennifer Roberts
Reader in Health Economics
ScHARR, University of Sheffield
Regent Court
Sheffield, S1 4DA
Email: J.R.Roberts@shef.ac.uk

Abstract

The primary focus of this paper is the role of health in determining retirement behaviours in the UK; a country that is experiencing population ageing combined with increasing early exit of older workers from the labour market. For older workers, health status can be an important factor in the decision to retire.

This paper investigates this phenomenon using 11 waves of the British Household Panel Survey. We follow individuals as they retire using transition models. By providing an appropriate counterfactual (the same individual before and after retirement), this allows us to overcome the problems of endogeneity and unobservable individual heterogeneity, which plague investigation of these causal relationships.

While there is a plethora of evidence on the importance of financial incentives to the retirement decision there is much less that considers the relationship between health and retirement and very little in the European context. This paper provides useful information about the relative importance of health within the UK pensions and benefits system.

Preliminary results suggest that the effect of health limitations on the retirement hazard is positive and statistically significant for men and women. The effect of health problems combined is positive but not statistically significant. Analysis of individual components reveals that `other' health problems are significant for men, whilst problems with sight are significant for women. Self-assessed health has no significant effect on the retirement hazard after accounting for the other (more objective) health measures. The survival analysis reveals some counterintuitive results for income and pension entitlement.

1. **Introduction**

This paper is part of a larger project which aims to investigate the role of socioeconomic factors, health and household structure on labour market transitions in the later stages of working life. Our focus here is on retirement behaviour and in particular the role of health in determining early retirement. We use data from the first 11 waves of the British Household Panel Survey (1991 to 2001).

Until recently, the UK (along with most developed countries) experienced a trend towards earlier retirement, especially among males. Along with more generous social security systems and the introduction of early retirement options in pension systems, early retirement plans were advocated to 'solve' unemployment problems, with early retirees making room for younger unemployed individuals. More generous health and disability insurance systems also contributed to the trend by enabling individuals in bad health to drop out of the labour market without facing severe financial consequences. This 'disability route' into retirement has been identified as an important phenomenon of the labour market in the UK (Blundell et al 2002).

There is some evidence that the trend towards ever younger retirement has levelled off (Disney and Hawkes 2003). This may be related to reductions in the basic state pension and growing reliance on private and occupational pensions which are increasingly unreliable as a source of income in older age. However, an increase in wealth (particularly housing equity) of those facing the retirement decision today as compared to some decades ago may help to explain why employment rates among older people are still below 1970s levels.

A recent survey for the Dept for Work and Pensions (Humphrey et al 2003) explored the factors affecting labour market participation among 2800 people aged between 50 and 69. There were three important sub-groups among those of working age who were not in employment: (i) those people (largely male) who were not seeking work due to ill-health (50%); (ii) those who were not seeking work because they were financially secure and did not want to work (22%); (iii) those people (largely female) who were not seeking work because they were looking after the home or family. Health had a significant impact on continued labour market participation for all respondents, and 20% had been forced to retire or leave a job because of ill-health.

The relationship between health and retirement is complex and it is important to consider life cycle earnings and consumption patterns and how these are related to long-term health and short-term changes in health. While poor health may necessitate early exit from the labour market, experience of long-term poor health may also mean that the individual is less financially equipped to enter into early retirement.

In order to understand these relationships it is necessary to use longitudinal data. In this way we can track people from work into retirement and by providing an appropriate counterfactual (the same individual before and after retirement), we can seek to overcome the problems of endogeneity (reverse causality) and unobservable individual heterogeneity (confounding), which plague investigation of these causal relationships. Given the dearth of evidence on health and retirement in the UK this paper makes a significant contribution to the knowledge base in this area.

Section 2 outlines the theory and existing evidence on health and retirement. Section 3. describes our data, variable definitions and methods of analysis. The results are reported in Section 4 and discussed in Section 5.

2. Background Literature and Theory

While there is much evidence on the causal effects of income and wages on health (Benzeval and Judge, 2001; Smith 1999; Backlund, 1999, 1996; Ettner, 1996;), there is less concerning the (reverse) impact of health on income and the decision of early retirement. For workers, health impacts on wages due to its affect on productivity and the potential for employers to discriminate against workers with health problems regardless of their productivity. For older workers there is an important additional causal route because health (their own and that of their spouse) can be an important factor in the decision to retire, and as such health will be closely related to income. We believe it is the combined effect of health and wealth that is of primary importance in the retirement decision.

The existing literature has in general found a positive impact of health on wages and/or income (Wu 2003; Madden 2004; Sundberg 1996). However, the vast majority of work uses cross-sectional data. In addition the existing literature tends to focus on prime age adults; hence the findings may not be applicable to older workers.

For the UK, Walker and Thompson (1996) use the first three waves of the British Household Panel Survey (BHPS) to estimate a model of hourly wages including measures of disability. Applying OLS and procedures to account for selective participation, they find that disability reduces years of schooling, wages and the probability of labour force participation. Disability appears to mainly affect participation rather than wages, and once the endogeneity of schooling has been accounted for the effect of disability on wages is very small.

Contoyannis and Rice (2001) have estimated the effect of self-assessed general and psychological health on hourly wages using data from six waves of the BHPS. Their results show that reduced psychological health reduces the hourly wage for males, while excellent self-assessed health increases the hourly wage for females.

Economic theory on the relationship between retirement and health (Lazear, 1986) states that agents have preferences over current and future leisure which depend, inter alia, on current and expected health status. Poorer health reduces the probability of continued work because it may: increase the disutility of work; reduce the return from work (via lower wages); entitle the individual to benefits and other non-wage income that is contingent on not working. A possible counteracting affect is that poorer heath may increase consumption requirements (for example via increased health care costs) therefore necessitate higher income. However, if poorer health also reduces life expectancy then the annualised consumption available from existing wealth is raised, and this may still lead to earlier retirement.

While there is a plethora of evidence on the importance of financial incentives to the retirement decision (Lumsdaine and Mitchell, 1999) there is much literature that considers the relationship between health and retirement and very little in the UK context. Blundell et al (2002) use 2 waves of the UK Retirement Survey to assess the effect of pension incentives on the decision to retire. They also include self assessed health (the Severity Score) and this is one of the few non-pension related and age variables that is significant when they model the conditional probability of exit from employment (together with retirement status of spouse and whether or not they have an outstanding mortgage) However the potential endogeneity of SAH is ignored and health is treated as exogenous variable.

For the US, Bound et al. (1999) use 3 waves of the Health and Retirement Study to consider the retirement behaviour of men and women aged 50-62 in wave 1. Theoretically as health declines with age it requires adaptation or cessation of work activities. Whether and how workers respond to declines in health status depends on various factors like the nature of the decline, the expected persistence, the age at which it occurs, the worker's human capital, the economic situation and preferences for consumption and leisure. The authors estimate a health stock variable to deal with the potential endogeneity of self-assessed health to the retirement decision and find that health plays an important role with changes in health as important as the long-term level of health in determining the retirement age.

For the UK, Disney et al (2003) motivate their work with the observation that ill health is a major reason for retirement, especially among those with no occupational pension. They use similar techniques to Bound et al (1999) with the first 8 waves of the BHPS. There are some problems with their (non) definition of actual retirement but they find that better health is strongly associated with continued economic activity. In assessing the effects of the 1995 disability benefit reforms, which tightened eligibility conditions and reduced the incentives to retire via the disability route in the UK, they find no significant effects but admit that this could be due to methodological flaws.

For Germany, Siddiqui (1997a, 1997b) uses the GSOEP and finds strong and significant effects of health on retirement behaviour. Individuals reporting a chronic condition or a disability are four times more likely to leave employment at the earliest possible date than healthy individuals. Riphahn (1997) carries out similar analysis but distinguishes between unemployment and disability routes into retirement and shows that these routes cannot be considered substitutes but rather are driven by different forces. More recently Riphahn (1999) has investigated the dynamic effect of health shocks on the employment and income of older workers. She finds significant effects of a health shock on leaving employment and finds small reductions in individual and equivalent household income. Another interesting fact is that health shocks seem to happen more often to those individuals/households which are already at the lower end of the income distribution before the health shock occurs.

3. Data and Methods

The BHPS is a nationally representative longitudinal data set that collects information on the same individuals over time. The data are extremely rich, including information on many variables including socio-demographic information, activity status, health status, use of health care services, income and sources of income. The panel structure of the data set is particularly suited to analysing dynamic causal relationships rather than simply identifying associations between variables.

The sub-sample analysed here consists of those individuals who were aged 50 or over <u>and</u> had a full interview <u>and</u> were in the labour force (defined as employed, self-employed or unemployed) in <u>wave 1</u> of the survey. This is a sample of n = 1245 individuals, 520 females and 725 males, who are then followed through the eleven annual waves of the surveys to see if they leave the labour market due to retirement[1]. 751 individuals are present for all 11 waves, but others are lost due to attrition, and by wave 11 420 people have left the sample[2]. Also 506 observations are omitted because we only deal with complete sequences of

[1] There is a potential selection issue that is not dealt with here since people in very poor health may not have been working in wave 1 and hence cannot be observed to retire in our sample.

[2] Attrition may be endogenous (related to SAH) but we do not deal with this issue in the current version of the paper.

responses, so if the individual is not interviewed in a subsequent wave they are deemed to have left the sample at that point even if they return in later waves[3].

Duration models of retirement

The starting point for our analysis is the duration model stock-sampling approach of Jenkins (1995). This method represents the transition to retirement as a discrete time hazard model enabling us to estimate the effect of covariates on the probability of retirement. The covariates will include health status and various socio-economic characteristics such as age, sex, housing tenure and equity, education and pension entitlement.

The Jenkins (1995) approach relies on a reorganisation of the data set so that the unit of analysis is changed from the *individual* to the *time at risk of an event* (in this case retirement). This allows a complex sequence likelihood to be simplified to standard estimation for a binary outcome. The approach also controls for 'stock sampling' since we sample only those people who are in the labour market at wave 1. These individuals can then either stay in the labour force or exit into retirement. The simplest assumption is a single exit (retirement) from the initial state (working), modelled as a binary probit, but this can be extended to a competing risks model using an ordered probit or a multinomial logit; for example, adding the transition from full-time to part-time employment[4].

We use data for a stock sample of all individuals who are working in wave 1 ($t = \vartheta$). $t = 1$ is the earliest year at which an individual is at risk of retirement (initially we assume this is at age 50). At the end of the time period for which we have data some people will still be working (censored duration data, $\iota_i = 0$), and some will have retired (complete duration data, $\iota_i = 1$)[5]. $t = \vartheta + s_i$ is the year when retirement occurs if $\iota_i = 1$ and the final year of our data period if $\iota_i = 0$. Each respondent, i, contributes s_i years of employment spell data from the interval between sample selection and the final wave.

The distribution of durations is modelled via probabilities of retiring at each value of t. There is a one-to-one relationship between these hazard rates and the probabilities of having completed employment spell durations of different lengths, summarised by the survivor function.

The discrete time hazard rate is:

$$h_{it} = \mathrm{P}\,[\mathrm{T}_i = t \mid \mathrm{T}_i \text{ Ó} t \,;\, X_{it}]$$

X_{it} is a vector of covariates which may vary with time.
T_i is a discrete random variable representing the time at which the end of the spell occurs.

The main issue for modifying the sample likelihood to take account of stock sampling is that the probabilities for both those who complete spells (retire) and those who do not are conditioned on them not having retired by the beginning of the sample time period (wave 1). The conditional probability of observing the event history of someone with an uncompleted spell at interview is:

---

3 We are hoping to relax this restriction in subsequent work.

4 Competing risks models are not considered here but will be part of our future research with this data.

5 There is also the possibility of extending the analysis to consider a number of spells of employment and retirement, as we already know that a single and permanent exit to retirement is not always the case for many individuals.

$$prob(T_i > t \mid T_i > \vartheta - 1) = \prod_{t=\vartheta}^{\vartheta+s_i} (1 - h_{it})$$

and the conditional probability of observing the event history of someone completing a spell between the beginning period, ϑ, and interview is:

$$prob(T_i = t \mid T_i > \vartheta - 1) = h_{i\vartheta+s_i} \prod_{t=\vartheta}^{\vartheta+s_i-1} (1 - h_{it}) = \left(h_{i\vartheta+s_i} / (1 - h_{i\vartheta+s_i})\right) \prod_{t=\vartheta}^{\vartheta+s_i} (1 - h_{it})$$

Accordingly, the corresponding log-likelihood of observing the event history data for the whole sample is:

$$\log L = \sum_{i=1}^{n} \iota_i \log\left(h_{i\vartheta+s_i} / (1 - h_{i\vartheta+s_i})\right) + \sum_{i=1}^{n} \sum_{t=\vartheta}^{\vartheta+s_i} \log(1 - h_{it})$$

The sequence log-likelihood is difficult to maximise directly using standard maximum likelihood routines. Jenkins (1995) simplifies the likelihood by defining:

$y_{it} = 1$ if $t = \vartheta + s_i$ and $\iota_i = 1$, $y_{it} = 0$ otherwise.
For stayers $y_{it} = 0$ for <u>all</u> spell periods.
For exiters $y_{it} = 0$ for <u>all</u> periods <u>except</u> the exit period.

The likelihood can now be expressed as:

$$\log L = \sum_{i=1}^{n} \sum_{t=\vartheta}^{\vartheta+s_i} y_{it} \log\left(h_{it} / (1 - h_{it})\right) + \sum_{i=1}^{n} \sum_{t=\vartheta}^{\vartheta+s_i} \log(1 - h_{it})$$

The log-likelihood is therefore simplified to a more standard form for analysis of the binary variable y_{it} where the unit of analysis is now the spell period i.e. actual estimation is a simple binary model for the reorganised data set. The reorganised data set consists of multiple rows of observations foreach individual with as many rows as periods at risk. In this way the reorganised data set resembles a panel data set.

To complete the specification of the likelihood, an expression for the hazard rate is required. We specify a complementary log-log hazard rate which is the discrete-time counterpart of the hazard for an underlying continuous-time proportional hazards model (Prenticee and Gloeckler, 1978):

$$h_{it} = 1 - \exp(-\exp(X_{it}\eta + \chi(t)))$$

where $\chi(t)$ is the baseline hazard. We complete the specification by modelling $\chi(t)$ as a step function by using dummy variables to represent each period at risk. This non-parametric form for the baseline hazard leads to a semi-parametric specification of the discrete-time duration model[6]. The model can be generalised to account for unobserved heterogeneity uncorrelated with the explanatory variables (Narendrenathan and Stewart, 1993).All estimation is carried out in STATA (SE) v7. The discrete time proportional hazards models are estimated using the pgmhaz routine (Jenkins, 1998).

[6] An alternative is to specify a parametric form for the baseline hazard, most commonly, to assume this has a Weibull distribution by specifying $\chi(t) = \log(t)$.

Variable definitions
See Appendix 1 for a full list of variable definitions.

Health
The BHPS includes a measure of self assessed health status (SAH) where individuals are asked to rate their health on average over the last twelve months relative to someone of their own age; this is usually coded on a 5-point scale - very poor, poor, fair, good, excellent. In wave 9 the SF-36 health survey was included in the BHPS resulting in a changed SAH question and modification of response categories for that wave only. The wave 9 question asks for respondents 'general state of health' (without the age consideration) on the scale: poor, fair, good, very good, excellent. In order to achieve consistency over all 11 waves we follow the recommendations of Hernandez-Quevedo et al (2204) and recode SAH into a 4 category scale where 1 is *very poor or poor,* 2 is *fair*, 3 is *good or very good* and 4 is *excellent*.

The SAH variable is a subjective proxy for actual health, and is likely to be measured with error and endogenous to labour supply choices (Kerkhofs et al 1999); for example, individuals who are economically inactive may have an incentive to report poor health, and inactivity may contribute to their perception of their own health as poor. A number of empirical studies suggest that the resulting bias tends to exaggerate the negative impact of poor health on labour supply (e.g. Chirikos and Nestel, 1984; Anderson and Burkhauser, 1985). However, there is also evidence that SAH measures are reliable (Haveman et al; 1989, Stern, 1989) and some of this is in the context of health as a determinant of retirement behaviour in the US (Dywer and Mitchell, 1999).

As well as SAH we also use a number of arguably more objective health indicators available in the BHPS. Individuals are asked whether or not they have any of a list of specific health problems. This allows us to create a binary dummy variable to denote *any* health problems and also a number of binary dummies for the presence of specific problems with the following: arms, legs or hands, sight, hearing, skin conditions or allergies, chest/breathing, heart/blood pressure, stomach or digestion, diabetes, anxiety or depression, alcohol or drugs, epilepsy or migraine.

Two further binary health variables are available. The first measures self-reported functional limitations and is based on the question "does your health in any way limit your daily activities compared to most people of your age?[7]" The second asks respondents to report whether or not they are a registered disabled person.

Retirement and labour market status
As has already been noted in the literature (e.g. Bardasi et al 2002; Disney et al 1994) retirement is not a well-defined state. It is not always clear whether individuals in the relevant age group are economically inactive because they have retired or are simply unemployed for a period of time. This problem is exacerbated by pension entitlement because some individuals may associate retirement with the final and permanent exiting from work whereas others may not define themselves as retired unless they are actually in receipt of pension. Further social norms and routes into retirement via disability and unemployment complicate the self-reporting of labour market status for older workers. A DWP survey (Humphrey et al 2003) noted that after State Pension Age people appear to redefine their status, with the proportion

[7] This question is not asked in wave 9. In our analysis we assume wave 8 values hold in wave 9.

of people who described themselves as retired increasing from 26% of men aged 60-64 to 90% of men aged 65-69. They argue that this is not simply a function of actual retirement from paid work but is also about how people see themselves. This is more clearly reflected in the sudden drop in the numbers of people defining themselves as long-term sick or disabled after State Pension Age – from 27% to 1% in men, and from 16% to 4% in women.

In line with previous work, the definition of retirement used here is self-reported in answer to the question on job status in which individuals classify their status as one of the following: self-employed, employed, unemployed, retired, on maternity leave, caring for the family, in full-time education, long-term sick or disabled, or on a government training scheme. For the analysis reported here we assume that retirement is an absorbing (permanent) state[8].

Income and wealth

The main income variable used is the individual specific mean of log household income[9] across all waves in which individual is observed. In the models reported here we adapt this to the mean across all waves prior to retirement, in order to minimise endogeneity problems (as income will normally change significantly at retirement).

We also have information on pension entitlement, which distinguishes between people who have no occupational or private pension, an occupational pension, a private pension (pre 1988), a private pension (post 1988). Data on housing wealth[10] and housing tenure are also available[11].

Other socio-demographic variables

Other covariates include age, sex, marital status, educational attainment, regional dummies and a dummy for whether or not the individual has reached state pension age.

4. Results

Descriptive statistics for men and women in wave one are presented in Table 1. There are very few differences by sex; the only notable differences being (as expected) more women at state pension age, lower educational attainment for women and fewer women with private or occupational pensions. In terms of health in wave 1 most people report good or very good health with very few people reporting poor or very poor health[12]. Only very small numbers are registered disabled persons. While around two thirds of people report the presence of specific health problems only a small number report health problems that limit daily activities. The majority of people are home owners with remaining mortgages and most have housing equity of over £50,000.

---

8 Simple exploratory analysis of the BHPS for example shows that the assumption that retirement is a single and permanent exit from the labour market is not appropriate; of those people aged over 50 and retired in wave 1 some are participating in the labour market again in wave 2, and this pattern is repeated across all waves. We expect to relax this assumption at a later stage and investigate the sensitivity of our results

9 Household income consists of labour and non-labour equivalised real income, adjusted using the Retail Price Index and equivalised by the McClement's scale to adjust for household size and composition

10 We are grateful to Andrew Henley at University College, Aberystwyth for providing the results of his STATA programme that models housing equity using the BHPS data.

11 Due to complexities in the construction of the pension and housing equity variables we have been unable to include them in the current models. These variables will be used in future work.

12 While not reported here, average SAH deteriorates through time. This might suggest that respondents are not rating their own health 'compared to other people of your own age …'.

Information on labour market status in each wave is shown in Table 2, which also illustrates lost observations due to attrition. All 1245 people are in the labour force in wave 1 and this figure gradually decreases to 242 in the final wave. The number of people who classify themselves as retired increases from 108 in wave 2 to 545 in wave 11[13].

Survival analysis

As explained above, all individuals are in the labour market in wave 1 and in this first stage of the analysis retirement is considered the 'failure' outcome in a survival analysis. Figures 1a to 1f show actuarial life table survival curves for the 11 waves by various individual characteristics measured at wave 1. These provide simple descriptive assessments of the impact of the various variables on the probability of `survival' (non- retirement).

Figure 1a shows no clear relationship between retirement and self-assessed health measured on the 4 point scale. However, for men, health that limits daily activities (Figure 1b) does seem to increase the propensity for retirement, although there is no clear relationship for women. The presence of specific health problems appears to increase the probability of retirement for both men and women (1c). Figure 1d suggests that those in the lowest income quintile (INC1) appear to retire earlier than those in the other income quintiles. In Figure 1e we see that for both men and women those with only a state pension seem to retire earlier than those with either occupational or private pensions. In Figure 1f, there is a clear gradient in relation to housing tenure, in that those who have a mortgage remain in the labour market for longest, followed by those who rent and those who own their houses outright retire earliest.

This simple bivariate survival analysis can only give a partial picture of the relationship between retirement and the various factors considered here, so the next section discusses the results of the multivariate discrete time proportional hazards models.

Results for the discrete-time proportional hazard models are presented in Tables 3 to 5. Table 3 presents results for the full sample and Table 4 for men only[14]. Table 5 presents more detailed results for health problems broken down into its constituent component questions. Our discussion concentrates on the impact that the various specifications of the health variables have on the hazard of retirement. It worth noting however that the results, in general, show that the retirement hazard increases substantially after the state pension age (60 for women, 65 for men) is positive for individuals owning their house (although not significant) and negative for individuals with an outstanding mortgage on their property. This effect is significant at the 5% level for men but not for women. The estimated effect for individual-specific mean log household income is near zero for men and negative for women. However these effects are not significant.

In Tables 3 and 4, Model 1 reports results of specifying the impact of health in levels. In general the impact of the three self-assessed health variables (SAH-Ex, SAH-Good/VGood, and SAH-Fair) on the hazard of retirement are negative as expected (these are contrasted against a baseline of poor or very poor health) but are not significant. The impact of health limitations, health problems and disability are generally positive. Of these, health limitation is

---

13 As noted in footnote 5 retirement is not a permanent exit for some people so of the people classified as retired in each wave some may be working again in subsequent waves. However, in this analysis we treat first exit as permanent.

14 To conserve space the results for women are not reported but are commented on in the text. In addition estimated coefficients and associated standard errors are not reported for educational attainment, marital status and regional identifier variables.

significant (at less than the 1% level) for the full data and for men. The estimated effect of health limitations for women is much smaller than for men and is not significant.

The second model in each table specifies health effects as health shocks by conditioning on the corresponding first period measure of health. Again the only health effect that is significant is health limitations. The estimated impact of health limitations is larger for this specification compared to the model with health specified in levels form. For women, the effect of being disabled is large and positive and significant at the 5% level.

The third specification considers the impact of lagged health in the retirement decision. This model allows for an individuals' labour market status to adjust to changes in health and avoids any potential endogeneity between retirement and health. Again, we observe a large and positive effect for health limitations that is highly significant in all models including women. We also observe a gradient across the self-assessed health variables for men but the effects are not significant at the 5% level.

The final set of models specifies the impact of health as lagged health shocks by conditioning lagged health status on the set of initial period observations. These models exhibit negative estimated effects for self-assessed health and positive estimated effects for health limitations, disability and health problems (with the exception of disability for men). However, as with previous models only health limitations is significant.

Table 4 investigates in more detail health problems by specifying a set of dummy variables to capture the specific component parts of health problems. These results show that problems with sight, and other (not specified) health problems are significant in increasing the hazard of retirement.

5. Discussion

- Effects of health limitations on retirement hazard is large and positive and statistically significant for men and women.
- Counterintuitive results for income and pensions in survival analysis.
- No clear gradient for SAH in survival analysis.
- The effect of disability is small for men and large for women but needs to be interpreted with care due to small sample sizes.
- The effect of health problems combined on retirement hazard is positive but not statistically significant. Analysis of individual components reveals that `other' health problems is large, positive and significant for men, whilst difficulties in seeing (other than needing glasses to read normal size print) is positive and significant for women.
- Comparison of model specifications? Prefer health shocks, lagged health.

Future work

In future work we plan to extend the analysis presented here in three main ways. Firstly, by considering the effects of spouse's health, labour market status and pension entitlement on retirement. Secondly, by extending the model to consider other labour market transitions (like the move form FT to PT work) via a competing risks framework. Thirdly we will estimate frailty models allowing for individual heterogeneity in the hazard function. Fourthly we will investigate assumptions other than covariates acting proportionally on the baseline hazard. Finally, we will investigate a variety of methods of dealing with both sample selection and attrition in successive waves.

References

Anderson, Kathyrn H., Burkhauser, Richard V. "The Retirement-Health Nexus: A New Measure of an Old Puzzle." *Journal of Human Resources.*, 1985, *20* (3), 315-330.

Backlund, E., Sorlie, P. D., Johnson, N. J. "The shape of the relationship between income and mortality in the United States. Evidence from the National Longitudinal Mortality Study." *Ann Epidemiol*, January 1996, *6* (1), 12.

Backlund, E., Sorlie, P. D., Johnson, N. J. "A comparison of the relationships of education and income with mortality: the National Longitudinal Mortality Study." *Soc Sci Med*, November 1999, *49* (10), 1373-1384.

Bardasi, E., Jenkins, S. P., Rigg, J. A. "Retirement and the income of older people: a British perspective." *Ageing and society*, March 2002, *22* 131-159.

Benzeval, M., Judge, K. "Income and health: the time dimension." *Soc Sci Med*, May 2001, *52* (9), 1371-1390.

Blundell, R., Meghir, C., Smith, S. "Pension incentives and the pattern of early retirement." *Economic Journal*, March 2002, *112* 153-170.

Bound, John, Schoenbaum, Michael, Stinebrickner, Todd R., Waidmann, Timothy. "The dynamic effects of health on the labor force transitions of older workers." *Labour Economics*, June 1999, *6* (2), 179-202.

Butler, J. S., Anderson, Kathryn H., Burkhauser, Richard V. "Work and Health after Retirement: A Competing Risks Model with Semiparametric Unobserved Heterogeneity." *Review of Economics and Statistics.*, 1989, *71* (1), 46-53.

Chirikos, Thomas N., Nestel, Gilbert. "Economic determinants and consequences of self-reported work disability." *Journal of Health Economics*, August 1984, *3* (2), 117-136.

Contoyannis, P., Rice, N. "The impact of health on wages: evidence from the British Household Panel Survey." *Empirical Economics*, 2001, *26* (4), 599-622.

Disney R, Meghir, Costas, Whitehouse, Edward. "Retirement behaviour in Britain." *Fiscal Studies*, 1994, *15* (1), 24-43.

Disney R, Emmerson C, and Wakefield M. Ill-health and retirement in Britain: a panel data-based analysis. 2003. IFS Working Paper 03-02.

Disney R, Hawkes D. "Why has employment recently risen among older workers in Britain?," *The Labour Market Under New Labour*. 2003, 53-69.

Dwyer, D. S., Mitchell, O. S. "Health problems as determinants of retirement: are self-rated measures endogenous?" *Journal of Health Economics*, April 1999, *18* 173-193.

Ettner, Susan L. "New evidence on the relationship between income and health." *Journal of Health Economics*, 1996, *15* 67-85.

Haveman, R., Wolfe, B., Kreider, B., and Huang FM. Disability status as an unobservable: estimates frm a structural model. 1989. NBER Working Paper No 2831.

Hernandez-Quevedo C; Jones,Andrew M.; Rice,N. Reporting bias and heterogeneity in self-assessed health. Evidence from the British Household Panel Survey Health Economics Study Group, Glasgow 2004

Humphrey A, Costigan P, Pickering K, Stratford N, and Barnes M. Factors affecting the labour market participation of older workers . Research Report No 200. 2003. Department for Work and Pensions.

Jenkins, Stephen P. "Easy Estimation Methods for Discrete-Time Duration Models." *Oxford Bulletin of Economics and Statistics.*, 1995, *57* (1), 129-138.

Jenkins, S.P. Discrete time proportional hazards regression. STATA Technical Bulletin. STB-39, 1998.

Kerkhofs,M; Lindeboom,M; Theeuwes,J "Retirement, financial incentives and health". 1999 *Labour Economics* 6(2), 203-227.

Lancaster, Tony. *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press, 1990.

Lazear EP. "Retirement from the labour force," Ashenfelter O C, Layard R, *Handbook of Labour Economics Vol 1*. 1986.

Lumsdaine RL, Mitchell OS. "New developments in the economic analysis of retirement," Ashenfelter O C, Card D, *Handbook of Labor Economics Vol 3*. Elsevier, 1999, 3261-3307.

Madden D. "Labour market discrimination on the basis of health: an application to UK data." *Applied Economics*, 2004, *36* (5), 421-441.

Narendranathan,W Stewart,Mark B. How Does the Benefit Effect Vary as Unemployment Spells Lengthen? *Journal of Applied Econometrics* 1993, 8(4) 361-81

Prentice, R., Gloeckler, L. Regression analysis of grouped survival data with applications to breast cancer data. Biometrics 1978, 34: 57-67.

Riphahn, R. T. "Diasbility, retirement and unemployment: substitute pathways for labour force exit? An empirical test for the case of Germany." *Applied Economics*, 1997, *29* 551-561.

Riphahn, R. T. "Income and employment effects of health shocks. A test case for the German welfare state." *Journal of population economics,* August 1999, *12* 363-389.

Siddiqui S. "The impact of health on retirement behaviour: Empirical evidence from West Germany." *Health Economics*, 1997, *6* (4), 425-438.

Siddiqui S. "The pension incentive to retire: evidence for West Germany." *Journal of population economics*, 1997, *10* (4), 463-486.

Smith JP. "Health bodies and thick wallets: the dual relation between health and economic status." *Journal of Economic Perspectives*, 1999, *13* (2), 145-166.

Stern, S. "Measuring the effect of disability on labour force participation." *Journal of Human Resources.*, 1989, *24* 361-395.

Sundberg G. Essays on health economics. 1996. Dept of Economics, Uppsala University.

Walker I and Thompson A. Disability, wages and labour force participation: evidence from UK panel data. 1996. Dept of Economics, Keele University Working Paper 96/14.

Wu S. "The effects of health events on the economic status of married couples." *Journal of human resources*, 20043, *38* (1), 219-230.

Table 1 – Descriptive stats in wave 1.

		Men	**Women**
Age:	Mean (range)	57 (50-80)	56(50-76)
SAH	Very poor or poor	34	23
	Fair	115	93
	Good or very good	341	260
	Excellent	235	142
Health limits daily activities	Yes	64	42
	No	661	478
Presence of specific health problems	Yes	445	336
	No	280	183
Registered disabled person	Yes	19	5
	No	702	513
Housing tenure:	Owned outright	256	196
	Owned with mortgage	320	207
	Rented	149	117
Housing equity:	None (or negative)	158	126
	< £50,000	203	152
	>£50,000	364	242
Marital status	married or living as a couple	616	392
	widowed	14	51
	divorced or separated	47	53
	never married	48	24
Household Income	Individual-specific mean prior to retirement	9.66 (7.01 – 11.33)	9.64 (7.20 – 11.24)
Educational attainment	degree or higher degree	52	29
	HND or A level	114	62
	O level or CSE	143	110
	no qualifications	413	319
State pension age	Yes	62	113
	No	663	407
Pension entitlement	State pension only	293	288
	Occupational pension	186	97
	Private pension (post 1988)	46	26
	Private pension (pre 1988)	92	30

Table 2: Labour market status by wave

	1	2	3	4	5	6	7	8	9	10	11
Attrition *		139	189	245	280	307	324	340	369	390	420
self-employed	214	172	161	146	128	121	104	90	72	64	69
employed	923	705	592	494	429	377	335	293	243	207	168
unemployed	108	63	52	51	36	30	20	15	10	12	5
retired		108	180	225	285	322	390	430	481	521	545
LT sick, disabled		19	42	42	42	45	43	40	30	24	20
other		32	21	30	31	32	19	25	27	18	11
deaths		7	8	12	14	11	10	12	13	9	7
total	1245	1245	1245	1245	1245	1245	1245	1245	1245	1245	1245
in labour force**	1245	940	805	691	593	528	459	398	325	283	242

* includes those too old or infirm to participate in the survey. This is a very small number of people (17 in total).
** employed, self-employed and unemployed.

Appendix 1: Variable definitions

Variable	Description
Retire	Binary dependent variable, = 1 if respondent states they are retired, 0 otherwise.
SAH-Ex	Self-assessed health: 1 if excellent, 0 otherwise
SAH-Good/VGood	Self-assessed health: 1 if good or very good, 0 otherwise
SAH – Fair	Self-assessed health: 1 if fair, 0 otherwise
SAH – Poor/VPoor	Self-assessed health: 1 if poor or very poor, 0 otherwise (baseline category)
Health Limitations	Self-assessed health limitations: 1 if health limits daily activities, 0 otherwise
Health Disabilities	1 if respondent is registered disabled, 0 otherwise
Health Problems	Self-reported health problems: 1 if problem reported, 0 otherwise. There are also individual dummies for problems with: arms, legs or hands (arms), sight (see), hearing (hear), skin conditions or allergies (skin) chest/breathing (chest), heart/blood pressure (heart), stomach or digestion (stomach), diabetes (diabetes), anxiety or depression (anxiety), alcohol or drugs (alcohol), epilepsy (epilepsy), migraine (migraine) or Other (other).
Mean Log Hhinc	Individual-specific mean of log equivalised real household labour and non-labour income.
House Owned	1 if house owned outright, 0 otherwise
House Mortgage	1 if house has outstanding mortgage, 0 otherwise
House Rent	1 if house is rented, 0 otherwise (baseline cateory)
State Pension Age	1 if respondent is = > state pension age (60 years for women, 65 for men).
Widowed	1 if respondent is widowed, 0 otherwise
Divorced / Separated	1 if respondent is divorced or separated, 0 otherwise
Married / Couple	1 if respondent is married or living as a couple, 0 otherwise
Single	1 if respondent has never been married, 0 otherwise (baseline category).
Degree	1 if highest educational attainment is degree or higher degree, 0 otherwise.
HNDALev	1 if highest educational attainment is HND or A level, 0 otherwise.
OCSE	1 if highest educational attainment is O level or CSE, 0 otherwise.
Noqual	1 if no qualifications, 0 otherwise (baseline category).
Privpen	1 if respondent has made contributions to a private pension plan in last year, 0 otherwise.
Emplpen	1 if respondent is a member of an employers pension plan, 0 otherwise
Penstate	1 if respondent is reliant upon state pension, 0 otherwise (baseline category).
North West	1 if respondent resides in North West, Merseyside or Greater Manchester, 0 otherwise
North East	1 if respondent resides in North, South Yorkshire, West Yorkshire, North Yorkshire, Humberside or Tyne & Wear, 0 otherwise.
South East	1 if respondent resides in South East or East Anglia, 0 otherwise (baseline category)
South West	1 if respondent resides in South West, 0 otherwise
London	1 if respondent resides in Inner or Outer London, 0 otherwise
Midland	1 if respondent resides in East or West Midlands or West Midc , 0 otherwise
Scotland	1 if respodent resides in Scotland, 0 otherwise.
Wales	1 if respondents resides in Wales, 0 otherwise.

Table 3: Discrete-time proportional hazard models: Full sample

	Model 1 Health Levels		Model 2 Health Shocks		Model 3 Lag Health		Model 4 Lag Health Shocks	
	N = 6893		N = 6860		N = 5662		N = 5633	
	Coef	S.E.	Coef	S.E.	Coef	S.E.	Coef	S.E.
Men	.343	(.106)	.348	(.106)	.339	(.106)	.346	(.106)
Age	7.609	(1.593)	7.557	(1.593)	7.372	(1.617)	7.254	(1.619)
Age2	-11.042	(2.447)	-10.957	(2.447)	-10.677	(2.483)	-10.491	(2.486)
Age3	5.281	(1.245)	5.235	(1.245)	5.096	(1.263)	5.001	(1.265)
State Pension Age	1.355	(.138)	1.352	(.138)	1.320	(.137)	1.320	(.138)
Mean Log Hhinc	-.053	(.092)	-.041	(.092)	-.063	(.092)	-.051	(.092)
House Owned	.159	(.108)	.166	(.108)	.151	(.108)	.161	(.108)
House Mortgage	-.352	(.135)	-.338	(.135)	-.352	(.135)	-.342	(.135)
SAH – Ex	-.067	(.198)	-.083	(.209)				
SAH – Good/ VGood	-.227	(.181)	-.245	(.189)				
SAH – Fair	-.065	(.175)	-.097	(.180)				
Health Limitations	.450	(.126)	.478	(.133)				
Health Disabilities	.065	(.204)	.341	(.229)				
Health Problems	.106	(.102)	.107	(.110)				
Lag - SAH – Ex					-.159	(.203)	-.200	(.220)
Lag - SAH – Good/ VGood					-.173	(.185)	-.188	(.197)
Lag - SAH – Fair					-.064	(.179)	-.105	(.185)
Lag - Health Limitations					.589	(.128)	.660	(.139)
Lag - Health Disabilities					-.189	(.234)	.124	(.273)
Lag - Health Problems					.130	(.101)	.136	(.112)
SAH – Ex (0)			.039	(.250)			.093	(.254)
SAH – Good/ VGood (0)			.038	(.238)			.028	(.240)
SAH – Fair (0)			.141	(.234)			.147	(.233)
Health Limitations (0)			-.159	(.167)			-.259	(.170)
Health Disabilities (0)			-.662	(.360)			-.548	(.378)
Health Problems (0)			.001	(.100)			-.014	(.104)
Log-Likelihood	-1817.0		-1807.8		-1736.0		-1724.9	

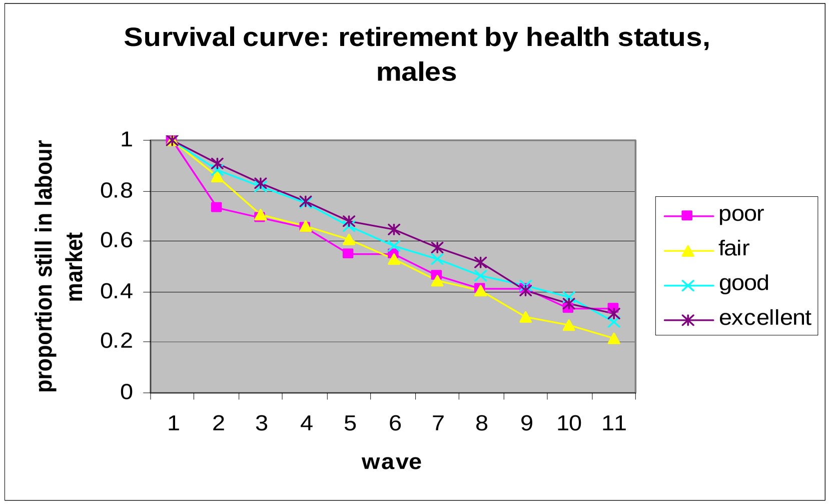
Table 4: Discrete-time proportional hazard models: Men

	Model 1 Health Levels		Model 2 Health Shocks		Model 3 Lag Health		Model 4 Lag Health Shocks	
	N = 3967		N = 3943		N = 3252		N = 3231	
	Coef	S.E.	Coef	S.E.	Coef	S.E.	Coef	S.E.
Age	12.097	(2.454)	12.131	(2.463)	12.129	(2.481)	12.010	(2.487)
Age2	-17.640	(3.743)	-17.687	(3.756)	-17.680	(3.783)	-17.499	(3.791)
Age3	8.467	(1.885)	8.491	(1.891)	8.485	(1.905)	8.396	(1.909)
State Pension Age	1.759	(.208)	1.750	(.209)	1.746	(.209)	1.729	(.210)
Mean Log Hhinc	.053	(.121)	.055	(.120)	.043	(.122)	.038	(.122)
House Owned	.171	(.152)	.186	(.153)	.160	(.152)	.168	(.152)
House Mortgage	-.387	(.187)	-.360	(.188)	-.381	(.188)	-.368	(.189)
SAH – Ex	.062	(.270)	.102	(.291)				
SAH – Good/ VGood	-.100	(.248)	-.065	(.265)				
SAH – Fair	-.034	(.237)	.022	(.252)				
Health Limitations	.640	(.166)	.688	(.183)				
Health Disabilities	-.046	(.257)	.143	(.299)				
Health Problems	.081	(.140)	.100	(.150)				
Lag - SAH – Ex					-.345	(.276)	-.351	(.306)
Lag - SAH – Good/ VGood					-.257	(.244)	-.208	(.277)
Lag - SAH – Fair					-.247	(.244)	-.294	(.264)
Lag - Health Limitations					.579	(.179)	.640	(.202)
Lag - Health Disabilities					-.330	(.299)	-.139	(.371)
Lag - Health Problems					.193	(.140)	.217	(.156)
SAH – Ex (0)			-.060	(.340)			.134	(.351)
SAH – Good/ VGood (0)			-.080	(.324)			-.010	(.337)
SAH – Fair (0)			.167	(.137)			.332	(.326)
Health Limitations (0)			-.149	(.233)			-.086	(.243)
Health Disabilities (0)			-.441	(.403)			-.296	(.433)
Health Problems (0)			-.052	(.137)			-.082	(.143)
Log-Likelihood	-972.3		-967.8		-914.6		-909.7	
Log-Likelihood	-817.3		-811.2		-793.0		-784.6	

Table 5: Discrete-time proportional hazard models: Health Problems

	Model 1 Health Problems All		Model 2 Health Problems Men		Model 3 Health Problems Women	
	N = 5662		N = 3252		N = 2410	
	Coef	S.E.	Coef	S.E.	Coef	S.E.
Men	.375	(.108)				
Age	7.558	(1.622)	12.202	(2.495)	13.378	(3.442)
Age2	-10.983	(2.491)	-17.817	(3.807)	-20.979	(5.419)
Age3	5.262	(1.267)	8.568	(1.918)	10.907	(2.834)
State Pension Age	1.345	(.138)	1.768	(.211)	1.343	(.227)
Mean Log Hhinc	-.036	(.093)	.039	(.124)	-.085	(.148)
House Owned	.145	(.108)	.162	(.153)	.126	(.164)
House Mortgage	-.376	(.136)	-.413	(.190)	-.340	(.207)
Lag - SAH – Ex	-.060	(.213)	-.265	(.294)	.122	(.323)
Lag - SAH – Good/ VGood	-.072	(.193)	-.167	(.267)	.021	(.296)
Lag - SAH – Fair	.019	(.183)	-.179	(.250)	.185	(.283)
Lag - Health Limitations	.563	(.132)	.537	(.184)	.626	(.199)
Lag - Health Disabilities	-.275	(.237)	-.440	(.309)	.208	(.398)
Lag - Health Problems - arms	.136	(.089)	.172	(.128)	.036	(.128)
Lag - Health Problems - see	.448	(.166)	.419	(.258)	.549	(.223)
Lag - Health Problems - hear	.029	(.121)	-.060	(.150)	.134	(.219)
Lag - Health Problems - skin	-.147	(.156)	-.047	(.222)	-.282	(.227)
Lag - Health Problems - chest	.065	(132)	.143	(186)	-.029	(.197)
Lag - Health Problems - heart	-.047	(.101)	-.040	(.142)	-.038	(.150)
Lag - Health Problems - stomach	-.050	(.166)	-.080	(.236)	-.164	(.246)
Lag - Health Problems - diabetes	-.354	(.259)	-.002	(.306)	-.668	(.524)
Lag - Health Problems - anxiety	.290	(.176)	.371	(.275)	.195	(.239)
Lag - Health Problems - alcohol	.335	(1.079)	1.104	(1.095)	-14.359	(1163.3)
Lag - Health Problems - epilepsy	-.150	(.521)	.712	(.737)	-.362	(.725)
Lag - Health Problems – migraine	.036	(.153)	-.337	(.317)	.182	(.178)
Lag – Health Problems - other	.492	(.160)	.538	(.221)	.414	(.246)
Log-Likelihood	-1724.6		-908.0		-785.7	

Figure 1a



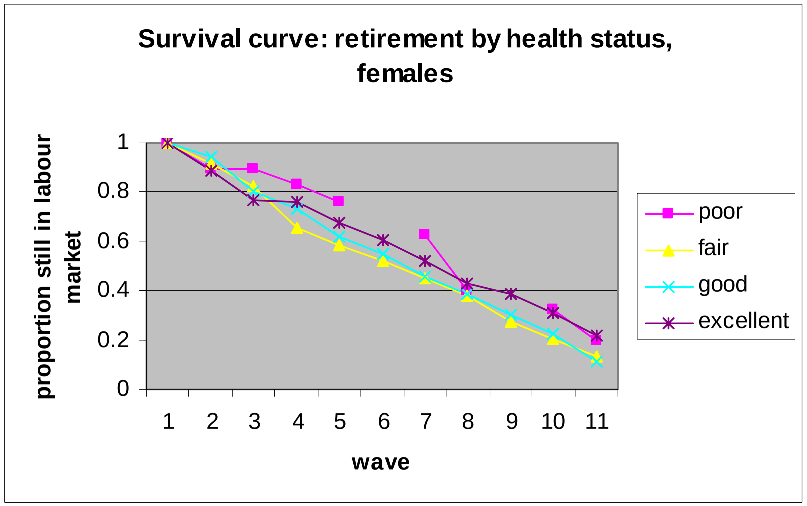
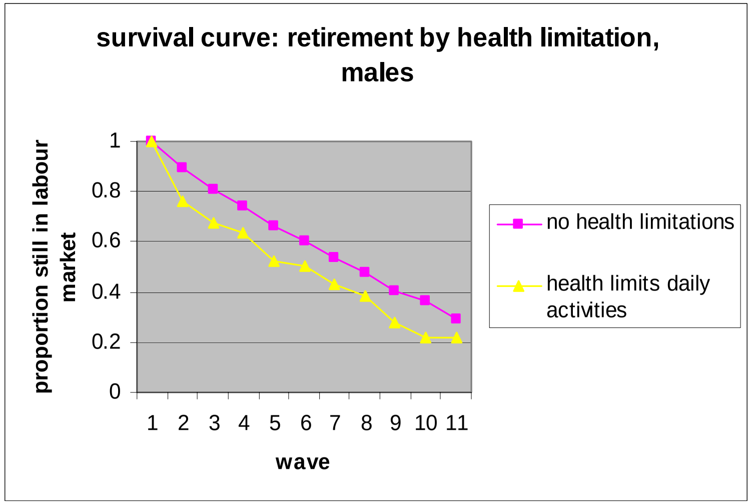


Figure 1b



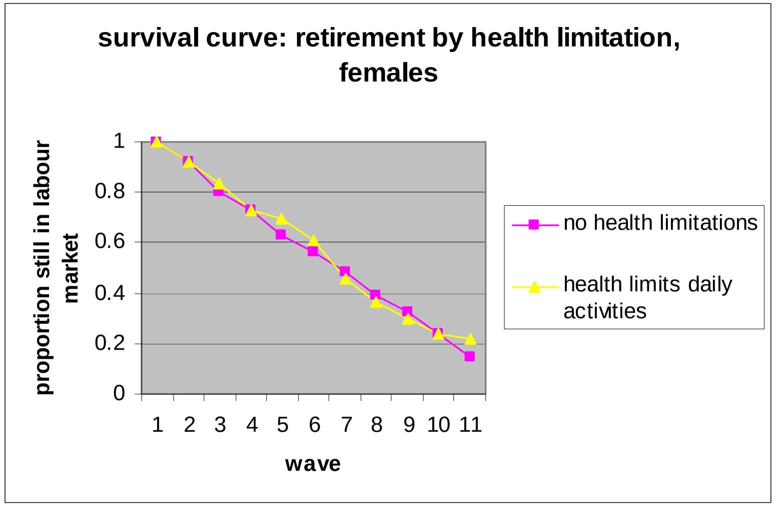


Figure 1c

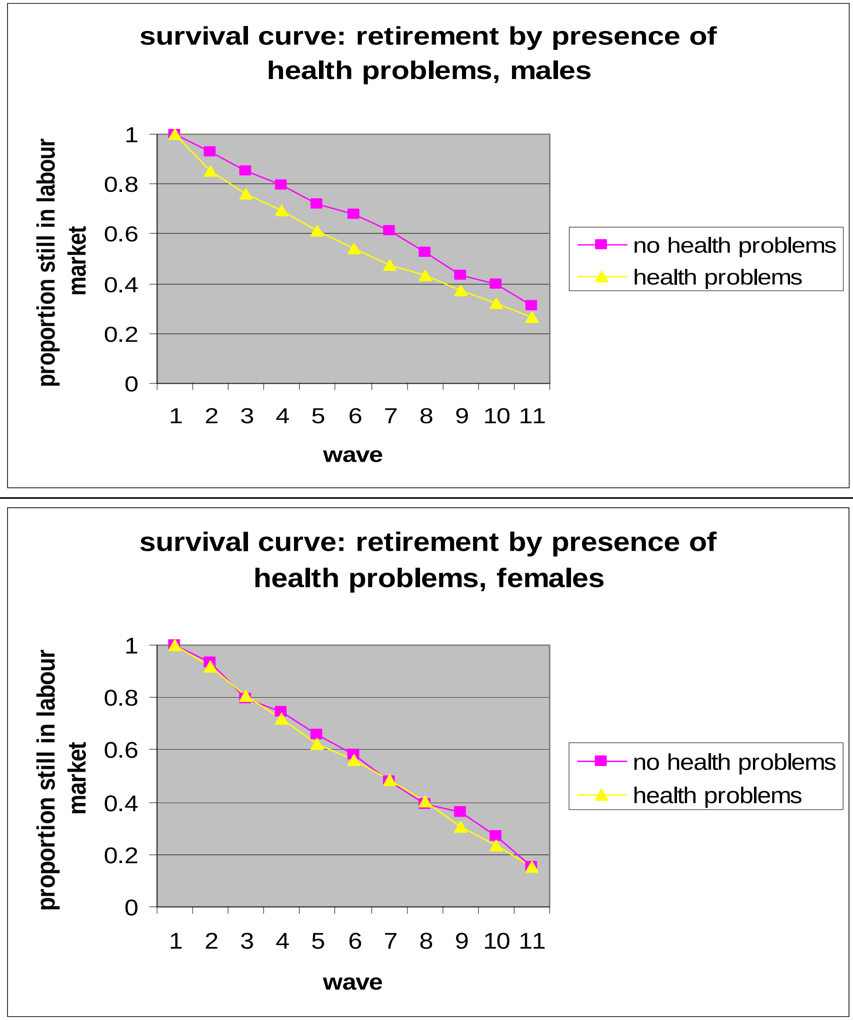


Figure 1d

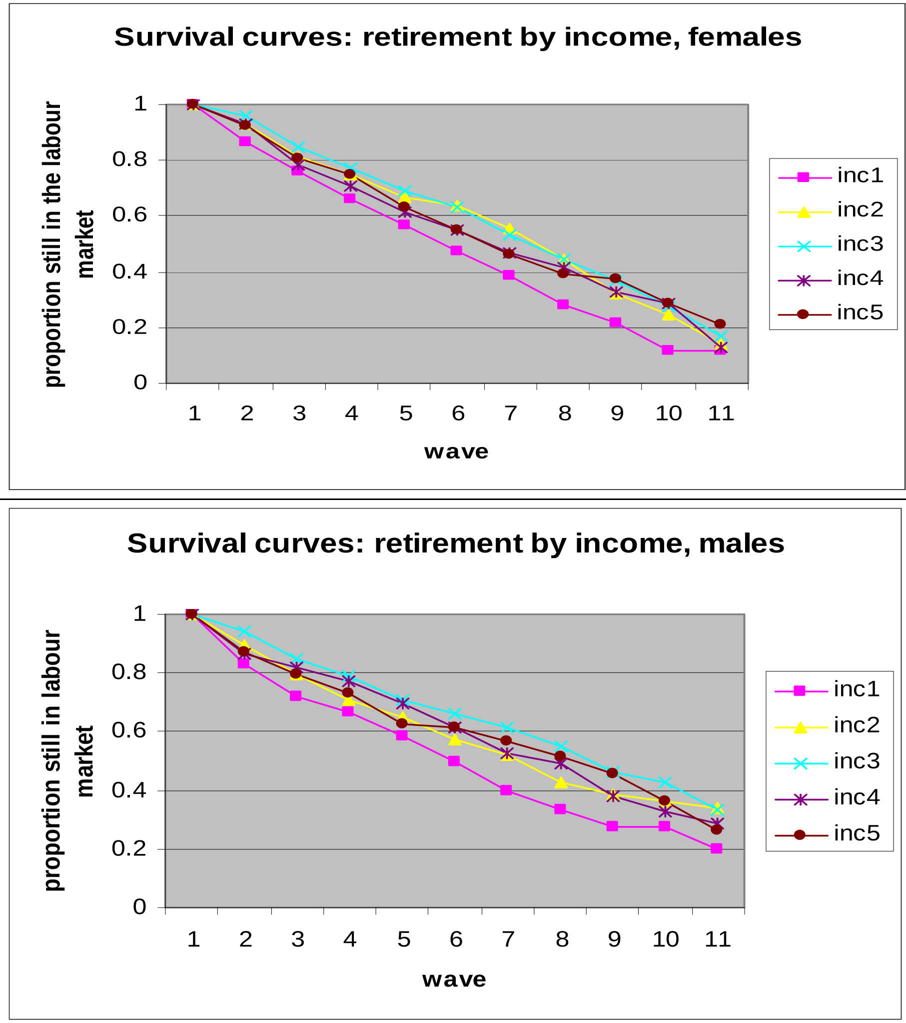


Figure 1e

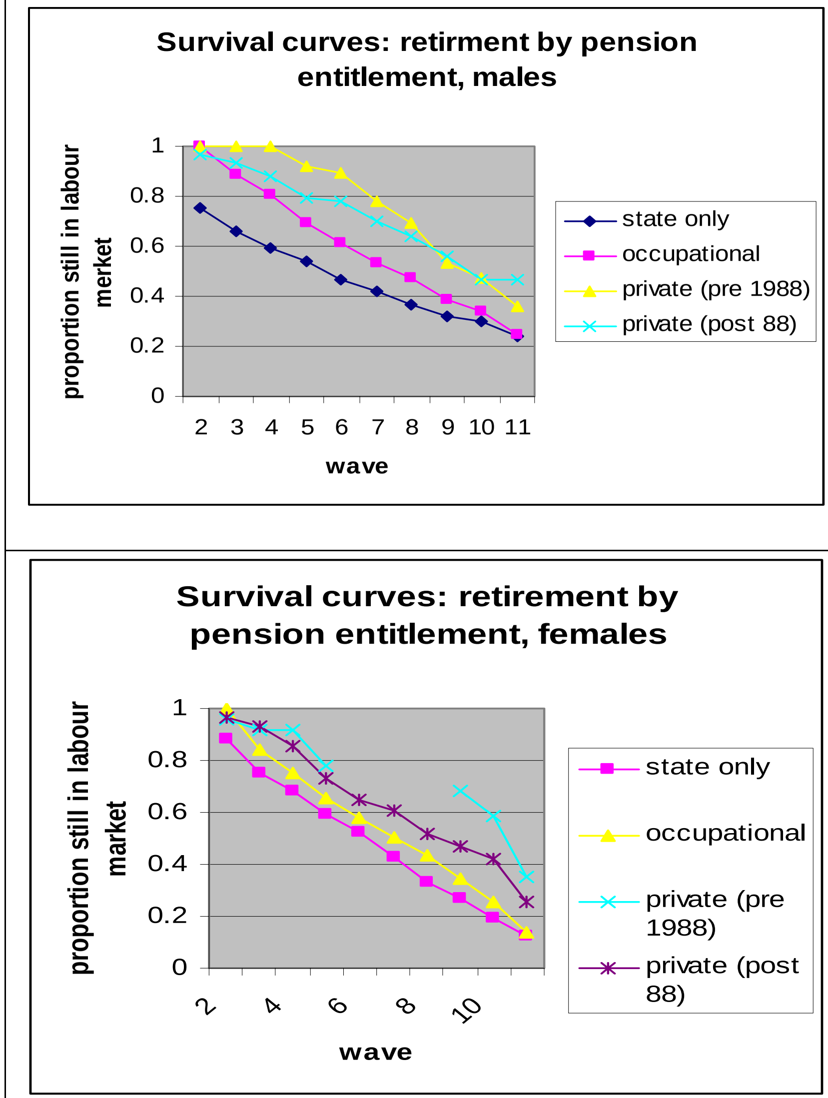


Figure 1f