Institute for Fiscal Studies

Modelling work, health, care and income in the older population

The IFS retirement simulator (RetSim)

IFS Working Paper W14/12

James Browne
Carl Emmerson
Katherine Heald
Andrew Hood

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Modelling Work, Health, Care and Income in the Older Population

The IFS retirement simulator (RetSim)

James Browne

Institute for Fiscal Studies

Carl Emmerson

Institute for Fiscal Studies

Katherine Heald

Institute for Fiscal Studies

Andrew Hood

Institute for Fiscal Studies

Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE

Published by

The Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE
Tel: +44 (0) 20-7291 4800
Fax: +44 (0) 20-7323 4780
Email: mailbox@ifs.org.uk
Website: http://www.ifs.org.uk

ISBN 978-1-909463-51-6

Contents

Preface	1
1. Introduction	2
2. Model Overview	4
Estimation and prediction	5
Simulation	8
Relationships and fertility	9
Ordering the modules	9
3. Mortality Model	11
Model specification	11
Demographics	12
Health and care	12
Income, wealth and work	13
Calibration	13
4. Health Model	16
Calculation of the health index	16
Model specification	17
Demographics	18
Health and care	18
Income, wealth and work	19
5. Care Models	21
Data cleaning and combining records	21
Splitting care receipt by type	22
Formal care	22
Splitting care provision intensity	23
Model specification: care receipt	24
Demographics	24
Health and care	25
Income, wealth and work	26
Model specification: care provision	27
Demographics	27
Health and care	28
Income, wealth and work	29

6. Labour Market Models 30
Earnings history 30
Earnings matching model 32
Age cut offs for participation 34
Model specifications 35
Transitions from no work 36
Transitions from part-time work 38
Transitions from full-time work 39
7. Disability Benefits 41
Benefit rules 42
Model overviews 43
Data cleaning 46
Modelling levels of claim 48
Disability living allowance 48
Attendance allowance 50
Incapacity benefits 50
Reform assumptions 50
Introducing personal independence payments 50
Introducing employment support allowance 53
Model forms and specifications 54
Incapacity benefits 55
Disability living allowance: younger adults 56
Disability living allowance: older adults 56
Attendance allowance 57
Levels of award 58
Disability living allowance 58
Attendance allowance 59
Carer's allowance 59
Calibration 60
8. Financial data 64
Property wealth 64
Mortgages 64
House price growth 65
Net property wealth 65
Income from property 65

Financial wealth 66
Income from financial assets 67
State pensions 68
Private pensions 68
Private pension income 69
9. Net Incomes and Poverty 71
Creating net incomes 71
Calculating average incomes and poverty rates 72
Appendix: Full Model Specifications 74
Mortality 74
Health 77
Care receipt 82
Care provision 89
Labour supply 95
Disability benefits 103
Glossary 110
References 111

Preface

The authors gratefully acknowledge funding from the Joseph Rowntree Foundation (project reference 1112004A), the IFS Retirement Saving Consortium, and the Economic and Social Research Council (ESRC) through the Centre for the Microeconomic Analysis of Public Policy at IFS (grant reference RES-544-28-5001). The IFS Retirement Saving Consortium comprises Age UK, Department for Work and Pensions, Financial Conduct Authority, HM Treasury, Institute and Faculty of Actuaries, Investment Management Association, Just Retirement and Money Advice Service.

The authors would like to thank James Banks, Daniel Chandler, Jonathan Cribb, Paul Johnson, Gemma Tetlow, and members of the IFS Retirement Saving Consortium for comments and advice. The authors are particularly grateful to Rowena Crawford and Zoe Oldfield for invaluable assistance with the construction of some of the data in the report. The views expressed and any remaining errors are the authors' alone.

The ELSA data are made available through the UK Data Archive (UKDA). ELSA was developed by a team of researchers based at the National Centre for Social Research, University College London, and the Institute for Fiscal Studies. The data were collected by the National Centre for Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments co-ordinated by the Office for National Statistics. The developers and funders of ELSA and the UKDA do not bear any responsibility for the analysis or interpretations presented here.

1. Introduction

The UK population is ageing, a process which brings with it a variety of concerns around the prospects for pensioner incomes and the appropriate design of taxpayer-funded support for pensioners. These concerns have led to substantial reforms in recent years. For example, the last government increased financial support for pensioners substantially and, following the recommendations of its Pensions Commission, legislated to introduce automatic enrolment into workplace-based pensions for most employees. In addition to introducing automatic enrolment, the current government has legislated to speed up the move to a single-tier state pension system, tasked the Dilnot Commission to review the funding of adult social care and, in its most recent Budget, announced a relaxation of the rules governing annuitisation of defined contribution pension schemes. The first change to the state pension age in over half a century is also under way, with the female state pension age rising gradually from age 60 since April 2010. These are all truly radical changes.

This changing structure of the population will have many economic implications, affecting the labour market, the demand for different goods and services (both publicly and privately provided), and the demand for, and provision of, informal care between family members. The comparatively rapid growth of the older population makes it increasingly important that public policies targeted at this group are well designed, both for those who benefit from these policies and for those who pay for them.

We cannot assume simply that the pensioner population a decade from now will look similar to today's population. There will not just be more pensioners but those retiring over the next few years will have experienced different economic conditions in their working lives, been subject to a different policy environment at different points in their lives, benefited from different technological and medical advances, and made different decisions about their savings than have today's pensioners.

This paper sets out the methodology, assumptions, and modelling specifications used to produce the outputs reported in Emmerson, Heald and Hood (2014), which aims to shed some light on how the demographic and financial circumstances of this group will change and to which this document is a supplement. The core demographic outputs are produced by RetSim, a dynamic microsimulation model that has been estimated from, and runs on, data from the English Longitudinal Study of Ageing (ELSA). The RetSim model takes individuals from ELSA who are aged 52 and over in 2010–11 and simulates them forwards in time to project outcomes for the population aged 65 and over in 2022–23. The net incomes of the simulated households are calculated using the IFS tax and benefit model, TAXBEN, a static microsimulation model of the tax and benefit system in each relevant year.

RetSim is a dynamic microsimulation model. This kind of model takes a group of people who are representative of the population to be modelled and 'ages' them, simulating their characteristics of interest in future time periods. We start with data from ELSA waves 1 to 5, and estimate the relationships between the outcomes we are interested in modelling, and other individual and family characteristics. Because ELSA is a longitudinal survey, interviewing the same people in a number of years, we can examine those relationships over time. We formalise these relationships in a set of regression equations, which for m the basis of the model. We then take ELSA wave 5 data and use what we've learned about these relationships, assuming that the associations we find in the data hold in the future, to predict the probability that each individual will have each outcome (e.g. a particular working status or health status) in two years' time. In some cases, namely the probability of dying and the probability of receiving disability benefits, we calibrate these probabilities to capture additional information to that which we see in ELSA. We then generate random numbers to give each individual a single state for each outcome, based on the probabilities we've calculated. After the first period we only have observed data on fixed characteristics like sex and education level, so we use the simulated outcomes as inputs into the models of future circumstances.

This report is structured as follows. Section 2 outlines the structure of the RetSim model, covering both the estimation and simulation stages of the modelling process and describing the data used. Sections 3 to 7 give details of the individual modules within RetSim, each of which is concerned with the simulation of a particular characteristic of interest such as health or working status. These sections include the technical details of the simulations as well as the model specifications and a summary of the marginal effects of the explanatory variables used in each case. Section 8 describes the modelling assumptions used to arrive at measures of gross income and net wealth in each period of the simulation. Section 9 gives an overview of the TAXBEN model which applies the relevant tax and benefit system to derive net income from gross income in each simulation period. The Appendix contains the full specification and full set of marginal effects for each regression model.

2. Model Overview

We examine the evolution of the circumstances of the population aged 65 and over in England from 2010–11 through to the early 2020s by simulating the future circumstances of the population which is aged 52 and over in 2010–11. Taking respondents to the fifth wave of ELSA in 2010–11 as our base population, we simulate changes in individuals' characteristics for each two-year[1] simulation period from 2010–11 to 2022–23.

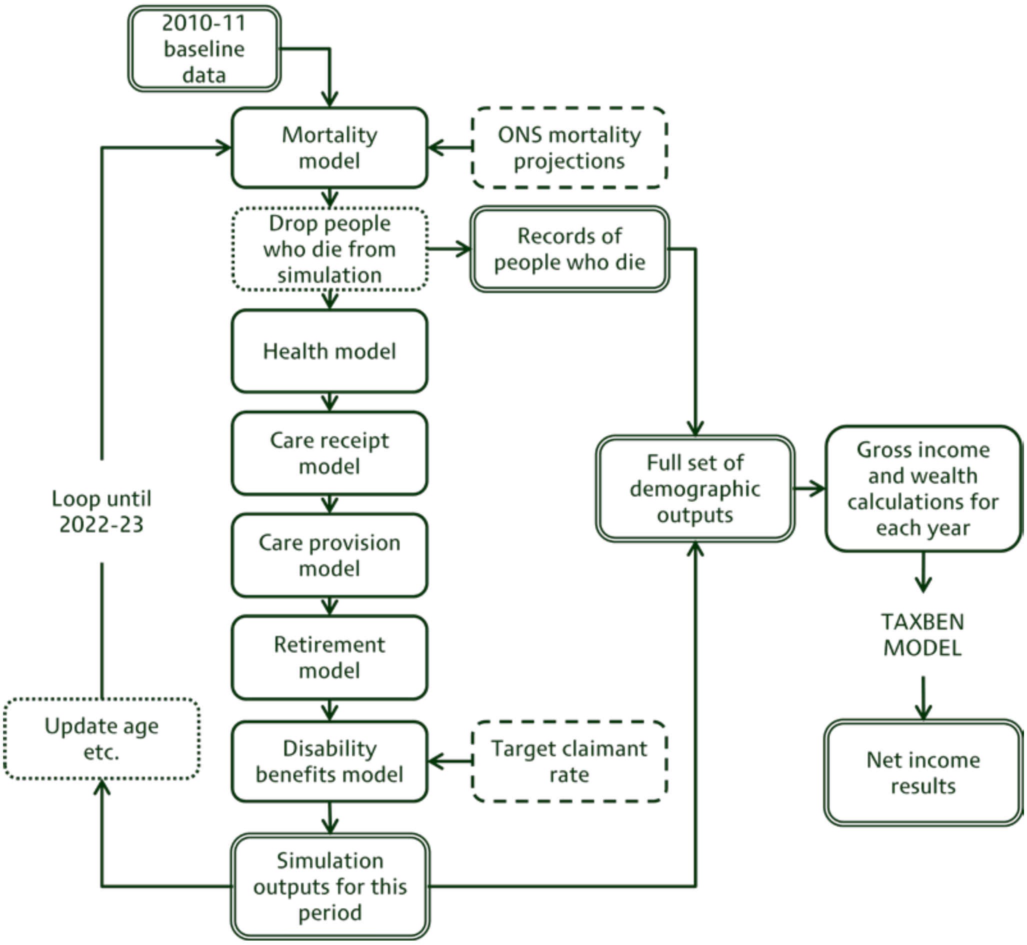
Although ELSA provides a huge amount of information on respondents, we do not attempt to capture every facet of the simulated individuals' circumstances in our model. Instead we focus on a few core characteristics which could reasonably be expected to have a significant impact on the economic situations of these people in the future. A significant by-product of this approach is that the detailed modelling of those core characteristics provides interesting outputs in its own right.

Our model uses data from the first five waves of ELSA to estimate mortality, and transitions between states of health, care receipt, care provision, paid work, and disability benefit receipt, conditional on a range of observed characteristics. The model comprises two main parts: the estimation stage involves predicting transition probabilities between these states over a two year period conditional on the individual's current circumstances, and the simulation stage combines these probabilities with randomisation to project the circumstances of the individuals observed in the 2010–11 data (the wave 5 cohort) forward to the early 2020s.

To simulate an individual from one period to the next we pass them through a series of modules, each modelling the evolution of one outcome. An overview of this structure is shown in Figure 1, with more detail given in sections 3 to 7. A 'simulation period' is a two year interval in the model, and the outcomes for one simulation period are simulated in one loop of the model shown in Figure 1.

[1] The length of the simulation period is dictated by the interval between observations in the ELSA data, which is a biennial survey.

Figure 1. Overview of the model



We allow the outcomes of the earlier modules to affect the outcomes of later modules within a simulation period. In its most obvious form that means that if a person dies in the mortality module then he is not passed through to the health module to obtain a health status. More generally, it means that we can allow people's health status at time $t + 2$ as well as their health at time t to affect whether (for example) they receive care or whether they work at time $t + 2$, but whether or not they are in paid work in period $t + 2$ cannot affect health until the next period ($t + 4$). We do not allow people's future status to affect their current status (e.g. $t + 2$ cannot affect t, and $t + 4$ cannot affect $t + 2$). The ordering of the modules is discussed in more detail below.

Estimation and prediction

ELSA, the English Longitudinal Study of Ageing, is a panel survey of the household population aged 50 and over in England. The survey covers 'core'

ELSA members – those in the sample eligible to complete the questionnaire – and their partners, and also contains information about the make-up of the households in which they live.

Because ELSA is a longitudinal dataset we can follow the same people over time and see how their circumstances change from one period to the next. We call any pair of observations (ELSA interviews, separated by two years) in which the same person is observed a 'transition'. We refer to the first period of the transition as 'time t' and the second period as 'time $t + 2$'.The basis of our microsimulation model is a set of regression equations which predict the evolution of the characteristics of interest: that is, what is the probability that an individual will be in each of a set of discrete states at time $t + 2$, given his circumstances at time t? These 'circumstances' are a range of characteristics which are fixed (e.g. sex), are assumed to be fixed (e.g. education level), are measured at time t (e.g. health or working status), or have already been predicted for time $t + 2$ (e.g. 'health at $t + 2$' is used in the care receipt regression).

We use five waves of data, taken at two-year intervals from 2002–03 to 2010–11, meaning that we can estimate the model on around 34,000 transitions and simulate the future circumstances of around 10,000 adults in around 7,000 households. We also have data on all deaths among those surveyed, and baseline data from the Health Survey for England for most of the respondents to first wave of ELSA, from which we take information on smoking status and socio-economic groups.

We use a series of regression models (probits, multinomial logits and probits,[2] and ordered probits, depending on the nature of the outcome of interest) to determine the probability of someone being in a particular state at time $t + 2$. In most cases the lagged value of this status – i.e. the status at time t – is included as an explanatory (right hand side) variable. One exception to this is in predicting working status, when we run separate regressions for each transition (equivalent to including lagged working status and its interaction with all other variables in the model).

The models are relatively parsimonious, because anything included must either be, or be assumed to be, unchanging (e.g. level of education, sex) or else be simulated into the future (e.g. health, working status). The detail of how we construct dependent (left hand side) variables for each outcome is discussed in

[2] We use multinomial probits in preference to logits, because they do not assume the independence of irrelevant alternatives (IIA) property, but in some cases we are not able to achieve convergence in the multinomial probit and instead use a multinomial logit for which convergence is more easily achieved.

the relevant model sections. The majority of the explanatory variables are drawn directly from the ELSA data and are self-explanatory (such as sex, age, or region). Most of these are fixed, and retain the same value throughout the simulation. The rest, such as age and time until SPA, are iterated deterministically in each simulation loop (i.e. in each circuit of the process in Figure 1). Additional explanatory variables that we create which require a little more explanation are described below.

Sex interacted versions of every explanatory variable (except sex itself) are created and included in the specifications. These are calculated by multiplying the explanatory variable by 0 for men and 1 for women, and thus allowing the explanatory variables to impact the dependent variable differently for men and women. This is equivalent to running all the regressions separately for men and women.

Baseline wealth quintiles are created within age bands and separately for couples and for singles. The 'wealth' in question is the total net wealth of the benefit unit at the first point at which the individual is observed in ELSA. These are 'baseline' quintiles because we do not allow them to vary throughout the simulation, implying a relatively fixed ordering of individuals by wealth throughout the simulation.

Baseline income quintiles are calculated as quintiles of the household's total equivalised net income, comprised of income from employment, self employment and benefits, at the first point the individual is observed in ELSA. Again, these are 'baseline' values in the same sense as are the wealth quintiles.

Self employment is indicated if the individual has ever reported being self employed in the waves of ELSA in which we observe him. We do not explicitly account for self employment in either the demographic simulation or the calculation of net income but we allow an observed history of self employment to impact both labour supply decisions, and the matching of earnings where data are missing.

Presence of a mortgage in the next period is defined on the basis of the outstanding mortgage term reported at baseline: we evolve this deterministically throughout the simulation, assuming that everyone takes the full mortgage term to pay off outstanding mortgages and doesn't take out an additional mortgage or extend an existing one.

IDAOPI (income deprivation affecting older people index) quintiles and **IMD** (indices of multiple deprivation) quintiles from 2004 are matched on to each household on the basis of geographical location (LSOA), with quintile 1 being the least deprived.

Simulation

Having estimated a model for a given characteristic (e.g. health status), we can create a predicted value for two years hence for each member of the wave 5 cohort. For the binary (probit) models, this will be a single probability p that the individual has a positive value for the dependent variable in the model. For the multinomial models this is instead a set of probabilities $p_1 \dots p_n$ corresponding to the n possible outcomes.

In the majority of cases, these probabilities are used directly. In the mortality and disability benefit modules, however, the probabilities are scaled to match external (ONS or DWP) data to match a projected real-world outcome (improvements in life expectancy) or to correct for under-reporting in the survey (disability benefit claimant rates). This process is discussed in more detail in the relevant sections below.

In each case, we generate a random number u drawn from a uniform distribution on the range 0 to 1, and compare this with the predicted probabilities.

In the binary case, individuals are given outcome 1 if $u \leq p$ and outcome 0 otherwise. In the multinomial case, individuals are given outcome j[3] such that $\sum_{i=0}^{j-1} p_i < u \leq \sum_{i=0}^{j} p_i$ [4]

Our simulation runs from 2010–11 to 2022–23. We take 2010–11 as the starting year because this is the most recent year for which we had ELSA data at the time of building the model. The sample in ELSA wave 5 is representative of the population aged 52 and over, and we are interested in predicting the future circumstances of those aged 65 and over: this dictates that the end year of the simulation be no later than 2022–23, being the final year in which we have a representative number of 65 year olds (as those aged 52 in 2010–11 are 64 in 2022–23 and 66 by 2024–25, the next year we could simulate). Alternatively the model could be used to look at, for example, the population aged 60 and over but at the cost of only being able to go forwards to 2018–19. Equivalently, for older individuals the model could be extended further forwards, for example the population aged 75 and over could be examined through to 2032–33.

---

[3] Where $p_0 = 0$

[4] For example, if the probabilities are $p_1 = 0.2$, $p_2 = 0.3$, and $p_3 = 0.5$, and $u = 0.6$, the individual is given outcome 3 because $u > p_1$, $u > p_1 + p_2$ and $u < p_1 + p_2 + p_3$. This ensures that the probability than we give person outcome i based on the random number we draw is equal to the probability of him having outcome i that we predict from the model.

Relationships and fertility

Our model simulates people from age 52 onwards, and the outputs focus on the population aged 65 and over. While marriage, divorce, and the arrival of new children are all possible within this group, they are relatively unimportant in comparison with a younger age group. For example, ONS data show that divorce rates in couples where the husband is aged 60 and over, despite recent increases, were still only 2.2 per 1,000 married people in 2010 compared to a population average of 11.0 and a peak of 22.5 for 30 to 34 year olds (Office for National Statistics, 2012). For this reason, we do not add complexity to our model by attempting to model fertility or relationship formation or dissolution (other than as a result of the death of one partner).[5]

There are certain circumstances where it would be helpful to have information about the arrival or ageing of a grandchild, which might trigger a change in the individual's caring responsibilities. However, we do not have the data to simulate this and instead use the existence of the respondent's children at baseline (which we assume to remain fixed – i.e. there are no further births and we do not allow for the death of existing children) to proxy for the possibility of grandchildren existing or being born.

Ordering the modules

The modules within the model are ordered as in the sections of this report: first mortality, then health, care receipt, care provision, working status, and finally disability benefit receipt.

We require some element of ordering within the structure because jointly estimating the entire model would be overwhelmingly complex. In addition, it is actively useful to be able to impose a within-period causal ordering in certain cases, such as between health and care receipt, and between working status and disability benefit receipt.

Most of this ordering is fairly uncontroversial: policy dictates that disability benefit receipt is dependent on working status, for example, and common sense dictates that whether someone is alive affects whether they have a health or

[5] Adding a model for relationship dissolution would be quite feasible. But without a model for partnership formation this would lead to the model underestimating the number of couples and overestimating the number of singles. Adding a partnership formation model would be more tricky since, in addition to deciding who re-partners, the model would also have to find an appropriate match for them.

working status. Certain aspects are a little more subtle, though: in particular, we assume that in any given period the decision to provide care is made before, and affects, the labour supply decision rather than vice versa. To the extent that this is a non-trivial decision, though, one should bear in mind that we allow lagged working status and care provision to affect both care provision and working status in the next period, which will reduce the influence of the module ordering.

The evolution of wealth and gross income, as discussed in section 8, takes place alongside the demographic simulation (rather than interacting with it), and the derivation of net income from gross income is a final step performed on each year's demographic and financial output to achieve the final outputs. We do, however, include baseline income and wealth quintiles in the model specifications, so a measure of financial circumstances is accounted for to the extent that individuals do not move between quintiles over the course of the simulation.

3. Mortality Model

The mortality module determines whether a simulated individual dies by time $t + 2$ subject to being alive at time t. Because this is the first module run in a simulation period, dying at $t + 2$ means that the individual is counted in summary outputs for other characteristics (health, care, etc.) at time t, but is not present in any such outputs at $t + 2$.

We have information on the deaths of ELSA sample members, taken from death records, which allow us to regress 'death in the next two years' on a range of time t characteristics. This model drives the distribution of death probabilities within age, sex and birth year cells, but we scale the mean probability of death within these cells to match the projections from the ONS life tables.[6] This is for two main reasons. First, the death rate amongst the ELSA sample is lower than in the population as a whole, and so in calibrating our model to the ONS probabilities we are adjusting for sampling error (and sample selection). Second, and more importantly, increasing life expectancy is likely to be a key driver of demographic trends over the next decade, and it is not possible to capture this using the ELSA data alone.

Model specification

The mortality model is a probit, with the dependent variable being a binary indicator for death within two years. The specification contains information on age, sex, couple status, the diagnosis before age 50 of a range of illnesses, education, care provision and receipt, health, home ownership, deprivation, region, childhood health, socio-economic group, disability benefit receipt, smoking, working status and baseline income and wealth.

The variables are interacted by sex, and sex itself is also included. The pseudo-R^2 statistic for this model is 27.6%.

The full specification, and the full set of marginal effects for the model, are shown in the Appendix to this paper, with the more significant or interesting results discussed below. The marginal effect of a binary explanatory variable is the percentage point change in the dependent probability that is observed when its

[6] More detail on the projected improvements in mortality is set out in section 3 of the results paper (Emmerson, Heald, & Hood, 2014).

value is 1 rather than 0. The differential impacts of variables for men and women are the result of a specification which interacts all explanatory variables with sex.

Demographics

The marginal effect of sex is statistically significant, with women being 4ppts less likely to die than men, holding all else constant.

The probability of death in a given period increases in age, with the effect becoming highly significant from the age of 75. If anything, the effect of age is more pronounced for women than for men, although the difference between the effects is not significant[7].

Being in a couple has a statistically significant effect, with men being 1.2ppts less likely to die if they are in a couple than if they are single. The effect is less pronounced for women, who are 0.5ppts less likely to die in a given period than are their single counterparts.

We observe no consistent or significant influence of education level or of deprivation indices on the outcome of the mortality model.

Health and care

Probability of death in a given period increases with worsening health, particularly for men, with the effect of the three poorest (of five possible) levels of health being statistically significant.

Receipt of care is highly significant, with receipt of informal care indicating a 2ppt increase in chance of death in the next two years for men, and formal care an increase of 3ppts, with effects of a similar magnitude for women. We see provision of care indicating a statistically significant reduction in chance of death in a given period of about 1ppt.[8]

While the receipt of disability benefits is not statistically significant, receipt of disability living allowance (DLA) is linked with an increased probability of death[9]

---

[7] An insignificant difference means the p value on the sex interacted version of the variable is insufficient to merit a significance star.

[8] Note that, similarly to all the marginal effects presented in this paper, this relationship is not necessarily causal (i.e. giving care doesn't cause you to live longer). Instead, it is more likely that a third factor, which is not picked up elsewhere in the model specification, is correlated with both increased care provision and increased longevity and is showing up in the marginal effect of care provision.

[9] Again, this relationship does not imply that being awarded a disability benefit directly increases the chance of death. Instead, the status of the individual as a disability benefit claimant is telling

in a given period for both sexes, receipt of incapacity benefit is linked to a reduced chance of death for men. Receipt of attendance allowance (AA) is linked to an increased probability of death for women.

Smoker status is highly significant, despite being recorded in 1998, with those smoking at that point (who may or may not be smoking at t) being 2ppts more likely to die in the next two years than non-smokers. While statistically insignificant, having been an ex-smoker in 1998 also confers an increased chance of death in the next two years.

Income, wealth and work

The effect of working status is highly statistically significant, with men in work being 2ppts less likely to die within two years than their non-working counterparts. Women in part-time work are 0.5ppts less likely, but those in full-time work are 0.4ppts more likely, to die in a given period than are female non-workers.

We observe no consistent or significant influence of baseline income or wealth on the outcome of the mortality model, suggesting that the correlation between mortality and wealth that we observe in the results of the simulation is explained by other factors correlated with both mortality and wealth.

Calibration

The probit model described above provides a probability of death for each simulated individual. We take the mean probability within each age, sex and birth year cell and compare it with the equivalent figure from the ONS life tables.[10] We then scale each individual probability from the model by the ratio of the mean and the ONS value. This allows us to bring the mean probabilities into line with the ONS projections while maintaining the variation in probabilities as determined by the regression.

---

us something about his health status – and thus his chance of death - that is not picked up in the health or care receipt indicators.

[10] ONS life tables give the probability of death in the next year, given age, sex and birth year, whereas we need the probability of death over the next two years. This is calculated as

$$P(Death\ by\ t+2) = P(Death\ by\ t+1) + P(Death\ by\ t+2 | No\ death\ by\ t+1)$$

where we can also take the latter probability directly from the life table (for the following year), given the person's birth year and sex, on the basis that the figures in the life table are already conditional on having survived to that year.

For example:

1. Person A has a probability of death, estimated from the regression model, of 0.02.

2. People of person A's age, sex and birth year have a mean estimated probability of death from the regression model of 0.03 (so person A is less likely to die than others of his age, sex and birth year – we get this variation within the group by including multiple explanatory variables in the regression, rather than just assigning probabilities directly from ONS life tables).

3. The ONS probability of death over the two years in question, for people of this sex and birth cohort, is 0.06 (due to a combination of people in ELSA being less likely to die than the ONS thinks people in the population are as a whole, and projected life expectancies having improved between the ELSA data collection and the simulation year).

4. We divide the ONS probability by the average ELSA probability to get a factor of $0.06/0.03 = 2$.

5. We multiply person A's probability of death by this factor, giving him a death probability of $0.02 * 2 = 0.04$.

6. We repeat this process for everyone else in person A's age, sex and birth year cell. This means that the average probability in that cell adjusts to match the ONS probability of 0.06, but person A's probability of death is still lower than average for his age and sex.

7. We repeat this process for everyone in every cell.

8. We then generate a uniform random number for person A (and for everyone else in the model), as described in the simulation section above, and use this to determine whether we simulate his death in this period or not.

This process creates an almost perfect match between the mean (adjusted) modelled probability and the ONS probability in each cell, and any variation from the ONS projections seen in the aggregate outputs is then due to the randomisation process by which each individual's probability is resolved into a single state. This means that the particular value from our mortality model is what it can tell us about the distribution of probabilities within age and sex cells, and the implications of that variation. These results are presented in Emmerson, Heald and Hood (2014).

Note that a key assumption underlying this methodology is that the improvements in mortality predicted in the ONS life tables are evenly distributed amongst individuals in an age / sex / birth year cell: that is, that improvements in longevity (defined as the proportional reduction in the probability of death) are not concentrated within a single socio-economic group, region, or wealth quintile, for example.

4. Health Model

The health module simulates a person's health at time $t + 2$. For simplicity's sake we represent health in each period as being at one level of a discrete variable, the construction of which is described below. This dependent variable strikes a balance between capturing the complex nature of health, which is a multidimensional concept including physical and mental wellbeing, and being a variable which we can robustly simulate throughout the model.

We select variables to construct this index which are, as far as possible, objective and which can also be measured reliably across all ages and compared between individuals. This rules out, for example, self-reported quality of health, a measure which is overwhelmingly influenced by the reference point of the individual.

Calculation of the health index

ELSA contains a wealth of information on both subjective and objective measures of health. We use some of these as explanatory variables in the regressions, such as leg length (as a proxy for childhood health and nutrition), self-reported childhood health, smoker status, and the diagnosis by age 50 of a range of conditions. However, our key health indicator is the five-level health status variable that we construct from raw ELSA data.[11]

The five-level health status variable is the condensed version of a health index derived as the sum of a number of binary flags, each indicating a particular health problem, as follows:

- Mobility problems, one flag for each of:
  - difficulty walking 100 yards
  - difficulty sitting for two hours
  - difficulty getting up from a chair
  - difficulty climbing one flight of stairs
  - difficulty reaching above shoulder level
  - difficulty lifting more than 10 lb

[11] This index has been developed by James Banks, Richard Blundell and James Browne for the purposes of predicting eligibility for disability benefits as part of an ongoing IFS project.

    - difficulty picking up a 5p coin from a table
- Eyesight (including when assisted by glasses or similar), one flag for each of:
    - difficulty recognising a friend across the street (distance)
    - difficulty reading newspaper print (close)
- Fair or poor hearing (including when assisted by hearing aid or similar)
- Urinary incontinence in the last 12 months
- Left previous job due to stress
- Score of at least 4 on clinical depression scale questions

People given a health status of 0 (best health) have no positive flags, health status 1 (good health) indicates one positive flag, 2 (OK health) indicates two or three flags, 3 (poor health) indicates four or five flags, and 4 (worst health) indicates six or more flags.

The five-level health status variable is used as an explanatory variable throughout the model, and is the dependent variable in the multinomial regression described in this section.

Model specification

This model is an ordered probit, a multinomial model which acknowledges an underlying ordering in the outcome states. This specification is appropriate because of the structure of the health index which underpins the discrete health statuses.

The specification contains information on age, sex, couple status, the diagnosis before age 50 of a range of illnesses, education, care provision and receipt, health, home ownership, deprivation, region, childhood health, leg length[12], socio-economic group, disability benefit receipt, smoking, working status, and baseline income and wealth.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo-R^2 statistic for the model is 26.6%.

[12] Leg-length is thought to be correlated with nutrition in childhood and hence later-life health.

In our description of this multinomial model we talk about the marginal effects of explanatory variables on the probability of being in the best or worst health group. These are the two extreme outcomes, with explanatory variables having similar effects on the probability of being in the intermediate categories.

Demographics

Sex has a large marginal effect but is not statistically significant. Women are 13ppts less likely than men to be in the best health group and 4ppts more likely to be in the worst health group.

The effect of age on health status is statistically significant from the age of 65 and highly significant from age 70, with men aged 90 and over being 23ppts less likely to be in the best health category, and 6ppts more likely to be in the worst health category, than men under 55. The same trend is observed, though somewhat less strongly, for women, with those aged 90 and over being 16ppts less likely to be in the best health group and 4ppts more likely to be in the worst health category than women under 55.

Those in couples, and particularly women in couples, are likely to be healthier than single people. The effect of being in a couple on being in the best health group is a 2ppt increase for men and 3ppt increase for women. There is an accompanying 1ppt decrease in the probability of being in the worst health group for both men and women in couples compared to their single counterparts.

The effect of higher levels of formal education on the probability of being in the best health group is positive, though not statistically significant: men with a degree are 2ppts more likely to be in the best health group than those with no qualifications. The effect for women is weaker but still positive.

We observe no consistent or significant influence of deprivation indices on the outcome of the health model.

Health and care

Lagged health has, unsurprisingly, a highly significant effect on future health. Men in the worst health category at time t are 61ppts, and women 63ppts, less likely to be in the best health group in the next period than those in the best health group at t. Even being in 'good' as opposed to 'best' health – recall that this is scoring 1 rather than 0 in the health index – makes both men and women 17ppts less likely to be in the best health group at $t + 2$. Similarly, those in the worst health group at time t are 17ppts more likely to remain in that state at $t + 2$ than are those in the best health group to move into it, and those in the poor health group are 13ppts more likely to move into it than are those in the best health.

We see early (by age 50) diagnosis of diabetes and arthritis having a statistically significant negative impact on the likelihood of being in the best health group, and a significant positive impact on the chance of being in the worst health group.

The significant impact of childhood health on future health status, with poor or worst childhood health conferring a 5 to 6ppt decrease in the chance of men being in the best health group, provides additional evidence for the persistence of health status. There is a similarly sized negative impact for women, and a significant positive impact of around 1 to 2ppts for both sexes on the chance of being in the worst health group.

Both lagged provision and lagged receipt of care are significant in predicting future health, with both having a negative effect on the probability of being in the best health group. Providing care reduces the probability by 2ppts for men and 1ppt for women, and receiving care indicates a reduction in the probability of up to[13] 12ppts. Similarly, providing care increases the chance of being in the worst health group by up to 1ppt and receiving care indicates an increase in the probability of worst health of up to 3ppts: recall that this is over and above the effect of lagged health.

Having been either a current or a regular smoker in 1998 has a significant effect, increasing the probability of worst health by up to 1ppt. Additionally, lagged receipt of disability benefits indicates a significantly increased probability of worst health with the biggest effect, an increase of 4ppts, being observed for men in receipt of DLA.[14]

Income, wealth and work

Lagged working status has a positive effect on the probability of being in the best health status, conferring an increase of around 2ppts. There is a smaller (around 1ppt) negative impact on the chance of being in the worst health group.

The effect on health of being in any baseline wealth quintile other than the lowest is positive, and being in the highest wealth quintile has a significant effect: an

[13] 'Up to' because of different effects by sex and from receiving formal or informal care. Full details are given in the Appendix. 12ppts relates to a man receiving formal care. A woman receiving informal care sees a reduction of 6ppt.

[14] Again, the effect of IB receipt on men is to lower their probability of being in the worst health: this is similar to the result we saw for mortality. In the absence of a state pension age dummy in these regression specifications it is plausible that the receipt of IB is acting in this role to some extent, and that different state pension ages for men and women are making the interpretation of this marginal effect somewhat ambiguous.

increase in the probability of best health of around 4ppts for both sexes and a decrease in the probability of worst health of around 1ppt. The effect of baseline income quintiles is similar, though not significant and of a smaller magnitude.

5. Care Models

In this section we consider two models: one for receipt of care and one for provision of care. As discussed below we define care relatively loosely, being assistance with day-to-day tasks with which the recipient has difficulty. We then split care receipt into formal (provided by professional staff) and informal (provided by friends or family who may or may not be part of the household and who may or may not receive carer's allowance) and split care provision into low and high intensity as discussed below.

Both single people and people in couples are covered by the same model, but the ELSA data show that by far the majority of carers are in couples, and that much of this activity is people caring for their partner. This means that our most interesting results and interactions below are in the context of partners caring for each other, but the model allows for all forms of care (i.e. to people within and outside of the household, and to older and younger people, including caring for grandchildren) by anyone in the model.

We place the care receipt model before the care provision model, meaning that the within-period decision about care provision at $t + 2$ is influenced by the individual's own care receipt at $t + 2$. This decision is made on the basis that care receipt will be strongly correlated with care need, over which the individual has minimal control, which in turn affects the ability, if not the inclination, to provide care.[15]

Data cleaning and combining records

Our aim at this stage is to determine from the raw ELSA data which people give and receive care. We combine responses on all caring roles (e.g. caring for a partner, a parent, a grandchild or a friend) as the first stage in determining who provides care, and responses on all types of care receipt (i.e. informal care, from relatives or others, and formal care, from professional staff) in determining who receives care.

We then compare these two flags: if a person says that they receive care from their partner, we ensure that the partner has a positive care provision indicator.

---

[15] Note that we do not use this opportunity to use partner's receipt of care as an explanatory variable in the care provision regression, because of concerns about endogeneity. Instead, we use partner's lagged care receipt and partner's $t + 2$ health status.

Similarly, if a person reports giving care to their partner we ensure that the partner's care receipt indicator is positive. The former of these two cases is significantly more common than the latter, and accounts for around 40% of all people ultimately flagged as providing care. This is likely to be due to the way in which the relevant questions are asked: care providers are asked about 'active provision of care', whereas care recipients are asked whether they have difficulty completing any of a list of everyday tasks and then who, if anyone, helps them to complete them. It is quite possible that the care provision question is interpreted as a stricter test than the care receipt question. There may also be an effect of traditional gender roles within couples here, particularly given the birth cohorts that we are modelling, with what we define as care not being recognised as such by some of the wives providing it or some of the husbands receiving it. This in turn could potentially lead to under-reporting of care receipt (especially among men) and under-reporting of care provision (especially among women) despite our efforts to correct for it.

Splitting care receipt by type

Having determined who receives care we use the information on who helps those people with the everyday activities with which they struggle in order to classify this as formal or informal care. We model these two types of care as separate outcomes in a multinomial model.

Formal care

Our definition of formal care is any care provided by professional staff, for example from carers provided by the local authority or a charity, or paid for privately.

ELSA is unique in the UK in surveying people not only in private households but also in residential care homes if they have moved into them since starting the survey. Note that, although ELSA contains data on individuals who move into care homes, we do not include these individuals in our model.[16] This leads to some inconsistencies in the modelling of their partners in that those partners are treated in the model as single households. However, there are only 66

[16] There are a number of arguments against including these individuals in the simulation. The most compelling is that we have minimal information on their finances, and their incomes are in no way comparable to those of the rest of the population. To make an adjustment to account for this in our net income results would be disproportionately complex and time consuming. In addition, we are aiming to model the household population in England, of which these individuals no longer form a part. ELSA gives zero weights to these individuals.

institutional residents in the wave 5 data, of whom only 25 are in a couple, meaning that this is relatively immaterial.[17]

There is also a possibility that, by excluding from our model the possibility that an individual leaves the private household population other than by death, we could be overstating the number of recipients of formal care in the simulation (i.e. people who would, in reality, move into a care home are being reported as being in receipt of formal care in the household population), although the results of the simulation do not suggest that this is an issue.

Our model only allows people to receive one of formal and informal care in a given period, with formal care taking precedence if, in the data, the individual reports receiving both.

Splitting care provision intensity

We model care provision as an ordered probit, which allows us to differentiate between high (35 or more hours per week of care) and low intensity (up to 35 hours per week of care) provision. Caring for a few hours per week rather than a substantial number of hours per week is likely to affect labour supply decisions in a different way, and by splitting the care provision into high and low intensity we can capture these different effects. In addition, by modelling care provision of up to 35 hours per week and of 35 hours or more per week separately, we model one of the criteria used to determine eligibility for carer's allowance, something used in our benefits modelling later in the model.

As described above, we infer care provision for a number of individuals on the basis of their partner's questionnaire. We assume that all of these cases are instances of low intensity care, on the assumption that individuals providing 35 hours or more of care per week would be likely to view themselves as having a caring role, and to have reported this (and the hours spent caring) in their own interviews.

We interact the intensity of lagged care provision with sex and couple status (i.e. we allow for the impact of lagged care provision on current care provision to vary by whether the individual is a single man, a single women, a man in a couple or a

[17] It is also an unfortunate reality that the majority of care home residents at time t are unlikely to be alive at $t + 2$, so any inconsistency is to some extent self-limiting. BUPA data suggest a survival probability of just 55% in the first year following admission to a care home, with 50% mortality by 462 days (Forder & Fernandez, 2011).

women in a couple) when using it as an explanatory variable in the care provision model.

Model specification: care receipt

This model is a multinomial probit, a specification which does not imply an ordering in the multiple outcome states.[18]

The specification contains information on age, sex, couple status, the diagnosis before age 50 of a range of illnesses[19], education, care provision and receipt, presence of children at baseline[20], health, home ownership, deprivation, region, socio-economic group, disability benefit receipt, working status and baseline income and wealth.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix.

Demographics

The effect of sex on the probability of receiving care is not statistically significant, but women have a 0.7ppt higher chance of receiving informal care, and a 0.4ppt lower chance of receiving formal care, than men.

The probability of receiving care generally increases with increasing age with the effect becoming statistically significant, and more pronounced, from age 75 for informal care and age 80 for formal care. Men aged 85 to 89 are 8ppts more likely, and women are 7ppts more likely to receive informal care than those aged under 55. This jumps to 14ppts (men) and 8ppts (women) for those aged 90 and over. The relative magnitudes of the effects switch between sexes for formal care, with men aged 90 and over being 6ppts more likely than those under 55 to receive formal care, and women 8ppts more likely.

Being in a couple confers a highly significant increase of 11ppts on men and 10ppts on women in the probability of receiving informal care, and a highly significant reduction of 2ppts (men) and 3ppts (women) in the probability of receiving formal care. This amounts to a 9ppts lower chance of men receiving no

[18] Also, importantly, it does not require that the independence of irrelevant alternatives (IIA) holds: we suggest that the absence of a formal care option would not have a proportionally equal effect on the probability of receiving no care or informal care, for example, and so we choose this specification over a multinomial logit.

[19] In this case we do not use the full set of conditions used in the health regressions, because of smaller sample sizes in the formal care group.

[20] In this case, children indicate a possible source of care provision.

care and a 7ppts lower chance of women receiving no care, holding all else (including health status and age) constant.

Higher levels of formal education at baseline are associated with an increase the probability of receiving formal care by up to 2ppts and decrease in the probability of receiving informal care by up to 4ppts, with the effect of a degree being statistically significant. Note that, overall, those with a degree are therefore less likely to receive any form of care than those with no qualifications, although that overall effect is not statistically significant.

We observe no consistent or significant influence of deprivation indices on the outcome of the care receipt model.

The effect of living in certain geographical regions appears to be significant: compared to the reference case of the North East, people living in all regions have a reduced probability of receiving informal care. The difference between living in the North East and in the East or West Midlands, the South East, or the South West is statistically significant. Those living in East or West Midlands and the South East are also significantly less likely to receive any care.

Having children at baseline, who are a potential source of care provision, increases the probability of receiving informal care by 3ppts for women (the effect for men is smaller at 0.4ppt) and decreases the probability of receiving formal care by a statistically significant 2ppts for men and 1ppt for women.

Health and care

Next period health (i.e. the health we have predicted for the period in which we are now predicting care) and lagged care receipt (i.e. care at t and $t-2$) both have large and significant impacts on the probability of receiving care. This is the first time in this report that we demonstrate the use of modelled outcomes in the right hand side of our regression models. What we mean here is that, standing at time t and predicting care receipt at $t+2$, we find that health at $t+2$, which we've already predicted, and care receipt at t, both have a significant effect on care receipt at $t+2$.

Those in the worst health at $t + 2$ are up to 29ppts more likely than those in the best health to receive informal care at $t+2$, and even being in good health as opposed to best health at $t+2$ increases the probability by up to 12ppts. The effect on formal care is smaller, but the overall effect of worst health at $t+2$ is a decrease of up to 35ppts in the probability of receiving no care at all. Lagged health is also included in the specification, but its effect is small in comparison to that of next period health.

Lagged provision of care decreases the probability of a man receiving any care by a statistically significant 3ppts. While having a partner in the worst health at $t + 2$ increases the probability of receiving formal care by up to 2ppts the reduction in the probability of receiving informal care of up to 3ppts leads to an overall reduction in the probability of receiving any care by up to 2ppts, the reduction in the capacity of the partner to provide care being the dominant factor.

Care receipt is relatively persistent, with not only care receipt at t but also care receipt at $t - 2$ having a significant effect on care at $t + 2$. Receiving informal care at t increases the probability of receiving informal care at $t + 2$ by up to 14ppts, and receiving informal care at $t - 2$ increases the probability at $t + 2$ by up to a further 7ppts. Formal care receipt has a similar predictive effect on receipt of formal care – up to 6ppts if receiving at t and up to a further 3ppts if receiving at $t - 2$ – but it also has a predictive effect on receipt of informal care.[21] Formal care receipt at t significantly increases the probability of informal care receipt at $t + 2$ by up to 10ppts. While the equivalent effect of lagged informal care receipt on formal care probabilities is positive, the magnitudes of the effects are much smaller and are not statistically significant.

Lagged receipt of disability benefits has a material positive impact on the probability of receiving any care, with DLA receipt being the strongest (and only statistically significant) predictor, raising the probability by up to 9ppts.

Income, wealth and work

Lagged working status has a negative impact on the probability of receiving any care, with full-time work decreasing the probability of receiving care at $t + 2$ by a highly significant 5ppts.

Perhaps surprisingly, we observe no consistent or significant influence of baseline income or wealth on the outcome of the care receipt model.[22]

[21] Note that the variable definition and model specification allow individuals to receive just one of formal and informal care in any period.

[22] This may well be because there are a number of effects acting in different directions. The reduction in need for (formal) care captured by increased wealth (i.e. any effect over and above that captured by education, health level, etc.) is counteracted by an increased ability to pay and perhaps, therefore, a lower bar for choosing to access formal care. There is also the issue that some households who receive formal care at baseline could have significantly reduced their wealth at the point at which it is measured in order to pay for this, shifting some individuals between contemporaneous wealth quintiles in a way which is not material in our other models.

Model specification: care provision

This model is an ordered probit, a multinomial model which acknowledges an underlying ordering in the outcome states.

The specification contains information on age, sex, couple status, education, the intensity of care provision interacted with couple status, care receipt, presence of children at baseline[23], health, home ownership, deprivation, region, socio-economic group, working status, baseline income and wealth, and whether the partner dies by $t + 2$.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo-R^2 statistic for the model is 24.8%.

As above, the references to high and low intensity care provision refer to caring for 35 hours per week or more, and for fewer than 35 hours per week, respectively.

Demographics

Sex has a material, but statistically insignificant, effect on the probability of providing care with women 2ppts more likely to provide high intensity care, and 7ppts more likely to provide any care, than men.

The probability of providing care generally decreases with increasing age. Men aged 85 to 89 are 5ppts less likely to give care than those aged under 55, and men aged 90 and over are 16ppts less likely. Women aged 85 to 89 are 15ppts less likely to give care than those aged under 55, but those aged 90 and over are only 10ppts less likely. This is likely to be due to the prevalence of care provision within couples, and the relative life expectancies of men and women.

The effect of being in a couple is complex in this model, because of the way the explanatory variables are interacted. This is explored below but, generally, those in couples give more care and the effect is significant.

The probability of providing care increases with higher levels of formal education, with the effect of 'some qualifications' being significant and the effect being to increase the probability of provision by up to 2ppts. The same pattern is

[23] Indicating both a source of care for a partner, potentially decreasing the probability of providing care, and the possibility of grandchildren, potentially increasing the probability of providing care.

seen in the probability of providing high intensity care, with those with a degree being 1ppt more likely to do so than those with no qualifications.

There is no clear effect of deprivation on the provision of care.[24]

Health and care

While an individual's own health and care receipt at $t+2$ has a material, and often statistically significant, effect on the likelihood of providing care, it is care provision at time t and the characteristics of their partner (if they have one) which have the biggest and most significant effects.

Those in poorer health are increasingly less likely to provide care, with the effect becoming highly significant in the poor and worst health groups: being in the worst health group reduces the probability of providing any care by 7ppts for women and 8ppts for men. Men receiving formal care at $t+2$ are 5ppts less likely to give care, but men receiving informal care are a highly significant 4ppts more likely to give care. Women in receipt of either form of care are 1ppt more likely to provide care. Although this may feel slightly counterintuitive, this is likely to reflect the interdependence of health and care needs of people in couples.

Those whose partner receives care at time t are up to 5ppts more likely to provide any care at $t+2$. A larger and highly significant effect is that those whose partner is in the worst health at $t+2$ are 31ppts (men) or 28ppts (women) more likely to give care at $t+2$. Even having a partner in good, as opposed to best, health increases the probability by 10ppts for men and 6ppts for women.

Because of the way we interact care provision and couple status in this model, it is more complex to interpret these marginal effects than those in other models. Table 1 shows the marginal effect of being in any combination of sex, care provision and couple statuses on the probability of giving any care at time $t+2$. The couple status is as at time $t+2$, with 'bereaved' individuals being those whose partner was alive at t but has died by $t+2$. The reference case is a single male providing no care at t.

The figures in the table show, for example, that bereaved individuals are much less likely to provide care than are their single counterparts, even though the

[24] Although statistically insignificant, there is a large impact of deprivation (IMD) quintile on the probability of men giving care, with those in the most deprived quintile being 3ppts more likely to provide care than those in the least deprived. However, being in the two most deprived quintiles of the IDAOPI measure has a statistically significant impact and reduces the probability of providing care by up to 4ppts. There is a strong correlation between quintiles of IMD and IDAOPI, suggesting that interpreting these marginal effects is not straightforward.

absolute probability may still be relatively high (particularly for women). Care provision is more persistent if the care provided at time t was at a high intensity. Conditional on care provision at time t, those in couples are actually less likely to provide care at $t + 2$ than are their single counterparts suggesting that, although those in couples provide relatively more care, that care provision is less persistent.

Table 1. Marginal effects of sex, lagged care provision and couple status on care provision at t+2 (ppt change)

	Male			Female		
	Single	**Couple**	**Bereaved**	**Single**	**Couple**	**Bereaved**
No care at t	-	5.6	-19.7	7.5	14.8	2.3
Low care at t	23.5	17.9	3.9	29.4	29.3	16.8
High care at t	38.3	30.9	5.6	48.0	40.3	27.8

Income, wealth and work

Being in full-time work at time t decreases the probability of providing care at $t + 2$ by a highly significant 4ppts for men and 3ppts for women, with a smaller negative effect for part-time work.

We observe no consistent or significant influence of baseline income or wealth on the outcome of the care provision model.

6. Labour Market Models

The labour market module simulates individuals' labour supply decisions: whether they work full-time (over 30 hours per week), part-time (up to 30 hours per week), or not at all. We run three separate regression models, one for each current working status, which essentially allows us to interact current working status with all other explanatory variables in the model. It also allows us to impose different age restrictions on each transition: we allow people to move within (i.e. from full-time to part-time work and vice versa) and out of the labour market up to the age of 79 (after which they must retire), but to move into the labour market only up to the age of 69. The basis of these cut offs is discussed below.

Earnings history

In order to model the gross, and ultimately net, income of the pensioner population we need to have information on their earnings. Workers report their earnings data in ELSA, and so we can use this information straightforwardly for as long as the individual remains in the working status (full-time, part-time, or no work) that he was in at the time of reporting the earnings.[25] However, our model allows people to move into employment states in which they have not been observed in the data and we need a methodology for allocating earnings to that work.

Broadly, our aim is to determine full-time and part-time earnings for each person, to be applied when he is simulated as working at that intensity. We also use the natural logarithm (log) of full-time earnings as some measure of 'potential' earnings as an explanatory variable in the labour supply models: this means that we also have to construct full-time earnings in the estimation data (2002–10). The methodology relies on the following principles:

- Where possible, we use an individual's reported earnings in preference to matched earnings from another record: so, if someone has been observed in full- but not part-time work, we would scale his full-time earnings to derive part-time earnings (or vice versa), rather than matching on part-time earnings from another record.

[25] Note that our model does not allow for the possibility of an individual moving, during the simulation, to a higher or lower paid job of the same intensity (full-time or part-time): he can only change his wage by moving between states. We make an assumption about real wage growth for the youngest people in our model as shown in Figure 2.

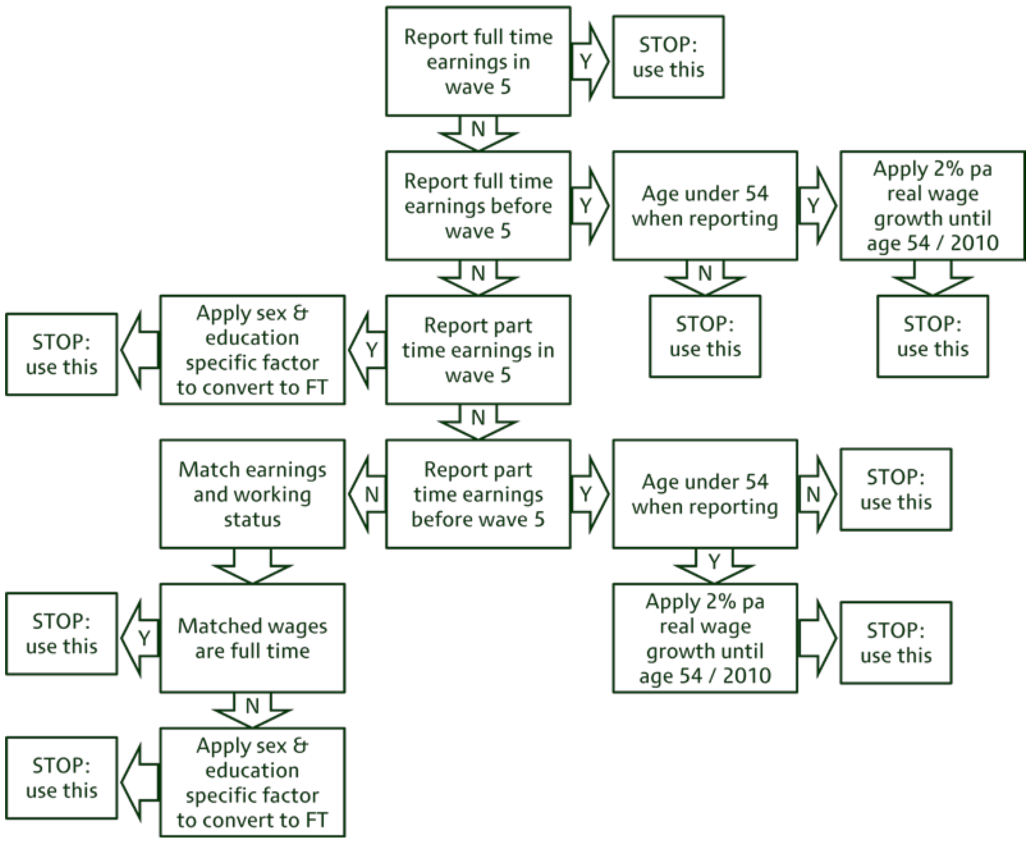
- Where earnings are reported in multiple years, we use the earnings reported in the year closest to the year in which they are missing.

- We only match earnings to individuals who have never been observed in work, and we use the model described below to do this.

- Throughout the simulation we assume that individuals experience real earnings growth of 2% p.a. before the age of 55, and no real earnings growth thereafter. We increase nominal earnings by CPI when constructing gross incomes for each year, as described in section 8.

- The process for determining full-time earnings for the simulation population is shown in Figure 2. The process for determining part-time earnings is almost identical, with the roles of part and full-time earnings information reversed. In both cases we use the most recent earnings information when multiple historical earnings have been reported, and we move all earnings information to the same price year, using the CPI, before calculating scaling factors or performing propensity score matching.

Figure 2. Construction of full-time earnings for 2010–11



In order to obtain both part- and full-time earnings estimates when the individual has reported earnings from only one of these states, we calculate sex and education specific scaling factors from cases in the ELSA data where individuals move between these two states and report income in each case. This approach yields the factors shown in Table 2, which are calculated as the median of the ratio of $t + 2$ to t earnings across the transition in question for the relevant sex and education group.

As we would intuitively expect, all the full-time to part-time scaling factors are less than one. However, not all of the part-time to full-time factors are greater than one, and few of them are substantially greater. We suggest that this is because of the age group we are considering: very few people, especially men, are moving from part to full-time work at this point in the lifecycle, and those who are could conceivably be transitioning from a high intensity part-time job (for example consultancy or some form of self employment) to a less intense and more regular, but less well paid, full-time job.

The differences in scaling factors between men and women in the same education group are likely due to differences in the number of hours worked in each state, as well as differences in the type of work and the absolute value of the earnings in the calculations.

Table 2. Factors for scaling full- and part-time earnings

Sex	Education at baseline	Full- to part-time scaling factor	Part- to full-time scaling factor
Male	No qualifications	0.62	1.06
Female	No qualifications	0.75	1.11
Male	Some qualifications	0.53	1.00
Female	Some qualifications	0.70	1.18
Male	Degree	0.67	0.99
Female	Degree	0.69	1.33
Male	All	0.57	1.01
Female	All	0.71	1.18

Source: Authors’ calculations from wave 1–5 ELSA data

Earnings matching model

As shown in Figure 2 there are circumstances in which we have no information at all on an individual’s earnings, having never observed them in work, and so we match on earnings from a similar individual. In order to do this we model working status in the next period as an ordered probit. We create a ‘matched to’ group of people who are out of work at t and who also report no $t + 2$ earnings (i.e. the people we need data for), and a 'matched from' group of people who are also not working at t but who report earnings at $t + 2$. We use the underlying linear prediction from the regression model[26] as an indication of propensity to earn in the next period conditional on current characteristics, and calculate this for everyone in both groups. This means that we can compare propensity scores to find the closest match for each 'matched to' individual from the 'matched from' group, and give the 'matched to' person the 'matched from' person's earnings.[27]

We retain information on whether the matched earnings were reported as full- or part-time earnings by the 'matched from' person, and use this in conjunction with the factors in Table 2 to determine a full- and a part-time earnings figure for the recipient of the match.

The regression model used to generate the propensity scores contains information on: age, membership of DB and DC pension schemes[28], education, self employment[29], care receipt, intensity of care provision, presence of an outstanding mortgage, health, home ownership, deprivation, partner's working status, health, region, socio-economic status, DLA receipt, time since last worked, whether below SPA at next period, whether a partner is below SPA, whether a DB scheme member is below the Normal Retirement Age (NRA) for their scheme, and baseline wealth quintile (but not income quintile, because of endogeneity). As with our other models, we interact all these variables with sex and include sex as an explanatory variable. We do not include lagged working status because everyone we need to match to is out of work, and everyone we match from is in work, but we do include partner's lagged working status.

The pseudo-R^2 statistic for the model is 41.7%. The explanatory variables with statistically significant effects on the outcome are: the indicators for being below NRA and SPA and for a partner being below SPA; DC scheme membership; health, informal care receipt, and high intensity care provision; home ownership and

[26] Although Stata's post-estimation prediction produces, and we generally use, discrete probabilities for each possible outcome state, these are based on an underlying linear measure. It is this linear measure that we use directly in the propensity score matching.

[27] This propensity score matching uses the psmatch2 command in Stata, described in the help documentation and here: http://repec.org/bocode/p/psmatch2.html.

[28] 'Membership' here is defined as whether the individual is still contributing to a pension scheme and is measured in the ELSA data and then assumed to remain fixed.

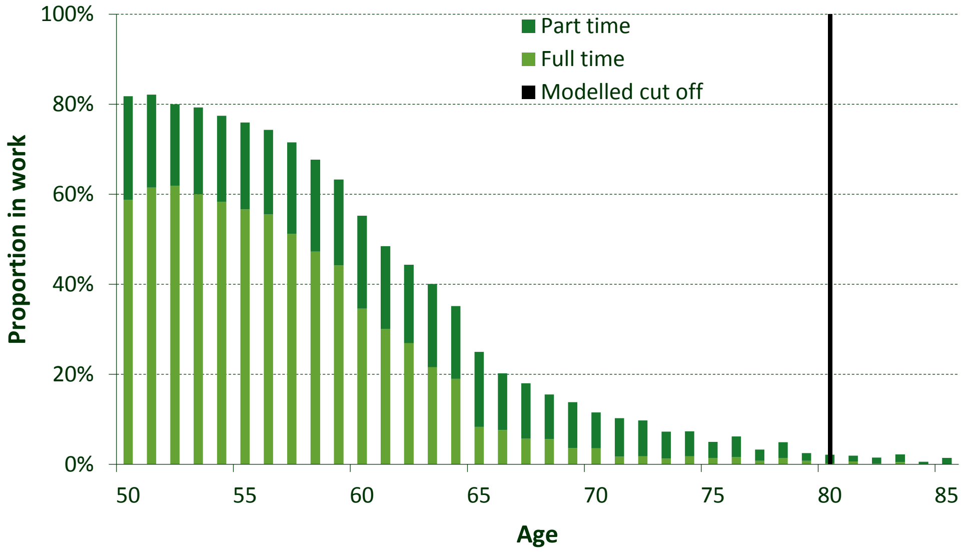
[29] This records whether the individual was ever observed in self employment in the five waves of ELSA.

outstanding mortgages; age; time since last worked, whether ever self employed, and partner's working status; and DLA receipt.

Age cut offs for participation

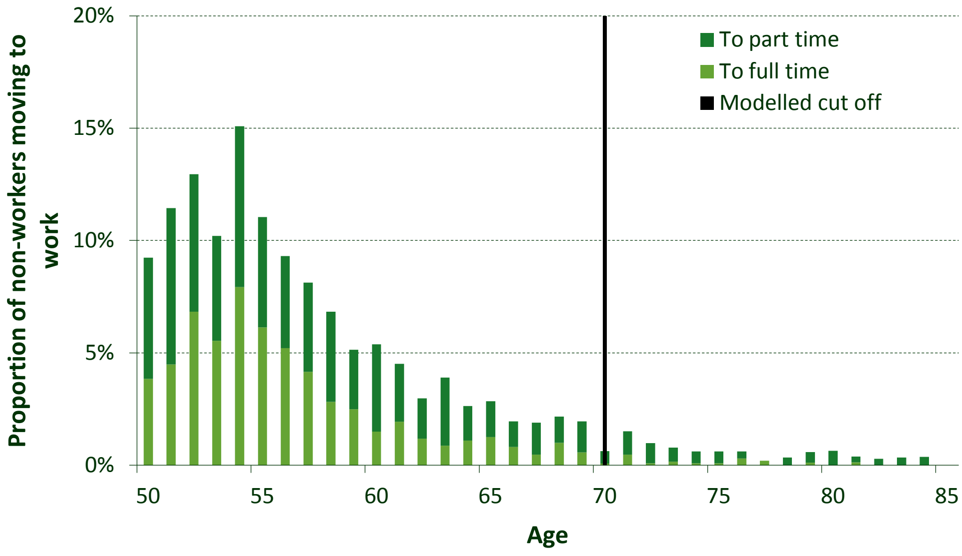
We do not allow people to remain in the labour market indefinitely, not least because of the paucity of data on which to base the regression models for the very oldest section of the labour market. Instead, we allow people to enter the labour market (i.e. move from no work to some form of work) up to the age of 69 and to move or remain within the labour market (in full- or part-time work) up to the age of 79. This decision is supported by the ELSA data underlying the models, shown in Figure 3 and Figure 4, which show a very small minority of people being constrained by these limits.

Figure 3. Proportion of people in paid work by age



Source: ELSA waves 1–5

Figure 4. Proportion of non-workers moving into work by age



Source: ELSA waves 1–5

Model specifications

We construct three multinomial logit models, one for individuals in each working status at time t, to predict the probability that they take each of the three possible working statuses at $t+2$. We do not use an ordered model because that would imply a consistent preference ordering between working statuses that is unlikely to remain consistent within individuals as they age, or across all individuals in our simulation population. We construct three models as this is equivalent to interacting current working status with all variables: so the effect of a given variable can change depending on whether the transition in question is into a higher or lower intensity working status, for example.

There are certain key differences in the modes used to simulate working status, and the model used to match earnings that we described previously:

- We use an ordered probit to generate a propensity score and a multinomial logit to generate probabilities for simulation, because we use them for two different purposes. When we simulate labour supply decisions there is no clear preference ordering between working statuses which holds well for all individuals in the model at all ages. When we generate a propensity score, however, we are aiming to capture information about some latent propensity to earn, based on a range of characteristics, which we can more easily conceptualise as being a continuous variable underlying realised, ordered, working statuses.

- We use a single regression model to generate a propensity score but three regression models, conditioned on current working status, to estimate the simulation probabilities. This is simply because everyone for whom we need to generate earnings data is currently not working.

The specifications for the three models used to simulate working status are almost identical, with the exception that time since last worked is included only for those who are out of work at t. They include information on age, membership of DB and DC pension schemes, education, self employment[30], care receipt, intensity of care provision, presence of an outstanding mortgage, health, home ownership, deprivation, partner's working status, health, region, socio-economic status, DLA receipt, whether below SPA at next period, whether a partner is below SPA, whether a DB scheme member is below the NRA for their scheme, baseline wealth quintile, and own and partner's potential full-time earnings.[31] Again, we interact all of these variables with sex and include sex as an explanatory variable.

Note that in each model, by construction, the marginal effect of an explanatory variable on being in no work, part-time work and full-time work must sum to zero.

Transitions from no work

The pseudo-R^2 statistic for this model is 28.5% and it is estimated on all individuals aged up to 67 at time t who were also out of work at time t. This allows people to move into the labour market until they are 69 (at $t + 2$).

Women are 8ppts less likely to move into full-time work but 5ppts more likely to move into part-time work than are men (so are 3ppts more likely to stay out of work). The effect of increasing age on both men and women is to make transitions into work less likely, with the effect on the probability of men moving into full-time work being statistically significant. The effect on women of being below SPA at $t + 2$ is to increase the probability of moving into work by 3ppts[32].

[30] Whether the individual was ever observed in self employment in the five waves of ELSA.

[31] Where the potential full-time earnings in the next period are negative (i.e. a self-employment loss) we set the log wage variable to zero, but we create an additional binary variable indicating a self-employment loss, which we also include as an explanatory variable.

[32] The effect works in the opposite direction for men, but the age variables (age, sex interacted age, and age squared) work in combination with the SPA effect and it is likely that this seemingly

Contributing to a DB pension scheme at baseline makes men 56ppts more likely to remain out of work, unless they are below the NRA, the marginal effect of which is to increase the probability of working by 61ppts, more than cancelling out the membership effect. The effect for women is less consistent, with both baseline contribution to a scheme and being below the NRA acting to decrease the probability of moving into employment (by 19ppts and 37ppts respectively). Contributing to a DC scheme, however, makes both men and women more likely to enter work: a 3ppt increase for men and a 4ppt increase for women. The effect on men moving into full-time work is statistically significant.

Being in a couple increases the probability of remaining out of work by 2ppts for men and 3ppts for women, although having a partner in work generally appears to increase the probability of entering work by up to 2ppts. Having a partner below SPA increases the probability of moving into work by 1ppt for men and 2ppts for women, suggesting that some form of joint retirement planning is implicitly captured by our model.

Having ever been observed in the ELSA data to have been in self employment increases the probability of returning to work by up to 20ppts, having a statistically significant effect on the probabilities of moving into both part and full-time work.

The effects of care provision are not large but are interesting. Men providing low intensity care are 0.6ppts more likely to remain out of work than those not providing care, but are also 0.3ppts more likely to move into full-time work. Women are 0.3ppts more likely to remain out of work and 0.3ppts more likely to move into part-time work. Providing high intensity care makes both men and women more likely to remain out of any type of work, the effect being 1ppt for men and 3ppts for women. This suggests that some individuals are responding to a partner's need for low level care by remaining out of work to provide care, while others may be increasing their propensity to do some form of work, perhaps to compensate for a partner's lost earnings. This is borne out in the effects of having a partner receiving informal care, which increases men's probability of remaining out of work by 2ppts and of moving into full-time work by 1ppt, and increases women's probabilities of moving into full-time work by 1.2ppts and part-time work by 0.4ppts.

---

paradoxical result is due to the complexity of the interactions between the variables and the differential SPAs for men and women.

A partner's receipt of formal care has a very large effect, making men 78ppts less likely to move into any form of work and making women 26ppts more likely to remain out of work but also 10ppts more likely to move into part-time work: again, conceivably in order to replace a partner's earnings.

Increasing levels of potential full-time earnings increase the probability that an individual will remain out of work. Because 'non work' in our model encompasses both unemployment and retirement or inactivity, it is conceivable that those who were observed in the past with higher earnings have voluntarily retired and have less need to re-enter the labour market. A partner's higher potential earnings do have a significant positive effect on the probability of a man moving into full-time work, however. Receipt of DLA makes men 3ppts less likely to enter work and women 6ppts less likely, and increasing time since last worked has a highly significant negative effect on the probability of both men and women entering work.

Having an outstanding mortgage significantly increases the probability of moving into full-time work for both men and women (by 1.6 and 1.4ppts respectively), and also increases the probability of women moving into part-time work.

Transitions from part-time work

The pseudo-R^2 statistic for this model is 11.2% and it is estimated on all individuals aged up to 77 at time t who were also in part-time work at time t. This allows people to move between full-time and part-time work until they are 79 (at $t + 2$).

Women are 34ppts more likely than men to remain in part-time work, and 8ppts more likely to move from part-time work to no work, making them a highly significant 42ppts less likely to move from part- to full-time work. This could suggest that part-time work is more likely to be an active choice for women but a second-best choice behind full-time work for men. However, increasing age makes men significantly less likely to transition from part to full-time work, implying that there is more active choice involved at older ages. Being below state pension age has a highly significant negative effect on the probability of leaving part-time work for no work, of 16ppts for both sexes. However, women below SPA are 6ppts more likely to move into full-time work than men in the same circumstances, who are more likely to remain in their part-time job.

Being below the NRA of a DB pension scheme reduces the probability of leaving part-time work by up to 10ppts, with scheme membership generally increasing the probability of leaving work by up to 4ppts. Membership of a DC scheme seems to have little effect on men (0.4ppt increase in probability of leaving work) but reduces women's probability of leaving work by 8ppts, in favour of either part- (+4.2ppts) or full- (+3.5ppts) time work.

Being in a couple increases the probability of leaving part-time work by 4ppts for men and 5ppts for women. Having a partner below SPA has a material effect on men's labour supply decisions, increasing the probability of moving into full-time work by 5ppts, at the expense of leaving work (-7ppts).

Having ever been self employed has a small positive effect (2 to 3ppts) on the probability of moving to full-time work, at the expense of both leaving work and remaining in part-time work.

Providing low intensity care at $t + 2$ makes women 2ppts more likely to leave all paid work, and men 4ppts more likely to stay in part-time work. Having a partner who receives informal care makes men more likely to either leave work or move to full-time work, and women more likely to stay in part-time work or move to full-time work. Having a partner who receives formal care, however, makes both men and women much less likely to leave work, a decrease of 9ppts for men and 26ppts for women, with both being most likely to remain in part-time work. Being in poor or worst health at $t + 2$ increases the probability of leaving work by up to 18ppts and has a larger effect on women than on men.

Having higher potential earnings makes women most likely to move into full-time work, and least likely to leave work altogether, but they make men least likely to move to full-time work and most likely to remain in part-time work. Having a partner with higher potential earnings, though, makes men significantly more likely to remain in part-time work and significantly less likely to leave work altogether. Receipt of DLA makes men 20ppts more likely to stay in part-time work and women 10ppts more likely to leave work.

Having an outstanding mortgage makes men 9ppts, and women 7ppts, less likely to leave work, both being almost equally as likely to remain in part-time work as to move into full-time work.

Transitions from full-time work

The pseudo-R^2 statistic for this model is 15.4% and it is estimated on all individuals aged up to 77 at time t who were also in full-time work at time t. Again, this allows people to move within the labour market until they are 79 (at $t + 2$).

Women are significantly (31ppts) less likely to remain in full-time work than are men, and are 26ppts more likely (again a significant effect) to move into part-time work. Increasing age has a significant positive effect on the probability of leaving full-time work in favour of no work. Being below SPA again has a significant effect, making both sexes 12ppts less likely to leave work and also less likely to move into part-time work (men by 7ppts and women by 4ppts).

The effect of DB pension scheme membership significantly increases the probability of leaving work by 12ppts for men and 3ppts for women, but being below the NRA almost completely cancels this out, with a negative effect of 11ppts for men and 4ppts for women. DC scheme membership has a significant positive effect of 6ppts for men and 2ppts for women on the probability of remaining in full-time work.

Being in a couple significantly increases the probability of moving from full- to part-time work, by 6ppts for men and 1ppt for women. Having a partner in part-time work also has a significant effect, increasing the probability of men moving to part-time work by 4ppts and women by 5ppts; it also makes women 5ppts more likely to leave work altogether. Having a partner in full-time work, however, significantly increases the probability of remaining in full-time work, an effect of 5ppts for men. Having a partner below SPA also significantly increases the probability of remaining in full-time work by 5ppts for men, reducing the probability of moving to part-time work by 4ppts.

Providing low intensity care decreases both sexes' probability of remaining in full-time work by 3ppts, with men 4ppts more likely to leave work altogether and women 2ppts more likely to move to part-time work. Providing high intensity care has a highly significant effect, reducing the probability of remaining in full-time work by 17ppts for men and 16ppts for women: men are 9ppts more likely to leave work completely and women 12ppts more likely. Counteracting this, the effect of having a partner who receives informal care increases the probability that men remain in full-time work by 3ppts and increases the probability that women move out of full-time work to either part-time or no work by 1ppt. A partner who receives formal care makes men 10ppts more likely to remain in full-time work and women 6ppts more likely to move into part-time work, with both being less likely (9ppts for men, 4ppts for women) to leave work altogether.

An individual's own poorer levels of health have a significant negative impact of up to 24ppts on the probability of remaining in full-time work. There is an accompanying significant increase of up to 20ppts in the probability of an individual leaving work altogether.

Higher earnings have a significant effect, making both men and women more likely to remain in full-time work, with the effect being stronger for women. The effect of partners' earnings is negligible. Receipt of DLA makes both sexes 7ppts more likely to leave work. It also increases the probability of moving to part-time work by 9ppts for women and 3ppts for men.

Having an outstanding mortgage has a significant positive impact on the probability of both men and women remaining in full-time work, an effect of 5ppts for men and 6ppts for women. Men are 3ppts less likely to leave work altogether and women 5ppts less likely.

7. Disability Benefits

We use the IFS tax and benefit model, TAXBEN, to determine receipt of means tested benefits on the basis of simulated gross income, but we directly simulate the receipt of disability benefits within the demographic simulation. This allows us to use information on working, health, and care status, as predicted by the model, to govern receipt.

The benefits we model are disability living allowance (DLA), attendance allowance (AA) and a range of incapacity benefits: incapacity benefit (IB), employment support allowance (ESA) and severe disablement allowance (SDA). We also model carer's allowance (CA), which is paid to people providing significant levels of care. More detail on the eligibility for these benefits is set out later in this section.

A number of factors make this element of the modelling somewhat complex. Data quality, both in terms of the accuracy and rate of reporting, is a concern, with many people reporting benefits to which they are not (by reason of age, mainly) entitled, many reporting benefit incomes which do not match award amounts for the benefit they are reported under, and an overall reported claimant rate which is significantly lower than the rate recorded by the Department for Work and Pensions (DWP), which administers the benefits. The complexity in the system presents a challenge, with the various levels of care and mobility award providing 11 possible DLA benefit amounts and the quality of the income data, as above, being insufficient to differentiate accurately between them. Finally, the reforms to the DLA and IB systems (being replaced with personal independence payments – PIP – and ESA respectively) take place within our simulation period, and we have made a number of assumptions on the basis of DWP data[33] to allow us to model the effects of these reforms. Our approach to tackling each of these issues is outlined in more detail in the sections below: in particular we adjust the modelled data to compensate for under-reporting.

For each benefit we perform the following steps:

- Rescale our in-sample prediction in 2010–11 to boost claimant rate from the reported level to the target level observed in the administrative data

[33] DWP claimant rates are calculated by dividing caseloads from the tabulation tool (http://tabulation-tool.dwp.gov.uk/100pc/tabtool.html) by census population numbers (http://www.nomisweb.co.uk/census/2011/DC1104EW/view/2092957699?rows=c_age&cols=c_sex). More detail is given in the calibration section.

- Simulate benefit receipt throughout the simulation period:
  - Predict change in claimant rate and use this to calculate new 'target' rate: this is the claimant rate that our simulation will match in the period in question
  - Predict benefit receipt at $t+2$, including boosting claimant rate to match the target rate
- (Where applicable) determine the level of benefit received by the modelled claimants
- (Where applicable) apply modelling assumptions to account for policy changes in the benefit system

The detail of the modelling of each benefit, including more detail on calibration of the claimant rate, modelling of reforms, and determination of levels of benefit, is given below.

Benefit rules

We model the following non-means tested benefits:

Disability living allowance is paid as a combination of two components: care and mobility. A person can be entitled to one or both of the components. The care component is paid at three levels: high, middle and low, and the mobility component is paid at two levels: higher and lower. Individuals can make new claims for DLA until the age of 64 (inclusive), and any DLA claimed at age 64 may continue to be claimed beyond that age. DLA is currently being reformed, and will ultimately be replaced by personal independence payments for claimants aged under 65.

Personal independence payments are being rolled out to replace DLA for claimants aged under 65. The reform makes a number of changes to the way the benefit is administered, including: the facility to review more frequently the awards made; a new, 'more objective' assessment; and the removal of automatic entitlement for certain conditions. It also changes the structure of the award system, replacing three levels of the care component and two levels of the mobility component with two levels of the daily living component and two levels of the mobility component (leading to eight possible combinations of components, in place of the current 11) (Department for Work and Pensions, 2014). DWP's impact assessment predicts that these reforms will lead to a 20% reduction in working age spending on DLA/PIP (Department for Work and Pensions, 2012).

Attendance allowance is payable to those aged 65 and over who otherwise satisfy the conditions for the middle or higher rates of the care component of DLA. Attendance allowance is thus paid at a higher and a lower rate. To date, no plans have been announced for changes to AA.

Carer's allowance can be received by anyone caring for an eligible person for at least 35 hours per week, while not earning more than £100 per week from paid work and not being in full-time education. The person being cared for must receive either attendance allowance or the middle or higher levels of the disability living allowance care component in order to be eligible.

Incapacity benefit is payable to those who are judged to be incapable of work and who have either paid or been credited with sufficient national insurance contributions or who became incapable of work in youth. It is paid at three rates: short term lower, short term higher, and long term The upper age limit for claiming IB is the SPA. As our model works on two-year transitions, and the lower and higher short term rates are payable only for the first and second six months of the claim respectively, we model all claimants as if they are receiving long-term IB. Incapacity benefit is currently being replaced by employment support allowance.

Severe disablement allowance is a legacy benefit: there have been no new claims allowed since April 2001. Existing claimants continue to claim, but IB replaced SDA in 2001 and ESA is now replacing IB. We allow existing claimants to continue to receive SDA in our model, but we do not make any adjustments for under-reporting other than to the total number of people claiming IB, ESA or SDA (so any under-reporting of SDA is essentially compensated for by over-stating the number of people on IB or ESA, but the small numbers of people claiming SDA make any other approach disproportionate).

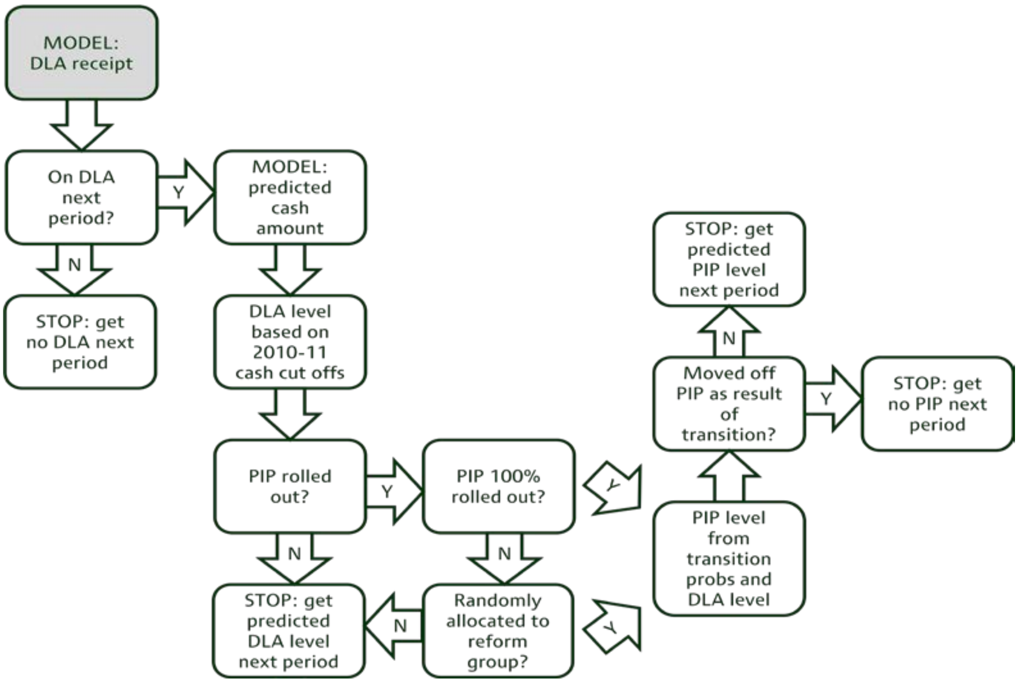
Employment support allowance is replacing IB and SDA: all new claims from October 2008 have been for ESA rather than IB, and the rollover of existing claimants from IB to ESA is underway. Similarly to IB, ESA is a benefit aimed at those judged to have 'limited capability for work' who are not entitled to jobseekers' allowance (JSA) or statutory sick pay (SSP). ESA claimants are placed in either the support group or the work related activity group (WRAG). Those in the WRAG must undertake work focused interviews and health assessments as a condition of receiving the benefit.

Model overviews

The technical detail of the disability benefit models, including the detail of the calibration and the basis of the reform assumptions, is set out below. The diagrams in this section offer an overview of the process by which individuals at each age can receive each disability benefit.

Those aged up to 63 in the current period are passed through the younger adult disability living allowance (DLA) model, which allows movement both on to and off the benefit to be observed up to age 65 (as we are predicting for time $t + 2$, so the restriction to age 63 at time t accomplishes this).[34] Having determined whether an individual receives DLA (a step which includes calibrating the claimant rate to be consistent with DWP claimant data) we determine the level of DLA received: again, this is discussed in more detail below. In years where the PIP (personal independence payment) reform has been implemented we then estimate the PIP level the individual receives, which can result in movement off the benefit. This process is set out in Figure 5.

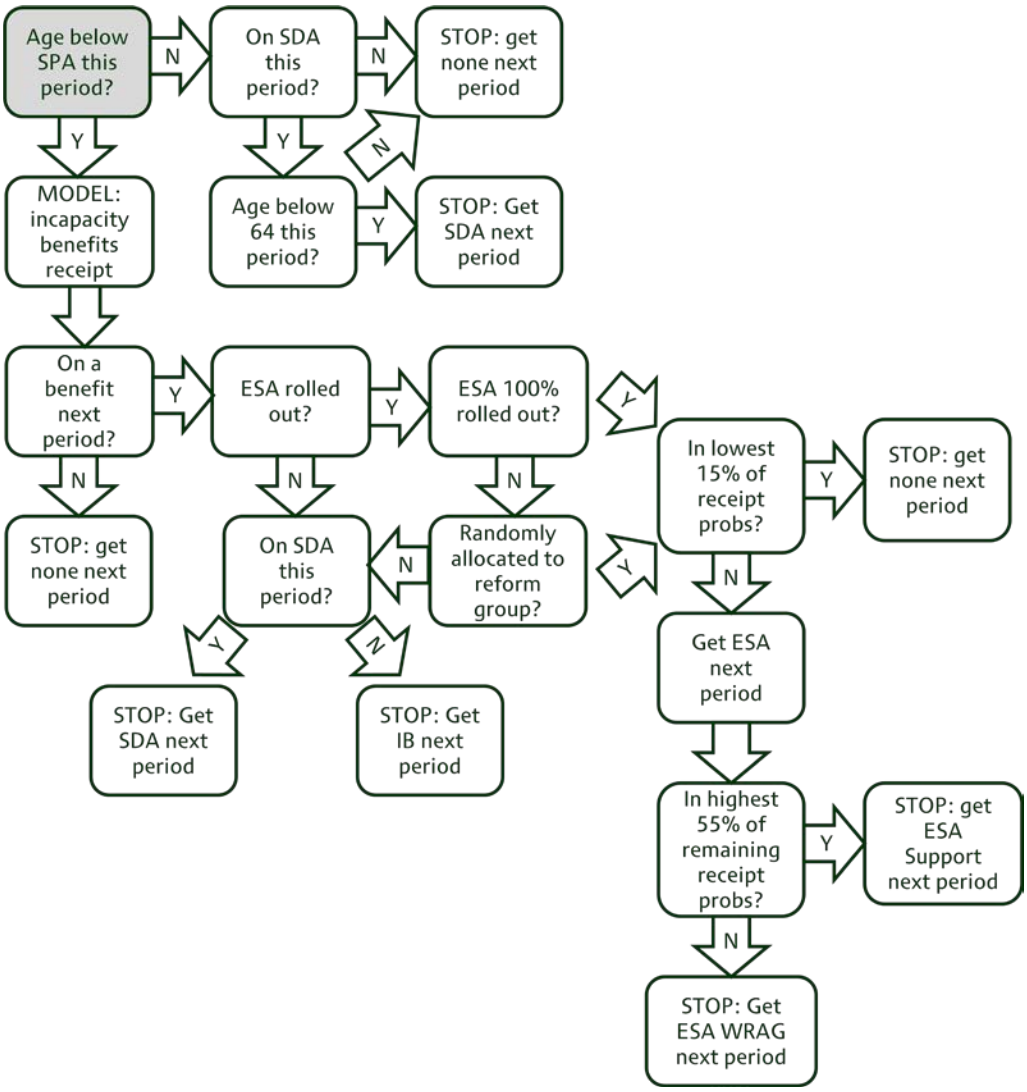
Figure 5. DLA receipt for younger (52–65) adults



[34] The policy rule actually prevents new claims after the age of 64. The reason we set the modelled cut off at 65 is that our model works on two year transitions. That means we observe some people only at odd ages and some only at even ages. If we set the cut off at 64, then we do not observe the 'odd age' individuals claiming above age 63, and we do not model movement on to the benefit for the 'even age' people at age 64. By setting the cut off at 65 we allow everyone the opportunity of claiming at age 64 at the cost of also allowing some new claims at age 65. While this is not a perfect representation of reality, it does ensure that new claimants aged below SPA, at least in the earlier years of the model, claim DLA rather than AA. We increase the lower age limit for AA to 66 for consistency.

Adults below SPA can claim an incapacity benefit: IB, ESA or SDA. By limiting the model to individuals aged up to and including SPA minus one year at time t we limit receipt of the benefits in the model to those aged up to SPA plus one year. Similarly to the argument set out above for setting the DLA cut off to age 65, this is slightly different to the policy rule limiting eligibility to SPA but is a modelling decision made to best compensate for the limitation of modelling over two-year simulation periods. This process is set out in Figure 6; the basis for the percentage figures used in modelling the reform to ESA is set out in the section on reform assumptions below.

Figure 6. Receipt of incapacity benefits

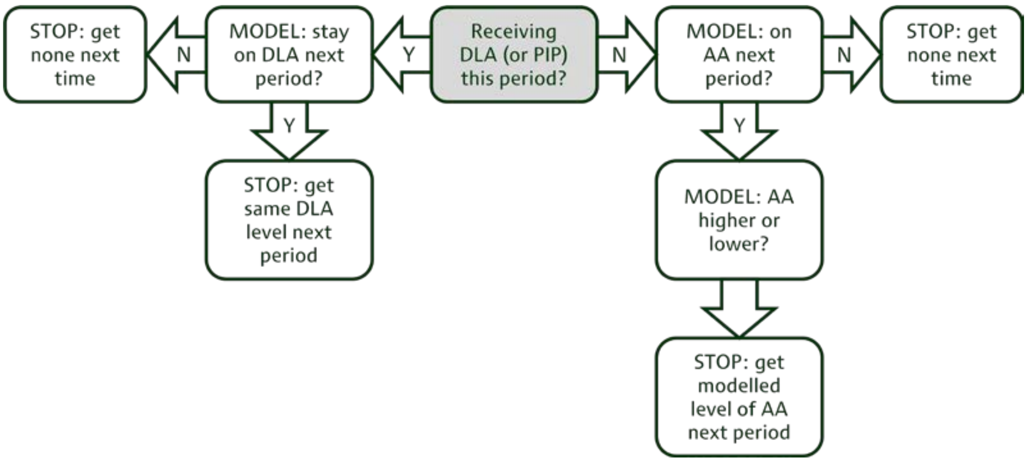


No new claims can be made for SDA: we allow claimants to retain their benefit, move off the benefit, or, if they are of an applicable age, to move from SDA to ESA.

ESA is rolled out in place of IB and SDA in line with the reform assumptions discussed below, and receipt of ESA results in an individual being placed in the Support or Work Related Activity groups. SDA can be claimed to age 65 regardless of SPA, so we allow anyone aged over SPA minus one year but below 63 in the current period (i.e. women aged 60–62 in certain years) to continue to receive any SDA they were receiving in the following period with no further modelling.

The process for those aged over 65 (over 63 in the current period) is simpler: this group can continue to receive DLA which was initially claimed before age 66, and any new claims for benefit from this age group are for Attendance Allowance (AA). No plans for reforms to AA have been announced, and the reforms to DLA do not directly affect those aged over 65. People who are modelled as moving off DLA are not allowed to move on to AA until the following period. This process is set out in Figure 7.

Figure 7. DLA and AA receipt for older (over 65) adults



Data cleaning

There are certain policy rules that we follow in cleaning the data and modelling the benefits. In cleaning the data we use for estimation, we enforce the following:

- Only those aged over 65[35] can claim AA. Anyone below this age reporting a claim of AA has the income reclassified as DLA.
- Only existing DLA claimants can continue to receive DLA after the age of 65. We use the panel aspect of ELSA to enforce this as far as possible, with those

[35] As set out above, this age is a compromise between the policy cut off of age 64 and the complexities involved in modelling across two year simulation periods.

who we observe starting a claim for DLA after the age of 65 having the reported income reclassified as AA. Where we observe an individual for the first time when they are aged over 65 and reporting receipt of DLA, we assume that this is reported correctly.

- People over SPA + 1 who report claiming IB have the income reclassified to whichever of DLA or AA is most appropriate, based on their age and their reported receipt of AA and DLA.

- As incapacity benefits can only be claimed by those who are not working or who are working for very few hours per week, we don't allow anyone in full time work to claim these benefits.

- Only those aged up to 65 can claim SDA: anyone over 65 has reported SDA income reclassified as whichever of AA and DLA is most appropriate.

- CA is paid (in 2010–11) to individuals earning no more than £100 per week and caring for at least 35 hours per week for someone who claims AA or the middle or higher rates of DLA. We clean the data to ensure that the earnings and caring hours requirements are satisfied, although we aren't able to test the benefits claimed by the person being cared for.

    - Around a third of people reporting income from CA do not report providing care (and aren't flagged in our model as providing care, on the basis of the partner's questionnaire). The majority of these do report receiving care, so we assume that this is misreporting of DLA or AA. We allocate the income to whichever of these the individual reports receiving, if either, and to the age-appropriate benefit (DLA up to age 65, AA for 66-plus) if neither is reported.

    - This leaves a handful of respondents who report receiving CA but neither give nor receive care. We set their income from CA and their indicator of CA receipt to zero.

As well as maintaining the age cut offs described above, we make the following assumptions in the simulation:

- Both men and women can claim SDA until age 65. That means that there are women who will be entitled to claim SDA after they leave the scope of the incapacity benefits model (up to SPA + 1). We assume that they remain on SDA until the age of 65 (or death, if sooner) if they were claiming it at SPA + 1.

- New claims for IB cannot be made after 2008: claims must be for ESA instead. We do not specifically differentiate between new and existing claims for IB, but we do ensure that the model reflects the observed speed of the rollout of ESA for the population aged 50 and over (see the section on reform assumptions, below) and that people who receive ESA in one period do not move back to IB the following period.

- We do not model statutory sick pay (SSP) in the simulation, primarily because it is a short-term benefit and we are working with two-year transitions. However, there are 25 people in wave 5 ELSA who report receiving SSP. We force their inclusion in the additional IB claimant group that we create in wave 5 to reach the DWP claimant rate, because of the similarity of the benefits and the relative likelihood of this group moving on to IB in the future.[36]

Modelling levels of claim

Our main motivation for modelling the receipt of disability benefits is that it will allow us to model family incomes more accurately. In order for this measure to be as precise as possible, we need to take account of the different amounts of income that receipt of each benefit can lead to. This section sets out the methodology and modelling assumptions that we use to do this.

In all cases the ultimate output from the demographic model is an indication of which, if any, named level of benefit an individual received - e.g. 'lower level AA' - rather than a cash amount. We then allocate the appropriate cash amount to the individual in the financial stage of the modelling.

Disability living allowance

As described above, there are 11 levels of DLA receipt resulting from the various possible combinations of care and mobility award.

For younger adults (those aged up to 65 in the model), we use a linear regression on log reported income from DLA to produce an ordering of DLA claimants according to the amount they are likely to report receiving. In the 2010–11 data we then use percentiles defined by DWP claimant data (Department for Work and Pensions, 2011) to group people into levels. For example, the data show that 9% of claimants receive the lowest cash amount[37] of DLA and so we allocate the

[36] There are 3 cases where SSP receipt is reported but we do not force IB receipt, because the individuals are beyond the age range for IB (there is no upper age limit for SSP).

[37] Note that there are two combinations - the lowest level of care claimed in isolation and the lowest level of mobility claimed in isolation - which attract identical cash amounts. For the purpose of the model we treat low care as the lowest level of claim and low mobility as the second

9% of simulated claimants with the lowest predictions from the DLA income regression to that combination of care and mobility award.

Having performed this classification in 2010–11 we then read off the predicted cash amounts which correspond to the cut offs between levels.[38] We hold these cash amounts fixed in the following simulation periods and use them to classify claimants into levels on the basis of their predicted cash amounts. This means that we only force the distribution to match DWP data in 2010–11 and allow the changing health and care needs of the population as projected by our model, among other factors, to influence the concentration of claimants in different care and mobility combinations in later periods.

The way that we translate DLA levels into post-reform PIP levels is set out in the section on reform assumptions below.

Older adults (those aged over 65 in the model) can only claim DLA if they were also claiming it as a younger adult. We assume that there is no change in level of DLA claimed between the final observation below age 66 and any future observations – that is, all older adults inherit and keep the DLA level that they claimed as younger adults. That means that we only need to allocate levels directly to older adults if they are already claiming DLA above the age of 65 in 2010–11.

Having run the in-sample regression in 2010–11 to achieve an appropriate claimant count we then order these older claimants by their probability of receiving DLA and allocate them to levels of award using percentiles drawn from the DWP claimant data (Department for Work and Pensions, 2011). This is a similar approach to that described above for younger adults, but it uses predictions of propensity to claim, rather than of reported income, and is only performed in one year.

---

lowest. This is coherent with the assumptions we make in modelling the reform from DLA to PIP, which essentially abolishes the lowest level of care award.

[38] These are not the same as the cash values of the award combinations as dictated by policy. Both the misreporting and noise in the data, and the small differences in cash amount between different award combinations, make classifying claimants directly on the reported benefit income in 2010–11 impractical. Instead we explicitly model reported income, which is often significantly different to any policy-dictated award amount, and use this as a measure of likely relative entitlement to DLA.

Attendance allowance

The structure of AA is much less complex than that of DLA, with only two possible award amounts in each year. This makes it much less complex to use the reported income data directly. We use the midpoint between the higher and lower claimant rates prevailing when each wave of ELSA data was collected as a cut off point, and classify claimants in the data into the higher and lower levels on the basis of the income from AA that they report. We then perform a within-period ordered probit regression, estimated on observations that we have classified in this way, to determine which level a known claimant receives in each simulation period.

Incapacity benefits

Our incapacity benefits regression covers claims of IB, ESA, and SDA. We take SDA receipt as reported in 2010–11, so a 2010–11 SDA claimant will go on receiving SDA for as long as he continues to claim an incapacity benefit or until SDA is reformed into ESA. The way that we allocate claimants between IB and ESA during the rollout, and between levels of ESA, is covered in the section on reforms below. We assume that all modelled incapacity benefit claims are for the long-term IB amount with no additions.

Reform assumptions

Over the course of the simulation it is anticipated that the reforms from ESA to IB and from DLA to PIP will be completed. Because our estimation sample contains data only on the legacy systems[39] we estimate the probability that an individual receives the legacy benefit and then apply reform assumptions to capture, as far as possible, the effect of the policy changes.

Introducing personal independence payments

The reform from DLA to PIP began, for new working age claimants, in 2013. The reforms were rolled out to existing working age claimants in pilot areas from late 2013, with the intention that claimants in all areas will be contacted from early 2015 and rollout will be complete by the end of 2017(Department for Work and Pensions, 2014). We simplify this process, assuming that no one is on PIP in

[39] Although ESA was introduced in late 2008, only ELSA wave 5 allows respondents to report ESA receipt. We therefore do not have information on transitions within the new system. We boost claimant numbers to meet pre-reform levels to counteract any effect of the reform captured in the data and then apply our reform assumptions on top of this, to avoid double-counting the impact.

2014–15. We then model PIP as 50% rolled out in 2016–17, and fully rolled out in 2018–19.

We capture in our modelling the change in the combination of levels of award, and the associated movement of some claimants off the benefit altogether. The effect of any increased objectivity of the assessments and the removal of automatic entitlement for certain conditions is captured to the extent that it is reflected in the figures predicting claims at each level in DWP's impact assessment(Department for Work and Pensions, 2012), on which we base our assumptions. However, we are not able to capture the effect that the change in frequency of assessment might have on the persistence of benefit claims over time.

The DWP impact assessment contains projected claimant numbers in each possible rate combination in 2015–16 under both the DLA and PIP frameworks. We use these figures to infer the probability that a person who receives a particular level of DLA under the legacy system (modelled as described above) receives each level of PIP under the reformed system. We limit the complexity of this calculation by making the following assumptions:

1. Movement between care levels is independent of movement between mobility levels, and vice versa. We test the validity of this assumption by comparing the PIP claimant levels it implies with those reported in the impact assessment, and find that it holds well.
2. Everyone receiving the highest rate of the care component under DLA receives the higher level of the daily living component under PIP. This assumption is supported by the data in the impact assessment, subject to assumption 5 below.
3. Everyone receiving the lowest rate of the care component under DLA receives no daily living component under PIP. This assumption is based on the removal of a level of care award and the reduction in numbers receiving any care component when DLA projections are compared with PIP projections.
4. People only move by one level – e.g. no one moves from high mobility under DLA to no mobility under PIP. For the purpose of defining levels, we equate higher DLA mobility with enhanced PIP mobility, lower DLA mobility with standard PIP mobility, high DLA care with enhanced PIP daily living and middle DLA care with standard PIP daily living, and we assume (as in point 3) that lower DLA care has essentially been abolished.
5. People either remain on the combination of benefits they received under DLA or move to a lower level of benefit when transitioning to PIP. Note that this assumes that the amended assessment process doesn't lead to

anyone getting a higher level of benefit, or anyone who didn't get DLA getting PIP.

This set of assumptions gives us the set of transition probabilities shown in Table 3.

Table 3. Transition probabilities (%) used to model reform of DLA

	PIP rate combination								
DLA rate combination	No PIP receipt	No daily living, standard mobility	Standard daily living, no mobility	No daily living, enhanced mobility	Standard daily living, standard mobility	Enhanced daily living, no mobility	Enhanced daily living, standard mobility	Standard daily living, enhanced mobility	Enhanced daily living, enhanced mobility
Low care, no mobility	100	0	0	0	0	0	0	0	0
No care, low mobility	70	30	0	0	0	0	0	0	0
Low care, low mobility	70	30	0	0	0	0	0	0	0
Middle care, no mobility	12	0	88	0	0	0	0	0	0
No care, high mobility	0	27	0	73	0	0	0	0	0
Middle care, low mobility	8	3	62	0	27	0	0	0	0
Low care, high mobility	0	27	0	73	0	0	0	0	0
High care, no mobility	0	0	0	0	0	100	0	0	0
High care, low mobility	0	0	0	0	0	70	30	0	0
Middle care, high mobility	0	3	0	8	24	0	0	65	0
High care, high mobility	0	0	0	0	0	0	27	0	73

Source: Authors' calculations based on tables 1 & 2 of the DWP impact assessment (Department for Work and Pensions, 2012)

Note: We calculate, based on the above assumptions, the probability of moving 'down' one care or mobility level given the starting level: e.g. the impact assessment has 1,040,000 people receiving high mobility under DLA, in combination with any or no level of care, and 760,000 under PIP. This implies a probability of moving from high mobility (to, under our assumptions, standard PIP mobility) of 27%. We calculate similar probabilities for each mobility or care transition, using the difference in total claimants under the two scenarios as the number of people claiming neither component under PIP and maintaining the assumption of 100% transition from low care to no daily living and 100% transition from high care to enhanced daily living. Where people claim both components of DLA we combine the probabilities for care and mobility, assuming independence as noted above, to derive the full set of transition probabilities.

In order to implement the reform, we first use the models described above to simulate receipt of DLA and, where receipt occurs, the level received. In years where PIP is implemented we then generate another random number from a uniform distribution and combine this with the transition probabilities above, conditional on the simulated DLA level, to generate a PIP level. For example:

1. Person A has a probability of receiving DLA, conditional on the characteristics listed in the model specification below, of 0.3

2. Person A is given a randomly generated number of 0.22, which is less than 0.3, and so we model him as claiming DLA

3. Because person A claims DLA we predict, conditional on a range of characteristics, the cash amount of DLA he would report receiving in 2010. We compare this to the level cut offs generated from the combination of DWP and ELSA data as described above and determine a level of DLA receipt. In this case, say the level generated is 4: that is, the middle rate of the care component of DLA and no mobility component.

4. Because person A would be claiming middle care and no mobility, we can see from Table 3 that he has a 12% chance of receiving no PIP under the reforms, and an 88% chance of receiving the lower level of the daily living component (and, as under DLA, no mobility).

5. We give person A another uniformly generated random number, this time 0.75. Because $0.12 < 0.75 < (0.88 + 0.12)$ we give person A the level associated with the probability of 0.88: that is, the lower level of the daily living component.

6. When person A is passed to the TAXBEN model he will be given the level of income associated with that combination of PIP receipt in the year in question.

Introducing employment support allowance

ESA was introduced in October 2008 in place of IB for new claimants. The transfer of existing claimants from IB to ESA is currently underway. The data we have from ELSA are primarily data on IB receipt rather than ESA receipt, so we model IB receipt throughout the simulation and apply assumptions on top of this to account for the reform. We estimate two key parameters to allow us to model the reform: the percentage of IB claimants who are not entitled to ESA as a result of changes to the entitlement rules, and the proportion of ESA claimants placed in the support, as opposed to the work related activity, group.

It appears a reduction in claimant numbers of about 18% was anticipated prior to the reassessment programme (McInnes, 2012), and around 40% of new claimants and 15% of reassessed incapacity benefit claimants (Department for Work and Pensions, 2014) are being found fit for work. We take the most conservative (in terms of reduction in taxpayer support) of the three numbers available and model a reduction in the claimant rate of 15%.

DWP statistics show the results of reassessments for existing IB claimants, adjusted for appeals where the outcome is known, indicating that the percentage of those entitled to claim being placed in the support group rose steadily from early 2012 from around 40% to over 70%. Figures for new claimants suggest that around 30% of ESA recipients were placed in the support group in the earlier years of the rollout, rising sharply in mid 2011 to around 55% (Department for Work and Pensions, 2014). We make an assumption that 55% of entitled claimants are placed in the support group and 45% in the work related activity group (WRAG).

In order to determine the speed of the rollout as it affects the population we are modelling[40], we take DWP claimant data from the start of each financial year (Department for Work and Pensions, 2011)[41], and combine this with the fact that rollout is anticipated to be completed by the end of 2014. We therefore model the reform as 16% rolled out at the start of 2010–11, 35% at the start of 2012–13 and 78% at the start of 2014–15.

Model forms and specifications

The models described below are, with the exception of those determining the level of AA or DLA receipt, probit models. These binomial models determine the probability that an individual receives the benefit in the following period. In the majority of cases the models allow movement both onto and off the benefit. The older adult DLA model is the exception, where existing claimants can remain on the benefit or move off it, but no new claimants can move on. The model predicting DLA income for those in receipt of DLA, which is used as an interim step in modelling the level of DLA receipt, is a linear regression. The model predicting the level of AA receipt for those in receipt of AA is an ordered probit model.

[40] By this we mean that we account for the likelihood that the age group we are modelling are split between new claimants and stock claimants in different proportions than is the working age population as a whole.

[41] The data are drawn from the time series function of the tabulation tool.

As described below, the probabilities are scaled before the randomisation step to bring them into line with DWP data on claimant rates. The marginal effects in this section show the impact of explanatory variables on the outcome before any calibration takes place, and so represent an upper bound on the variation observed in the final model.

Incapacity benefits

The dependent variable in this model is a binary indicator of receipt of any of IB, ESA or SDA at $t + 2$ and is run for all individuals aged up to SPA–1 at time t who are not in full-time work at $t + 2$.

The explanatory variables include information on age, couple status, education, care receipt, health, lagged receipt of incapacity benefits interacted with years from SPA, region, working status, and lagged receipt of DLA.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo R^2 statistic for the model is 55.3%.

Sex has a statistically insignificant effect on the probability of benefit receipt, but women are 5ppts less likely to be in receipt of incapacity benefits than are men. Age is included as a continuous variable, and each year of age above 50 reduces the probability of receipt by 0.3ppts. Those in couples are less likely to be in receipt of IB, a reduction of 2ppts for men and 3ppts for women.

Health level at $t + 2$ is highly significant, with even a drop from best to good health increasing the probability of receipt by 4ppts for men and 3ppts for women. This rises to 9ppts for men and 10ppts for women in the worst health group. Care receipt at $t + 2$ increases the probability of receipt by up to 5ppts, with the effect of informal care receipt being highly significant.

Being in part-time work at $t + 2$ reduces the probability of benefit receipt by up to 9ppts and is highly significant, and part-time work at t decreases the probability by up to a further 2ppts. However, being in full-time work at t increases the probability of receipt at $t + 2$ by 4ppts, because the group on which the regression is estimated are not in full-time work at time t: this variable therefore indicates a recent transition out of full-time work.

Lagged receipt of an incapacity benefit, when interacted with the number of years before the individual reaches SPA, has a highly significant positive effect on the probability of claiming IB, increasing it by 1ppt for each additional year's distance from SPA. Lagged receipt of DLA is also highly significant, increasing the probability of IB receipt by 9ppts for men and 8ppts for women.

Disability living allowance: younger adults

The dependent variable in this model is a binary indicator of receipt of any level of DLA at $t + 2$ and is run for all individuals aged up to age 63 at time t.

The explanatory variables include information on age, couple status, education, care receipt, health, region, working status, lagged receipt of an incapacity benefit, and lagged receipt of DLA.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo R^2 statistic for the model is 64.2%.

Sex has a statistically insignificant effect on the probability of benefit receipt, but women are 0.7ppts more likely to be in receipt of DLA than are men. Age is included as a linear variable, and each year of age above 50 reduces the probability of receipt by 0.1ppt for men and 0.3ppts for women. Those in couples are around 2ppts less likely to be in receipt of DLA than are single people, a statistically significant effect.

Health level at $t + 2$ is highly significant, with a drop from best to good health increasing the probability of receipt by 3ppts for men and 2ppts for women, rising to 9ppts (men) and 7ppts (women) in the worst health group. Care receipt at $t + 2$ is highly significant, and increases the probability of receipt by up to 5ppts.

Being in work at $t + 2$ reduces the probability of receipt by up to 7ppts and is highly significant.

Lagged receipt of IB has significant effect on the probability of DLA receipt, increasing it by 3ppts for men but reducing it by 2ppts for women. Lagged receipt of DLA itself has a highly significant positive effect of up to 10ppts for men and 13ppts for women.

Disability living allowance: older adults

The dependent variable in this model is a binary indicator of receipt of any level of DLA at $t + 2$ and is run for all individuals aged over 63 at time t who are in receipt of DLA.

The explanatory variables include information on age, couple status, care receipt, health, and working status. Lagged receipt of DLA is included implicitly in the criteria for the regression.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo R^2 statistic for the model is 11.2%.

Sex has a statistically insignificant , but very large, effect on the probability of benefit receipt, with women being 42ppts more likely to remain on DLA than are men. Age is included as a categorical variable with the oldest age group being those aged 75 and above, and the reference group being those aged 63 or 64 at t. Men are up to 32ppts more likely to remain on DLA if aged between 65 and 74, a highly significant effect, but are only 9ppts more likely to remain on DLA after age 75. Women are up to 8ppts more likely to remain on DLA between 65 and 74, and the difference in the effect between women and men aged 70 to 74 is statistically significant. Women aged 75 and over are 13ppts less likely to remain on DLA than are those in the reference group.

Men in couples are 1ppt more likely than singles to remain on DLA, but women in couples are 1ppt less likely. Neither effect is statistically significant.

The effect of health on continued receipt of DLA is inconsistent and generally not statistically significant. More coherent, though still not statistically significant, is the effect of care receipt, with receipt of formal care increasing the probability of remaining in receipt of DLA by 9ppts for men and 22ppts for women; the figures for informal care are 9ppts and 14ppts respectively.

Being in work at $t + 2$ reduces the probability of remaining in receipt by up to 23ppts. We do not interact working status with sex in this regression because all women in the regression are above SPA in the estimation sample and the sample sizes are too small.

Attendance allowance

The dependent variable in this model is a binary indicator of receipt of either level of AA at $t + 2$ and is run for all individuals aged over 63 at time t who are not in receipt of DLA.

The explanatory variables include information on age, couple status, education, care receipt, health, region, working status, and lagged receipt of AA.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo R^2 statistic for the model is 35.7%.

Sex has a statistically insignificant effect on the probability of benefit receipt, but women are 5ppts less likely to be in receipt of AA than are men. 85 to 89 is the only age band to have a positive effect on probability of receipt for men (compared with the reference case of being age 63 or 64 at t)[42]. All age bands for women confer an increased probability of receipt in comparison to the reference case, but there is no clear trend in the magnitude of the effects. Those in couples are less likely to be in receipt of AA, but the effect is only 0.3ppts for men and 1.1ppts for women, and is not statistically significant.

Health level at $t + 2$ is highly significant, with a drop from best to good health increasing the probability of receipt by 4ppts for men and 1ppt for women, rising to 10ppts (men) and 9ppts (women) in the worst health group. Being in the worst two health groups at t also has a significant effect, increasing probability of receipt at $t + 2$ by up to a further 4ppts.

Receipt of informal care at both t and $t + 2$ has a significant effect, increasing the probability of receipt by up to 9ppts in total. Receipt of formal care has a larger effect on women's receipt of AA than it does on men's, increasing women's probabilities by up to 11ppts and men's by up to 4ppts (taking the effect of receipt at t and $t + 2$ together).

Being in work at $t + 2$ reduces the probability of receipt by up to 5ppts, although this effect is not statistically significant.

Lagged receipt of AA has a highly significant effect, increasing the probability of receipt by 13ppts for men and 14ppts for women.

Levels of award

Several of the benefits that we model are available at different levels depending on the circumstances of the claimant. This section sets out how we differentiate between these levels in our modelling.

Disability living allowance

As described earlier in this section we use an estimated value of DLA income to classify DLA recipients into levels of receipt, each of which relates to a single combination of care and mobility awards. As set out in the section on reform assumptions, below, we then apply transition probabilities to these levels to model the impact of the PIP reform. The model described in this section is that used to predict reported DLA income.

[42] Note that this does not imply that there is no correlation between receipt of AA and age, just that there is not a significant amount of variation which is explained by age not, for example, by health status or care receipt.

We run this regression on all DLA claimants aged up to 65 who are claiming DLA in the current period. The dependent variable is log reported DLA income. The model is a linear regression and is run within-period for all those aged below 66 who are modelled as being in receipt of DLA. The pseudo-R^2 statistic for the model is 25.8%.

The explanatory variables which have a significant effect on the outcome of the model are age, formal care receipt, having children at baseline, being in the worst health group, being in receipt of IB, time since last worked, and being in full-time work. Being in the highest wealth quintile has a material, but not statistically significant, negative effect, as does being in the least deprived IDAOPI quintile.

Attendance allowance

We use an ordered probit to predict whether recipients of AA receive the higher or the lower level of the benefit.

The specification contains information on age, education, health, working status, region and sex. None of the marginal effects are statistically significant, but several are large. Women are 22ppts more likely than men to receive the higher level, for example, and individuals with a partner who is still in work are up to 19ppts more likely to receive it. There is a generally diminishing probability of receiving the higher level with increasing age, with recipients aged 90 and over being up to 16ppts less likely to receive the higher level than those aged under 70.[43]

Carer's allowance

We run this regression on all individuals who we model as caring for at least 35 hours per week at $t + 2$ who also meet the policy criteria for receipt of CA. We simplify these requirements to being that individuals must be out of work at $t + 2$ or having no more than £100 per week of gross income from earnings.[44] The dependent variable in this model is a binary indicator of receipt of CA.

[43] Again, this is likely to indicate that the probability of receipt is explained less by age than by other factors (e.g. health and care) as age increases, rather than that older people are less likely to receive high level AA.

[44] As discussed previously we are not able to capture the benefit receipt of the person being cared for. We measure gross, rather than net, earnings (the policy rule is defined on net earnings) because we do not have access to net earnings at this point in the model. However, an income of £100 per week is well below the personal allowance and the likelihood of a person with gross

The explanatory variables include information on age, couple status, lagged receipt of CA, region, and partner's receipt of DLA or AA. We implicitly include caring intensity and working status in the criteria for the regression.

The variables are interacted by sex, and sex itself is also included. The full specification, and the marginal effects of each variable, are listed in the Appendix. The pseudo R^2 statistic for the model is 31.6%.

Sex has a statistically insignificant but extremely large effect on the probability receipt of CA, with women being 82ppts more likely than men to be in receipt. The probability of receipt generally decreases with increasing age, with the marginal effects of the age groups being as large as 38ppts for the oldest group of women, although few of the effects are statistically significant. Those in couples are significantly less likely to be in receipt of CA than their single counterparts, with the effect being a reduction in probability of 19ppts for men and 6ppts for women. This is an interesting result in view of the result that those in couples generally provide more high intensity care than singles.

A partner's receipt of care confers a positive effect on the probability of receipt of CA, of up to 9ppts for men partnered with a recipient of formal care. Similarly, a partner's receipt of disability benefits has a positive effect, increasing the probability of receiving CA by up to 9ppts for partners of those receiving DLA. However, the one highly significant explanatory variable is lagged receipt of CA, which increases the probability of receipt by 22ppts for men and 17ppts for women.

Calibration

In common with other social surveys,[45] disability benefit receipt is under-reported in ELSA. In order to ensure that the model outputs, especially in terms of the income distributions and poverty results, are as accurate as possible we make adjustments to the modelled claimant rate to bring it more into line with DWP's reported disability, incapacity and caring benefit claimant rates.

The process we use for this is similar to that used in the mortality module, but is a little more complicated. Ideally, we would follow the same process, simply scaling up each individual's raw modelled probability of benefit receipt by a

earned income of more than £100 having net earned income of less than £100 (and so qualifying for the benefit in a way we do not allow) is negligible.

[45] See, for example, table B.1 of the IFS report on Living Standards, Poverty and Inequality in the UK, which shows that the Family Resources Survey picks up only 77% of DLA spending (Cribb, Hood, Joyce, & Phillips, 2013, p. 141).

factor capturing the difference between the average raw modelled probability and the equivalent figure from DWP's claimant rate data. However, unlike in the mortality module, the average probability of benefit claims derived from the uncalibrated ELSA model is strongly influenced by a relatively small number of individuals with relatively large claimant probabilities. If these probabilities are scaled by the appropriate factor, we end up with a large number of 'probabilities' which are greater than one, and so, when these are capped at one, the new claimant rate does not match the administrative data. To circumvent this problem, we perform the process iteratively until the overall claimant level is close to the DWP information. This means that we lose some of the variation in the model in exchange for a more accurate overall claimant rate.

Take the following example. We have five individuals with probabilities (from the regression model specified on the ELSA data) of receiving a given benefit of 0.05, 0.05, 0.1, 0.15 and 0.8, giving a mean probability of 0.23. Imagine that the claimant rate we are aiming for is 0.46. By the same logic as used in the mortality module, this implies that we should multiply each raw probability by 2 (= 0.46 / 0.23). This would give probabilities of 0.1, 0.1, 0.2, 0.3 and 1.6. But since probabilities are bounded at 1, our actual mean claimant rate is the mean of 0.1, 0.1, 0.2, 0.3 and 1, i.e. 0.34 – significantly below our target rate.

We therefore implement an iterative step: we now calculate a new multiplying factor, being the ratio of the target value (0.46) and the new mean modelled value (0.34), giving a factor of 1.35. Multiplying all the probabilities again (and again capping them at a maximum value of 1) we get 0.135, 0.135, 0.27, 0.405 and 1, with a mean of 0.39. This process is repeated until the mean value becomes arbitrarily close to the target value (subject to sufficient numbers of individuals in the age and sex cell, and sufficient variation in the raw probabilities).

Having achieved a mean probability of receipt which matches our target claimant rate, we then apply the randomisation process that we apply in all modules to the adjusted probabilities to determine which individuals are modelled as actually receiving the benefit.

In the explanation above, we refer to the target rate as being derived from DWP claimant rate data. In fact, the process by which we derive it is somewhat more subtle.

In each year we need a 'target' probability to which to calibrate our model - or, more accurately, a set of target probabilities, one for each age and sex cell. In 2010–11, where we perform in-sample prediction on the baseline ELSA sample to boost the claimant rate, we can take the target rate almost directly from DWP claimant data: the numbers of claimants in the age sex cells are published for

England, and we combine this information with 2011 Census data to determine the claimant rate.[46] Similarly, we can take 2012 DWP data for the first year of the simulation. The one exception to this is the IB claimant rates, where we use 2008 claimant data for both 2010 and 2012 to avoid double-counting the effect of the reforms to the system on the claimant rate.

In the years following 2012, we use a more complex approach. There are no publicly available projections of future claimant rates at the level we would require, and it would be too limiting, in the sense that we would lose all the information from our model about the way in which the population's evolving health and labour supply is affecting claims, to assume a fixed claimant rate over the course of the model.

The process below, based on a toy example of three simulation periods, shows how we arrive at a target rate for each year:

1. We model claims reported in ELSA on the basis of a range of characteristics, making no adjustment to the claimant rate.

2. We run the simulation without making any adjustment to the reported claimant rate in any year. This is the shadow model, and it gives us a series of modelled claimant rates over the course of the simulation based only on ELSA data. Let the claimant rates the shadow model predicts in each period be 4%, 5% and 5.5%.

3. We use the shadow claimant rate in each year to work out the percentage increase in the claimant rate between periods. Between period one and period two it is 25% and between period two and period three it is 10%, based on the dummy figures in this example.

4. We start in period one (2010–11) with a target claimant rate from administrative data: say this is 10%.

5. We calculate the target claimant rate in period two by scaling the period one target rate by the shadow model's predicted change in the claimant rate: this gives us a target rate in period two of $10\% * (1 + 20\%) = 12\%$.

6. We calculate the target claimant rate in period 3 by scaling the period two target rate by the period two to period three change in the shadow model, getting a period three target of $12\% * (1 + 10\%) = 13.2\%$

[46] Data are generated by DWP's tabulation tool (Department for Work and Pensions, 2011). In each year and for each benefit we take the February data.

Having derived a target rate for every year, we run the ‘active’ model, scaling up the predicted probabilities to match the target rate as each period is simulated.

By combining the two models, scaling up the mean probabilities from the active model to match the previous period’s target rate adjusted by the percentage change in the claimant rate predicted by the shadow model, we achieve three important things: a claimant rate which is closer to reality than a pure model from the ELSA data would allow us to achieve, a claimant rate that evolves in a way that draws on the information our detailed model of health and care can provide, and a consistent life history for each person.[47]

[47] This approach may seem unduly complex. It is worth noting that we tried two other methods of meeting the claimant rate in each year, neither of which was successful:

1. Match the reported claimant rate in period 1, then run the model and increase the target rate in period two by the change in mean predicted probability between the two periods (i.e. essentially do what we describe above with two separate models but all in one go). This doesn’t work because a lot of the power in the model comes from reported receipt in the last period and, having boosted the incidence of receipt in period one, we thus inflate the predicted probabilities for the following year. Applying that growth to the target claimant rate strengthens that feedback process and the rate grows unfeasibly rapidly over time, to a rate of about 40% claimants in 2022–23 for younger males on DLA.

2. Match the claimant rate in the first iteration, then use the power that comes from lagged receipt in the model to carry that information through to later years. Unfortunately, this effect wears off in one or two periods, and we drop back down to an ELSA-level claimant rate. This could be because the ‘extra’ people who we put onto the benefit to match the target rates look less like the claimants reported in ELSA, so the model pushes them back off the benefit more quickly.

8. Financial data

Having run the demographic simulation we run an additional set of modules which calculate the gross income and wealth for each person and family (benefit unit) in each year. We consider income from earnings, pensions (both state and private), financial investments, and buy-to-let property, and we model wealth held in financial investments and property. The future pension incomes that we calculate are based on the assumption that individuals in work continue to accumulate pension wealth. This section outlines the assumptions used to evolve these income streams and wealth holdings between 2010–11 and 2022–23. We consider the modelling of wealth first in each section, as this generally gives rise to an income stream.

Property wealth

Property wealth in our model is made up of main residence wealth (the value of the house that the family lives in) and other property wealth (the value of any second or holiday homes, or buy-to-let properties).

We take the gross value of the main residence, the net value of any other property, and data on mortgages (including total mortgage debt and mortgage outstanding on the main residence) directly from ELSA[48], and use these to construct measures of gross and net main residence and other property wealth.

Mortgages

Having determined the number of mortgages held by the household, the outstanding capital on each, and the repayment term remaining for each, we calculate the monthly mortgage repayment as

$$repayment = \frac{capital * r * (1+r)^N}{(1+r)^N - 1}$$

where r is the annual interest rate (assumed in the model to be 6%) and N is the number of compounding periods (assumed to be the outstanding term of the mortgage). This calculation ensures that the mortgage capital is repaid at, but not before, the end of the mortgage term.

[48] We make some general assumptions to correct for inconsistencies in the data to arrive at a set of variables holding the value of each outstanding mortgage and the term remaining (such that no one holds negative mortgage debt) and gross and net values of the properties which are consistent with the mortgages outstanding (and are such that no one holds negative gross property wealth).

House price growth

We apply regional house price growth rates to gross main residence wealth, based on the region in which the household lives. We apply the equivalent England figure to other property wealth, as we cannot identify where the other property is located.

We allow gross property wealth to evolve in line with ONS data on regional house prices in the first few years of the simulation, and then combine this with OBR forecasts about future house price trends for later periods. Essentially, we allow regional house price growth rates after 2013–14 (Office for National Statistics, 2014) to experience the same percentage point change year-on-year as the OBR predicts for the UK until 2018 (Office for Budget Responsibility, 2014). We extrapolate the OBR numbers beyond 2018 to continue this process for the remaining years of the simulation. This means that the variation between growth rates in different regions in the later years of the simulation is a fixed effect equal to the difference in growth rates in 2013, the last year for which outturn data are available.

Net property wealth

We calculate net property wealth at time $t + 1$ as:

$$net_{t+1} = net_t + k_{t+1} + \Delta gross_{t+1}$$

where net is the net property value at the end of the period, k is the capital repaid within the period, and $\Delta gross$ is the change in gross wealth as a result of the change in house prices.

We calculate the capital repaid as

$$k_t = repayment - (K_{t-1} * r)$$

where K is the outstanding capital on the mortgage at the end of the period: i.e. as the element of the repayment which is not going towards paying of the newly accumulated interest.

Income from property

We take income from property as reported at in 2010–11, and grow this at the rate we use for gross other property values: i.e. our calculated growth rate for England as a whole. We imply, in making this assumption, that the growth in gross house prices is a reasonable proxy for the growth in rents.

Financial wealth

We model three components of net financial wealth: savings, investments, and non-mortgage debt. These variables are measured at the benefit unit level.

We use data from successive waves of ELSA to determine a factor by which to multiply each wealth holding, in order to estimate net financial wealth in each simulation period. The factors are determined as the mean percentage change between waves of ELSA in total wealth holdings within the wealth group and age band. We define the age band on the age of a reference person[49] in the benefit unit. We define the wealth bands on the basis of percentiles of the wealth distribution by age group as given in the tables below.

Note that these figures are the change in wealth holdings, not the return to the wealth holdings, and they are changes across two year periods rather than annual changes.

By using percentage changes we implicitly assert that anyone with a zero holding retains a zero holding for the rest of the simulation. This includes those with zero holdings in 2010–11 and those who pay off their debt or spend their wealth to reach zero in any simulation period.

Table 4. Factors for the growth of savings wealth

	Savings band			
	Quartile 1	**Quartile 2**	**Quartile 3**	**Quartile 4**
50-59	2%	0%	0%	13%
60-69	0%	0%	1%	8%
70-79	2%	1%	1%	5%
80+	7%	1%	0%	6%

We see that growth in savings wealth is generally negligible, with the exception of the wealth held in the richest quartile.

Table 5. Factors for the growth of investment wealth

	Investment band		
	None	**Below median**	**Above median**
50-59	0%	-4%	12%
60-69	0%	-1%	8%

[49] Taken to be the oldest person in a couple, or the only person in a single person benefit unit.

	Investment band		
	None	Below median	Above median
70-79	0%	-1%	12%
80+	0%	3%	0%

We see a similar pattern in the changes to investment wealth as we do in the changes to savings wealth, with the larger changes being concentrated at the top of the wealth distribution. Those with lower non-zero financial wealth appear to generally be reducing their wealth holdings, with the exception of the oldest age band, and those at the top end are increasing theirs.

Table 6. Factors for the growth of non-mortgage debt

	Debt band		
	None	Below median	Above median
50-59	0%	5%	6%
60-69	0%	6%	7%
70-79	0%	4%	10%
80+	0%	22%	12%

Debts are presented in the data as positive figures, so a negative percentage would equate to the paying off of debt. We see that all groups with non-zero debts increase their debt on average, with large percentage changes in some groups. However, is should be borne in mind that the debt holdings of the 80 and over age band with below median debt is in the order of a few hundred pounds, so the large percentage changes are not a cause for concern (or an indication that we are modelling older people as having access to substantially more credit than they would enjoy in reality).

Income from financial assets

We take reported income from financial assets in 2010–11 and grow it at the same rate as we grow wealth holdings: i.e. we assume that, if your wealth increases by x%, so too does your income from that wealth. For the purposes of the modelling we classify income from stocks and bonds as being from 'investments' and income from TESSAs, National Savings accounts, and other bank accounts as being from 'savings', growing them accordingly.[50]

State pensions

We assume that everyone in the model draws their state pension as soon as they reach state pension age, which we know to be an assumption that holds for the vast majority of individuals (Crawford & Tetlow, 2010), and that the decision to draw a private pension income is conditional on age and working status as described below.

Over half of the simulated individuals are already claiming state pension in 2010–11, but for those who are not we need to make a number of assumptions about their NI history to estimate the level of pension they will be entitled to when they reach SPA: for example, we assume that those who are currently working in 2010–11 have been in work since leaving full-time education, and those who are not in work were in work between leaving education and leaving their previous job. These criteria are by no means perfect, but they allow us to make consistent and reasonable assumptions about the life history of the individuals in the model.

Recent reforms to the state pension system mean that anyone reaching state pension age on or after 6th April 2016 will receive the single-tier state pension, as long as they satisfy the minimum qualifying period. We assume in our modelling that this system is in place from 2016–17.

Where people are already reporting a state pension income in 2010–11 we use this reported data in our modelling of gross income. However, we clean the data rather than taking it straight from the survey, to correct for common misreporting of other benefits (primarily disability and incapacity benefits) as state pension income.

Private pensions

In order to calculate an individual's DC pension fund - and hence future pension income - at the point at which he retires, we need to know not only at what rate the current fund will appreciate (assumed to be 2.5% real), but also how much the he will contribute to the fund during the simulation if he continues working.

---

[50] ELSA documentation defines savings as 'money invested in "safe" assets such as bank accounts, savings accounts, and cash ISAs'. It defines investments as 'money invested in "risky" assets, such as shares, bonds, stocks and shares ISAs or life insurance ISAs'.

From ELSA we know the value of contributions currently being made. For individuals who report these contributions as a monetary amount, we assume that in future years they contribute the same (nominal) cash amount. For individuals who report their contributions as a percentage of salary, we assume that in future years they contribute the same fraction of salary. Banks, Emmerson and Tetlow (2007) provide evidence that such assumptions would have been appropriate between 2002–03 and 2004–05.

For individuals who do not report the employee or employer contribution to their pension, we impute these conditional on sex, age and whether the scheme is an employer pension.

Private pension income

In modelling receipt of private pensions we need to make an assumption about the point at which people choose to draw their pension income in relation to their current age and working status. Figure 8 shows data from ELSA wave 5 on the proportion of people in each age group and working status who are receiving income from their private pension.

On the basis of this information we make the following simplifying assumptions (subject to membership of a private pension scheme):

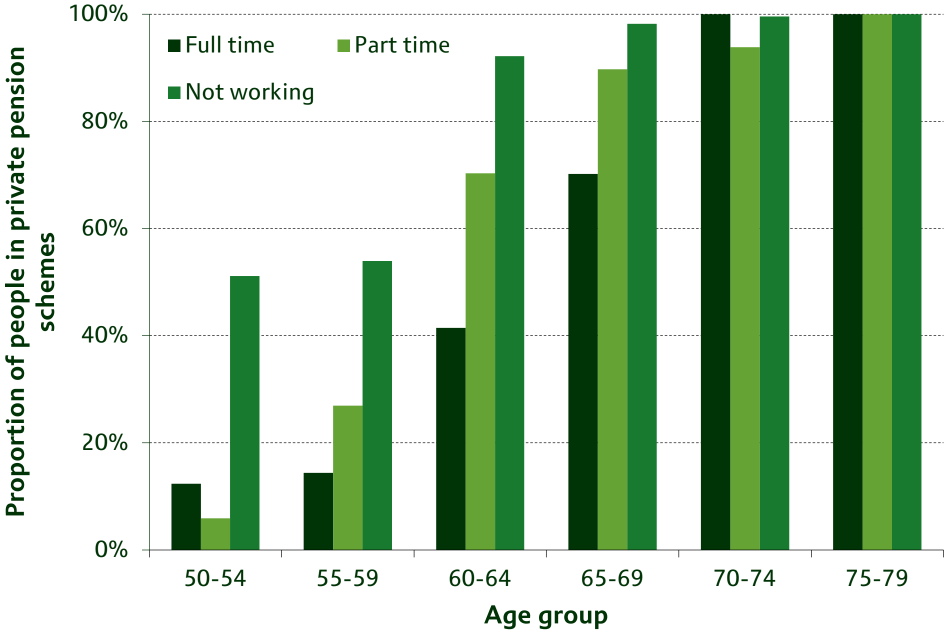
- Anyone aged 65 or over draws their private pension. This is because the majority of individuals with a private pension in this age band report receiving their private pension, and this is true of those working full-time as well as those working part-time and those not in paid work.

- Anyone aged 60 to 64 and in part-time work draws their private pension. This is because the majority of individuals with a private pension in this age band who are working part-time report receiving their private pension, whereas this is not true of those in this age group with a private pension who work full-time.

- Anyone below age 65 who is not working draws their private pension. This is because the majority of those with a private pension who are aged under 65 who are not in paid work report receiving their private pension.

These criteria imply that no working people below age 60 will draw a pension. As shown in Figure 5 only a minority of those who have a private pension and are in paid work and aged under 60 report receiving income from a private pension. We trigger pension receipt for anyone who meets the criteria in the first year of the model (2010–11) or who moves into a state which meets the criteria (either by ageing or changing working status) for the first time in a simulation period. Once someone has started to draw the pension he will continue to receive it until death, regardless of future labour market transitions.

Figure 8. Proportion of pension scheme members drawing a pension income by age and working status



Source: ELSA wave 5 data

In order to calculate the private pension income that an individual will receive in each simulation period, conditional on the year of retirement, we combine reported information from ELSA on pension scheme membership, pension contributions, and pension wealth with a series of single-life annuity rates.[51]

If an individual reports in 2010 that he receives an income from a private pension, we continue this throughout the simulation. If the amount he reports receiving is less than his full entitlement, we increase his income to the full entitlement when he meets one of the criteria above.

We do not explicitly limit members of DB pension schemes to meeting the NRA for their scheme before drawing a pension but, as discussed in the exploration of the marginal effects of the models, the variable indicating whether such people have met the NRA has a strong influence on labour supply decisions.

[51] Annuity rates were taken on 29th April 2014 from the Money Advice Service tables available at http://pluto.moneyadviceservice.org.uk/annuities, for a variety of retirement ages; the second best available rate was used in each case.

9. Net Incomes and Poverty

Creating net incomes

Using the IFS tax and benefit microsimulation model, TAXBEN, we can, using the gross incomes from the RetSim model, calculate the benefits and tax credits that individuals and households are entitled to, and the taxes that they are liable to pay, under hypothetical tax and benefit systems.[52] Hence, using the current default rules for annually uprating tax thresholds and benefit and tax credit amounts, and taking account of direct tax and benefit reforms that were announced in and before the 2014 Budget, we can simulate net household incomes in future years according to what the tax and benefit system will look like in those future years under current policies. This includes modelling the introduction of policies that are not yet in place, such as the introduction of a transferable income tax allowance for married couples.

We also model the gradual rollout of Universal Credit over the simulated period, in line with the timetable implied by Office for Budget Responsibility (2013). In 2014–15, we assume that no-one in our simulated population receives Universal Credit. In 2016–17, we assume that 5% of means-tested benefit claimants are new claimants, and so receive Universal Credit. Of the remaining 95%, we assume 25% have been rolled onto Universal Credit with transitional protection (they cannot lose in cash terms). From 2018–19 onwards, we assume Universal Credit is fully in place and all transitional protection has expired.

As discussed in sections 7 and 8 of this report, the receipt of disability benefits and the state pension is modelled by RetSim. For the remaining (relatively unimportant) non means-tested benefits, such as widows' and bereavement benefits, we use reported receipts directly from the ELSA data. Entitlement to means-tested benefits is calculated by TAXBEN. Throughout our modelling, we assume full take-up of means-tested benefits. This will lead us to overstate the total amount of benefit income received by the modelled population – for example, the take-up of pension credit in 2009 to 2010 was 73-80% by expenditure (62-68% by caseload).[53] As a result, our figures for average incomes will be somewhat too high, although the projected future trajectory should be

[52] See Giles and McCrae (1995) for details of the TAXBEN model.

[53] https://www.gov.uk/government/collections/income-related-benefits-estimates-of-take-up--2

broadly unaffected.[54] As discussed below, the calibration of our poverty line ensures that the assumption of full take-up does not affect the poverty level.

Calculating average incomes and poverty rates

Until this point in the modelling we have not weighted the data, either in conducting the simulation or in presenting the results (in Emmerson, Heald and Hood (2014) chapters 3 and 4). This is for a number of reasons: the fact that we include age, sex and region in all the regression specifications minimises the need to weight the estimation sample; the mechanics of the calibration process for mortality and, in particular, disability benefit receipt would be made significantly more complex by the inclusion of weights; and the presentation of the results by age and sex minimises the possibility of confusion as a result of compositional effects. When we present gross income and wealth results at a benefit unit level we also leave the results unweighted, because ELSA weights are given at an individual, not a family level. This makes it difficult to present family-level results (like housing wealth or family level gross income) in an informative way.

When calculating percentiles of the net income distribution, we do so using the weighted data. This is because the cross-sectional weights provided with the ELSA data make the sample representative of the 65 and over population of England, and so projected changes in statistics calculated on the weighted data should bear a closer resemblance to changes in the population as a whole. Using the weighted data reduces our effective sample size, since some of the individuals simulated are not core ELSA sample members, and so have a zero weight. Of the 9,645 individuals modelled, 1,074 have a zero weight, and so have no impact on our projections for net incomes and absolute poverty.

If net family income is to reflect the standard of living that individuals experience, and if we are to compare these incomes across different family types, then some method is required to adjust incomes for the different needs that different families face. We make this adjustment using the OECD equivalence scale shown in Table 7. Since children are relatively rare in our simulated population, the key assumption here is that a single individual needs two thirds of the income of a couple to enjoy the same standard of living.

[54] The effect of any increase or decrease in the generosity of means-tested benefits will be slightly overstated.

Table 7. Modified OECD equivalence scale

	Equivalence scale
First adult	0.67
Spouse	0.33
Child aged under 14	0.20
Child aged 14 and over	0.33

Our model projects a different measure of income to that used in calculating official poverty statistics: for example we can only model family income, not household income, and our incomes assume full take-up of means-tested benefits. In producing our absolute income poverty projections, we adjust for such differences by selecting as our poverty line the income level which gives the same poverty rate in our modelled population in 2010–11 as the official data (the Family Resources Survey) suggest for that age group in England.

Appendix: Full Model Specifications

In the main body of this document we present and discuss the more material or more interesting drivers of the outcomes of the models, as well as summarising the model specifications. In this section, for completeness, we present the full specifications and marginal effects for each model.

Recall that marginal effects show the change in percentage chance of an individual experiencing the specified outcome, given a positive value for the explanatory variable in question, holding all else constant. Almost all the variables listed are binary indicators, where 1 is the positive value (indicating that the individual possesses the characteristic); the rest are continuous variables (age, log earnings, time since last worked, years from SPA). The differential impacts of variables for men and women are the result of a specification which interacts all explanatory variables with sex. In interpreting the marginal effects for both sexes, the figure given should be compared with zero (i.e. compared against the figure for the reference case), not with the corresponding marginal effect for the opposite sex.

The significance stars in each table represent p values of < 0.001 (***), < 0.01 (**) and < 0.05 (*). A significance star in the female column indicates a significant difference in the effect on females to the effect on males, and the lack of a significance star in that column therefore does not mean that the explanatory variable has no significant impact on women.

Mortality

The marginal effects presented below show the change in the probability of death in the next two years (death by time $t + 2$, conditional on survival to time t) resulting from a positive value for the explanatory variable, holding all else constant.

Table A 1. Effect on probability of death in next two years

	Men		Women	
Age 55 - 59	0.36		1.29	
Age 60 - 64	0.20		2.01	
Age 65 - 69	0.99		3.60	
Age 70 - 74	1.59		4.23	
Age 75 - 79	3.57	***	5.74	
Age 80 - 84	4.83	***	6.49	
Age 85 - 89	5.01	***	8.20	

	Men		Women
Age 90 +	6.97	***	9.51
Baseline wealth quintile 2	0.55		-0.04
Baseline wealth quintile 3	0.32		-0.01
Baseline wealth quintile 4	0.32		0.14
Baseline wealth quintile 5	0.27		0.10
In a couple at t + 2	-1.16	***	-0.53
Angina diagnosis by age 50	0.43		-1.95
Heart attack diagnosis by age 50	-0.17		3.49
Diabetes diagnosis by age 50	0.59		2.02
Stroke diagnosis by age 50	-3.83		
Arthritis diagnosis by age 50	-0.59		-0.34
Cancer or malignant tumour diagnosis by age 50	2.32		2.40
Mental health diagnosis by age 50	-1.04		-0.22
Some qualifications at baseline	0.21		-0.45
Degree at baseline	0.37		0.47
Receives informal care at t	1.69	***	1.33
Receives formal care at t	2.93	***	2.55
Provides any care at t	-0.73	*	-0.62
Good health at t	0.50		-0.06
OK health at t	1.01	**	0.01
Poor health at t	2.21	***	0.57
Worst health at t	2.22	***	0.58
Good health at t - 2	-0.19		0.63
OK health at t - 2	-0.49		0.47
Poor health at t - 2	-0.35		0.47
Worst health at t - 2	-0.63		0.38
Health information missing at t - 2	-1.55	***	-0.50
Owns home (with or without mortgage)	-0.82		-0.47
IDAOPI quintile 2	-0.07		-0.27
IDAOPI quintile 3	0.38		-0.39
IDAOPI quintile 4	0.11		-1.06
IDAOPI quintile 5	-0.10		-0.79
IMD quintile 2	-0.35		0.93
IMD quintile 3	0.16		0.74
IMD quintile 4	-0.48		1.05
IMD quintile 5	-0.36		1.12
North West	0.46		-0.97
Yorkshire and The Humber	0.61		-0.25
East Midlands	0.41		-0.94

	Men		Women	
West Midlands	0.32		-0.63	
East of England	0.30		-0.46	
London	-0.22		-1.38	
South East	0.97		-0.29	
South West	0.66		-0.67	
Good childhood health	0.24		0.39	
OK childhood health	-0.42		0.44	
Poor childhood health	-0.33		-0.09	
Worst childhood health	-1.67		0.65	
Childhood health variable	-0.94		2.35	
Childhood health information missing	4.56	***	4.59	
Socio-economic classification: intermediate	-0.18		-0.09	
Socio-economic classification: manual, routine	-0.16		0.07	
Socio-economic classification: missing	-1.23		-0.31	
Sex			-3.83	*
On AA at t	-0.03		1.20	
On DLA at t	1.16		1.67	
On IB at t	-1.20		0.03	
Current smoker in 1998	2.00	***	1.74	
Ex-smoker (occasional) in 1998	0.19		0.26	
Ex-smoker (regular) in 1998	0.76		0.37	
Smoker information missing	-1.06		-2.13	
In part-time work at t	-1.99	***	-0.49	
In full-time work at t	-1.98	***	0.36	
Baseline income quintile 2	0.00		-0.09	
Baseline income quintile 3	0.33		-0.59	
Baseline income quintile 4	0.07		-0.40	
Baseline income quintile 5	-0.39		-0.01	
Baseline income quintile missing	0.37		1.10	

Health

The marginal effects presented below show the change in the probability of having the best (or worst) health status in two years' time resulting from a positive value for the explanatory variable, holding all else constant.

This model is an ordered probit, and the same specification therefore governs the probability of being at any of the five possible levels of health, although we present only two of the five sets of marginal effects here.

Table A 2. Effect on the probability of being in best health

	Men		Women
Age 55 - 59	-1.62		0.15
Age 60 - 64	-2.74		-2.18
Age 65 - 69	-3.81	**	-2.47
Age 70 - 74	-5.34	***	-4.75
Age 75 - 79	-10.67	***	-8.98
Age 80 - 84	-13.89	***	-11.66
Age 85 - 89	-19.64	***	-11.78
Age 90 +	-22.63	***	-16.20
Baseline wealth quintile 2	1.34		1.15
Baseline wealth quintile 3	1.66		2.96
Baseline wealth quintile 4	1.43		2.45
Baseline wealth quintile 5	3.93	**	3.93
In a couple at t + 2	1.72		2.64
Angina diagnosis by age 50	-2.97		-2.50
Heart attack diagnosis by age 50	-3.82		-0.49
Diabetes diagnosis by age 50	-6.71	***	-3.94
Stroke diagnosis by age 50	-10.15		-3.70
Arthritis diagnosis by age 50	-5.49	***	-5.87
Cancer or malignant tumour diagnosis by age 50	-6.37		-3.45
Mental health diagnosis by age 50	-2.75		-4.60
Some qualifications at baseline	1.08		0.54
Degree at baseline	2.03		0.34
Receives informal care at t	-9.32	***	-6.39
Receives formal care at t	-12.18	***	-7.05
Provides any care at t	-2.07	**	-0.56
Good health at t	-17.10	***	-16.66
OK health at t	-30.96	***	-29.86
Poor health at t	-48.18	***	-46.28

	Men		Women
Worst health at t	-61.34	***	-62.71
Good health at t - 2	-8.70	***	-10.01
OK health at t - 2	-14.50	***	-14.83
Poor health at t - 2	-15.23	***	-19.96
Worst health at t - 2	-22.17	***	-27.73
Health information missing at t - 2	-6.51	***	-11.60
Owns home (with or without mortgage)	1.45		-0.46
IDAOPI quintile 2	0.11		0.12
IDAOPI quintile 3	-0.14		-0.99
IDAOPI quintile 4	-0.20		-0.42
IDAOPI quintile 5	-0.06		-0.09
Leg length	-0.05		0.02
Leg length information missing	-2.22	*	-2.12
IMD quintile 2	0.78		-0.45
IMD quintile 3	0.08		-1.05
IMD quintile 4	-1.71		-0.01
IMD quintile 5	-2.77		-1.70
North West	-1.18		0.54
Yorkshire and The Humber	-1.60		1.47
East Midlands	-2.17		-2.20
West Midlands	-2.24		0.14
East of England	-3.44		-0.71
London	-3.98	*	-0.93
South East	-1.80		0.08
South West	-3.01		-0.43
Good childhood health	-2.02	*	-2.02
OK childhood health	-2.30	*	-3.10
Poor childhood health	-5.22	***	-4.57
Worst childhood health	-6.14	**	-4.77
Childhood health variable	-4.49		3.32
Childhood health information missing	-4.03	***	-2.06
Socio-economic classification: intermediate	-2.62		1.38
Socio-economic classification: manual, routine	-0.89		0.06
Socio-economic classification: missing	0.95		-0.07
Sex			-13.20
On AA at t	-7.65	***	-6.00
On DLA at t	-13.89	***	-10.64
On IB at t	4.84	*	-0.67
Current smoker in 1998	-3.32	***	-1.59
Ex-smoker (occasional) in 1998	-1.36		-1.24

	Men		Women
Ex-smoker (regular) in 1998	-2.44	**	-1.22
Smoker information missing	-2.20		2.66
In part-time work at t	2.15		2.23
In full-time work at t	2.37		2.05
Baseline income quintile 2	0.14		0.15
Baseline income quintile 3	0.26		-0.25
Baseline income quintile 4	0.90		0.96
Baseline income quintile 5	1.44		1.67
Baseline income quintile missing	1.43		1.74

Table A 3. Effect on the probability of being in worst health

	Men		Women
Age 55 - 59	0.44		-0.04
Age 60 - 64	0.75		0.60
Age 65 - 69	1.04	**	0.68
Age 70 - 74	1.46	***	1.30
Age 75 - 79	2.92	***	2.46
Age 80 - 84	3.80	***	3.19
Age 85 - 89	5.38	***	3.23
Age 90 +	6.20	***	4.44
Baseline wealth quintile 2	-0.37		-0.31
Baseline wealth quintile 3	-0.45		-0.81
Baseline wealth quintile 4	-0.39		-0.67
Baseline wealth quintile 5	-1.08	**	-1.08
In a couple at t + 2	-0.47		-0.72
Angina diagnosis by age 50	0.81		0.69
Heart attack diagnosis by age 50	1.04		0.13
Diabetes diagnosis by age 50	1.84	***	1.08
Stroke diagnosis by age 50	2.78		1.01
Arthritis diagnosis by age 50	1.50	***	1.61
Cancer or malignant tumour diagnosis by age 50	1.75		0.94
Mental health diagnosis by age 50	0.75		1.26
Some qualifications at baseline	-0.30		-0.15
Degree at baseline	-0.56		-0.09
Receives informal care at t	2.55	***	1.75
Receives formal care at t	3.33	***	1.93
Provides any care at t	0.57	**	0.15

	Men		Women
Good health at t	4.68	***	4.56
OK health at t	8.48	***	8.18
Poor health at t	13.19	***	12.67
Worst health at t	16.80	***	17.17
Good health at t - 2	2.38	***	2.74
OK health at t - 2	3.97	***	4.06
Poor health at t - 2	4.17	***	5.47
Worst health at t - 2	6.07	***	7.59
Health information missing at t - 2	1.78	***	3.18
Owns home (with or without mortgage)	-0.40		0.13
IDAOPI quintile 2	-0.03		-0.03
IDAOPI quintile 3	0.04		0.27
IDAOPI quintile 4	0.05		0.12
IDAOPI quintile 5	0.02		0.02
Leg length	0.01		-0.01
Leg length information missing	0.61	*	0.58
IMD quintile 2	-0.21		0.12
IMD quintile 3	-0.02		0.29
IMD quintile 4	0.47		0.00
IMD quintile 5	0.76		0.47
North West	0.32		-0.15
Yorkshire and The Humber	0.44		-0.40
East Midlands	0.59		0.60
West Midlands	0.61		-0.04
East of England	0.94		0.20
London	1.09	*	0.25
South East	0.49		-0.02
South West	0.82		0.12
Good childhood health	0.55	*	0.55
OK childhood health	0.63	*	0.85
Poor childhood health	1.43	***	1.25
Worst childhood health	1.68	**	1.31
Childhood health variable	1.23		-0.91
Childhood health information missing	1.10	***	0.56
Socio-economic classification: intermediate	0.72		-0.38
Socio-economic classification: manual, routine	0.24		-0.02
Socio-economic classification: missing	-0.26		0.02
Sex			3.62
On AA at t	2.09	***	1.64
On DLA at t	3.80	***	2.91

	Men		Women
On IB at t	-1.32	*	0.18
Current smoker in 1998	0.91	***	0.43
Ex-smoker (occasional) in 1998	0.37		0.34
Ex-smoker (regular) in 1998	0.67	**	0.33
Smoker information missing	0.60		-0.73
In part-time work at t	-0.59		-0.61
In full-time work at t	-0.65		-0.56
Baseline income quintile 2	-0.04		-0.04
Baseline income quintile 3	-0.07		0.07
Baseline income quintile 4	-0.25		-0.26
Baseline income quintile 5	-0.39		-0.46
Baseline income quintile missing	-0.39		-0.48

Care receipt

The marginal effects presented below show the change in the probability of receiving formal or informal care in two years' time resulting from a positive value for the explanatory variable, holding all else constant.

This model is a multinomial probit, and the same specification therefore governs the probability of receiving either type of, or no, care.

By construction, the sum of the effects in the three tables in this section is zero for each explanatory variable. We include the 'no care' table for completeness because it neatly summarises the impact of different variable on the probability of receiving any care: indeed, the increased and decreased probabilities of receiving any care reported in the model specification section are, respectively, the decreased and increased probabilities of receiving no care presented in this table.

Table A 4. Effect on the probability of receiving no care

	Men		Women
Age 55 - 59	-1.51		-1.01
Age 60 - 64	-1.44		-1.67
Age 65 - 69	-2.04		-3.22
Age 70 - 74	-4.39	**	-4.78
Age 75 - 79	-7.32	***	-7.91
Age 80 - 84	-11.15	***	-9.70
Age 85 - 89	-13.42	***	-13.49
Age 90 +	-20.29	***	-16.36
Baseline wealth quintile 2	-0.17		0.82
Baseline wealth quintile 3	0.29		1.07
Baseline wealth quintile 4	0.38		1.11
Baseline wealth quintile 5	1.05		2.58
In a couple at t + 2	-8.69	***	-6.76
Angina diagnosis by age 50	-2.31		-3.74
Arthritis diagnosis by age 50	-1.73		-1.10
Cancer or malignant tumour diagnosis by age 50	-0.10		-1.28
Mental health diagnosis by age 50	-1.41		-0.95
Some qualifications at baseline	0.22		0.70
Degree at baseline	1.64		0.74
Receives informal care at t	-14.76	***	-14.26
Receives formal care at t	-16.07	***	-11.26
Receives informal care at t - 2	-7.54	***	-7.22
Receives formal care at t - 2	-1.97		-5.76

	Men		Women
Care receipt information missing at t - 2	-4.67	***	-5.62
Provides any care at t	3.40	**	-0.13
Has children (inc grown-up) at baseline	1.37		-1.85
Good health at t	-0.74		-0.42
OK health at t	-2.11		-0.82
Poor health at t	-0.83		-0.65
Worst health at t	-3.05		0.04
Good health at t + 2	-11.06	***	-13.36
OK health at t + 2	-19.09	***	-22.04
Poor health at t + 2	-28.05	***	-30.35
Worst health at t + 2	-31.08	***	-35.17
Owns home (with or without mortgage)	-0.43		-0.88
Partner receives informal care at t	-0.25		0.28
Partner receives formal care at t	0.26		-2.51
Partner good health at t	0.50		1.00
Partner OK health at t	-0.24		0.50
Partner poor health at t	-0.98		1.19
Partner worst health at t	-1.13		2.08
Partner good health at t + 2	-1.25		-0.55
Partner OK health at t + 2	-1.16		-1.08
Partner poor health at t + 2	-0.01		-0.85
Partner worst health at t + 2	1.80		1.42
IMD quintile 2	0.20		-0.53
IMD quintile 3	0.47		-0.70
IMD quintile 4	0.08		-1.12
IMD quintile 5	-0.62		-2.21
North West	2.40		1.39
Yorkshire and The Humber	1.11		0.98
East Midlands	3.84	*	2.10
West Midlands	4.01	**	3.18
East of England	-0.49		1.17
London	3.26		2.64
South East	4.10	**	3.48
South West	2.26		2.15
Socio-economic classification: intermediate	0.37		-0.10
Socio-economic classification: manual, routine	-0.62		0.68
Socio-economic classification: missing	0.54		1.29
Sex			-0.30
On AA at t	-3.82		-4.95

	Men		Women
On DLA at t	-8.70	***	-5.53
On IB at t	4.54	*	-0.59
In part-time work at t	2.60		3.21
In full-time work at t	4.67	***	5.20
Baseline income quintile 2	-1.65		1.35
Baseline income quintile 3	-0.25		0.52
Baseline income quintile 4	-1.48		0.28
Baseline income quintile 5	-0.49		-0.40
Baseline income quintile missing	-0.45		-0.17

Table A 5. Effect on the probability of receiving informal care

	Men		Women
Age 55 - 59	0.46		0.78
Age 60 - 64	-0.24		0.66
Age 65 - 69	0.54		0.82
Age 70 - 74	2.69		1.48
Age 75 - 79	5.06	**	3.78
Age 80 - 84	7.72	***	3.69
Age 85 - 89	7.94	***	6.89
Age 90 +	14.38	***	7.96
Baseline wealth quintile 2	0.63		-1.03
Baseline wealth quintile 3	-0.07		-1.28
Baseline wealth quintile 4	-0.12		-1.83
Baseline wealth quintile 5	-1.28		-3.31
In a couple at t + 2	10.99	***	9.72
Angina diagnosis by age 50	2.29		5.82
Arthritis diagnosis by age 50	1.83		1.68
Cancer or malignant tumour diagnosis by age 50	-1.90		1.04
Mental health diagnosis by age 50	0.43		0.10
Some qualifications at baseline	-0.49		-1.22
Degree at baseline	-3.63	***	-1.45
Receives informal care at t	14.48	***	13.80
Receives formal care at t	10.42	***	5.26
Receives informal care at t - 2	6.85	***	6.30
Receives formal care at t - 2	-1.06		2.51
Care receipt information missing at t - 2	3.75	***	3.69
Provides any care at t	-1.74		0.97
Has children (inc grown-up) at baseline	0.40		2.75
Good health at t	0.27		0.06

	Men		Women
OK health at t	1.27		0.96
Poor health at t	-0.02		1.03
Worst health at t	2.14		0.43
Good health at t + 2	9.72	***	11.95
OK health at t + 2	18.00	***	19.32
Poor health at t + 2	24.07	***	25.90
Worst health at t + 2	26.54	***	29.08
Owns home (with or without mortgage)	0.33		0.90
Partner receives informal care at t	-0.25		-0.51
Partner receives formal care at t	-3.30		1.29
Partner good health at t	-1.20		-1.04
Partner OK health at t	-0.58		-0.20
Partner poor health at t	0.41		-1.12
Partner worst health at t	0.22		-3.53
Partner good health at t + 2	1.05		0.63
Partner OK health at t + 2	0.43		0.17
Partner poor health at t + 2	0.19		-0.75
Partner worst health at t + 2	-2.94		-3.37
IMD quintile 2	0.32		1.21
IMD quintile 3	-0.08		0.81
IMD quintile 4	0.28		1.65
IMD quintile 5	0.32		2.14
North West	-2.95		-0.38
Yorkshire and The Humber	-1.99		0.10
East Midlands	-5.67	***	-1.76
West Midlands	-5.09	***	-1.93
East of England	-0.78		-0.55
London	-2.93		-1.19
South East	-5.18	***	-2.63
South West	-3.65	*	-0.98
Socio-economic classification: intermediate	-1.20		0.21
Socio-economic classification: manual, routine	1.08		0.09
Socio-economic classification: missing	1.50		-0.61
Sex			0.73
On AA at t	2.15		3.87
On DLA at t	7.42	***	4.03
On IB at t	-2.31		0.31
In part-time work at t	-0.27		-2.91
In full-time work at t	-3.10	*	-4.35

	Men	Women
Baseline income quintile 2	1.35	-1.39
Baseline income quintile 3	0.42	-1.08
Baseline income quintile 4	0.25	-1.52
Baseline income quintile 5	0.32	-1.44
Baseline income quintile missing	-1.52	-2.11

Table A 6. Effect on the probability of receiving formal care

	Men		Women
Age 55 - 59	1.04		0.23
Age 60 - 64	1.68		1.01
Age 65 - 69	1.50		2.40
Age 70 - 74	1.70		3.30
Age 75 - 79	2.26		4.13
Age 80 - 84	3.43	**	6.01
Age 85 - 89	5.48	***	6.60
Age 90 +	5.91	***	8.41
Baseline wealth quintile 2	-0.46		0.21
Baseline wealth quintile 3	-0.22		0.21
Baseline wealth quintile 4	-0.25		0.72
Baseline wealth quintile 5	0.23		0.74
In a couple at t + 2	-2.30	***	-2.96
Angina diagnosis by age 50	0.02		-2.07
Arthritis diagnosis by age 50	-0.09		-0.58
Cancer or malignant tumour diagnosis by age 50	2.00		0.25
Mental health diagnosis by age 50	0.98		0.85
Some qualifications at baseline	0.27		0.51
Degree at baseline	1.99	***	0.72
Receives informal care at t	0.28		0.46
Receives formal care at t	5.66	***	6.00
Receives informal care at t - 2	0.69		0.92
Receives formal care at t - 2	3.02	***	3.25
Care receipt information missing at t - 2	0.92		1.92
Provides any care at t	-1.66	*	-0.83
Has children (inc grown-up) at baseline	-1.77	***	-0.90
Good health at t	0.47		0.35
OK health at t	0.84		-0.14
Poor health at t	0.85		-0.38
Worst health at t	0.91		-0.48
Good health at t + 2	1.34		1.41

	Men		Women
OK health at t + 2	1.10		2.71
Poor health at t + 2	3.98	***	4.45
Worst health at t + 2	4.54	***	6.09
Owns home (with or without mortgage)	0.10		-0.02
Partner receives informal care at t	0.49		0.23
Partner receives formal care at t	3.04	**	1.21
Partner good health at t	0.70		0.04
Partner OK health at t	0.82		-0.30
Partner poor health at t	0.57		-0.07
Partner worst health at t	0.91		1.45
Partner good health at t + 2	0.20		-0.08
Partner OK health at t + 2	0.72		0.91
Partner poor health at t + 2	-0.18		1.60
Partner worst health at t + 2	1.14		1.94
IMD quintile 2	-0.52		-0.67
IMD quintile 3	-0.39		-0.11
IMD quintile 4	-0.35		-0.53
IMD quintile 5	0.30		0.07
North West	0.55		-1.01
Yorkshire and The Humber	0.88		-1.08
East Midlands	1.83		-0.34
West Midlands	1.08		-1.25
East of England	1.26		-0.61
London	-0.33		-1.45
South East	1.08		-0.85
South West	1.38		-1.17
Socio-economic classification: intermediate	0.83		-0.10
Socio-economic classification: manual, routine	-0.46		-0.77
Socio-economic classification: missing	-2.04	**	-0.68
Sex			-0.43
On AA at t	1.67	**	1.07
On DLA at t	1.28		1.50
On IB at t	-2.23	*	0.28
In part-time work at t	-2.32		-0.30
In full-time work at t	-1.57		-0.85
Baseline income quintile 2	0.31		0.04
Baseline income quintile 3	-0.17		0.57
Baseline income quintile 4	1.23		1.25
Baseline income quintile 5	0.17		1.84

	Men	Women
Baseline income quintile missing	1.97	2.29

Care provision

The marginal effects presented below show the change in the probability of providing low or high intensity care (under 35 hours per week, or 35 or more hours per week) in two years' time resulting from a positive value for the explanatory variable, holding all else constant.

This model is an ordered probit, and the same specification therefore governs the probability of providing either intensity of, or no, care. Again, by construction, the sum of the marginal effect of a variable on providing no care, low intensity care, and high intensity care is zero.

Table A 7. Effect on the probability of providing no care

	Men		Women
Age 55 - 59	0.33		0.45
Age 60 - 64	0.02		2.14
Age 65 - 69	0.81		2.72
Age 70 - 74	0.19		2.86
Age 75 - 79	2.93		5.70
Age 80 - 84	2.37		8.00
Age 85 - 89	5.39		14.95
Age 90 +	15.78		9.65
Baseline wealth quintile 2	-0.30		0.26
Baseline wealth quintile 3	0.99		-0.37
Baseline wealth quintile 4	1.65		0.06
Baseline wealth quintile 5	1.97		-0.20
Provides low intensity care at t: in couple	11.21	***	7.45
Provides high intensity care at t: in couple	13.03	**	15.09
In a couple at t + 2	-5.59	***	-7.30
Some qualifications at baseline	-1.87	**	-1.25
Degree at baseline	-2.26		-1.04
Receives informal care at t	1.99		3.16
Receives formal care at t	4.24		6.19
Receives informal care at t + 2	-3.77	***	-0.94
Receives formal care at t + 2	4.61		-0.98
Provides low intensity care at t	-23.50	***	-21.92
Provides high intensity care at t	-38.29	***	-40.56
Has children (inc grown-up) at baseline	1.64		-0.35
Good health at t	0.51		0.13
OK health at t	-0.44		-0.29

	Men		Women
Poor health at t	1.24		-1.71
Worst health at t	3.40		1.98
Good health at t + 2	0.31		0.01
OK health at t + 2	1.50		0.62
Poor health at t + 2	4.79	***	2.49
Worst health at t + 2	7.51	***	6.77
Owns home (with or without mortgage)	-1.63		1.13
IDAOPI quintile 2	0.45		-0.03
IDAOPI quintile 3	-0.12		0.12
IDAOPI quintile 4	3.02	*	0.01
IDAOPI quintile 5	4.09	**	1.35
Partner dies by t+2 if alive at t	19.65	***	5.15
Partner receives informal care at t	-1.46		-0.49
Partner receives formal care at t	-5.28		-3.83
Partner good health at t	-1.84		-1.70
Partner OK health at t	-2.43	*	-1.96
Partner poor health at t	-4.90	***	-1.18
Partner worst health at t	-3.06		-2.05
Partner good health at t + 2	-9.74	***	-5.52
Partner OK health at t + 2	-19.53	***	-14.46
Partner poor health at t + 2	-26.67	***	-23.24
Partner worst health at t + 2	-31.32	***	-27.72
IMD quintile 2	-1.79		-0.19
IMD quintile 3	-2.02		0.61
IMD quintile 4	-2.31		0.73
IMD quintile 5	-3.01		-0.09
North West	-1.16		-1.18
Yorkshire and The Humber	-1.47		-1.04
East Midlands	0.35		-0.18
West Midlands	-0.22		-1.36
East of England	-2.44		-0.43
London	-1.74		-0.04
South East	-0.36		0.24
South West	-0.25		-0.02
Socio-economic classification: intermediate	0.59		1.00
Socio-economic classification: manual, routine	-0.60		1.13
Socio-economic classification: missing	-1.45		0.18
Sex			-7.48
In part-time work at t	1.25		1.71
In full-time work at t	4.16	***	3.17

	Men		Women
Baseline income quintile 2	0.91		-1.07
Baseline income quintile 3	0.54		-0.82
Baseline income quintile 4	0.23		-1.16
Baseline income quintile 5	0.81		0.80
Baseline income quintile missing	1.23		1.07

Table A 8. Effect on the probability of providing low intensity care

	Men		Women
Age 55 - 59	-0.24		-0.33
Age 60 - 64	-0.01		-1.56
Age 65 - 69	-0.59		-1.99
Age 70 - 74	-0.14		-2.09
Age 75 - 79	-2.14		-4.17
Age 80 - 84	-1.73		-5.85
Age 85 - 89	-3.94		-10.92
Age 90 +	-11.53		-7.04
Baseline wealth quintile 2	0.22		-0.19
Baseline wealth quintile 3	-0.72		0.27
Baseline wealth quintile 4	-1.21		-0.04
Baseline wealth quintile 5	-1.44		0.15
Provides low intensity care at t: in couple	-8.18	***	-5.44
Provides high intensity care at t: in couple	-9.51	**	-11.02
In a couple at t + 2	4.08	***	5.33
Some qualifications at baseline	1.37	**	0.91
Degree at baseline	1.65		0.76
Receives informal care at t	-1.46		-2.31
Receives formal care at t	-3.10		-4.52
Receives informal care at t + 2	2.75	***	0.69
Receives formal care at t + 2	-3.37		0.71
Provides low intensity care at t	17.16	***	16.01
Provides high intensity care at t	27.97	***	29.62
Has children (inc grown-up) at baseline	-1.20		0.26
Good health at t	-0.38		-0.09
OK health at t	0.32		0.21
Poor health at t	-0.91		1.25
Worst health at t	-2.48		-1.44
Good health at t + 2	-0.23		-0.01

	Men		Women
OK health at t + 2	-1.09		-0.46
Poor health at t + 2	-3.50	***	-1.82
Worst health at t + 2	-5.49	***	-4.94
Owns home (with or without mortgage)	1.19		-0.82
IDAOPI quintile 2	-0.33		0.02
IDAOPI quintile 3	0.09		-0.09
IDAOPI quintile 4	-2.20	*	-0.01
IDAOPI quintile 5	-2.99	**	-0.98
Partner dies by t+2 if alive at t	-14.35	***	-3.76
Partner receives informal care at t	1.06		0.36
Partner receives formal care at t	3.86		2.80
Partner good health at t	1.34		1.24
Partner OK health at t	1.77	*	1.43
Partner poor health at t	3.58	***	0.86
Partner worst health at t	2.24		1.50
Partner good health at t + 2	7.11	***	4.03
Partner OK health at t + 2	14.26	***	10.56
Partner poor health at t + 2	19.48	***	16.97
Partner worst health at t + 2	22.88	***	20.24
IMD quintile 2	1.31		0.14
IMD quintile 3	1.48		-0.45
IMD quintile 4	1.68		-0.54
IMD quintile 5	2.20		0.07
North West	0.84		0.86
Yorkshire and The Humber	1.07		0.76
East Midlands	-0.25		0.13
West Midlands	0.16		0.99
East of England	1.78		0.32
London	1.27		0.03
South East	0.26		-0.18
South West	0.18		0.01
Socio-economic classification: intermediate	-0.43		-0.73
Socio-economic classification: manual, routine	0.44		-0.83
Socio-economic classification: missing	1.06		-0.13
Sex			5.46
In part-time work at t	-0.91		-1.25
In full-time work at t	-3.04	***	-2.31
Baseline income quintile 2	-0.66		0.78
Baseline income quintile 3	-0.39		0.60
Baseline income quintile 4	-0.17		0.85

	Men	Women
Baseline income quintile 5	-0.59	-0.58
Baseline income quintile missing	-0.90	-0.78

Table A 9. Effect on the probability of providing high intensity care

	Men		Women
Age 55 - 59	-0.09		-0.12
Age 60 - 64	0.00		-0.58
Age 65 - 69	-0.22		-0.73
Age 70 - 74	-0.05		-0.77
Age 75 - 79	-0.79		-1.54
Age 80 - 84	-0.64		-2.16
Age 85 - 89	-1.45		-4.03
Age 90 +	-4.26		-2.60
Baseline wealth quintile 2	0.08		-0.07
Baseline wealth quintile 3	-0.27		0.10
Baseline wealth quintile 4	-0.45		-0.01
Baseline wealth quintile 5	-0.53		0.05
Provides low intensity care at t: in couple	-3.02	***	-2.01
Provides high intensity care at t: in couple	-3.51	**	-4.07
In a couple at t + 2	1.51	***	1.97
Some qualifications at baseline	0.51	**	0.34
Degree at baseline	0.61		0.28
Receives informal care at t	-0.54		-0.85
Receives formal care at t	-1.14		-1.67
Receives informal care at t + 2	1.02	***	0.25
Receives formal care at t + 2	-1.24		0.26
Provides low intensity care at t	6.34	***	5.91
Provides high intensity care at t	10.33	***	10.94
Has children (inc grown-up) at baseline	-0.44		0.10
Good health at t	-0.14		-0.03
OK health at t	0.12		0.08
Poor health at t	-0.33		0.46
Worst health at t	-0.92		-0.53
Good health at t + 2	-0.08		0.00
OK health at t + 2	-0.40		-0.17
Poor health at t + 2	-1.29	***	-0.67
Worst health at t + 2	-2.03	***	-1.82

	Men		Women
Owns home (with or without mortgage)	0.44		-0.30
IDAOPI quintile 2	-0.12		0.01
IDAOPI quintile 3	0.03		-0.03
IDAOPI quintile 4	-0.81	*	0.00
IDAOPI quintile 5	-1.10	**	-0.36
Partner dies by t+2 if alive at t	-5.30	***	-1.39
Partner receives informal care at t	0.39		0.13
Partner receives formal care at t	1.42		1.03
Partner good health at t	0.50		0.46
Partner OK health at t	0.66	*	0.53
Partner poor health at t	1.32	***	0.32
Partner worst health at t	0.83		0.55
Partner good health at t + 2	2.63	***	1.49
Partner OK health at t + 2	5.27	***	3.90
Partner poor health at t + 2	7.19	***	6.27
Partner worst health at t + 2	8.45	***	7.47
IMD quintile 2	0.48		0.05
IMD quintile 3	0.55		-0.17
IMD quintile 4	0.62		-0.20
IMD quintile 5	0.81		0.02
North West	0.31		0.32
Yorkshire and The Humber	0.40		0.28
East Midlands	-0.09		0.05
West Midlands	0.06		0.37
East of England	0.66		0.12
London	0.47		0.01
South East	0.10		-0.07
South West	0.07		0.01
Socio-economic classification: intermediate	-0.16		-0.27
Socio-economic classification: manual, routine	0.16		-0.31
Socio-economic classification: missing	0.39		-0.05
Sex			2.02
In part-time work at t	-0.34		-0.46
In full-time work at t	-1.12	***	-0.85
Baseline income quintile 2	-0.24		0.29
Baseline income quintile 3	-0.15		0.22
Baseline income quintile 4	-0.06		0.31
Baseline income quintile 5	-0.22		-0.21
Baseline income quintile missing	-0.33		-0.29

Labour supply

We create three models for labour supply decisions, conditional on the individual's current working status. We model nine 'transitions' between states when the decision to remain in the same state is included in each case.

By construction, the marginal effects within each group of three transitions (the three possible $t + 2$ statuses, given the fixed time t status) for each explanatory variable sum to zero.

Table A 10. Marginal effects: transitions from no work

	None to none			None to part-time			None to full-time		
	M		F	M		F	M		F
Age	0.4		0.3	-0.1		-0.3	-0.3	***	0.0
Age squared	0.0			0.0			0.0		
Baseline wealth quintile 2	-0.2		2.1	-0.2		-2.0	0.3		-0.1
Baseline wealth quintile 3	1.1		1.7	-0.9		-1.6	-0.2		-0.1
Baseline wealth quintile 4	-0.7		3.4	-0.3		-3.3	1.0		-0.1
Baseline wealth quintile 5	-0.3		1.6	1.2		-1.7	-0.9		0.1
Below NRA (for DB scheme) at t + 2	-61.5		37.0	70.4		-85.9	-8.9		48.9
Below SPA at t + 2	1.4		-3.1	-0.8		0.8	-0.7		2.3
Belongs to a DB scheme at baseline	56.3		18.9	-65.5		16.1	9.2		-35.0
Belongs to a DC scheme at baseline	-3.4		-4.4	0.8		3.7	2.6	***	0.7
In a couple at t + 2	1.9		2.8	-1.2		-0.6	-0.8		-2.1
Some qualifications at baseline	-1.8		-0.5	2.3		0.1	-0.5		0.4
Degree at baseline	-3.0		-4.0	3.7		2.2	-0.6		1.8
Ever self employed	-18.2	***	-19.5	10.4	***	9.1	7.9	***	10.4
Receives informal care at t + 2	-0.5		-0.1	1.0		-0.5	-0.5		0.6
Receives formal care at t + 2	81.5		30.0	-51.6		3.3	-29.9		-33.3

	None to none			None to part-time		None to full-time		
	M		F	M	F	M		F
Provides low intensity care at t	-0.2		0.2	1.1	-0.2	-0.9		0.1
Provides high intensity care at t	3.7		2.0	-0.5	0.1	-3.2		-2.2
Provides low intensity care at t + 2	0.6		0.3	-0.9	0.3	0.3		-0.6
Provides high intensity care at t + 2	1.0		2.6	-0.2	-1.8	-0.7		-0.8
Has outstanding mortgage at t + 2	-0.8		-1.8	-0.7	0.4	1.6	**	1.4
Good health at t + 2	-2.0		0.2	1.1	0.5	0.9		-0.7
OK health at t + 2	4.1		2.7	-3.5	-1.4	-0.6		-1.3
Poor health at t + 2	10.2	**	4.2	-7.1	-1.0	-3.1		-3.2
Worst health at t + 2	6.4		11.6	-3.6	-6.0	-2.8		-5.5
Owns home (with or without mortgage)	-0.1		-1.8	0.2	2.1	-0.1		-0.4
IDAOPI quintile 2	-0.4		0.3	0.1	0.3	0.3		-0.6
IDAOPI quintile 3	0.1		0.6	0.1	-0.2	-0.2		-0.4
IDAOPI quintile 4	1.7		-2.0	0.0	0.3	-1.7		1.7
IDAOPI quintile 5	1.0		0.2	0.3	-1.6	-1.4		1.5
Log potential full-time earnings at t + 2	0.3		0.7	0.1	-0.5	-0.5	**	-0.3
Partner below SPA at t + 2	-1.4		-2.0	0.3	1.0	1.1		1.0
Partner receives informal care at t + 2	2.2		-1.7	-3.3	0.4	1.1		1.2
Partner receives formal care at t + 2	78.0		26.3	-50.5	9.8	-27.5		-36.1
Partner good health at t + 2	-0.9		-0.6	0.8	0.8	0.2		-0.3
Partner OK health at t + 2	-1.3		-1.4	2.1	0.5	-0.8		0.8
Partner poor health at t + 2	-1.5		0.9	0.9	-0.8	0.7		-0.1
Partner worst health at t + 2	2.2		2.3	-1.9	-2.0	-0.3		-0.3
Partner log potential full-time earnings at t + 2	-0.1		0.0	0.0	0.0	0.1	**	0.0
Partner negative potential full-time earnings at t + 2	1.8		1.8	-2.8	-2.2	1.0		0.4

	None to none		None to part-time		None to full-time	
	M	F	M	F	M	F
Partner in part-time work at t	-1.6	-1.2	2.2	0.6	-0.5	0.7
Partner in full-time work at t	1.2	-1.8	-0.6	0.9	-0.6	0.9
IMD quintile 2	-1.8	-0.1	1.4	0.0	0.3	0.1
IMD quintile 3	-2.6	-0.4	2.0	-0.7	0.5	1.0
IMD quintile 4	-3.8	1.4	2.7	-0.6	1.1	-0.8
IMD quintile 5	0.1	-0.1	-1.8	0.1	1.7	-0.1
North West	1.3	-1.4	-1.5	0.8	0.1	0.6
Yorkshire and The Humber	1.6	-1.5	-1.1	-1.2	-0.5	2.7
East Midlands	0.7	-4.2	-0.5	0.3	-0.2	3.9
West Midlands	1.5	-1.1	-0.1	0.9	-1.4	0.2
East of England	0.3	-4.2	0.0	1.8	-0.3	2.4
London	2.8	-2.3	-3.9	-0.5	1.1	2.7
South East	2.9	-2.7	-2.6	0.9	-0.3	1.7
South West	1.4	-2.2	-0.8	1.4	-0.6	0.8
Socio-economic classification: intermediate	-2.1	-0.8	1.1	0.6	0.9	0.2
Socio-economic classification: manual, routine	-2.2	0.0	2.2	0.6	-0.1	-0.6
Socio-economic classification: missing	-2.0	1.4	0.3	0.1	1.7 *	-1.4
Sex		2.8		5.3		-8.1
On DLA at t	3.2	5.6	-1.6	-3.0	-1.6	-2.6
Time since last worked (years)	0.7 ***	0.4	-0.4 **	-0.2	-0.3 ***	-0.2
Negative potential full-time earnings at t + 2	1.8	6.3	3.9	-3.6	-5.7	-2.7

Table A 11. Marginal effects: transitions from part-time work

	Part-time to none		Part-time to part-time		Part-time to full-time	
	M	F	M	F	M	F
Age	-0.2	0.5	0.9	-0.2	-0.8 **	-0.4
Age squared	0.0		0.0		0.0	

	Part-time to none			Part-time to part-time			Part-time to full-time	
	M		F	M		F	M	F
Baseline wealth quintile 2	-5.6		-1.5	8.7		6.6	-3.0	-5.1
Baseline wealth quintile 3	-3.9		-2.9	3.9		6.9	0.0	-4.1
Baseline wealth quintile 4	-5.9		-0.3	7.8		2.8	-1.9	-2.4
Baseline wealth quintile 5	0.6		-0.2	3.6		4.4	-4.2	-4.2
Below NRA (for DB scheme) at t + 2	-9.3		-10.2	11.6		6.7	-2.3	3.5
Below SPA at t + 2	-15.7	***	-16.3	15.8	**	10.6	0.0	5.7
Belongs to a DB scheme at baseline	4.2		2.2	-6.2		0.0	2.0	-2.2
Belongs to a DC scheme at baseline	0.4		-7.6	-0.6		4.2	0.2	3.5
In a couple at t + 2	3.9		4.7	-1.8		-3.1	-2.1	-1.7
Some qualifications at baseline	2.7		-0.6	-4.7		1.3	2.0	-0.8
Degree at baseline	1.3		-3.5	-2.0		3.2	0.8	0.3
Ever self employed	-1.8		-2.0	-0.4		-0.7	2.2	2.6
Receives informal care at t + 2	11.1	*	5.8	-4.9		0.3	-6.3	-6.1
Receives formal care at t + 2	17.7		6.0	-36.7		-18.5	19.0	12.5
Provides low intensity care at t	4.6		-0.3	-3.6		0.4	-1.0	-0.1
Provides high intensity care at t	-2.8		2.2	6.4		-2.0	-3.6	-0.2
Provides low intensity care at t + 2	-0.9		1.7	4.0		-1.3	-3.1	-0.4
Provides high intensity care at t + 2	11.7		13.1	-5.5		-7.0	-6.2	-6.1
Has outstanding mortgage at t + 2	-9.0		-6.5	4.7		2.8	4.3	3.7
Good health at t + 2	-0.3		0.4	0.4		-2.3	-0.1	1.9
OK health at t + 2	-0.9		0.7	4.8		-3.2	-3.8	2.5
Poor health at t + 2	10.3		12.0	-17.9		-8.1	7.6	-3.9
Worst health at t + 2	6.9		18.4	-3.7		-15.1	-3.1	-3.4
Owns home (with or without mortgage)	1.9		1.4	2.9		-0.4	-4.8	-1.1
IDAOPI quintile 2	2.6		-0.5	-2.1		3.7	-0.6	-3.2
IDAOPI quintile 3	-0.5		-0.3	-0.3		1.9	0.9	-1.7
IDAOPI quintile 4	-6.1		-0.1	11.3		3.0	-5.2	-3.0

	Part-time to none			Part-time to part-time			Part-time to full-time	
	M		F	M		F	M	F
IDAOPI quintile 5	-4.0		-4.1	7.2		3.1	-3.1	1.0
Log potential full-time earnings at t + 2	0.1		-3.8	0.7		1.1	-0.8	2.7
Partner below SPA at t + 2	-6.7		1.2	1.3		-2.7	5.4	1.5
Partner receives informal care at t + 2	4.0		-5.8	-11.6		4.7	7.6	1.1
Partner receives formal care at t + 2	-9.0		-26.1	5.8		20.9	3.2	5.2
Partner good health at t + 2	-3.3		-3.7	6.9		3.6	-3.6	0.1
Partner OK health at t + 2	-3.7		-2.1	5.5		3.4	-1.7	-1.3
Partner poor health at t + 2	-6.3		0.4	10.6		2.1	-4.3	-2.5
Partner worst health at t + 2	-13.5		-2.1	17.6		1.8	-4.1	0.2
Partner log potential full-time earnings at t + 2	-0.4	***	-0.1	0.5	***	0.2	-0.1	-0.1
Partner negative potential full-time earnings at t + 2	-6.8		0.6	7.8		-3.0	-0.9	2.4
Partner in part-time work at t	-2.0		3.2	-0.4		-4.7	2.5	1.6
Partner in full-time work at t	3.5		1.4	-2.3		0.0	-1.2	-1.5
IMD quintile 2	3.1		1.1	-1.4		-3.2	-1.8	2.1
IMD quintile 3	7.3		-0.9	-5.5		-3.4	-1.8	4.3
IMD quintile 4	10.5		2.3	-12.9		-4.6	2.4	2.3
IMD quintile 5	14.7		2.5	-10.2		-3.0	-4.5	0.5
North West	-3.1		-2.0	6.1		-0.7	-3.0	2.6
Yorkshire and The Humber	0.4		-5.8	3.4		3.1	-3.8	2.7
East Midlands	-0.5		-4.8	0.9		-1.4	-0.4	6.3
West Midlands	6.1		-0.4	-7.3		-3.0	1.2	3.4
East of England	-2.6		-3.2	4.0		1.5	-1.4	1.7
London	0.5		-1.4	5.0		-3.7	-5.5	5.2
South East	1.2		-3.5	2.5		0.1	-3.7	3.4
South West	1.5		-4.8	1.7		-2.0	-3.2	6.9

	Part-time to none				Part-time to part-time				Part-time to full-time			
	M		F		M		F		M		F	
Socio-economic classification: intermediate	-1.4		-2.8		5.5		6.1		-4.1		-3.4	
Socio-economic classification: manual, routine	1.3		-6.4		-1.0		8.2		-0.3		-1.8	
Socio-economic classification: missing	-3.9		-3.2		3.1		2.7		0.9		0.5	
Sex			7.5				34.4				-41.8	***
On DLA at t	-10.1		9.7		20.0		-7.8		-9.9		-1.9	
Negative potential full-time earnings at t + 2	17.7		-17.9		-7.1		6.9		-10.6		11.1	

Table A 12. Marginal effects: transitions from full-time work

	Full-time to none				Full-time to part-time				Full-time to full-time			
	M		F		M		F		M		F	
Age	1.4	***	1.1		0.4		0.0		-1.8	***	-1.1	
Age squared	0.0	***			0.0				0.0	**		
Baseline wealth quintile 2	1.5		-3.8		2.9		-0.6		-4.4		4.3	
Baseline wealth quintile 3	0.8		1.3		0.3		-0.2		-1.1		-1.1	
Baseline wealth quintile 4	1.5		-0.9		1.5		-1.5		-3.0		2.5	
Baseline wealth quintile 5	2.9		0.6		2.5		-0.4		-5.4		-0.1	
Below NRA (for DB scheme) at t + 2	-10.9	***	-4.2		-5.2		-3.7		16.1	***	7.9	
Below SPA at t + 2	-12.0	***	-12.3		-7.5	***	-3.9		19.5	***	16.2	
Belongs to a DB scheme at baseline	12.1	***	3.2		3.0		-0.4		-15.1	***	-2.8	
Belongs to a DC scheme at baseline	-2.5		0.7		-3.4	*	-2.4		5.9	***	1.7	
In a couple at t + 2	-1.4		2.9		5.7	**	1.3		-4.4		-4.2	
Some qualifications at baseline	-1.4		-0.9		2.3		-1.2		-0.9		2.2	
Degree at baseline	-1.8		-1.1		7.1	***	0.7		-5.3		0.4	
Ever self employed	-0.4		2.6		1.4		2.6		-1.0		-5.1	
Receives informal care at t + 2	6.4	***	9.7		1.6		-1.2		-8.1	***	-8.5	

	Full-time to none			Full-time to part-time			Full-time to full-time		
	M		F	M		F	M		F
Receives formal care at t + 2	8.0		35.0	6.9		-134	-14.9		98.5
Provides low intensity care at t	-2.1		3.0	0.1		-1.9	2.0		-1.1
Provides high intensity care at t	0.9		-0.2	-6.0		0.5	5.1		-0.3
Provides low intensity care at t + 2	3.5		0.3	-0.4		2.4	-3.1		-2.7
Provides high intensity care at t + 2	9.0		11.8	7.6		3.8	-16.6	***	-15.5
Has outstanding mortgage at t + 2	-3.4	***	-5.1	-1.3		-0.9	4.7	***	6.0
Good health at t + 2	1.5		-2.4	1.3		1.8	-2.8		0.7
OK health at t + 2	6.8	***	-1.4	-2.7		1.9	-4.1		-0.5
Poor health at t + 2	12.3	***	5.2	-0.9		-1.1	-11.4		-4.1
Worst health at t + 2	20.3	***	8.2	4.0		-0.8	-24.2	***	-7.4
Owns home (with or without mortgage)	0.1		1.8	5.3		2.4	-5.4		-4.2
IDAOPI quintile 2	-1.7		-2.5	1.9		-1.0	-0.1		3.5
IDAOPI quintile 3	-1.8		-2.8	-2.0		-1.9	3.8		4.6
IDAOPI quintile 4	-3.1		-2.3	-0.2		-1.9	3.3		4.2
IDAOPI quintile 5	-5.2		-6.2	-0.3		-1.9	5.5		8.1
Log potential full-time earnings at t + 2	-0.9		-2.4	-0.6		-1.9	1.5	*	4.3
Partner below SPA at t + 2	-0.8		0.5	-4.1	***	-0.4	5.0	***	-0.1
Partner receives informal care at t + 2	-2.0		0.9	-1.4		0.5	3.4		-1.4
Partner receives formal care at t + 2	-9.4		-4.0	-0.4		5.6	9.8		-1.6
Partner good health at t + 2	-2.1		2.1	0.5		-0.1	1.6		-2.0
Partner OK health at t + 2	-0.7		1.3	0.2		-2.2	0.5		0.9

	Full-time to none				Full-time to part-time				Full-time to full-time			
	M		F		M		F		M		F	
Partner poor health at t + 2	0.6		-0.5		0.9		-3.4		-1.5		3.9	
Partner worst health at t + 2	-5.9		-13.3		4.8		-0.9		1.1		14.2	
Partner log potential full-time earnings at t + 2	0.0		-0.2		-0.1		0.0		0.1		0.1	
Partner negative potential full-time earnings at t + 2	-0.2		0.9		1.1		1.3		-0.9		-2.1	
Partner in part-time work at t	0.2		4.9		4.2	**	4.5		-4.4	*	-9.4	
Partner in full-time work at t	-3.8	*	-0.8		-1.0		0.6		4.8	*	0.3	
IMD quintile 2	0.1		0.5		-0.2		0.2		0.1		-0.6	
IMD quintile 3	1.4		1.0		0.7		-0.3		-2.0		-0.7	
IMD quintile 4	4.5		2.6		1.5		-0.8		-6.0		-1.8	
IMD quintile 5	3.8		1.3		4.1		1.8		-7.9		-3.1	
North West	-0.7		2.3		5.7		1.5		-5.0		-3.8	
Yorkshire and The Humber	-0.5		2.8		0.8		5.9		-0.3		-8.7	
East Midlands	-1.3		0.6		4.5		4.4		-3.3		-5.0	
West Midlands	-0.7		1.6		2.7		0.8		-1.9		-2.4	
East of England	-0.7		2.5		3.6		3.9		-2.9		-6.4	
London	-1.3		0.7		4.5		3.9		-3.2		-4.6	
South East	-0.3		4.2		3.2		1.2		-2.9		-5.4	
South West	-0.7		-0.7		6.8		4.5		-6.1		-3.7	
Socio-economic classification: intermediate	-2.6		-0.2		3.7		-3.0		-1.1		3.1	
Socio-economic classification: manual, routine	0.7		-0.6		-2.5		-1.5		1.8		2.1	
Socio-economic classification: missing	1.0		-3.5		-2.2		0.1		1.2		3.4	
Sex			5.5				25.7	***			-31.2	***
On DLA at t	6.5		7.1		2.7		9.1		-9.1		-16.2	
Negative potential full-time earnings at t + 2	0.8		-14.8		-0.2		-4.9		-0.6		19.6	

Disability benefits

The marginal effects presented below show the change in the probability of receiving the specified benefit in two years' time resulting from a positive value for the explanatory variable, holding all else constant. In some cases, as described in section 7, explanatory variables are omitted from the model specification because we condition on them before running the model (e.g. only those who are not in full time work at $t+2$ are given a probability of receiving incapacity benefits, and so full time work at $t+2$ is not included in the specification).

Table A 13. Effect on the probability of receiving an incapacity benefit

	Men		Women
Age	-0.33	***	-0.30
In a couple at t + 2	-1.91		-3.34
Some qualifications at baseline	-0.43		-1.04
Degree at baseline	-3.28	*	-2.51
Receives informal care at t	-0.69		-0.45
Receives formal care at t	-6.33		0.04
Receives informal care at t + 2	3.21	***	4.03
Receives formal care at t + 2	3.11		4.82
Good health at t	1.69		1.71
OK health at t	2.25		1.08
Poor health at t	0.90		0.46
Worst health at t	1.67		0.25
Good health at t + 2	4.01	***	3.05
OK health at t + 2	5.97	***	6.20
Poor health at t + 2	8.21	***	7.72
Worst health at t + 2	9.09	***	9.86
Receives IB at t, interacted with years from SPA	1.40	***	1.32
North West	-1.14		-0.42
Yorkshire and The Humber	-2.03		0.60
East Midlands	-0.02		-0.81
West Midlands	-2.29		1.88
East of England	-2.53		-0.45
London	-1.02		0.93
South East	-2.34		-0.21
South West	-1.26		0.94
Sex			-4.58
On DLA at t	9.19	***	7.65
In part-time work at t	-2.09		1.25

	Men		Women
In full-time work at t	3.83	***	4.47
In part-time work at t + 2	-6.27	***	-9.38

Table A 14. Effect on the probability of receiving DLA (younger adults)

	Men		Women
Age	-0.14	*	-0.28
In a couple at t + 2	-1.78	***	-1.86
Some qualifications at baseline	-0.56		-0.27
Degree at baseline	-1.81		-1.00
Receives informal care at t	0.91		0.37
Receives formal care at t	3.09		0.03
Receives informal care at t + 2	4.07	***	3.78
Receives formal care at t + 2	4.81	**	5.22
Good health at t	0.44		0.52
OK health at t	1.21		-0.04
Poor health at t	0.22		0.65
Worst health at t	-0.11		1.07
Good health at t + 2	2.61	***	1.56
OK health at t + 2	4.15	***	2.64
Poor health at t + 2	5.97	***	4.24
Worst health at t + 2	9.43	***	6.78
Receives DLA at t	9.75	***	12.88
Partner in part-time work at t + 2	1.40		-3.27
Partner in full-time work at t + 2	0.04		-0.75
North West	0.89		-0.10
Yorkshire and The Humber	0.59		-0.10
East Midlands	1.94		-1.20
West Midlands	0.84		0.83
East of England	-0.18		-0.71
London	0.91		-0.51
South East	0.60		-0.43
South West	0.68		0.00
Sex			0.70
On IB at t	3.17	***	-1.78
In part-time work at t	0.55		0.28
In full-time work at t	1.78	**	0.35
In part-time work at t + 2	-4.07	***	-4.90
In full-time work at t + 2	-6.52	***	-3.95

DLA income is imputed (for the purpose of classifying recipients into levels) once receipt of DLA at $t + 2$ has been determined. This regression should therefore be viewed as modelling an income at time t, which is why there are no $t + 2$ explanatory variables.

Table A 15. Effect on reported income from DLA (younger adults)

	Men		Women
Age	-2.20	**	-0.99
Age squared	0.11		
Baseline wealth quintile 2	0.52		-3.08
Baseline wealth quintile 3	8.55		-3.53
Baseline wealth quintile 4	-4.64		1.84
Baseline wealth quintile 5	-18.15		-3.54
Below SPA at t + 2	2.82		3.52
Angina diagnosis by age 50	10.50		20.30
Heart attack diagnosis by age 50	-15.93		4.89
Diabetes diagnosis by age 50	-5.20		1.00
Stroke diagnosis by age 50	27.82		9.19
Arthritis diagnosis by age 50	-3.41		6.27
Cancer or malignant tumour diagnosis by age 50	14.79		18.95
Mental health diagnosis by age 50	-1.93		-5.70
Some qualifications at baseline	-4.21		-2.37
Degree at baseline	-3.34		-0.28
Receives informal care at t	9.04		-0.32
Receives formal care at t	25.18	**	25.18
Receives informal care at t - 2	2.82		8.84
Receives formal care at t - 2	26.29		21.98
Care receipt information missing at t - 2	3.62		6.06
Provides any care at t	-6.18		1.29
Has children (inc grown-up) at baseline	11.87	**	1.70
Good health at t	-8.28		5.83
OK health at t	0.63		4.81
Poor health at t	9.24		17.33
Worst health at t	17.88	**	14.83
Good health at t - 2	-9.19		-2.30
OK health at t - 2	-10.48		-4.15
Poor health at t - 2	-0.01		-9.42
Worst health at t - 2	1.79		4.28
IDAOPI quintile 2	-8.42		3.30
IDAOPI quintile 3	-4.77		1.27

	Men		Women
IDAOPI quintile 4	-7.22		10.33
IDAOPI quintile 5	-19.88		16.23
Leg length	0.01		-0.08
Leg length information missing	-11.12	**	16.31
Partner in part-time work at t	-4.45		0.62
Partner in full-time work at t	-4.43		-5.26
IMD quintile 2	-5.55		7.81
IMD quintile 3	-7.84		-0.16
IMD quintile 4	-5.24		-4.69
IMD quintile 5	-0.23		-18.24
North West	2.67		-1.45
Yorkshire and The Humber	-0.33		-3.19
East Midlands	-1.56		-3.46
West Midlands	5.05		-5.52
East of England	0.72		2.59
London	-4.66		-3.01
South East	11.81		5.72
South West	6.04		0.41
Sex			-28.08
On IB at t	62.30	***	67.08
Current smoker in 1998	-4.81		8.21
Ex-smoker (occasional) in 1998	-10.10		2.42
Ex-smoker (regular) in 1998	-2.93		8.11
Smoker information missing	-2.90		0.12
Time since last worked (years)	1.38	***	-0.33
In part-time work at t	2.75		-17.01
In full-time work at t	46.98	***	26.87

Table A 16. Effect on the probability of receiving DLA (older adults)

	Men		Women	
Age 65 - 69	25.52	***	7.56	
Age 70 - 74	31.48	***	6.05	*
Age 75 +	8.58		-13.14	
In a couple at t + 2	1.14		-0.96	
Receives informal care at t + 2	8.96		13.60	
Receives formal care at t + 2	8.85		21.60	
Good health at t	-7.63		-11.90	
OK health at t	-13.39		-2.41	
Poor health at t	-4.93		-4.96	

	Men		Women
Worst health at t	-8.49		-1.51
Good health at t + 2	23.31		-14.67
OK health at t + 2	16.57		-5.20
Poor health at t + 2	29.38	**	-7.31
Worst health at t + 2	23.63		-5.05
Sex			41.52
In part-time work at t + 2	-20.09		
In full-time work at t + 2	-23.21		

Table A 17. Effect on the probability of receiving AA

The level of AA received is modelled once receipt of AA at $t + 2$ has been determined. This regression should therefore be viewed as modelling an outcome at time t, which is why there are no $t + 2$ explanatory variables.

	Men		Women
Age 65 - 69	-2.79		1.87
Age 70 - 74	-1.20		2.31
Age 75 - 79	-0.10		3.55
Age 80 - 84	-0.61		3.58
Age 85 - 89	1.21		1.82
Age 90 +	-3.07		1.57
In a couple at t + 2	-0.29		-1.07
Some qualifications at baseline	-0.95		-1.39
Degree at baseline	-3.34	**	-2.92
Receives informal care at t	2.91	***	1.91
Receives formal care at t	1.36		4.22
Receives informal care at t + 2	6.23	***	6.49
Receives formal care at t + 2	2.88		6.72
Good health at t	1.15		-0.30
OK health at t	1.68		2.04
Poor health at t	4.24	***	1.87
Worst health at t	4.38	**	2.63
Good health at t + 2	3.93	***	1.17
OK health at t + 2	4.91	***	3.03
Poor health at t + 2	8.36	***	6.59
Worst health at t + 2	10.40	***	8.57
On AA at t	12.81	***	14.24
North West	0.24		3.17

	Men	Women
Yorkshire and The Humber	-0.72	-1.24
East Midlands	-1.39	-0.59
West Midlands	1.39	1.27
East of England	-1.95	0.89
London	-0.32	-0.85
South East	-1.72	0.08
South West	-0.26	0.96
Sex		-4.67
In part-time work at t	-4.98	-2.53
In full-time work at t	-1.52	0.39

Table A 18. Effect on the probability of receiving higher level AA, given any receipt

	Men	Women
Age 70 - 74	0.17	-10.56
Age 75 - 79	-0.24	-5.26
Age 80 - 84	-1.20	-13.72
Age 85 - 89	-9.53	-17.54
Age 90 +	-14.81	-16.34
Some qualifications at baseline	-4.51	1.62
Degree at baseline	11.17	-5.52
Receives informal care at t	9.56	-0.54
Receives formal care at t	7.27	3.58
Good health at t	-1.14	-10.23
OK health at t	2.21	-3.26
Poor health at t	1.25	2.16
Worst health at t	1.17	3.47
Partner in part-time work at t	14.67	18.73
Partner in full-time work at t	18.65	0.76
North West	6.39	2.39
Yorkshire and The Humber	9.27	10.71
East Midlands	5.27	5.81
West Midlands	12.94	4.30
East of England	13.74	4.21
London	3.30	-1.90
South East	11.33	-1.06
South West	-0.86	-6.34
Sex		22.38
In part-time work at t	-8.78	

	Men		Women	
In full-time work at t	-10.10			

Table A 19. Effect on the probability of receiving CA, given high intensity care provision and earnings under £100/week

	Men		Women	
Age 55 - 59	4.25		-11.29	
Age 60 - 64	-2.44		-27.06	**
Age 65 - 69	-12.98		-24.93	
Age 70 - 74	-25.02	*	-29.67	
Age 75 +	-8.15		-38.35	**
In a couple at t + 2	-19.43	**	-6.17	
Receives carers' allowance at t	22.16	***	16.72	
Partner receives informal care at t + 2	3.81		1.11	
Partner receives formal care at t + 2	8.82		4.54	
Partner receives AA at t + 2	5.47		4.40	
Partner receives DLA at t + 2	8.56		7.64	
North West	76.39		1.22	
Yorkshire and The Humber	75.84		-0.12	
East Midlands	75.46		-0.66	
West Midlands	58.70		0.62	
East of England	76.29		-9.48	
London	73.40		-2.85	
South East	73.04		2.57	
South West	75.87		-3.39	
Sex			82.04	

Glossary

CA	carer's allowance
DB	defined benefit (pension scheme)
DC	defined contribution (pension scheme)
DLA	disability living allowance
DWP	Department for Work and Pensions
ELSA	English Longitudinal Study of Ageing
ESA	employment support allowance
IB	incapacity benefit
IFS	Institute for Fiscal Studies
LFS	Labour Force Survey
NI	National Insurance
NRA	normal retirement age (associated with DB pension scheme)
ONS	Office for National Statistics
PIP	personal independence payment
RetSim	the IFS retirement simulator – a dynamic microsimulation model
SDA	severe disablement allowance
SPA	state pension age
SSP	statutory sick pay
TAXBEN	the IFS tax and benefit model

References

Banks, J., Emmerson, C., & Tetlow, G. (2007). *Better prepared for retirement? Using panel data to improve wealth estimates of ELSA respondents.* London: Institute for Fiscal Studies.

Crawford, R., & Tetlow, G. (2010). Employment, retirement and pensions. In J. Banks, C. Lessof, J. Nazroo, N. Rogers, M. Stafford, & A. Steptoe, *Financial cisrumstances, health and well-being of the older population in England: The 2008 English Longitudinal Study of Ageing* (pp. 11-75). London: Institute for Fiscal Studies.

Cribb, J., Hood, A., Joyce, R., & Phillips, D. (2013). *Living Standards, Poverty and Inequality in the UK: 2013.* London: Institute for Fiscal Studies.

Department for Work and Pensions. (2011, Feb). *Disability Living Allowance cases in payment: average weekly amount of benefit (pension age)*. Retrieved March 21, 2014, from DWP Tabulation Tool: http://tabulation-tool.dwp.gov.uk/100pc/dla/carepay/ccmobpay/ccclient/a_carate_r_carepay_c_ccmobpay_p_ccclient_pension_age_feb11.html

Department for Work and Pensions. (2011, February). *Disability Living Allowance cases in payment: average weekly amount of benefit (working age)*. Retrieved March 21, 2014, from DWP Tabulation Tool: http://tabulation-tool.dwp.gov.uk/100pc/dla/carepay/ccmobpay/ccclient/a_cawklyamt_r_carepay_c_ccmobpay_p_ccclient_working_age_feb11.html

Department for Work and Pensions. (2014, March 27). *Employment and Support Allowance Outcomes of Work Capability Assessments, Great Britain: supplementary tables.* Retrieved May 13, 2014, from GOV.UK: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/297937/esa_wca_140327.xls

Department for Work and Pensions. (2012, May). *Impact Assessment: Disability Living Allowance Reform.* Retrieved May 8, 2014, from GOV.UK: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220176/dla-reform-wr2011-ia.pdf

Department for Work and Pensions. (2014, April 11). *Simplifying the welfare system and making sure work pays: Introducing Personal Independence Payment.* Retrieved May 8, 2014, from GOV.UK: https://www.gov.uk/government/policies/simplifying-the-welfare-system-and-making-sure-work-pays/supporting-pages/introducing-personal-independence-payment

Department for Work and Pensions. (2011, Feb). *Tabulation tool*. Retrieved May 2014, from GOV.UK: http://tabulation-tool.dwp.gov.uk/100pc/tabtool.html

Emmerson, C., Heald, K., & Hood, A. (2014). *The Changing Face of Retirement: Future patterns of work, health, care and income among the older population.* London: Institute for Fiscal Studies.

Forder, J., & Fernandez, J. L. (2011). *Length of stay in care homes, report commissioned by Bupa Care Services, PSSRU Discussion Paper 2769.* Canterbury: PSSRU.

Giles, C., & McCrae, J. (1995). *TAXBEN: the IFS microsimulation tax and benefit model.* London: Institute for Fiscal Studies.

HM Treasury. (2014). *Budget 2014.* London: Her Majesty's Stationery Office.

McInnes, R. (2012, July 13). *ESA and Incapacity Benefit Statistics.* Retrieved May 13, 2014, from Parliament.uk: www.parliament.uk/briefing-papers/sn01420.pdf

Office for Budget Responsibility. (2013). *Economic and Fiscal Outlook.* The Stationery Office Limited.

Office for Budget Responsibility. (2014, March 19). *Economic and Fiscal Outlook Charts and Tables - March 2014 (chart 3.13).* Retrieved May 27, 2014, from Economic and fiscal outlook – March 2014: http://budgetresponsibility.org.uk/pubs/March_2014_EFO_Economy_Supplementary_Tables.xls

Office for National Statistics. (2014, April 15). *House Price Index, February 2014: Monthly & Quarterly Tables 1 to 19 (table 10).* Retrieved May 27, 2014, from House Price Index, February 2014: http://www.ons.gov.uk/ons/rel/hpi/house-price-index/february-2014/rft-monthly-and-qtly-february-14.xls

Office for National Statistics. (2012, December 20). *Table 3a: Divorces in England and Wales 2011.* Retrieved 5 7, 2014, from ONS.GOV.UK: http://www.ons.gov.uk/ons/rel/vsob1/divorces-in-england-and-wales/2011/rft-divorces--number-of-divorces-age-at-divorce-and-marital-status-before-marriage.xls