Insurance, Efficiency and the Design of Public Pensions[*]

Cormac O'Dea[†]

March 19, 2018

Abstract

Government pension spending in advanced economies can be divided into two types: (1) Social Security-style benefits that depend on working-life earnings and (2) means-tested old-age income floors. Spending on the former type of benefit is substantially greater than spending on the latter in most countries. Using an estimated lifecycle model that accounts for these, as well as endogenous labor supply, private savings and realistic uncertainty, this paper investigates the optimal combination of the two types of pensions. For countries that provide public pensions that depend on career-average earnings, I show that large (revenue-neutral) welfare gains can be obtained by increases in means-tested old-age income floors. While such transfers cause costly distortions, these are more than offset by the value of the insurance provided. The optimality of greater means-tested support is specific to older individuals: I find that such support to younger households should be much lower. These results imply that governments should provide strong work incentives for the young, but provide pensions with good insurance properties for the old.

JEL Classification: D91, E21, D14

Keywords: Social Security; Means-testing; Pensions; Lifecycle; Savings; Household Finance

1 Introduction

Public pension payments to retirees are one of the most costly activities carried out by governments – accounting for over 8% of GDP on average in OECD countries in 2014. These payments are provided

---

through one or both of two types of scheme: 1) Social Security-style benefits that depend on earnings during working life and 2) means-tested income floors to the elderly which depend only on wealth accumulated by old age. Payments of the latter type play only a small role in the social insurance systems of most countries: for example in the US for every dollar spent on Supplemental Security Income to the aged (which provides a means-tested income floor), 140 dollars is paid out in Social Security benefits.

This paper evaluates the optimal balance of these different types of provision for pensioners using an estimated lifecycle model that contains both these types of public pension, as well as realistic uncertainty, endogenous labor supply and each of pension and non-pension private saving. The purpose of the estimated model is to both value the insurance provided by different counterfactual public pension schemes and to quantify their costs and the costs of the distortions that they induce. In implementing the optimum, the trade-off that policymakers face is that old-age means-tested income floors provide help to the elderly when they need it most, giving valuable insurance against low lifetime earnings, poor investment returns and longevity, but they may provide strong disincentives to work and save. Progressive Social-Security style public pensions can be less distortionary, but have less attractive insurance properties.

Heterogenous preference parameters are estimated using rich survey data on wealth holdings and labor supply of the elderly that has been linked to administrative records on survey respondents' history of lifetime earnings. The estimated model is used to obtain the ex-ante optimal level of a means-tested income floor for the elderly. I find that levels of these typically provided in advanced economies (e.g. that provided by Supplemental Security Income in the US) are too low: increases funded by any of reducing the generosity of Social Security, increasing tax rates or, especially, reducing the extent of private pension subsidies deliver large increases in welfare. The positive welfare effect from a more generous means-tested income floor comes from a reduction in the variance of lifetime consumption and increases in leisure time, partially offset by lower average lifetime consumption (due to the now-larger distortionary means-tested payment).

Similar counterfactual analysis applied to *working-age* means-tested floors suggest that these (e.g. food stamps) should not be increased. Optimal levels of means-tested income floors provided to those of working-age are found to be much lower than optimal old-age income floors. This is due to the fact that, relative to old-age means-tested income floors, those offered to working-age households have a greater negative effect on labor supply, and the productivity of those whose behaviour they distort is much greater. The cost of providing a means-tested income floor to the young is therefore substantially greater than the cost of providing it to the old. These results point to the value of providing strong incentives to work to the young alongside pensions with good insurance properties to the old.

The first branch of the literature to which this paper relates considers role of means-testing in the op-

timal design of social insurance. Golosov and Tsyvinski (2006) show that an optimal Disability Insurance (DI) system will condition benefit receipt on household asset holdings; Rendahl (2012) finds a similar result for Unemployment Insurance (UI). The mechanism in these papers is that households planning to falsely claim DI or shirk in their employment search while claiming UI will want to accumulate assets to maintain consumption levels which will only be partially insured by DI/UI payments. Asset-testing the payments reduces the capacity for planning to lie/shirk and therefore mitigates the distortions that social insurance provision can imply. Introducing a retirement motive for saving, however, might mitigate (or overturn) the strong results in those papers. If households want to save for retirement, asset-testing DI or UI will lead to some combination of lesser social insurance against disability/unemployment (for households who save in spite of the asset test) and sub-optimal saving for retirement (for households who do not save because of the asset test). Considering the retirement savings motive also raises, though, the question considered by this paper, that is whether old-age Social Insurance payments – the largest part of most social insurance systems – should be asset-tested, either by asset-testing Social Security or complementing Social Security with more generous asset-tested income floors.

There is evidence to suggest that increases in means-tested benefits would be welfare-increasing. Braun et al. (2016) show that a 33% increase in the generosity of means-tested social insurance programs in the US (Medicaid, Supplemental Social Security Income, food stamps and a number of smaller programs), would be welfare-increasing if funded through the payroll tax. The current paper adds to theirs by finding the optimal means-tested income floor rather than exploring the welfare-implications of ad-hoc changes, by estimating preference parameters and by considering the interaction of private pension subsidies and public pensions. Sefton and Van De Ven (2009), who do search for an optimal old-age means-tested floor (albeit in a model where there are no private pensions, where preference parameters are not estimated and where there is no endogenous labour supply of those over the age of 65 and so potentially in receipt of the transfer), find that enhancing the generosity of the means-tested component of the UK public pension system would be welfare-improving. Huggett and Parra (2010) and Golosov et al. (2013), in calibrated models, both find that making Social Security payments more progressive (albeit not by means-testing them) would be welfare-enhancing.

The second branch of the literature to which this paper relates concerns the design of public pensions. Following Auerbach and Kotlikoff (1987), this literature[1] has focussed on how Social Security might be made more affordable in light of the pressures imposed on it by changing demographics. Solutions that have been heavily studied include raising payroll taxes, delaying eligibility ages and reducing the generosity of benefits. The literature has, however, generally neglected to study the possibilities afforded

---

[1]A partial list of contributions is De Nardi et al. (1999), Conesa and Krueger (1999), Kotlikoff et al. (1999), Huggett and Ventura (1999), Nishiyama and Smetters (2007) and Conesa et al. (2009).

by the means-testing of benefits, in part as such transfers are known to reduce labour supply and crowd out private saving.[2] An exception is Kitao (2014) who studies four options to make Social Security sustainable – one of which involves an extreme form of means-testing whereby all Social Security benefits, after a small disregard, are withdrawn at an effective tax rate of 100%. This reform is rejected as "[due to] the large negative effects on economic activities and fiscal burden, it is unlikely to be a viable option for social security reform". I show, however, that less extreme reforms can be welfare-improving, notwithstanding the distortions that they induce.

The third literature to which this paper relates is the (small) literature which considers private pensions and their effect on behavior. Nishiyama (2011) investigates the budgetary and welfare properties of tax-deferred savings vehicles (e.g. 401k plan). Blau (2016) looks at the extent to which different types of private pension crowd out non-pension saving. These papers do not consider how private pensions and their tax treatment interacts with design issues around public pensions, as this paper does.

The key question that this paper seeks to address is whether means-tested income floors, which currently play a relatively minor role in most old-age social insurance systems, should be expanded. The analysis points to a greater role for means-tested income floors which provide, at an acceptable cost, valuable insurance to households against low working life earnings, poor investment returns and longevity. Pensions which amplify lifetime earnings risks (such as career-earnings related public pensions or private pension subsidies) are substantially less preferred.

The rest of the paper proceeds as follows. To motivate the modelling choices which will come later, Section 2 very briefly discusses some typical features of public and private pensions. Section 3 outlines the model used in the paper before Section 4 details the estimation procedures, gives parameter estimates and discusses model fit. Section 5 uses the model to find the optimal level of a means-tested old-age income floor. Section 6 concludes.

2 Pensions

This section briefly describes some typical features found in international pension systems - first describing the system of public pensions, and then private pensions. The aim of this section is to introduce some terminology that will be important throughout the rest of the paper and motivate some of the modelling decisions.

---

[2]Neumark and Powers (1998, 2000) show that SSI reduces pre-retirement saving and labour. Hubbard et al. (1995) first showed the role that income floors play in reducing wealth accumulation in a lifecycle model. De Nardi et al. (2010) show that, when there is a risk of large wealth shocks (such as those that come from healthcare needs at old ages), such floors can affect the behaviour even of those who have very high permanent income.
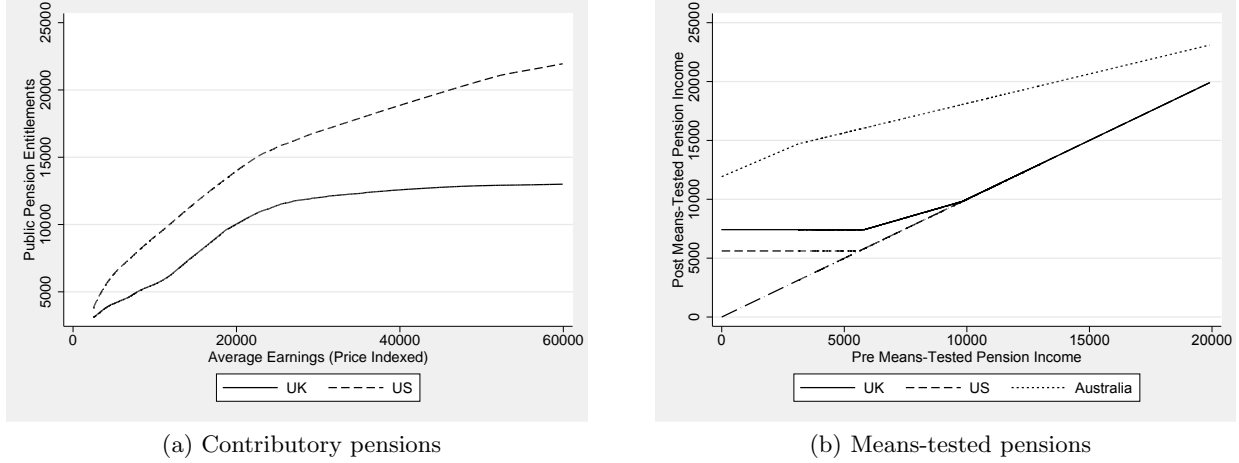
(a) Contributory pensions   (b) Means-tested pensions

Figure 1: Public Pensions in the UK, US and Australia

2.1 Public pensions

Public pensions can be either 'contributory' (they depend on earnings during working life) or can be 'means-tested' (they depend on income and assets in retirement). Examples of the former are Social Security in the US and the State Pension in the UK. Examples of the latter are Supplemental Security Income (SSI) in the US, Pension Credit (PC) in the UK and the Australian Age Pension. Figure 1 illustrates these. Figure 1(a) shows how Social Security (US) and State Pension (UK) pension entitlements vary with average working life earnings for a sample of individuals born between 1935 and 1950 (the data used here will be discussed further in Section 4).[3] Figure 1(b) shows, for each of the US, the UK and Australia, how income including means-tested transfers varies with income excluding it. For the poorest pensioners both SSI and PC top pension income up to a minimum level. This is initially withdrawn at an effective tax rate of 100% in both countries, though benefits over a certain quantity are withdrawn at a lower effective tax rate of 40% in the UK. In Australia a small amount of income is disregarded in applying the income test, after which the Age Pension is withdrawn at an effective tax rate of 50%.[4]

The conceptual difference between these two types of pension is whether earnings during working life (on the horizontal axis in the left-hand graph) or income in retirement (on the right-hand graph) determine the level of entitlement.

---

[3]In both countries, this illustrative analysis is carried out at an individual level and does not take into account benefits that are earned on the basis of a spouse's contributions.

[4]These figures use the values for the 2013 systems. US and Australian dollars are converted to pounds using average exchange rates for that year. The US figures are an average of the total (federal and state) entitlements for the states whose payments are delivered by the federal Social Security Administration. The x-axis here represents *pension* income - in the US and UK a certain level of employment income can be earned without affecting entitlement to the benefit. Each country also has different rules regarding how assets holdings affect entitlement. These are not discussed here.

2.2 Private pensions

Private pensions can be grouped into two broad types - Defined Benefit (DB) pensions and Defined Contribution (DC) pensions. DB pensions pay a fraction of some function of earnings – for example, career average earnings or final earnings. DC (401k-style) pensions are investment accounts owned by the individual that can be used to purchase an annuity or otherwise provide an income in retirement. DB pension income can be thought of as a deterministic function of earnings, while DC pension income is a stochastic function of contributions into a pension fund.

Private pensions are subsidised by the tax system in many countries. The form of these vary internationally but typically involves some form of tax-deductibility of contributions into pension funds (that is, such payments can be made out of gross earnings). If pension income is subjected to lower rates of tax than earned income, this tax deferral should be thought of as a subsidy (and will incentivise households to save in a pension - either by substituting consumption from during working life to retirement, or substituting towards pension saving from non-pension saving). Figure 2 illustrates that such favourable taxation is commonplace. It shows, for a selection of OECD countries, the average tax rate (black bar) on a worker earning average earnings and the average tax rate (grey bar) on a pensioner with a pension equal to average earnings. In most countries, average taxes on the latter are lower (and often substantially so). The lower burden of taxation on pension income comes in a number of forms – more generous tax deductions for the elderly, the ability to take some pension income tax-free and to pension income being exempt from payroll taxes.[5] The favourable treatment of pension saving, relative to non-pension saving, costs approximately 1.1% of GDP in the UK and 0.9% of GDP in the US relative to a benchmark where pension income is taxed similarly to earnings (see Appendix B for details on calculation of these figures).

---

[5]The extent to which one should consider the last of these a subsidy depends on whether payroll taxes are levied on income that is paid into a pension fund. In the US, payroll taxes must be paid on contributions made into pension funds while in the UK payment into pension funds can be exempt from such taxes. This means that earnings saved in a pension in the UK are not subject to payroll taxes either when earned or when drawn down.
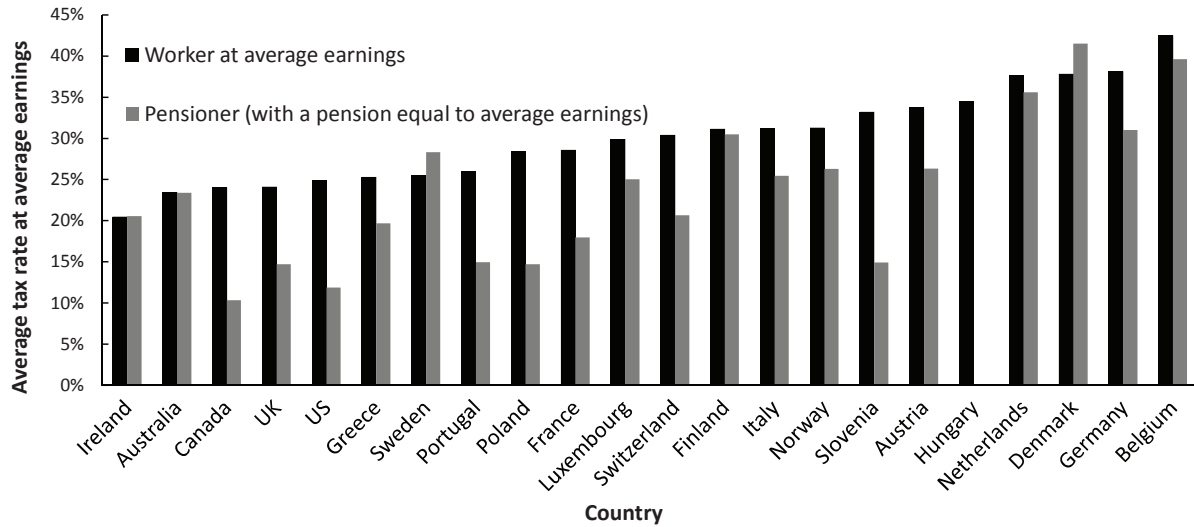
Figure 2: Average tax rates on earnings and pensions

3 Model

Before detailing the model I provide a short summary of its key features. The decision-making unit is a household that maximises an intertemporal utility function by choosing each of labour supply, consumption, pension saving and non-pension saving. Households are exposed to risk over: i) whether they get an employment offer, ii) productivity (which determines their earnings if in work), iii) the investment returns they earn on any DC pension wealth and iv) mortality. Government partially insures households against these risks by levying a progressive income tax, providing unemployment insurance and giving two types of payment to pensioners: a contributory pension and a means-tested income floor. Households are heterogenous in their access to private pensions: some are offered a Defined Benefit pension through their employer, while all can save in a (risky) Defined Contribution pension should they wish to.

The rest of this section discusses in turn the utility function, modelled pension and non-pension assets, the role of government and the household's maximisation problem. Table 9 in Appendix A gives a summary of all the variables introduced in this section.

3.1 Household composition, utility and decision

Household composition All households contain a married couple of age t who start their working life at age 20. Mortality is stochastic. Household composition (h) takes a value of 1, 2, 3 or 4 indicating, respectively, that both spouses are still alive, only the male is alive, only the female is alive or that both

spouses are dead.

Heterogeneity Households are one of four types (indexed by j). These types are each pairwise combination of low/high education and having access to a DB pension or not. Household types are determined before the start of working life and are fixed for life. Different types have different productivity processes and face different risks over employment (allowing for the fact those who typically have DB pensions (e.g government employees) could face less labour market risk than those working in the private sector). Preference parameters differ across types.

Utility Households get utility from consuming, from leisure and from leaving bequests. The period utility function, given in equation (1), is non-separable in consumption (c) and leisure (l). Consumption has a weight of ν and the coefficient of relative risk aversion on utility is γ.[6]

$$u(c,l) = \frac{(c^\nu l^{1-\nu})^{1-\gamma}}{1-\gamma} \tag{1}$$

Households value bequests through a warm-glow bequest function, of a form used by Nardi (2004) and French (2005), and given in equation (2). a^b are assets bequeathed, θ determines the importance of bequest motives to households and K is a constant that ensures that the marginal utility of leaving a zero bequest is finite.

$$b(a^b) = \theta \frac{(a^b + K)^{(1-\gamma)\nu}}{1-\gamma}. \tag{2}$$

Decisions The period in the model is a year. In each period households make four decisions. They decide i) employment at the extensive margin ii) non-housing consumption (c^{nh}), iii) how much, if anything, to contribute to their DC pension (dc) and iv) how much non-pension saving to do.

Employment and earnings The labour supply behaviour of only one household member - the 'principal earner' - is modelled. The labor supply of the second earner is exogenous. This main earner chooses whether or not to supply labour if offered a job. The probability of not getting a job offer ($ue_t = 1$) evolves according to a conditional Markov process where the probability of unemployment, $\pi_1(\tilde{e})$, is conditional on current productivity ($\tilde{e}$).

---

[6]The coefficient of relative risk aversion on consumption is

$$-\frac{\partial^2 U}{\partial c^2} \frac{c}{\frac{\partial U}{\partial c}} = -(\nu(1-\gamma) - 1).$$

When employed, earnings are equal to productivity, whose log (equation (3)) is the sum of a deterministic component (a quadratic in age) and a stochastic component (u).

$$\ln \tilde{e}_{it} = \delta_0 + \delta_1 t + \delta_2 t^2 + u_{it} \tag{3}$$

The evolution of u in periods following an employment offer (given in equation (4)) follows an AR(1) process with innovations distributed normally. The variance of these innovations differs in the first and in subsequent periods.

$$
\begin{aligned}
u_t &= \rho u_{t-1} + \xi_t \\
\xi_1 &\sim N\left(0, \sigma_\zeta^2\right) \\
\xi_t &\sim N\left(0, \sigma_\xi^2\right) \ \forall \ t > 1
\end{aligned}
\tag{4}
$$

In periods following a period of unemployment, the stochastic component of productivity is drawn from a distribution E.

The labour supply behaviour of the second earners is exogenous. Households receive a fixed payment (e^s) up to a retirement age for the second earner (t^{ret_s}).

Consumption Consumption is the sum of non-housing (c^{nh}) and housing consumption (c^h). The former is a choice, the latter is the product of a preference parameter representing the rental value of housing ($r^{houscon}$) and gross housing wealth (gh, which is an exogenous function of non-pension wealth and age):

$$c_t^h = r^{houscon} gh(a, t). \tag{5}$$

The function $gh()$ is given in Appendix C.1.

3.2   Assets

Households accumulate wealth to insure themselves against unanticipated falls in their income (for example, due to bad productivity draws or unemployment) and to provide consumption when they retire. They can save in up to three assets. These are, a Defined Benefit pension (for those types eligible), a Defined Contribution pension and non-pension wealth. These assets are now discussed in turn.

Defined Benefit pensions Two of the four household types accrue entitlements to DB pensions while working. They must make pension contribution (db_t) from their earnings at each age up to 65. This is

set at a fixed proportion (ϑ) of pre-tax earnings. Once they reach the age of 65, they receive a taxable pension that is a type-specific function of career-average earnings at the age of 64: $pp_t^{db} = db(ae_{64}, j)$.

Defined Contribution pensions Households can, each period, pay into a Defined Contribution (i.e. 401k-style) pension. The evolution of the stock of wealth in the DC fund depends on flows into the fund (dc -which is tax-deductible and so can be made out of *gross* income) and the return on the fund in each year (ϕ):

$$DC_{t+1} = (1 + \phi_{t+1})(DC_t + dc_t) \tag{6}$$

The return on DC funds is assumed to be *iid* and normally distributed with a mean of $\bar{\phi}$ and a variance of σ_ϕ^2.

DC wealth is decumulated from the age of 65. At this age a quarter of the fund is taken as a (tax-free) cash lump sum - this conversion of never-taxed pension wealth into non-pension wealth is one feature that makes saving in private pensions incentivised by the tax system.[7] The remaining three-quarters of the stock of DC wealth must be used to purchase a (taxable) life annuity. The lump sum ls_{65}^{dc} is given by $(0.25)DC_{65}$ and the stream of pension income at each age after 65 is given by:

$$pp_t^{dc} = q(0.75)DC_{65}$$

where q is an annuity rate that is actuarially fair up to the deduction of a fixed proportion to account for the administrative costs and profits of the annuity-providers.

Non-pension assets Households can save and accrue non-pension wealth (a) which accumulates according to the following inter-temporal budget constraint:

$$a_{t+1} = (1 + r_t)(a_t + y_t - c_t^{nh} - dc_t - db_t) \tag{7}$$

where r_t is the return on non-pension wealth and y is household income (the sum of gross earnings, unemployment insurance payments, public pension payments, private pension payments and interest less taxes). The tax function is discussed in the next subsection.

---

[7]Such tax-free lump sums are permissible in the UK.

3.3 Government

The government levies taxes and provides unemployment benefits and pensions. The modelled system is a stylised version of the prevailing UK system.

Taxes   The household tax function is fully detailed in Appendix D.7; the discussion here focusses on the tax treatment of private pensions. Private pensions are treated favourably through a combination of tax-deductibility of payments into pension funds as well as three features of how pension income is taxed. The first is the option to take part of the pension in a tax free lump sum, noted above. The second is that in the UK (as in the US and Canada) there are more generous income tax deductions for those over the age of 65 than younger individuals. Finally, after the age of 65 payroll taxes are not levied on any income.[8]

Unemployment benefits   Unemployment shocks are assumed to be verifiable by the government. Affected households receive an unemployment payment (ui) irrespective of their accumulated assets. Those who get an offer but who simply choose not to work can receive an asset-tested payment (ui^{mt}) if they are sufficiently poor.[9]

Public pensions   Those aged over 65 are entitled to two payments. The first is a Social-Security style public pension, payable from the age of 65 until death which is modelled as a function of career-average earnings at the age of 64 and household composition: ($ss_t = ss(ae_{64}, h)$). The government also provides a means-tested income floor to those over the age of 65 ($mtif(y_t, a_t, t, h)$) that depends on income, assets (which are assumed to generate a flow of income), age and household composition. This plays the role of Supplemental Security Income in the US and Pension Credit in the UK. The form of these functions was illustrated in Figure 1, and the precise form they take in the model is given in Appendix D.7.

3.4 State variables and the household maximisation problem

This section gives the household's maximisation problem, making explicit the state variables of the Dynamic Programming problem.

---

[8]This effectively reduces the two main rates of tax (including payroll taxes) from 32% and 42% to 20% and 40%. The treatment of private pensions and such social insurance contributions differs in the UK and US. In neither country are social contributions levied on private pension income. However, in the US, payroll taxes are levied on earnings paid into a private pension whereas in the UK they can be made exempt from NICs.

[9]This can be thought of as playing the role that SNAP (Food Stamps) play in the US. In the UK system it is 'income-based job-seekers allowance' – payable at a particularly low level and designed to protect against destitution.

3.4.1 State variables

The state variables are household type (j), age (t), non-pension wealth (a), whether unemployed in the current period (ue), productivity $(\tilde{e})$, DC pension wealth (DC), income from the DC pension (pp^{dc}), household composition (h) and average earnings (ae)[10]. The set of state variables is: $\mathbf{X}_t = \{j, t, a_t, ue_t, \tilde{e}_t, DC_t, pp^{dc}_t, h_t, ae_t\}$. There is uncertainty over the investment return (ϕ) earned on the DC fund, over whether an employment offer is received (ue), over productivity $(\tilde{e})$, and, due to stochastic mortality, over household composition (h). Below, the joint distribution of the first three of these will be denoted as $F(\phi, ue, \tilde{e})$. The distributions of ue and $\tilde{e}$ at age $t+1$ depend on their values at age t. s^m_{t+1} and s^f_{t+1} give, respectively, the probability that a man and a woman will survive to age $t+1$ conditional on them having survived to age t.

3.4.2 Household maximisation problem and value functions

Household's problem after the age of 65 Equation (8) gives the maximisation problem and associated value function of a household aged 65 or over with both spouses still alive. Such households have already annuitised their DC wealth. While in receipt of the pension, they can still choose to supply labour (whether the principal household earner works – which determines leisure (l)). They also choose their non-housing consumption (c^{nh}).

$$
\begin{aligned}
V_t(\mathbf{X}_t|h_t=1) = \max_{c^{nh}_t, l_t} \Bigg(\quad u(c_t, l_t) \quad &+\beta s^m_{t+1} s^f_{t+1} \int V_{t+1}(\mathbf{X}_{t+1}|1) dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t) \\
&+\beta s^m_{t+1}(1 - s^f_{t+1}) \int V_{t+1}(\mathbf{X}_{t+1}|2) dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t) \\
&+\beta(1 - s^m_{t+1})(s^f_{t+1}) \int V_{t+1}(\mathbf{X}_{t+1}|3) dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t) \\
&+(1 - s^m_{t+1})(1 - s^f_{t+1}) b(a^b_{t+1}) \Bigg)
\end{aligned}
$$

$$
\text{s.t.} \quad c_t = c^{nh}_t + c^h_t
$$

$$\text{and the intertemporal budget constraint in equation (7)}$$

Household's problem before age 65 The maximisation problem and associated value function faced by a household (again with both spouses still alive) which is aged less than 65 and so has not annuitised its DC wealth is given in (9). The problem differs from that of the post-annuitisation problem as there is now one additional choice variable – how much to contribute to the DC pension (dc) – and there are

---

[10]DC wealth (DC) is a state variable only up to the age of 65, DC income (pp^{dc}) is a state variable only after the age of 65. Up to and including the age of 65, the state variable ae represents average earnings up to the previous year. From the age of 66 onwards, it represents average earnings at ages up to and including 64.

now two intertemporal budget constraints (equations (6) and (7) - which relate respectively to DC wealth non-pension wealth).

$$V_t(\mathbf{X}_t|h_t = 1) = \max_{c_t^{nh}, dc_t, l_t} \left(\quad u(c_t, l_t) \quad +\beta s_{t+1}^m s_{t+1}^f \int V_{t+1}(\mathbf{X}_{t+1}|1)dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t) \right. \tag{9}$$

$$+\beta s_{t+1}^m (1 - s_{t+1}^f) \int V_{t+1}(\mathbf{X}_{t+1}|2)dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t)$$

$$+\beta(1 - s_{t+1}^m)(s_{t+1}^f) \int V_{t+1}(\mathbf{X}_{t+1}|3)dF(\phi_{t+1}, ue_{t+1}, e_{t+1}|ue_t, e_t)$$

$$\left. +\beta(1 - s_{t+1}^m)(1 - s_{t+1}^f)b(a_{t+1}^b) \right)$$

$$\text{s.t.} \quad c_t = c_t^{nh} + c_t^h$$

and the intertemporal budget constraints in equations (6) and (7)

There are no analytical solutions to the problems outlined in (8) and (9). Solutions are obtained numerically - using methods discussed in Appendix G.

4 Estimation and results

4.1 Estimation

Estimation of the model parameters follows a two-step procedure.[11] In the first step, some parameters are estimated outside the model, or are set with reference to the literature. In the second step, preference parameters and earnings processes are estimated using the method of simulated moments. Both these steps will be described below. Before that, the next subsection briefly introduces the main data source, defines the sample used and describes how household types are characterised.

4.1.1 Data, sample and definition of types

The main data used in this paper come from linked survey and administrative data. The survey data is the English Longitudinal Study of Ageing (ELSA) - a biennial longitudinal survey that contains a representative sample of the English private household population aged 50 and over. ELSA is one of a number of international 'ageing surveys' - modelled on the Health and Retirement Study (HRS) in the US. ELSA contains detailed data on demographics, labour market circumstances, earnings and the level and composition of wealth holdings. The first wave of ELSA covered 2002/03 and data from the first five waves are used in this paper.

---

[11]This two-step procedure is widely applied in papers that develop and estimate structural lifecycle models. See, amongst others, Gourinchas and Parker (2002), French (2005), Low and Pistaferri (2015), Lee et al. (2015), and Blundell et al. (2016).

ELSA respondents were asked for their National Insurance number (equivalent to Social Security number in the US) and permission to link to their history of National Insurance contributions. Data on these contributions allows a panel of earnings in each year of working life for ELSA respondents to be obtained. 80% of individuals consented to the link. Details on how I convert these data into a panel of earnings is given in Appendix E.1. These earnings data are used, in a manner described below, to estimate earnings processes, while the survey data yields moments of assets and employment which are used to estimate preference parameters.

A sample of couples is selected in which the primary earner in the couple (the member who has the highest lifetime earnings) was born between 1935 and 1950. There are 2,364 such households in the data. Those who never married and those who are divorced are not included. Only those couples with linked National Insurance data and where National Insurance contributions were made in at least 5 years are included in the sample. A number of additional sample restrictions are imposed. Households where the education of primary earner is not recorded are excluded, as are those where either member of the couple didn't fully complete the survey, and those where the sum of years of self-employment carried out by either member of the couple is greater than or equal to 5.[12] After applying these exclusions, 1,121 couples remain (47.4% of the original sample).

The model contains households of four types who differ in their education (low or high) and whether they have access to a DB pension. Households are characterised as low (high) education if the primary earner left school at (remained after) the age of 15 which was the compulsory schooling age for this cohort. The split of the sample into DB/non-DB types is complicated by the fact that many households have small amounts of DB wealth (because, for example, they worked for a year or two at some stage in their career in a job that provided a DB pension). A household is defined as being a DB pension type household if the primary earner spent at least one-third of working years contributing to a DB scheme (years accruing DB pension rights is recorded in the administrative data). Table 1 gives the proportion of the sample in each of the four types.[13]

4.1.2 Parameters estimated/set outside the model

This subsection gives details of the parameters set or estimated outside the model.

---

[12]The administrative data can be used to calculate the number of years spent in self-employment, but not the income from that employment.

[13]Some caution should be exercised in interpreting these as population shares. The sample is not fully representative of the original population-representative sample. Those with less education were less likely to give permission to link to the administrative records and so are under-represented. It is not possible to assess representativeness on access to the DB pension as information on the latter is not available for those who did not give permission to link to the administrative data. More detail on differences between the sample used here and the full ELSA sample is given in Appendix E.4.

Table 1: Proportion of each types in sample

	No DB Pension	DB Pension
Low Education	16%	32%
High Education	11%	41%

Stochastic component earnings process   The data generating process for the earnings data is given in equation (10). This differs from the model's earnings process given in equation (3) in Section 3 in two ways. First, the error term is η_{it} – the sum of the stochastic component of earnings (u_{it}) and serially uncorrelated measurement error ($m_{it} \sim N(0, \sigma_m^2)$). The dependent variable is denoted $\tilde{e}^{data}$ rather than $\tilde{e}$ to indicate that earnings are measured with error. Second, the coefficients are not the true coefficients of the earnings process ($\delta_0, \delta_1, \delta_2$) and are denoted as ($\bar{\delta}_0, \bar{\delta}_1, \bar{\delta}_2$) to indicate that this equation gives earnings observed in the data rather than potential earnings for everyone (including those who don't accept a job offer). Direct estimation of these parameters (by, for example, running a simple regression) will yield coefficients that are biased due to non-random selection into employment. These biased coefficients will be used within the model's estimation procedure to estimate the true parameters of the earnings process (discussed in Section 4.1.3). The parameters of the data generating process for η_{it} ($\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2$) are estimated outside the model using a standard approach (see, for example, Guvenen (2009) or Low et al. (2010)) by choosing those values that minimise the distance between the empirical covariance matrix of estimated residuals ($\hat{\eta}_{it}$) for ages up to 50 and the theoretical variance covariance matrix of $\eta_{it} = u_{it} + m_{it}$.

$$\ln \tilde{e}_{it}^{data} = \bar{\delta}_0 + \bar{\delta}_1 t + \bar{\delta}_2 t^2 + \underbrace{\eta_{it}}_{u_{it}+m_{it}} \tag{10}$$

Table 2 gives the estimates of these parameters for each type. Highly educated households have shocks to earnings that have a variance almost twice the level of those with low education. The two high education types have very similar earnings process estimates. However, the processes of those without a DB pension differ by education: those without a DB pension have shocks to their earnings process that have higher variance and lower persistence than those with such a pension.

Table 2 also shows (in the final row) the unemployment rate for each type (I define a household as unemployed in the data if the primary earner is recorded as having annual earnings of less than that provided by the UK's unemployment insurance level). The full Markov transition matrices and the probability distribution over productivity after an unemployment shock are shown in Appendix D). There are significant differences between groups in rates of unemployment – those without a DB pension (who are drawn disproportionately from the private sector) have substantially higher rates of unemployment

than those in DB careers (who are drawn disproportionately from those who work for the government).

Table 2: Earnings process estimates

	Type			
	Low Ed		High Ed	
	No DB	DB	No DB	DB
ρ	0.970	0.993	0.995	0.990
σ_ξ^2	0.007	0.005	0.010	0.012
σ_ζ^2	0.073	0.053	0.110	0.112
σ_m^2	0.053	0.027	0.057	0.023
Unemployment rate	0.13	0.05	0.15	0.03

Coefficient of relative risk aversion In related papers that also use a utility function that is non-separable in consumption and leisure, the choice for γ, the coefficient of relative risk aversion on utility, has usually been between 2 and 4. My main results are for $\gamma = 3$, but results are also given for $\gamma = 2$ and $\gamma = 4$.[14]

Other parameters set outside the model The other preference parameters set outside the model are K, L, and hrs. K is the parameter that determines the curvature of the bequest function - which is set at £650,000.[15] L is the hours endowment - it is set equal to an (annual) value of 5,824 (16 hours a day to be divided between work and leisure). hrs is the number of hours worked when employed and is set equal to 1,840 (40 hours a week for 46 weeks a year).

The other parameters set outside the model relate to the model's assets. Details are given in Appendix D: the distribution of returns on Defined Contribution pensions are discussed in Appendix D.1, the functions relating career-average earnings to defined benefit pension income and public pension income are given in Appendix D.3 and D.4. Appendix D.5 discusses the return on non-pension wealth.

4.1.3   Method of simulated moments estimation

Method of simulated moments estimation of preference parameters   Four preference parameters, each of which differ by type (j), are estimated using the method of simulated moments and data on employment, wealth and portfolio composition. These parameters are the discount factor (β_j), the

---

[14]In the literature that values social insurance over the lifecycle and uses a non-separable (in consumption and leisure) period utility function, a choice of $\gamma = 4$ is the most common (e.g. Auerbach and Kotlikoff (1987), Kotlikoff et al. (1999), Conesa et al. (2009), Nishiyama (2011)). As will be shown below, the results in this paper (which suggest a greater role for mean-testing) are strengthened when this level of risk aversion is used - that is the choice of $\gamma = 3$ is a conservative one given the results.

[15]French (2005) sets this at $500,000 in 1987 prices which, when converted to 2012/13 prices and converted to pounds sterling using the average exchange rate in that year, is approximately £650,000.

weight on consumption in the utility function (v_j), the consumption flow value of housing ($r_j^{houscon}$) and the weight on bequests (θ_j). The moments used are the proportion of men in work at each age between 52 and 75, mean non-pension holdings between the ages of 52 and 90[16] and (for those types without DB wealth) DC pension wealth between the ages of 52 and 75.[17] Wealth moments are top-coded in both data and simulations at the 95th percentile to mitigate the impact of the very wealthy. The parameters are estimated using standard GMM techniques.

It is worth noting which aspects of variation in the data will bear most heavily on the identification of particular estimated parameters. Total wealth (the sum of pension and non-pension wealth) contributes substantially to the identification of the discount factor (β) - the greater the holdings of wealth, the more patient are households and the higher will be the estimate of β. The trajectory of wealth late in life (the extent to which it is retained rather than consumed) contributes to the identification of the strength of the bequest motive (θ) - the greater the extent to which wealth is retained rather than consumed, the more households value the leaving of bequests and the higher will be the estimate of θ. The split in wealth between non-pension and pension wealth contributes to the identification of the consumption flow value of housing ($r_j^{houscon}$). The higher is non-pension wealth relative to pension wealth, the more value households place on the housing consumption flow, and the higher will be $r_j^{houscon}$. Finally, labour supply profiles pin down the relative weight of consumption in the utility function (v). The greater the extent to which older individuals remain at work, the higher will be their preference for leisure and the lower will be v – their (relative) preference for consumption.

Deterministic component of earnings process   To estimate the deterministic component of the earnings process of the principal earner[18] (the parameters of equation (3)) the first step is to obtain the parameters in equation (10) - the regression of observed earnings on a quadratic in age. Due to non-random selection into the labour market, these parameters $\{\bar{\delta}_0, \bar{\delta}_1, \bar{\delta}_2\}$ are not those of the true productivity process. These biased parameters are used within the model to estimate the true parameters using a method introduced by French (2005). Briefly, the approach involves (i) first solving the model and simulating behaviour using this (biased) profile. (ii) With the simulated data (where both accepted and rejected wage offers are observed), the bias is calculated at each age. (iii) These biases are used to

---

[16]While the cohort born between 1935 and 1950 form the basis for all other moments used in estimation, calculating these moments involves using data from individuals born before 1935. Data from older cohorts is used here as moments from the phase of life where wealth is being (or not being) decumulated is important to help identify the strength of the bequest motive. Age, period and cohort effects are estimated using the method of Deaton and Paxson (1994).

[17]The reason that moments on work after the age of 75 are not used is that the numbers are very low and don't change much - and so the additional moments do not provide much additional information. Similarly, after the age of 75, all pension wealth in the data is being decumulated in a mechanical manner (as the annuity stream becomes less valuable as fewer years of receipt are left in expectation), and so additional moments representing the pension wealth of the very old are not used.

[18]The exogenous earnings of the secondary earner are given in Appendix D.6.

'correct' the earnings process fed into the model in step (i). (iv) The corrected earnings process can then be fed back into the model which is solved and behaviour is simulated once again. Steps (ii) to (iv) are then repeated until convergence.[19]

4.2 Estimates and model fit

Table 3 gives the estimates of the preference parameters. In keeping with most papers that have estimated discount rates by educational attainment (see, for example, Dohmen et al. (2010) and Alan and Browning (2010)), those with more education are found to be more patient. This result is driven by the fact that in the data those with more education accumulate more wealth as a proportion of earnings than those with less. Figure 16 in Appendix E.3 illustrates this: it shows the ratio of mean total wealth (including public pension wealth), at its lifecycle peak. This ratio is 19 and 22 for the two low education types, but is substantially higher at 25 and 27 for the two high education types.

The estimates of ν (the consumption weight in the utility function) imply that, relative to those with less education, those with more education place a greater weight on consumption relative to leisure. The estimated values of ν can be given a tangible interpretation by considering their implication for the proportion of consumption they imply must be replaced on exiting work to keep marginal utility constant. These replacement rates are 69.0% and 68.7% for the two low education types but of 75.5% and 74.3% for the two high education types. The lower preference for leisure relative to consumption for the higher education groups (driven by later exits from the labour market, given the consumption possibilities afforded by their accumulated wealth) can perhaps be explained by their work tasks being less onerous, especially at older ages, making leisure time is less important.

The results on $r^{houscon}$ show that those with a DB pension get a greater consumption value from their housing wealth than those without. The estimates of the strength of the bequest motive (θ) can be given an intuitive interpretation by calculating their implication for the marginal propensity to consume in the last period of life (when death by next period is certain). The estimated parameter implies a marginal propensity to consume out of final period wealth of 3.8%, 3.5%, 6.1% and 5.3% for the four types.

---

[19]French (2005) notes that if the value function were concave, it would be possible to prove that this iterative procedure is a contraction and so a unique fixed point would exist. The value function here (as in French's paper) is not concave - however, using a number of starting values, it appears that unique fixed points for each type have been found.

Table 3: Preference parameter estimates

Type	Low education		High education	
	No DB	DB	No DB	DB
β	0.978	0.970	0.999	0.986
	(0.002)	(0.002)	(0.003)	(0.002)
ν	0.422	0.417	0.516	0.499
	(0.004)	(0.002)	(0.013)	(0.004)
$r^{houscon}$	0.022	0.031	0.027	0.036
	(0.001)	(0.002)	(0.013)	(0.002)
θ	0.020	0.018	0.059	0.053
	(0.002)	(0.001)	(0.013)	(0.002)
$\chi^2 stat.$	785	1171	230	653
(df)	83	59	83	59

The degrees of freedom in the χ^2 test differ across types as DC wealth moments are not used in the estimation procedure.

While the model is formally rejected by χ^2 tests of overidentification for each type (shown, with associated degrees of freedom in the bottom two rows of the table), Figures 3 (labour supply), 4 (non-pension wealth) and 5 (DC pension wealth) show that the model simulations replicate the economically-important features of the data. These include the decline in labour supply at older ages, the fact that households tend not to decumulate their non-pension wealth in the UK but rather retain it very late into the lifecycle, and the fact that those with Defined Benefit pensions accumulate very little Defined Contribution pension wealth.
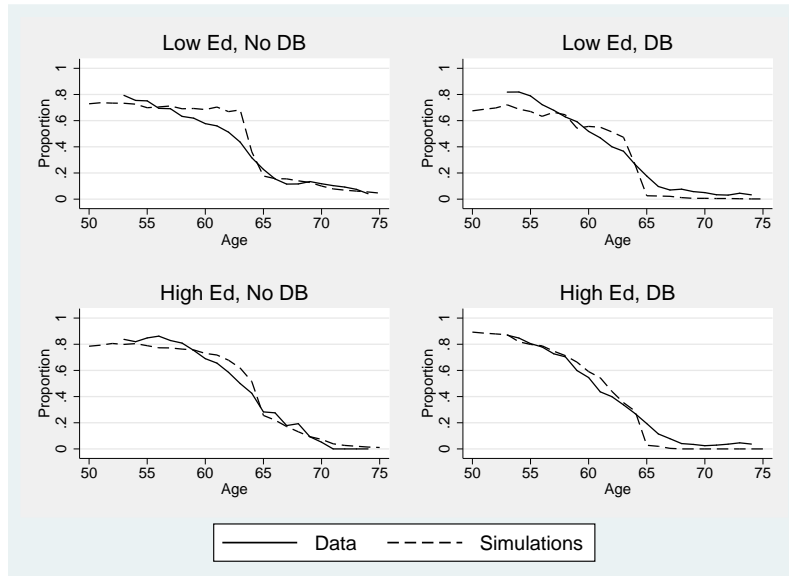
Figure 3: Fit: Labour supply
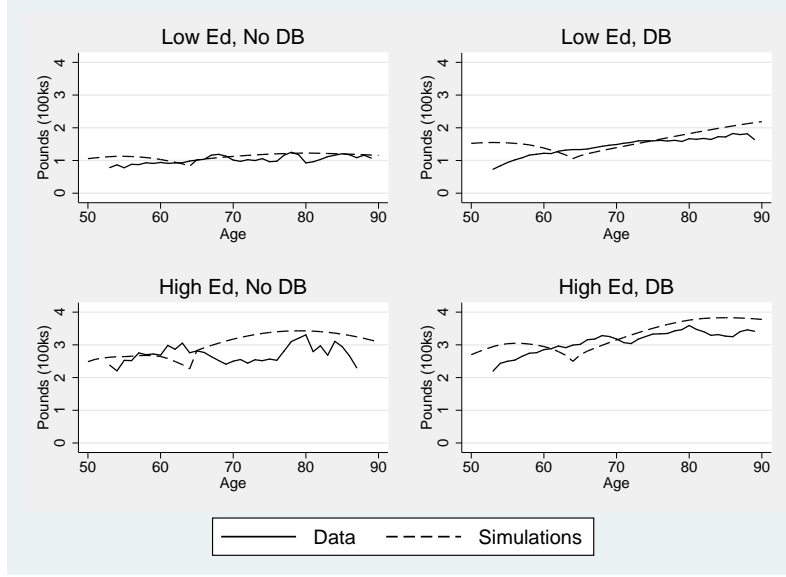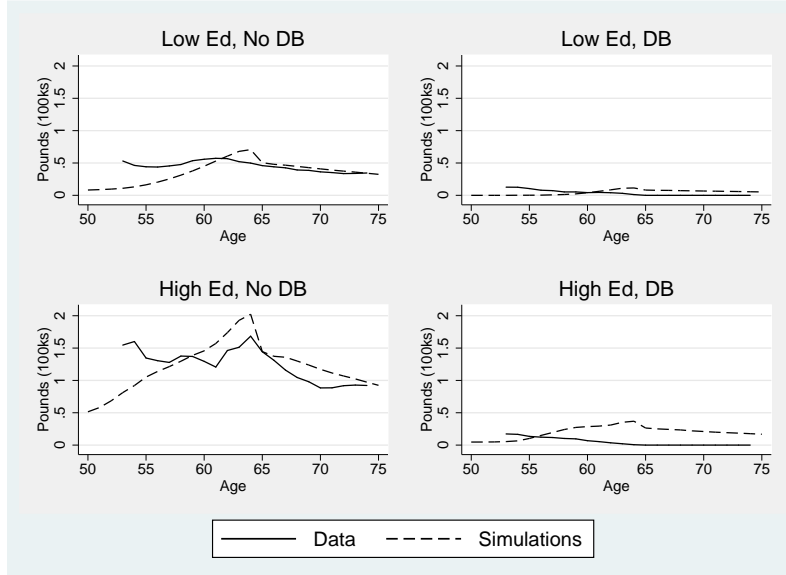
Figure 4: Fit: Non-pension wealth

Figure 5: Fit: DC pension wealth

The data profiles in Figures 3 to 5 are the moments that have been used to estimate the preference parameters. Appendix H shows how model simulations compare to data not used in estimation. Figures 17 and 18 show the 25th and 75th percentiles of modelled wealth (the sum of Defined Contribution pension wealth and non-pension wealth) and Figure 19 shows consumption over the lifecycle. That the match is also close in those figures indicates that the model can capture the heterogeneity that exists in wealth accumulation across households and consumption growth over the lifecycle.

5 Counterfactual analysis

This section uses the estimated model to compute the ex-ante optimal means-tested old-age income floor. Before turning to that, the manner in which welfare changes are measured is briefly outlined (with a longer treatment given in Appendix I).

5.1 Measuring Welfare Changes

The value function $V_t^j(\mathbf{X}_t)$, given in equation (9), expresses expected utility at age t for type j as a function of realised state variables. $\mathbb{E}_0[V_1^j(\mathbf{X}_1)]$ is therefore expected lifetime utility for type j before any uncertainty is realised, where the expectation operator is over $\mathbf{X}_1^j$ – the vector of initial state variables. Expected utility can also be expressed as a function of the optimal choices over those objects which give households utility (consumption, leisure and leaving bequests). Let $\mathfrak{s} = (\mathfrak{s}_1, \mathfrak{s}_2, \ldots, \mathfrak{s}_T)$ be the set of possible states of the world at every age. Each element in $\mathfrak{s}_t$ gives the history of realisations of the stochastic variables up to and including age t. Define household policy functions for each of consumption, leisure and bequeathed assets as $c(\mathfrak{s}), l(\mathfrak{s}), beq(\mathfrak{s})$. These functions, which give optimal behaviour as a function of the state of the world, are obtained by solving the Dynamic Progamming problem. The expected utility function ($\mathbb{E}_0[V_1^j(\mathbf{X}_1)]$) can be expressed as a function of these:

$$\tilde{V}_0^j\big(c(\mathfrak{s}), l(\mathfrak{s}), beq(\mathfrak{s})\big) \tag{11}$$

A policy change (for example a reform of the pension system) leaves the function $\tilde{V}_0^j$ unchanged but it will take different values of the arguments as households re-optimise in response to the reform. Denoting the post-reform policy functions as $c^{post}(\mathfrak{s}), l^{post}(\mathfrak{s})$ and $beq^{post}(\mathfrak{s})$, the new level of expected utility is:

$$\tilde{V}_0^j\big(c^{post}(\mathfrak{s}), l^{post}(\mathfrak{s}), beq^{post}(\mathfrak{s})\big)$$

The difference in welfare induced by the reform can be expressed by finding the proportionate change (Δ_j) in all *pre-reform* quantities that yield *post-reform* expected utility to households:

$$\tilde{V}_0^j\bigg((1+\Delta_j)c(\mathfrak{s}), (1+\Delta_j)l(\mathfrak{s}), (1+\Delta_j)beq(\mathfrak{s})\bigg) = \tilde{V}_0^j\bigg(c^1(\mathfrak{s}), l^1(\mathfrak{s}), beq^{post}(\mathfrak{s})\bigg) \tag{12}$$

This is a 'consumption-leisure-bequest' equivalent variation.[20]

---

[20]Common in the related literature (see for example Conesa et al. (2009), Low et al. (2010) or Braun et al. (2016)) is to express utility differences as a consumption equivalent variation (CEV) - the proportionate increase in (only) consumption in each state of world that would obtain the expected utility post-reform. This is not a sensible measure when a social welfare function averages utilities across households with different preferences for consumption. For a given utility difference, the CEV will tend to be larger the lower is the consumption weight on utility (as the less valuable is consumption to an agent the more additional consumption that will be needed to obtain a particular level of expected utility). Those household types

In evaluating any change from baseline policy, the social welfare function will be an equally-weighted average of these Δs (this applies a utilitarian social welfare function is applied along with an assumption of a population containing equal proportions of these types[21]):

$$W = \frac{1}{4}\Sigma_{j=1}^4 \Delta_j \tag{13}$$

5.2 Optimal income floor

The estimated model can be used to solve for the optimal level of the old-age income floor. To illustrate the types of changes that this experiment considers, Figure 6 illustrates how the mapping from pre-floor retirement income to post-floor retirement income would change with each of a 25% decrease and a 25% increase in the income floor.[22]
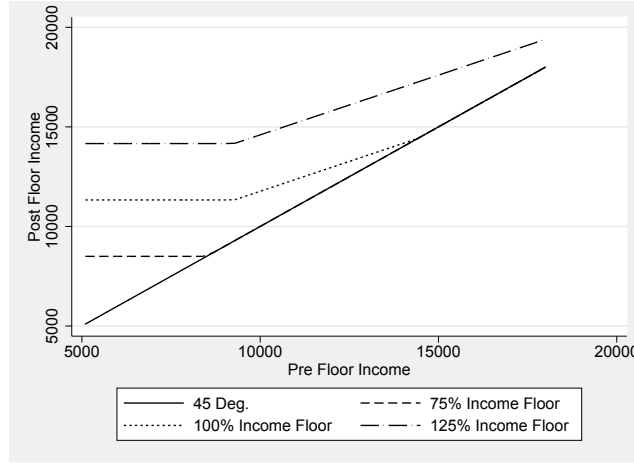

Figure 6: Illustration of changes to income floor

Government's problem   The government's problem, outlined formally in Appendix 5.2, is to maximise the social welfare function (equation (13)) by choosing the level of the income floor. This must be done subject to no change in the government budget balance over the lifecycle of this cohort. Different instruments are used (in different experiments) to balance the budget - these are i) a proportional tax/subsidy on all income, ii) changes in the contributory public pension, iii) changes in the UK's basic rate of tax[23] and iv) equal proportionate changes in both the UK's main income tax rates[24].

Table 4 reports the results of government's problem when the budget is balanced using the basic rate of tax (results using other instruments balancing the budget are shown later). The policy that maximises

---

with the lowest weights on consumption would bear most heavily on social welfare function. To avoid the government's objective function having this characteristic, a consumption/leisure/bequest equivalent variable (Δ) is used instead.

[21]Recall the caveat around interpreting the shares in Table 1 as population shares.

[22]In both cases, the point at which the effective marginal tax rate falls from 100% to 40% is left unchanged.

[23]This is a tax rate levied at 20% on earnings between approximately the 15th and 85th percentile of positive earnings.

[24]These two rates are 20% and 40%.

Table 4: Optimal income floor - budget balance using basic rate of tax

	Type			
	Low Ed		High Ed	
	No DB	DB	No DB	DB
Change in income floor	60.5%			
Change in basic rate of tax	$20\% \rightarrow 23.4\%$			
Change in lifetime welfare $(100\bar{\Delta})$	0.16			
Change in lifetime welfare $(100\Delta_j)$	0.86	-0.04	0.36	-0.54

social welfare increases the income floor by 60.5% (bringing it from £11,300 for couples to just over £18,000 and from £7,420 for singles to £11,870). This requires an increase in the basic rate of tax from 20% to 23.4% to balance the budget. Expected lifetime utility increases by 0.16%. This is the average across all four types – and so is the increase in expected utility before household type is realised. The change in expected utility once a household's type (but no other uncertainty) is realised is given in the final row of the table. There are welfare gains for both of the household types who don't have a DB pension. Among those who have a DB pension, the lesser educated households experience a very small fall in expected utility, with the higher educated households experiencing utility losses of just over half a percent of lifetime expected utility.

These welfare effects are generated by falls in consumption inequality and increases in leisure (the 'good news'), partly offset by falls in average consumption and in bequests left (the 'bad news'). Table 5 shows this (while Appendix J formally decomposes the total welfare effect into contributions coming from changes in the distributions of each of consumption, leisure and bequests). Taking the 'good news' first: household expected utility increases due to additional leisure time (row (1)), and falls in the variance of consumption (row (3)). The increase in utility coming from these is offset by the fact average consumption for the two highly educated types falls (row (2)) and bequests fall for each type as wealth at older ages falls (row (4)).

Grouping households of all types together, Figure 7(a) illustrates the change in the distribution of modelled wealth[25] induced by the move to the optimal system. It shows the 10th percentile, median and 90th percentile of wealth before the reform (solid line) and after it (dashed line). The 10th percentile is close to zero for most of the lifecycle and is zero by the age of 65 – that is the new income floor almost completely eliminates the incentive for those at the bottom to save. Wealth at the median is lower post-

---

[25]This is the wealth in the model that is accumulated endogenously: the sum of non-pension wealth and DC wealth; DB pension wealth is not included.

Table 5: Welfare effect – channels

	Quantity	Low Ed				High Ed			
		No DB		DB		No DB		DB	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
(1)	Average years of work	35.9	31.8	34.3	31.4	37.5	36.1	39.3	38.1
(2)	Average consumption (£1000s)	16.4	16.7	17.8	17.9	23.0	22.6	27.4	26.9
(3)	Var(log consumption)	0.11	0.07	0.10	0.07	0.14	0.12	0.12	0.12
(4)	Bequests (£1000s)	110	47	219	153	283	204	362	310

Row (3) shows the average (over years) of the annual variance of the log of consumption.
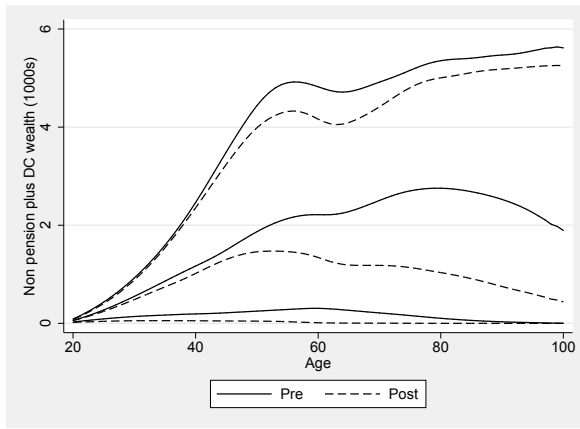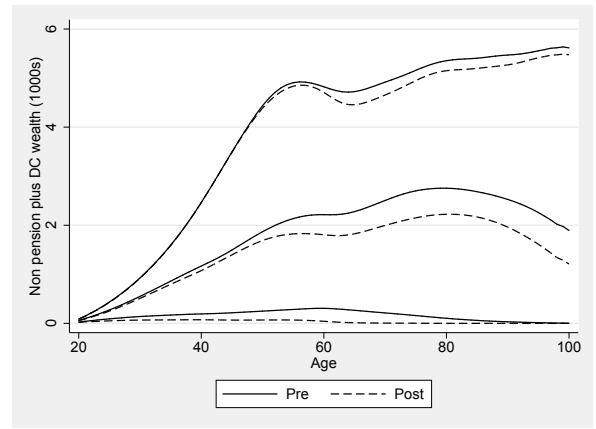
(a) Current tax treatment of private pensions   (b) Reformed tax treatment of private pensions

Figure 7: Modelled asset profiles under optimal systems

reform and is decumulated faster as households have a greater incentive to run down their wealth and rely on the more generous income floor. While wealth at the 90th percentile is lower over the whole of the lifecycle due to higher rates of income tax, it does not fall faster in old age as households who would have accumulated very high wealth stocks pre-reform are not attracted by the prospect of relying even on the more generous income floor.

The effects that this reform has on the wealth distribution are large (as were the effects of labour supply shown in Table 5). The next section shows how implementing an optimal income floor alongside changes in the treatment of private pensions can induce smaller distortions and yield substantially larger increases in welfare.

Changing tax treatment of private pensions   The counterfactual experiment reported above kept unchanged the tax treatment of private pension saving. The tax treatment in the UK has two particular features - i) tax-deductibility of contributions from earnings and ii) lower taxation of pension income

Table 6: Optimal income floor - reformed tax relief budget balance using basic rate of tax

	Type			
	Low Ed		High Ed	
	No DB	DB	No DB	DB
Change in income floor		34.1%		
Change in basic rate of tax		$20\% \rightarrow 18.6\%$		
Change in lifetime welfare ($100\bar{\Delta}$)		0.61		
Change in lifetime welfare ($100\Delta_j$)	0.96	0.59	0.58	0.33

relative to earnings. These features are expensive – they cost over 1.1% of GDP. Table 6 shows the optimal income floor when the second of these two features is removed.[26] This would move the UK to a system which retains tax-deductibility of private pension contributions, but taxes pension income and earnings equivalently. The additional tax revenue raised can be divided between a more generous old-age income floor and changes in the basic rate of tax. The system that optimises social welfare increases the income floor by 34.1% and reduces the basic rate of tax to 18.6%. The welfare gains here are equal to 0.61% of lifetime utility, substantially more than the gains of 0.16% of lifetime utility reported in Table 4 when the prevailing taxation of private pension saving is retained. Furthermore, in this case each of the four types gains from the reform. Figure 7(b) shows that the change in the distribution of modelled wealth induced by the move to this system is more modest than those when the optimal income floor, retaining private pension subsidies, is imposed.

It is worth exploring further the reason for the substantially higher increase in the income floor when private pension subsidies are retained (60.5%) compared to when they are removed (34.1%). Some intuition for what is driving this result can be obtained by looking directly at how the government budget balance responds to changes in the level of the income floor, keeping all other features of the tax system unchanged. Figure 8 illustrates this. It shows the change in government balance per household (on the vertical axis) for different proportional changes in the income floor (horizontal axis). Changes both allowing for and not allowing for behavioural responses are shown. In both cases reducing the generosity of the floor improves the government budget balance, but the improvement is greater when the fiscal implications of behavioural responses are not accounted for. This is in spite of the fact that removing the income floor leads to greater labour supply and earnings and therefore increases the tax base. The

---

[26]In particular, the changes made are to remove the tax-free lump sum, to levy payroll taxes (National Insurance Contributions) on income used to contribute to a pension fund, and to remove the fact that standard income tax deductions are higher for those over the age of 65.

explanation for this lies in the tax advantages associated with private pension saving. Reductions in the old-age income floor lead households to self-insure by increasing their (subsidised) private pension saving. The additional government spending on subsidising this self-insurance partially offsets the revenue saved by removing the income floor.
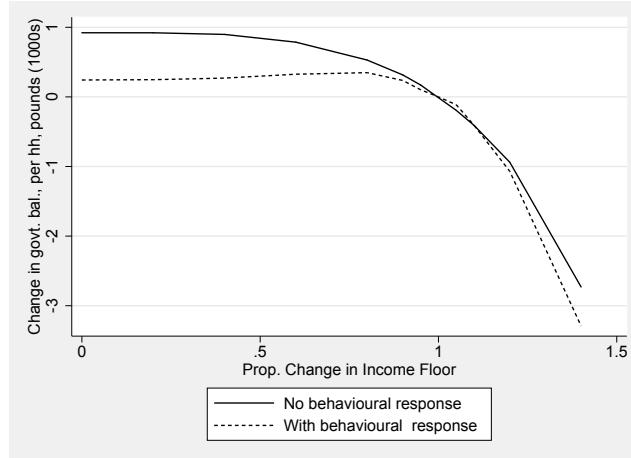

Figure 8: Fiscal implications of changing the income floor

Old-age income floors and private pension subsidies therefore act as substitutes in the government budget constraint: when the generosity of the former is increased (reduced), government spending on the latter falls (rises). Therefore, there are two distinct reasons for the very large (60.5%) optimal increase in the income floor when private pension subsidies are left in place. The first is straightforward - it is that income floors provide valuable insurance to households at a cost that households are prepared to pay. The second is that increasing income floors moves government spending on pensions from private pension subsidies towards income floors, which households prefer.

Balancing the budget through other means   The results so far have been for the case when the government budget is balanced using changes in the basic rate of tax. Table 7 shows equivalent results when the budget is balanced through reductions in the contributory public pension (panel (1)), through equal proportionate increases in both the basic and higher rates of tax (panel (2)), and through a proportional tax/subsidy on all net income (panel (3)). In all panels, the optimal system under the current tax treatment of private pensions and the reformed system that treats earnings and pension income equivalently is shown.

When the budget is balanced using the social-security style contributory public pension, the result is an extreme one - the optimal system involves abolishing the Social Security-style pension and using the resulting funds to substantially increase the income floor. This is how public pensions are delivered in Australia: a means-tested pension is provided but there is no contributory public pension and so the

richest elderly receive no pension from the state.

When both higher and basic rates of tax are used (increasing both in equal proportional terms) to balance the budget (panel (2)), the optimal increase in the income floor is smaller than in the base case. This is due to the fact that the labour supply effects of increases in taxation are found to be larger when higher earners are more heavily taxed.

The case where the budget is balanced by taxing/subsiding all income by a fixed proportion (panel (3)) and private pension subsidies are removed is the only scenario examined where the optimal design problem suggests a *lowering* of the income floor. The reason for this is that funding any increase in the income floor in this manner would involve reducing consumption possibilities in all states of the world, including, for example, those periods of working life when incomes are very low and the marginal utility of consumption is highest.[27] The utility cost of such taxes are large and the optimal policy is to reduce the income floor by 16.9%, allowing an increase in 1.8% of net income in each period. If private pension subsidies are left in place an increase in the income floor is optimal (in this case, income floors are less expensive as part of their cost is mitigated by reductions in household take-up of private pension subsidies).

Table 7: Balancing budget through other instruments

| | (1) | | (2) | | (3) | |
| | Cont. Pub. Pen. | | Both tax rates | | Prop. tax (subsidy) | |
Private pension treatment	Base	Reform	Base	Reform	Base	Reform
Change in income floor	67.9%	89.2%	40.3%	30.7%	38.9%	-16.9%
Change in contributory pension	-100.0%	-100.0%	-	-	-	-
New basic rate of tax	-	-	21.6%	18.3%	-	-
New higher rate of tax	-	-	43.2%	36.6%	-	-
Proportional tax (subsidy)	-	-	-	-	0.7%	-1.8%
Change in lifetime utility (100Δ)						
Average (all types)	0.29	0.70	0.14	0.64	0.10	0.67
Low Ed, No DB	1.32	1.77	0.66	0.93	0.59	0.31
Low Ed, DB	0.18	0.45	0.07	0.61	0.03	0.77
High Ed, No DB	0.24	0.98	0.13	0.59	0.08	0.71
High Ed, DB	-0.60	-0.40	-0.30	0.42	-0.30	0.87

Sensitivity to risk aversion assumption and other private pension arrangements   To investigate the sensitivity of the results presented here, Table 8 shows the outcomes from some alternative

---

[27]Households in these states of the world are less affected by increases in the other tax rates as the first £8,000 of income is exempted from income tax.

design problems. Column (1) reproduces the results from Table 6, that is, it shows the optimal income floor when a) private pension income is taxed equivalently to earnings and b) the budget is balanced using the basic rate of tax. Columns (2) and (3) show the optimal income floor when alternative values for the coefficient of relative risk aversion ($\gamma = 2$ and $\gamma = 4$ respectively) are used.[28] These can be compared to the baseline results in column (1) which are for $\gamma = 3$. The optimal level of the income floor increases with risk aversion but even with a coefficient of relative risk aversion of 2[29], the current retirement income floor is found to be too low.

Column (4) assesses the relevance of the results for future cohorts - for whom Defined Benefit pensions will be less prevalent than for the cohort whose behaviour is used to estimate preference parameters. The exercise conducted removes DB pension entitlement from those types who held them and removes their obligation to pay for DB pension. All households retain the ability to save in a DC pension. It is not clear *a priori* whether the optimal income floor will be higher or lower in a post-DB world. Households who previously had a DB pension are now exposed to investment risk on their DC funds (a risk against which they were previously insured by their employer). This will make income floors, which provide insurance against these risks, more valuable. However, their pension saving will now be more elastic (previously their DB pension wealth was assumed not to vary with changes in the policy environment), which will increase the distortions induced by, and taxes needed to fund, changes in the income floor. The latter effect dominates and the optimal increase in the retirement income floor is 22.1% (compared with 34.1% in the DB world).

The final column removes the tax deductibility of contributions into private pension savings. This means that income tax must be paid on all income, including income paid into a pension fund, in the period that it is earned.[30] The results here are very similar to those given in column (1) – the increase in the optimal income floor is slightly higher, the basic rate of tax is slightly lower, and the change in welfare is of a similar magnitude.

---

[28]The other preference parameters are re-estimated with the new levels of risk aversion imposed.

[29]This coefficient of relative risk aversion on utility implies a coefficient of relative risk aversion on consumption for the different types that varies between 1.4 and 1.5, which is at the lower end of values that have been estimated.

[30]In the US context that would mean only allowing pension saving in Roth-type accounts, in the UK it would mean pension saving would have the tax treatment of Individual Savings Accounts (ISAs).

Table 8: Additional Results

	(1)	(2)	(3)	(4)	(5)
	Reform	$\gamma = 2$	$\gamma = 4$	Post DB	No tax deferral
Change in income floor	34.1%	7.8%	67.5%	22.1%	36.0%
Basic tax rate	18.6%	18.3%	21.6%	18.7%	17.7%
Change in lifetime utility (100Δ)					
Average (all types)	0.61	0.32	0.63	0.41	0.69
Low Ed, No DB	0.96	0.32	1.49	0.62	1.17
Low Ed, DB	0.59	0.38	-0.05	0.42	0.41
High Ed, No DB	0.58	0.34	1.11	0.29	0.82
High Ed, DB	0.33	0.24	-0.05	0.31	0.38

An all-age income floor?   This paper argues that there is scope for revenue-neutral welfare-increasing increases in the income floor paid to those in old age. It is worth investigating the extent to which these results suggest a role for a more generous income floor for all ages. The advantages of an all-age income floor are similar to those for the elderly - they transfer resources from good states of the world (high income/high consumption states) to bad (low income/low consumption states). The costs are a diminished incentive to supply labour and a crowding-out of household saving.

There currently exists an all-age income floor for those not working in the UK that is less generous than the old-age income floor[31] which was denoted ui^{mt} in Section 3.3. I use the framework that generated the results in Table 4 to assess the optimal level of this all-age income floor. This suggests that the current all-age income floor (already substantially lower than the old-age income floor) is too high – the optimal level is 16% lower than the prevailing level. Such a reform would allow a reduction of 0.5% in the basic rate of tax. This is due to the fact that means-tested working-age income floors have a greater negative effect on labour supply, and the productivity of those whose behaviour they distort is much greater than the productivity of those whose behaviour responds to an old-age means-tested income floor. The optimality of lowering the working-age income floor implies that this paper's main results, which suggest the provision of more insurance to households, is not an automatic implication of the concavity of the household utility function and the set-up of the model. The analysis emphasises the value of providing good incentives to work for the young alongside pensions with good insurance properties for the old.

---

[31]Known as 'income-based job-seekers allowance' it guarantees (in the 2012/13 tax and benefit system) an income of just under £5,800 to couples. One can think of this as roughly analogous to SNAP (Food Stamps) in the US.

6 Conclusion

Providing public pensions is one of the costliest activities undertaken by governments in the developed world and a variety of different approaches are observed internationally. Many provide replacement rates that are either proportional to, or vary progressively with, career-average earnings while others provide a payment only to those who have income below a certain level. This paper investigates how governments should structure their public pension schemes and whether they should complement such schemes by subsidising private pensions.

The analysis suggests that private pension subsidies (such as tax rates on pensions that are lower than those on earnings) and pensions that are related to career-average earnings should be replaced with a combination of higher means-tested income floors in retirement and lower taxes on earnings. Private pension subsidies and earnings-related pensions project lifetime earnings risk into retirement. Means-tested income floors, on the other hand, provide valuable insurance against such earnings risks, as well as against investment risk and longevity. While there are distortions induced by providing a means-tested income floor to the elderly, these are more than offset by the value of this insurance. Means-tested income floors provided to those of working-age cause distortions with greater costs and should be substantially lower. Old-age is therefore a good time to provide insurance against lifetime productivity risk – calling into question the rationale for policies that, by amplifying those risks, do the opposite of this.

A   Parameter definitions

Table 9 summarises the parameters that enter the model and which are introduced in the body of the paper (excluding the appendices).

Table 9: Parameter definitions

Preference Parameters		State variables	
β	Discount factor	$\mathbf{X}$	Vector of all state variables
ν	Consumption weight in utility function	j	Household type
$r^{h_{houscon}}$	Value of housing in utility function	t	Age
θ	Weight on bequest	a	Non-pension wealth
K	Determinant of curvature of bequest	ue	Unemployment shock
γ	Coeff. of Rel. Risk Aversion (util.)	$\tilde{e}$	Productivity
		DC	DC pension wealth
		pp^{dc}	DC pension income
Labour market		h	Household composition
π_0	Prob. remaining in unemployment	ae	Average earnings
$\pi_1(\tilde{e})$	Prob. entering unemployment		
$\{\delta_i\}_{i=0}^2$	Params. of det. component of earn. proc.		
$\{\bar{\delta}_i\}_{i=0}^2$	Biased (due to selection) $\{\delta_i\}_{i=0}^2$	**Household choices**	
u	Stoch. component of earn. proc.	c^{nh}	Non-housing consumption
ρ	Autoregressive param. in stoch. earn.	l	Leisure
ξ	Innovation to autoregressive component	dc	DC pension contributions
σ_ζ	Variance of first innovation stoch. earn		
σ_ξ	Variance of subsequent innovations	**Utility function and arguments**	
m	Measurement error in earnings	$u()$	Single period utility function
σ_m^2	Variance of measurement error	c^h	Housing consumption
η	Stoch. earnings plus meas. err. $(u+m)$	c	Consumption $(= c^h + c^{nh})$
e	Primary earner earnings	L	Endowment of hours
e^s	Secondary earner earnings	hrs	Hours of work when employed
t^{ret_s}	Retirement age of secondary earner	$b()$	Bequest function
e^{data}	Primary earnings in data (inc. meas. error)	a^b	Assets bequeathed
$E()$	Post-unemployment productivity dist.	$gh(a,t)$	Function giving gross housing wealth
		$V()$	Value function
Assets		**Taxes and transfers**	
a	Non pension wealth	y	Net income
r_t	Return on pension wealth	ss	Public pension income
DC	DC wealth stock	$mtif$	Means-tested income floor
ϕ	DC wealth return realisation	ui	Unemp. insurance to those with no offer
$\bar{\phi}$	DC wealth mean return	ui^{mt}	Means-tested unemp. insurance.
σ_ϕ^2	DC wealth return variance		
ls^{dc}	DC tax free lump sum	**Counterfactual Analysis**	
q	Annuity rate (net of admin load)	$\mathfrak{s}_t$	History of stochastic realisations to period t
z	Administrative load	Δ	Cons, Leis, Bequest equiv. variation (CLBEV)
db	DB pension contribution	$\bar{V}_0()$	Value Function as function of policy func.
ς	DB contribution earnings proportion	$c(\mathfrak{s})$	Consumption policy function
pp^{db}	DB pension income	$l(\mathfrak{s})$	Leisure policy function
$db(ae_{64})$	DB income function (on average earnings at 64)	$beq(\mathfrak{s})$	Bequest policy function
Other			
s_{t+1}	Surv. prob. to $t+1$, (cond. on surv to t)		
$F()$	Dist. over $ue, \tilde{e}, \phi$		

B   Appendix to section 2

United Kingdom   The cost of tax relief on private pension saving less the tax revenue raised on private pensions was £21bn in 2013/14 (HMRC (2015)) or 1.1% of GDP.

United States   The cost of tax expenditures on private pensions was \$152bn in 2014 (or 0.9% of GDP). This is the sum (from Table 3 of US Department of the Treasury (2013)) of tax expenditures on Defined Contribution Plans, Defined Benefit Plans and Individual Retirement Accounts.

C   Model supplemental details

C.1   Housing consumption

Consumption, an argument of the period utility function, is the sum of non-housing consumption (a choice variable) and housing consumption, which is given by the product of $r^{houscon}$ (a parameter to be estimated) and gross housing wealth (re-producing equation (5)):

$$c_t^h = r^{houscon} gh(a, t).$$

The function (gh) which gives gross housing wealth as a function of the state variables non-pension wealth (a) and age (t) will be given in equation (16) below. Building up to it, the calculation is introduced in two steps. The first is to split non-pension wealth into *net* housing wealth and cash. The second is to convert net housing wealth to gross housing wealth. Taking those two steps in turn:

1. Splitting non-pension wealth into net housing wealth and cash   Net housing wealth and cash are given by:

$$
\begin{aligned}
a_t^c &= s^c(a_t, t) a_t \\
a_t^h &= \left(1 - s^c(a_t, t)\right) a_t
\end{aligned}
$$

where the share held in cash (s^c) is not a choice but is a function that depends on two state variables. The function contains a quadratic in age (t), a quadratic in assets (a) and an indicator for being at or over the age of 65[32] embedded in a Normal CDF to constrain it to be between 0 and 1:

$$s^c(a_t, t) = \Psi(\omega_0 + \omega_1 t + \omega_2 t^2 + \omega_3 a_t + \omega_4 a_t^2 + \omega_5 \mathbb{1}[t \geq 65]) \tag{14}$$

---

[32]This last variable is included as, at the time of receipt of private pensions (modelled as at age 65), the cash lump-sum will increase the share of wealth that is held in cash form.

2. Obtaining gross housing wealth from net housing wealth Households obtain consumption from their gross housing wealth, which is equal to their net housing wealth plus mortgage outstanding (*mort*). To obtain gross housing wealth from net housing wealth, it is necessary to know the leverage ratio (the ratio of mortgage debt to gross housing wealth). Knowledge of the leverage ratio (*lev*) allows gross housing wealth to be calculated[33]

$$gh(a,t) = \frac{1}{(1 - lev)} a^h$$

where $lev(t)$ is the leverage ratio at age t. The leverage ratio is not a modelled choice, but is set to be a (quadratic) function of age:

$$lev(t) = \mu_0 + \mu_1 t + \mu_2 t^2 \tag{15}$$

These two steps together give the function for gross housing wealth:

$$gh(a,t) = \underbrace{\frac{1}{(1 - lev(t))} \underbrace{(1 - s^c(a,t))(a)}_{\text{Net hous. wealth}}}_{\text{Gross hous. wealth}} \tag{16}$$

The estimated functions that give the share of wealth held in housing (equation (14)) is illustrated in Figures 11 and 12 while the function giving the leverage ratio (equation (15)) is given in Figure 13.

D Estimation and parameterisation - supplementary details

D.1 Unemployment probabilities and re-employment productivity distributions

Primary household earners in the data are defined as unemployed if they had less earnings in a year than the level provided by Jobseekers Allowance (the main benefit to the unemployed in the UK in 2012/13 (£3,692)). Data on primary earners up to the age of 50[34] is used to calculate the probabilities in the Markov transition matrix. These are given in Table 10. Table 11 shows the probability distribution over productivity deciles on exiting a state of unemployment.

---

[33]The simple calculations are:
$$\frac{1}{1 - lev} a^h = \frac{1}{1 - \frac{mort}{gh}} a^h = \frac{gh}{a^h} a^h = gh$$

[34]These employment probabilities represent the risk of involuntary unemployment. The maximum age is chosen to limit the effect of individuals choosing not to work, which occurs to a much greater extent among those close to conventional retirement ages.

Table 10: Unemployment probabilities

	Type			
	Low Ed		High Ed	
	No DB	DB	No DB	DB
	π_0			
	0.771	0.696	0.775	0.665
Productivity Decile		$\pi_1(\tilde{e})$		
1	0.227	0.099	0.201	0.063
2	0.049	0.026	0.024	0.016
3	0.027	0.010	0.054	0.01
4	0.014	0.013	0.032	0.006
5	0.004	0.008	0.029	0.003
6	0.007	0.012	0.038	0.002
7	0.013	0.003	0.020	0.004
8	0.017	0.003	0.005	0.001
9	0.007	0.002	0.004	0.003
10	0.010	0.000	0.010	0.000
	$\overline{ue}$			
	0.13	0.05	0.15	0.03

Table 11: Re-employment productivity distribution

	Type			
	Low Ed		High Ed	
	No DB	DB	No DB	DB
Productivity Decile				
1	0.130	0.174	0.118	0.238
2	0.043	0.005	0.029	0.032
3	0.022	0.013	0.037	0.020
4	0.005	0.017	0.011	0.008
5	0.002	0.013	0.008	0.008
6	0.007	0.017	0.013	0.012
7	0.010	0.013	0.003	0.008
8	0.002	0.003	0.000	0.004
9	0.000	0.003	0.005	0.000
10	0.007	0.000	0.000	0.004

D.2   Defined Contribution pension details

The mean and standard deviation of pension fund returns are based on an index known as the "DCisions index"[35]. This is an index of total fund return that reflects the asset allocation decisions made by leading

---

[35]http://www.ftse.com/Indices/FTSE_DCisions_Index_Series/

DC pension funds in their default investment strategies. This index provides information on returns stretching back to 1994. For years prior to 1994 when the DCisions index is not available, ϕ_t is estimated using the FTSE all-share index (on which data is available back to the early 1960s) and the ratio between the FTSE all-share index and the DCisions index over the period where both are available (1994 - 2010). The estimation of this is discussed further in Crawford and O'Dea (2014). The mean and standard deviation, respectively, of this time series are $\bar{\phi} = 3.97\%$ and $\sigma_\phi = 13.8\%$.

D.3 Defined Benefit Pensions

The function that relates average earnings at the age of 64 to defined benefit pension income in retirement is specified as a quadratic through the origin (equation (18)). The coefficients (db_1^j, db_2^j) are calculated using a regression of projected DB income on average earnings for a sample of those aged 60 to 64. Projected DB pension income is estimated using survey responses (see Banks et al. (2005) for more details on construction of this variable). Figure 9 shows the estimated relationship between earnings of the principal earner at the age of 64 and DB pension income for both education types.

$$
\begin{aligned}
db_t(ae_{64}, j) &= db_1^j ae_{64} + db_2^j ae_{64}^2 & \text{if} \quad ae_t < \hat{ae}_{db}^j & \quad (17) \\
&= db_1^j \hat{ae}_{db}^j + db_2^j (\hat{ae}_{db}^j)^2 & \text{if} \quad ae_t \geq \hat{ae}_{db}^j & \quad (18)
\end{aligned}
$$

where $\hat{ae}_{db}^j = \frac{-db_1^j}{2db_2^j}$, the point at which the quadratic starts to decrease.
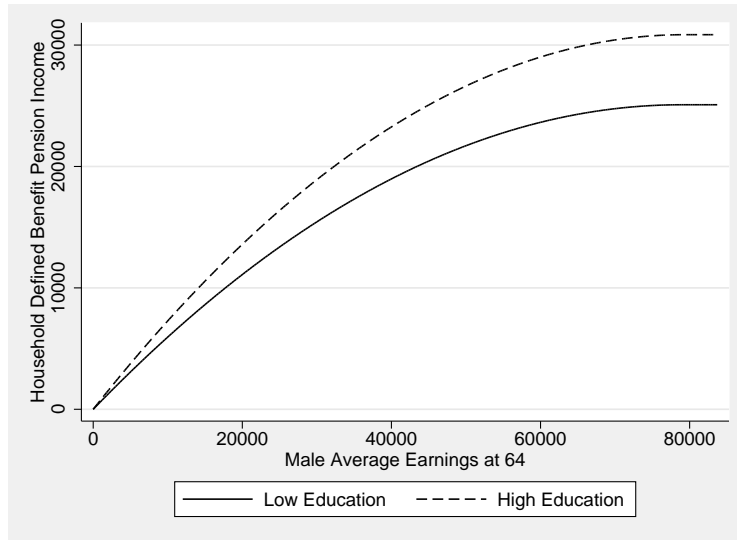

Figure 9: Modelled Defined Benefit pension entitlements

The proportion of earnings that employees with DB pensions must pay (ς) is set at 5% - close to the average paid by government employees in the UK.

D.4 Public pensions

The function that relates the principal earner's career-average earnings at the age of 64 to the Social Security style public pension (known in the UK as the state pension) income in retirement is specified as a quadratic through the origin (equation (20)). It is estimated using the sub-sample of those aged 60 to 64, who are close to their maximum level of public pension accumulation. These entitlements are calculated using the rules of the state pension system and the history of contributions (see Bozio et al. (2010) for more details). Figure 10 illustrates this relationship. This figure differs from Figure 1 which introduced the UK contributory public pension system in two ways. First, it maps the principal earner's career-average earnings into *household* pension entitlements rather than *individual* pension entitlements. Second, the estimation is parametric rather than non-parametric.

$$
\begin{aligned}
ss_t(ae_{64}, j) &= ss_1 ae_{64} + ss_2 ae_{64}^2 && \text{if} \quad ae_t < \hat{a}e_{ss} && (19) \\
&= ss_1 \hat{a}e_{ss} + ss_2 (\hat{a}e_{ss})^2 && \text{if} \quad ae_t \geq \hat{a}e_{ss} && (20)
\end{aligned}
$$

where $\hat{a}e_{ss} = \frac{-ss_1}{2ss_2}$, the point at which the quadratic starts to decrease.
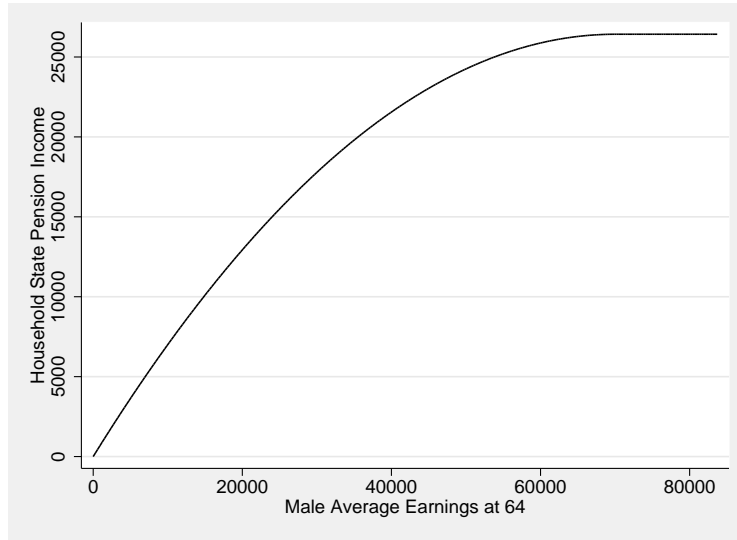

Figure 10: Modelled State Pension entitlements

D.5 Non-pension wealth return

The return on non-pension wealth (r_t) is a function of age and the level of non-pension wealth. Non-pension wealth can be thought of as the i) financial wealth plus ii) gross housing wealth less iii) mortgage wealth:

$$a_t = a^c + \underbrace{gh - mort}_{a^h} \tag{21}$$

where:

1. a^c is cash.

2. gh is gross housing wealth

3. $mort$ is mortgage debt

These three objects do not enter the model separately (households simply make a decision of how much non-pension wealth to hold), but the parameterisation of the return on non-pension wealth (r_t) depends on them.

The return on a given quantity of non-pension wealth (a_t) is:

$$r_t a_t = r^c a_t^c + r^{gh} gh_t - r^{mort} mort_t \tag{22}$$

where:

1. r^c is the return on cash

2. r^{gh} is the return on gross housing wealth

3. r^{mort} is the interest rate on mortgage debt

The three asset components and three rates of return are now treated in turn:

Components of non-pension wealth

1. **Cash**: $s^c(a_t, t)$ – the share of non-pension wealth held in cash form — was discussed above and is given by a function (equation (14)), the parameters of which are estimated using data from the Wealth and Assets Survey. Figure 11 shows the share of wealth held in cash by the level of non-pension wealth (holding age fixed at 50) for each type. The share has a U-shape with age – those with the least and the most wealth hold the most cash while those with middle levels of wealth
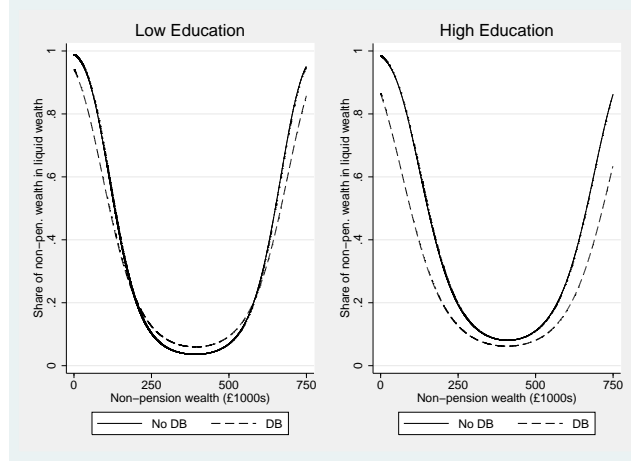
Figure 11: Share of non-pension wealth held in cash (equation (14)), by non-pen wealth and type
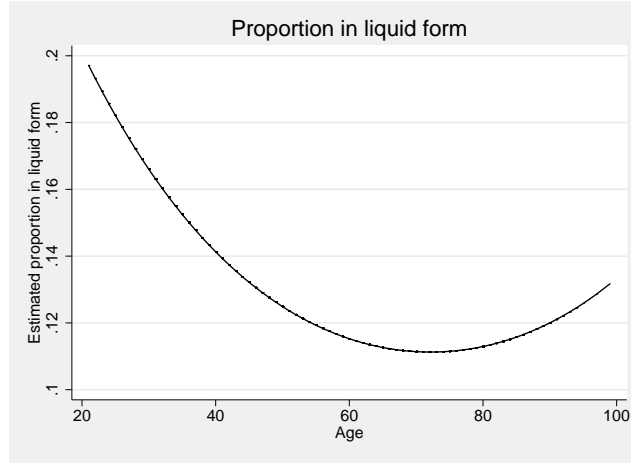

Figure 12: Share of non-pension wealth held in cash (equation (14)), by age

hold larger shares of their wealth in housing. Figure 12 shows the predicted share of wealth by age, holding wealth fixed at £200,000 (this is shown only for those with high education but without DB wealth - the pattern for the other types is very similar). The pattern by age is also U-shaped, though comparison of the scales of the two figures shows that the differences by age are of a more modest magnitude than the differences by wealth holdings.

2. **Gross housing wealth**: The function that returns gross housing wealth was introduced in Appendix C.1 and is given in equation (16), reproduced here:

$$gh(a,t) = \underbrace{\frac{1}{\big(1 - lev(t)\big)} \underbrace{\big(1 - s^c(a,t)\big)(a)}_{\text{Net hous. wealth}}}_{\text{Gross hous. wealth}}$$

The leverage ratio is a quadratic in age (and was given in equation 15). Its parameters are estimated
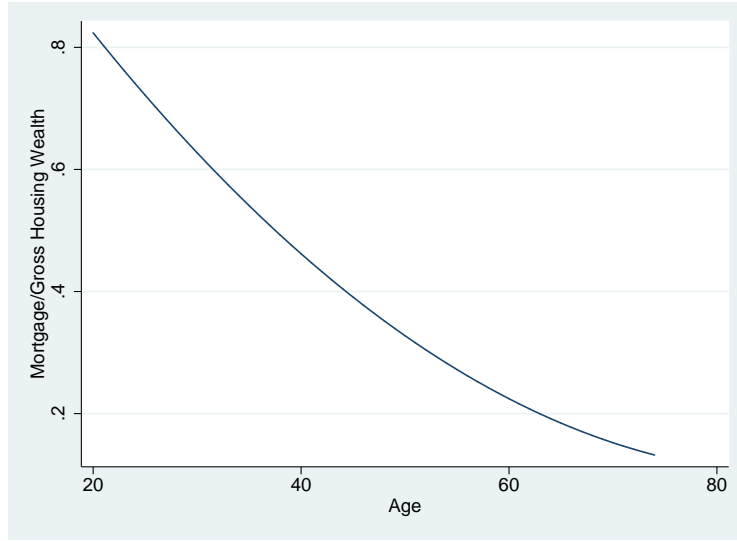
Figure 13: Leverage ratio (equation (15))

using data from the British Household Panel Survey. Figure 13 illustrates this function and shows that the leverage ratio (mortgage outstanding divided by gross housing wealth) declines with age.

3. **Mortgage debt**

   Mortgage debt is equal to gross housing gh less net housing wealth.[36]

Rates of return

1. **Return on cash**: The return on cash (r^c) is set at 1.6% – the average real return on cash balances between 1952 and 2012 (see Table 1 of Barclays Capital (2012)).

2. **Return on gross housing**: Gross housing accrues capital gains at a constant annual rate (r^{gh}). This is set to 2.8% – the average real appreciation of house prices from 1975 to 2013 calculated using data from Nationwide Building Society (2014).

3. **The mortgage interest rate**: The mortgage interest rate (r^m) is set at 3.5%. This is an average real rate from 1975 to 2013 calculated using historical interest rate data from Bank of England (2013).

---

[36]Or, equivalently, it is a function of the leverage ratio and net housing wealth:

$$mort = \frac{lev}{1 - lev} a^h$$

Summary

The implications of these for the return on net housing wealth ($\frac{r^{gh}gh - r^{mort}mort}{a^h}$) is illustrated in Figure 14. The function is slightly negative at the start of life[37] as the the mortgage rate is greater than the return on gross housing wealth which for young, highly leveraged households implies a negative return. As households age, the leverage ratio falls to zero, and the rate of return on housing converges to 2.8% – the rate of return on gross housing wealth.
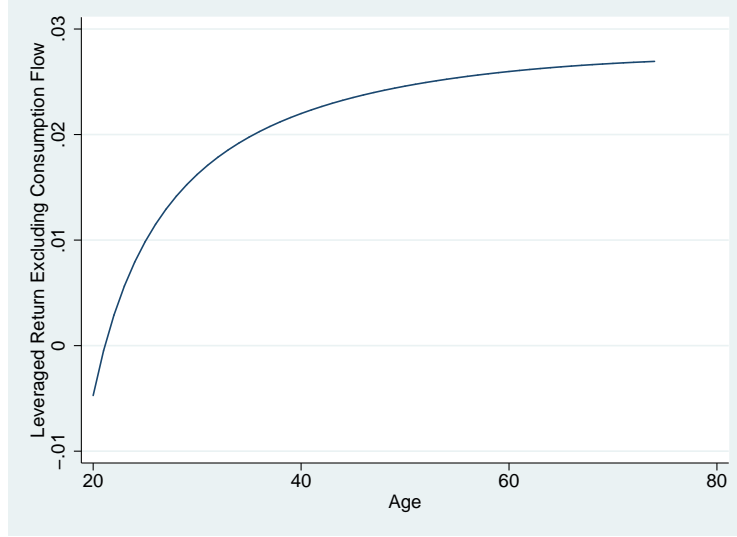

Figure 14: Leveraged return on net housing wealth

D.6 Earnings and retirement of the secondary earner

Earnings (e^s) of the secondary earner are set equal to their average earnings in the data over all years (including years when these earnings were zero). For each of the four types, these values are, respectively, £5,121, £5,282, £6,840 and £7,684 per year. Households get this level of earnings until the household's primary earner is aged $t^{ret_s} = 63$. In the data used in this paper, main earners are typically men, and secondary earners are typically women. The median age gap between men and women is 3 years. The age of 60 has been the focal retirement age for women in the UK for many years - this occurs at a male age (or 'household age') of 63.

D.7 Taxes and benefts

This section outlines the household tax and benefit function used in the model. The components of the tax and benefit system that are modelled are: Income Tax, National Insurance contributions, Working Tax Credit, Jobseekers Allowance (payments to the unemployed), the UK's public pension (including the

---

[37]This does not mean that households have no incentive to accumulate this form of wealth at these ages – recall that ownership of housing will also give them housing consumption, which is valued through the utility function.

Basic State Pension and Additional State Pensions) and Pension Credit (which delivers an income floor in old-age). The tax and benefit function $\tau()$ is a function of earnings of both members of the couple (e_t, e_t^s), interest income (r), whether the household has been subject to an unemployment shock (ue), public pension income (ss), private pension income (pp), unannuitised DC wealth (DC), cash assets (a^c), age (t), household composition (h) and average earnings (ae):

$$\tau^h(e, e^f, r, ue, ss, pp, cont, DC, a^c, t, h, ae) \tag{23}$$

Bequests are also taxed - the bequest taxation function which returns net bequests (a^b) is shown below.

The components of the tax function are now outlined:

Income tax   Income tax is levied on the sum of earnings, interest income, state and private pension income, less any contributions to private pensions. Tax is levied at the individual level. Taxable income for the principal and secondary earner respectively is (non-earned income is assumed to accrue to the principal earner):

$$
\begin{aligned}
ti &= e + r + ss + pp - dc - db \\
ti^s &= e^s
\end{aligned}
$$

Under the 2012/13 system, which is the basis for the tax function used here, income is taxed in three bands: the first tranche of income is untaxed, the second is taxed at 20% and the third at 40%.[38] The thresholds that define the bands vary with age, with a more generous treatment of older individuals. The equations below, together with Table 12, summarise the income tax system used in the model.

$$
\begin{aligned}
it(ti, t) &= 0 & \text{if } ti \leq \kappa_1^t \\
&= 0.2(ti - \kappa_1^t) & \text{if } \kappa_1^t < ti \leq \kappa_2^t \\
&= 0.2(\kappa_2^t - \kappa_1^t) + 0.4(ti - \kappa_2^t) & \text{if } \kappa_2^t < ti
\end{aligned}
$$

---

[38]An additional band, introduced in 2010 at 50% and reduced in 2012 to 45% is not modelled.

Table 12: Income tax thresholds

	Age		
	< 65	$65-74$	≥ 75
κ_1	8,105	10,500	10,600
κ_2	42,470	44,870	44,970

Employee National Insurance Employee National Insurance contributions are levied on earnings (capital income and other forms of income are exempt) and only on those aged less than the state pension age. Under the 2012/13 system it is levied at a rate of 12% of earnings between the 'Primary Threshold' ($pthresh$ - £7,952) and the 'Upper Earnings Limit' (uel - £42,511) and at 2% on earnings above that level. National Insurance is levied on a measure of earnings that excludes payments into Defined Contribution and Defined Pension funds ($e^{ni} = e - dc - db$).

$$ni(e,t,dc,db) = 0 \qquad\qquad\qquad\qquad \text{if } t \geq 65 \text{ or } e < pthresh$$

$$ni(e,t,dc,db) = 0.12(e^{ni} - pthresh) \qquad\qquad \text{if } t < 65 \text{ and } e \geq pthresh \text{ and } e^{ni} < uel$$

$$ni(e,t,dc,db) = 0.12(uel - pthresh) + 0.02(e^{ni} - uel) \text{ if } t < 65 \text{ and } e^{ni} \geq uel$$

Jobseekers' Allowance Unemployment shocks are assumed to be verifiable by the government. Affected households receive an unemployment payment (ui) irrespective of their accumulated assets. This is set equal to the level paid by the UK's 'contribution-based jobseekers' allowance level' and is paid to unemployed individuals who are looking for work. In 2012/13 an unemployed individual was entitled to £3,692

All those out of work (including those who get an offer but who simply choose not to work) can receive an asset-tested payment (ui^{mt}) if they are sufficiently poor. The structure of this follow's the UK's 'income-based' jobseekers allowance.

Working Tax Credit The UK has an in-work low-income subsidy (similar to the US Earned Income Tax Credit) known as the Working Tax Credit (WTC). The full operation of the tax (which depends on hours of work) is discussed in Pope and Roantree (2014). A slightly simplified version is modelled here. Single (couple) households in work with income (net of pension contributions) of less than £6,420 receive a maximum of £2,710 (£4,660). Earned income (again, net of pension contributions) in excess of £6,420 results in WTC being withdrawn at a rate of 41%.

State Pensions State pension entitlements are modelled as an (estimated) function of average earnings at the age of 64. The function is illustrated in Figure 10.

Pension Credit Pension Credit (PC) plays an important role in this paper – providing an income floor to the elderly. It was introduced in Section 2 and entitlement as a function of 'notional' income was illustrated in Figure 1. Notional income, which is used to assess entitlement to PC, is the sum of actual and imputed streams of income. Notional income includes any earnings, interest income, private pension or state pension income as well as an imputed stream of income from non-pension wealth imp:

$$y^{notional} = e + r + pp + sp + imp$$

imp is calculated as 10% (annually) of the stock of non-pension wealth with two exemptions. Wealth held in the principal private residence (family home) does not generate a stream of notional income (so, in the model, only cash, a^c, counts), nor does the first £10,000 of other non-pension wealth.

$$imp = 0.1 \max((a^c - 10,000), 0)$$

How the benefit operates can best be seen (for a single individual) from Figure 1. Formally, once notional income is calculated, entitlement in the model is given by:

$$
\begin{aligned}
pc(y^{notional}&(e, e^s, r, ss, pp, DC, a^c, t)) \\
&= \max(GC - \min(y^{notional}, SC) - t(max(y^{notional} - SC, 0)), 0) \quad \text{if } t \geq 65 \\
&= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } t < 65
\end{aligned}
$$

where GC is the 'Guarantee Credit Level' - the minimum income guaranteed to all pensioners in retirement, SC is the 'Savings Credit Threshold', the income level up to which PC is withdrawn at an effective tax rate of 100% and t is the taper rate (currently 40%) - the effective tax rate applied on notional income over SC. Table 13 gives the values of GC and SC for both singles and couples.[39]

Net earnings and income taxes/benefits Net taxes and benefits are:

---

[39]Pension Credit, as modelled above, reflects the system as it operates for recipients over the age of 65. In reality, households can be entitled to a less generous version of the benefit between the ages of 60 and 65 (see Section 4.4.3 of Hood and Oakley (2014)) for more details. Modelled households are not considered to be eligible for PC until the age of 65.

Table 13: Pension Credit parameters

| | Household type | |
	Singles	Couples
GC	7,400	11,300
SC	5,800	9,200

$$\tau^h(e, e^s, r, ue, sp, pp, cont, DC, a^c, t, h, ae) =$$

$$it(ti^m(e, r, sp, pp, cont), t) + it(ti^f(e^f), t)$$

$$+ ni(e, t, cont) + ni(e^f, t, 0)$$

$$- wtc(e, cont)$$

$$- jsa(ue, a_t)$$

$$- sp(ae, t)$$

$$- pc(y^{notional}(e, e^f, r, sp, pp, DC, a^c, t))$$

Bequest taxation Bequests to those other than a spouse are taxed in the UK. Bequests in the model occur only when the second and final member of the couple has died. That part of an estate over a certain threshold is taxed at a rate of 40%. That threshold is £650,000 (each spouse can leave up to £325,000 tax free).

$$\tau^b(A) = 0.4 \max\left((A - 650,000), 0\right) \tag{24}$$

Employer National Insurance contributions Employers must remit a National Insurance contribution assessed on their employee's earnings. This payment is not taken into account by employees in their decision making process and so does not enter the household tax and benefit function τ^h, nor does it play a role in the model solution. However, in Section 5, when considering changes in government balance as a result of changes in the design of pension systems, changes in revenue from employer NI (assuming *employer* behaviour is unchanged) are included.

Employer National Insurance Contributions (τ^{emp}) have a similar form to employee National Insurance contributions. Under the 2012/13 system they are levied at a rate of 13.8% of earnings above the 'Secondary Earnings Threshold' (£7,488).

$$\tau^{emp}(e, t) = 0.138 \max\left((e - 7,488), 0\right) \quad \text{if} \quad t < 65 \tag{25}$$

E   Data

E.1   Earnings data

To estimate the parameters of equations that determine earnings capacity (given in equations (3) and (4)), panel data on earnings for each of the four types is needed. This data comes from two sources - earnings data from before 2002 (when the survey started) is calculated using linked administrative data. This is combined with survey data from ELSA for years after 2002. The earnings data from these administrative and survey sources are now discussed.

Administrative data   The National Insurance (NI) data are the administrative record of individuals' National Insurance contributions, and the dataset that is used by the UK government to establish individuals' rights to claim contributory benefits such as the state pension. This data is used to estimate ELSA respondents' history of earnings. The NI records cover the years 1948 to 2003, though there are different levels of information for each of three sub-periods: 1948-1974, 1975-1996 and 1997 to 2003. **1997 to 2003**: Taking the most recent period first, the NI records contain uncensored data on annual earnings as, in these years, employers were required to report the total earnings of their employees.

1975 to 1996: For the middle period the NI records contain data on employee National Insurance contributions. National Insurance payments in that interval were levied as a proportion of earnings between two values which are known as the Lower Earnings Limit (LEL) and the Upper Earnings Limit (UEL). For the period under consideration these values have been located at approximately the 8th and 80th percentile of the distribution of (positive) earnings. This data on NI contributions therefore allow us to calculate earnings, subject to right-censoring at the UEL and conditional on there being some earnings above the LEL.

To predict censored earnings in the years 1975 to 1996, the first step is to obtain the coefficients of a fixed-effect Tobit on earnings from 1975 to 2003 with the censoring point in each year up to 1996 equal to UEL (from 1997 there is no censoring). These coefficients are used to predict earnings for those who are affected by the censoring. The fixed-effect Tobit, when the length of the panel is fixed, is known to yield inconsistent results due to the incidental parameters problem (see Neyman and Scott (1948) for a general discussion of this problem). However Greene (2004) investigates, using Monte Carlo methods, its properties and finds that parameters of the fixed effects Tobit model are little affected by this problem even with panel of lengths substantially shorter than the panel used here (which has length 29). Further, Figure 15 shows a plot of selected quantiles of earnings through time using the censored and imputed data prior to 1997 and the uncensored data from 1997 onwards. This shows only a very small discontinuity in 1997.

Figure 15: Selected quantiles of earnings
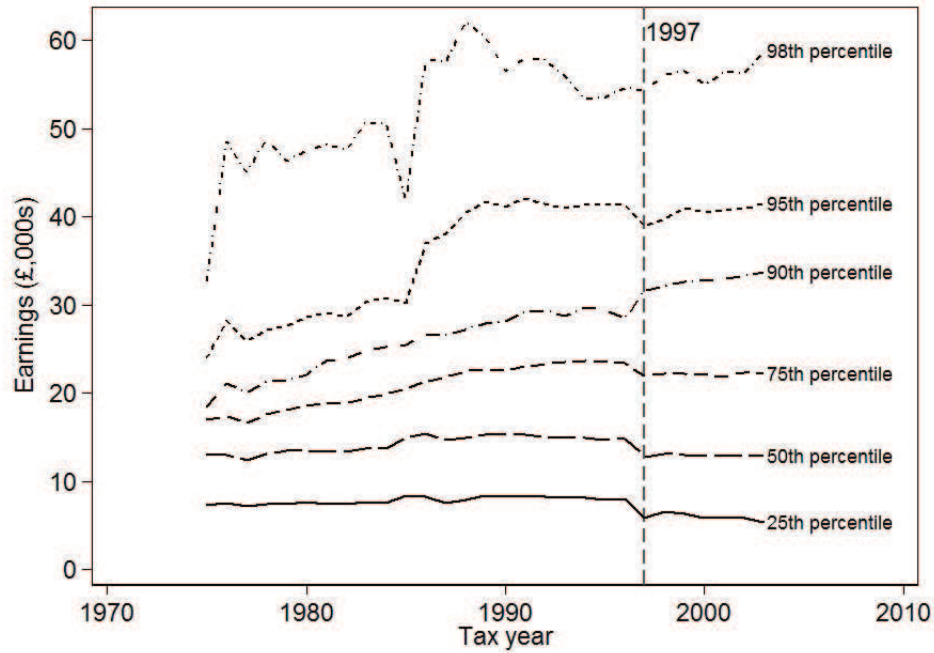
Prior to 1975: Before 1975 the NI records contain only data on the number of weeks that an individual earned above the LEL (and therefore paid NI contributions) and not the level of earnings. (This is because during this period the level of earnings was not relevant to the accrual of rights to working-age or retirement state benefits.)

To simulate earnings before 1975, we first obtain potential earnings if in work by 'backcasting' the fixed effect from Tobits described above. The fixed effect gives a measure of 'permanent' earnings (in 1975 earnings terms - as the the dummy for that year is the one which is excluded). For years before 1975, this level is adjusted for average economy-wide earnings growth and individual level earnings growth given an individual's age, sex and education level. Having obtained this measure of 'potential earnings' in each year, the next step is to predict the years in which the individuals were working. The NI data records how many weeks the individual made NI contributions between 1948 and 1975. Men are assumed to have worked those weeks immediately prior to 1975 (therefore any periods not working were at the start of working life). To take account of the diminished propensity for women to work after having children, it is assumed that they worked those weeks from the point of leaving full-time education (therefore any periods not working were immediately prior to 1975). The combination of the estimates of potential earnings in a particular year for each individual and the years in which they were working yields the earnings estimates for years prior to 1975.

The discussion above relates only to earnings in employment and not income earned in self-employment.

National Insurance contributions are levied on self-employment income- but in a different manner than on earnings. As a result, the NI records enable us to identify years in which self-employment income was earned, but not the level of that income. The measure of earnings therefore excludes income from self-employment.

E.2 Wealth measures in the English Longitudinal Study of Ageing

The ELSA data contains detailed information on the components of household wealth. The main components of non-pension wealth (the analogue in the data to the state variable a) are net financial wealth (cash, stocks and shares less any outstanding financial debt), net primary housing wealth (gross housing wealth less any associated mortgages), other net property wealth, business wealth and physical wealth (land, antiques and collectibles). Moments of the sum of these (that is moments of non-pension wealth) are used in the method of moments procedure to match simulated non-pension wealth (a).

Defined Contribution wealth is equal to the accrued fund value. Moments of this are used in the method of moments procedure to match simulated Defined Contribution wealth (DC).

The model requires estimates of the relationship between average earnings in working life and each of state pension and Defined Benefit pension. Average earnings for sample members is obtained using the linked administrative data. Data on projected state pension income are obtained from Bozio et al. (2010) who calculate them using the rules of the state pension system and the same administrative data on contributions used here. Projected DB pension income is estimated using survey responses (see Banks et al. (2005) for more details). Only those aged 60 to 65 are used in the calculation of these processes, so most accumulation of pension wealth has been done. However, the calculation of each needs to tackle the fact that when some of those in the sample that is used here are interviewed, their actual pension income will depend on labour market outcomes that have not yet been realised. The assumption that underlies the estimates used here is that individuals who are not working when interviewed do not return to the labour market, whereas those that are working when interviewed continue to work until their current pension's normal retirement age (for those with DB pensions) or their state pension age (for those without).

The model splits households into DB and no DB types with the latter type not having any DB wealth. The reality is somewhat less stark; very few households of the cohort studied here have no DB pension wealth at all (which would occur if neither spouse ever worked in a job that provides a DB pension). Therefore some of the households who are in the 'No DB' sample have some DB wealth. My approach is to treat this wealth as if it were DC wealth. To convert future DB *income* (which is what the data contain) into DC *wealth* the following calculation (for a household surveyed at age t) is applied:

$$DC_t = \underbrace{\frac{db_{65}}{q_{65}}}_{A} \underbrace{\frac{yw_t}{yw_{65}}}_{B} \underbrace{(1 + \phi)^{t-65}}_{C}$$

The first term (A) is projected DB income at the age of 65 divided by the annuity rate at 65 – this is a measure of DB wealth at the age of 65. The second term (B) accounts for the fact that, by age t, individuals will not have accrued all the DB pension wealth in A and scales it down by a factor equal to the ratio of the number of years worked in period t (yw_t) to the number of years assumed to be worked at the age of 65 (yw_{65}). Finally, the product of A and B is wealth that will not be realised until the age of 65, so to obtain its value in period t it is be discounted back using the mean return on DC wealth (term C).

E.3 Wealth profiles

Figure 16 gives mean wealth profiles (as a proportion of type-specific mean career-average earnings).

Figure 16: Total Wealth (ratio to type-specific average lifetime earnings)
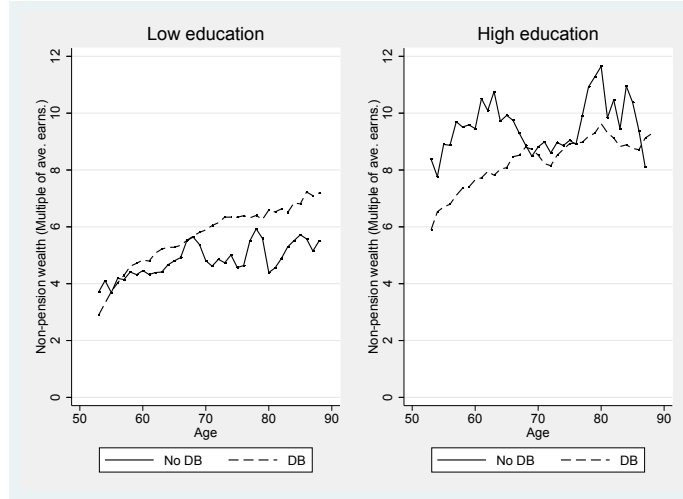
E.4 Representativeness of sample

Table 14 compares some characteristics of those households in the sample and those not in the sample. Mean household incomes and age are very similar between groups, though those in the sample have lower wealth levels and have (slightly) higher median levels of education.

Table 14: Representativeness of sample

	Not in sample	In sample
Mean household annual income	25,168	25,376
Mean age (of male)	58.7	58.6
Mean liquid wealth	139,013	94,189
Mean net housing wealth	159,175	141,155
Median age left full time education	15	16

F   Estimation of stochastic component of earnings

There are four parameters stochastic components of the earnings data generating process (10). These are $(\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2)$, respectively: the coefficient of autocorrelation in stochastic contribution to earnings, the variances of the first and subsequent innovations to that component, and the variance of measurement error.

These are estimated by standard minimum distance methods (see, e.g. Guvenen (2009), Low et al. (2010)) - by selecting those values of the parameters that bring two objects as close as possible (in a metric to be made explicit below). These are, first, the empirical variance-covariance matrix of $\hat{\eta}_{it}$ (the residuals estimated from equation (10)) and second, the theoretical variance covariance matrix of $(\eta_t = u_t + m_t)$ - the sum of the stochastic component of earnings and measurement error in earnings (in equation (10)). Before making explicit the distance to be minimised, it is worth outlining the latter (theoretical) variance-covariance matrix.

Noting that u and v are mutually-independent, the variance and autocovariances of η are:

$$var(\eta_t) = var(u_t) + var(m_t) \tag{26}$$

$$cov(\eta_t, \eta_{t+j}) = cov(u_t, u_{t+j}) + cov(m_t, m_{t+j}) \tag{27}$$

The initial variance of u and subsequent variances are respectively given by (the latter defined recursively):

$$var(u_1) = \sigma_\zeta^2$$

$$var(u_t) = \rho^2 var(u_{t-1}) + \sigma_\xi^2 \quad \forall \ t \geq \ 2$$

The autocovariance of order 1 and autocovariances of greater order of u are respectively given by (the

latter defined recursively):

$$cov(u_t u_{t+1}) = \rho var(u_t)$$

$$cov(u_t u_{t+j}) = \rho cov(u_t u_{t+j-1}) \quad \forall \ t \geq \ 2$$

The variance and autocovariance of the *iid* measurement error m are:

$$var(m_t) = \sigma_m^2$$

$$cov(m_t m_{t+j}) = 0$$

Equations (26) and (27) (with the terms they contain defined in the equations that follow) define the elements of the theoretical variance covariance matrix ($\boldsymbol{\Sigma}^{theor}$). Define $\mathbf{F}$ as the vector that gives the difference between each element in this matrix and the empirical variance covariance matrix ($\boldsymbol{\Sigma}^{data}$). $\mathbf{F}$ takes as arguments the four parameters of the stochastic component of earnings (including measurement error):

$$\mathbf{F}(\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2) = \boldsymbol{\Sigma}^{theor}(\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2) - \boldsymbol{\Sigma}^{data}$$

Define $\mathbf{A}$ as the variance covariance matrix of $\mathbf{F}$ (equivalently the variance covariance matrix of $\boldsymbol{\Sigma}^{data}$, calculated by bootstrapping). The parameters of the earnings process are selected to minimise the following objective function:

$$\min_{\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2} \mathbf{F}'(\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2) \mathbf{A}^{-1} \mathbf{F}(\rho, \sigma_\zeta^2, \sigma_\xi^2, \sigma_m^2)$$

G   Computational appendix

The following subsections discuss, respectively, the numerical procedures used to solve the household's decision problem (and simulate household behaviour) and those used to estimate preference parameters.

G.1   Model solution and simulation of optimal behaviour

This section outlines a) how the households' maximisation problem is solved to obtain decision rules and b) how these decision rules are used to simulate behaviour.

a) Solution

There is no analytical solution to the maximisation problem outlined. Decision rules are obtained numerically by iterating on the value function from the final period of life. Consider a household that has annuitised their pension and is aged 100. In this case, death is certain in the next period ($s_{t+1} = 0$) and the problem outlined in (8) reduces to:

$$V_{100}(\mathbf{X}_{100}|h_t = 1) = \max_{c_{100}^{nh}, l_{100}} \left(u(c_{100}, l_{100}) + b(a_{101}^b) \right) \tag{28}$$

subject to the intertemporal budget constraint on non-pension wealth.

Both the utility function $u(.)$ and the bequest function $b(.)$ are known and therefore at any point in $\mathbf{X}$ it is possible to carry out this maximisation (details given below) and obtain $c_{100}(\mathbf{X}_{100})$ and $l_{100}(\mathbf{X}_{100})$, the consumption and leisure policy functions and $V_{100}(\mathbf{X}_{100})$, the associated value function at those points. The knowledge of $V_{100}(\mathbf{X}_{100})$ at a subset of points in $\mathbf{X}$, combined with approximation methods (also discussed below), yields an approximation of $V_{100}(\mathbf{X}_{100})$ ($\hat{V}_{100}(\mathbf{X}_{100})$) at each point in $\mathbf{X}$.

With an approximation to $V_{100}(\mathbf{X}_{100})$ in hand, approximations to the 99th period policy and value functions can found ($\hat{c}_{99}(\mathbf{X}_{99})$, $\hat{l}_{99}(\mathbf{X}_{99})$, and $\hat{V}_{99}(\mathbf{X}_{99})$), again at a subset of points in the state space in that period, by solving the following functional equation:

$$
\begin{aligned}
\hat{V}_{99}(\mathbf{X}_{99}|h_t = 1) = \max_{c_{99}^{nh}, l_{99}} \Bigg(u(c_{99}, l_{99}) \;\; + \;\; & \beta s_{100}^m s_{100}^f \int \hat{V}_{100}(\mathbf{X}_{100}|1) dF(\phi_{100}, ue_{100}, e_{100}|ue_{99}, e_{99}) \\
+ \;\; & \beta s_{100}^m (1 - s_{100}^f) \int \hat{V}_{100}(\mathbf{X}_{100}|2) dF(\phi_{100}, ue_{100}, e_{100}|ue_{99}, e_{99}) \\
+ \;\; & \beta (1 - s_{100}^m)(s_{100}^f) \int \hat{V}_{100}(\mathbf{X}_{100}|3) dF(\phi_{100}, ue_{100}, e_{100}|ue_{99}, e_{99}) \\
+ \;\; & (1 - s_{100}^m)(1 - s_{100}^f) b(a_{100}^b) \Bigg)
\end{aligned}
\tag{29}
$$

This iterative process is repeated until the beginning of working life (age 20) is reached. For ages before 65, there is additionally, a decision to be made over how much to save in a DC fund.

Four features of the solution procedure will be detailed in the following discussion. These are i) the discretisation of the continuous variables, ii) the integration of the value function, iii) the approximation method for evaluating the value function at points outside the discretised state space and iv) how the optimisation is carried out.

Discretisation of state and control variables   The model has five continuous state variables that need to be discretised. These are productivity, average earnings, non-pension wealth, DC wealth and pension income. Earnings are placed on a grid (that has 10 elements) using a procedure of Tauchen (1986). Average earnings, non-pension wealth, DC wealth and pension income are discretised in a manner that gives smaller gaps between successive entries on the grid at lower levels. This is as the curvature of the value function (with respect to those state variables) will be greater at lower realisations of these states. 30 discrete points are used for each of cash assets, pension wealth and pension income and 15 discrete points are used for average earnings.

There are three control variables in the model. Whether to work, how much to save in a DC pension and how much to save in non-pension wealth. The first of these is naturally discrete. The other two (DC pension saving and consumption) are continuous but households are allowed to choose only from a finite set of options[40]. The proportion of earnings that is saved in a DC pension is restricted to take one of 6 values. These are 0%, 5%, 10%, 20%, 45%, 75%. At each maximisation, households can choose one of 100 consumption levels. These are 1%, 2% . . . 100% of their available resources.

Integration   Evaluation of the expectations in the households' problem involves integration of the value function over four stochastic variables. These are employment, earnings, survival and the return on DC funds. Realisations of survival, employment and earnings take one of a number of discrete outcomes – the first two as they are naturally discrete, the last as the procedure applied (Tauchen (1986)) allows earnings to take only a discrete subset of outcomes. Integration over the possible realisations of earnings and survival is therefore carried out by taking a weighted average of the value function realised at each possible outcome with the weights equal to the probability of that outcome. Realisations of the return on the DC fund are not restricted to a discrete subset. Integration over the distribution of possible outcomes is carried out using Gauss-Hermite quadrature with 10 nodes.

Approximation   It is required to evaluate the functions $V_t(.)$ at points in the state space other than those in the discrete sub-set of points in the discretised state space. Linear interpolation in multiple dimensions is used for this.

Optimisation   Optimisation is by grid search. There are up to 700 possible combinations of discrete choices in each period[41]. Utility (including expected utility in future periods) is calculated at each possible choice. That combination that yields the highest level of utility represents the households' optimal choice.

---

[40]The value function is not quasiconcave and so optimisation, as will be noted below, is implemented by grid search.

[41]This is less than $(2)(6)(100) = 1200$ ((work choices)(DC contribution choices)(Consumption choices)), as DC contributions are not allowed when an individual is not working.

b) Simulation

Once decision rules $\hat{c}_t(\mathbf{X}_t)$, $\hat{l}_t(\mathbf{X}_t)$, $\hat{dc}_t(\mathbf{X}_t)$ are obtained the behaviour of 5,000 households is simulated. The procedure is as follows:

1. Set initial values for state variables at the beginning of working life (age 20). Initial values are zero for non-pension wealth, DC wealth and average earnings. Employment and productivity are drawn randomly for each household. Pension income is zero 0 at the start of working life (and at all ages to 65. For household composition both members of the couple are assumed to be alive and in a couple at that age.

2. Using these values for the state variables and knowledge of the household's type optimal consumption, DC pension contribution, labour supply and annuitisation behaviour are obtained. Recall that optimal decisions have only been calculated for a particular subset of points in the state space. In most time periods, households will have realisations of the state variables that are off this grid. If the particular combination of state variables of the household is not on the grid, household labour supply and DC pension saving decisions are those that would be taken if they had that combination of state variables on the grid that is 'nearest' (in terms of absolute distance) to their realised state variables. Optimal consumption is then obtained by linear interpolation unless this would involve interpolating using grid points which imply different discrete choices, in which case optimal consumption is obtained by maximising the value function, conditional on the optimal discrete choice.

3. Once period 20 decisions are calculated, the new state variables for age 21 are obtained as follows:

   (a) Non-pension wealth in period 21 will follow from the consumption and saving decisions at age 20 along with the intertemporal budget constraint equation (7).

   (b) Defined Contribution pension wealth in period 21 will be given by the DC-wealth intertemporal budget constraint (6), accruing a return that is a random draw from the distribution of pension fund returns

   (c) Average earnings update according to the followings formula: $ae_{t+1} = \frac{tae_t + e_{t+1}}{t+1}$

   (d) Employment, productivity and household composition (mortality) are random draws from their (conditional on age 20 realisations) respective distributions.

4. Steps 2 and 3 are repeated to obtain optimal behaviour and state variables at each age up to age of 100.

G.2 Estimation of preference parameters

The procedure for estimating β, υ, $r^{h_{houscon}}$ and θ is to select them so that moments in the data are as close as possible to similarly-defined simulated moments. Define **H** as the vector that gives the difference between each of the modelled moments and those in the data. **H** takes the four preference parameters as arguments and can be calculated using simulated data obtained using the procedure outlined in the previous subsection. Define **C** as a diagonal matrix containing the variances of each of those moments. The preference parameters are selected to minimise the objective function given on the right hand side of:

$$\min_{\beta,\upsilon,r^{h_{houscon}},\theta} \mathbf{H}(\beta,\upsilon,r^{h_r},\theta)\mathbf{C}^{-1}\mathbf{H}(\beta,\upsilon,r^{h_r},\theta)$$

To keep notation to a minimum these objects are not subscripted by household type j, though each type has their own empirical moments and their own DGP for the simulated moments - and the estimation is carried out separately for each type. The use of the diagonal weighting matrix C rather that the 'optimal' weighting matrix (the full variance covariance matrix of H) is as the latter has been shown to have poor small sample properties (Altonji and Segal (1996)).

The minmisation is carried out using the Nelder-Mead simplex method (Nelder and Mead (1965)). The objective function has a number of local minima. By starting the algorithm at a variety of starting points, it appears that the parameters reported here obtain the global minimum.

H   Model Fit

Figure 17 and 18 show respectively the 25th percentile and the 75th percentile of endogenously modelled wealth (the sum of Defined Contribution wealth and non-pension wealth). These moments are not used in the estimation of the model parameters.

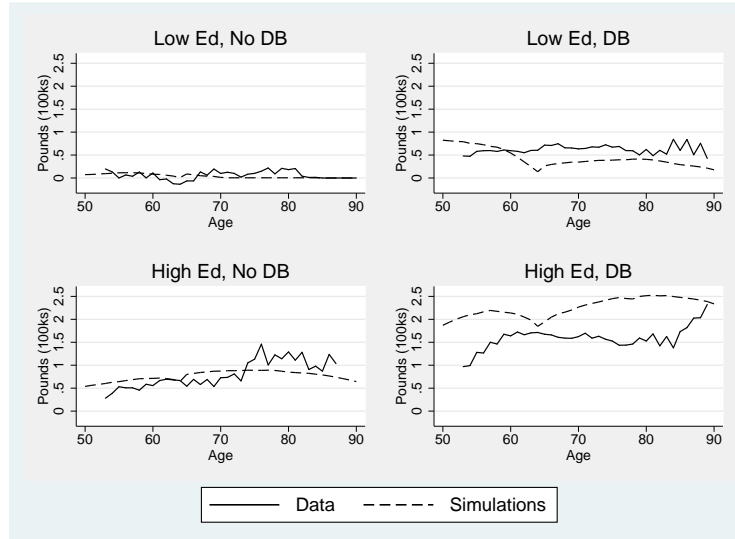Figure 17: Modelled (DC and non-pension) wealth, p25

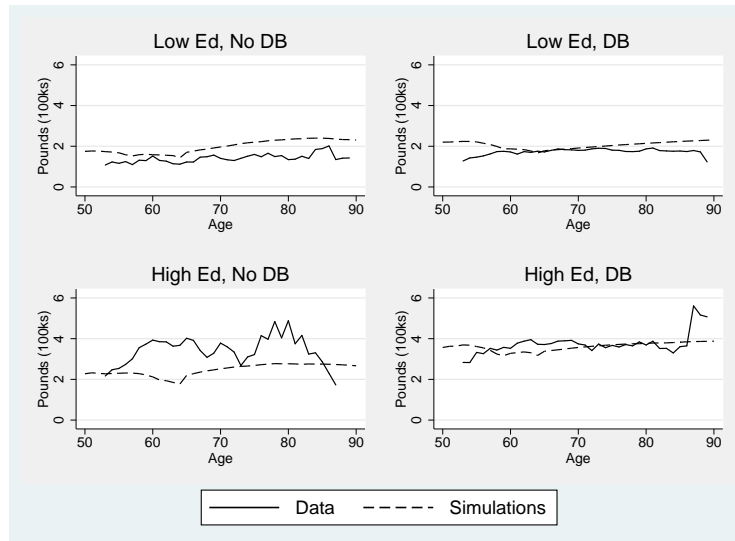Figure 18: Modelled (DC and non-pension) wealth, p75

Figure 19 shows mean consumption over the lifecycle. The consumption data (which is not used in the estimation procedure) is taken from the UK's Living Costs and Food Survey which doesn't contain information about pension holdings - therefore, the split given pools the pension types within education group.

Figure 19: Mean consumption over the lifecycle
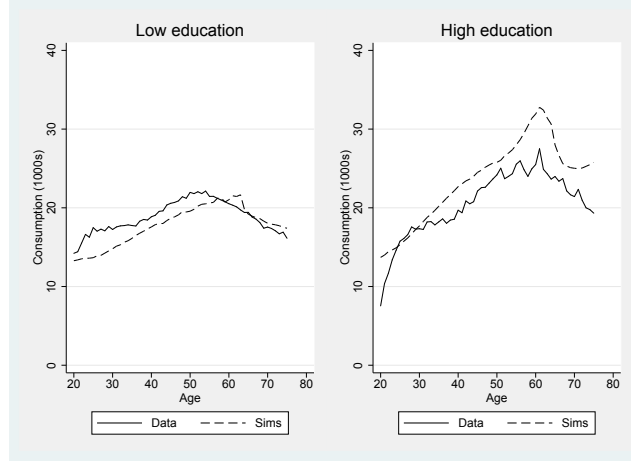
I Additional detail on design problem in counterfactual experiments

I.1 Expressing the value function as a function of policy functions

Households get utility either from the period utility function (equation (1)) or the bequest function (equation (2)), depending on household composition (whether there is a household member alive or whether the final household member has just died). I define, in equation (30) a function w which returns, depending on household composition the value given by the relevant utility function – the value of the period utility when at least one household member is alive (t with $h_t \leq 3$) and the value of bequests in the period immediately following death (t with $h_t = 4$ and $h_{t-1} < 4$). $\mathfrak{s}_t \in \mathfrak{S}_t$, which enters the function w is a history of realisations of all stochastic variables (unemployment shock, productivity, investment returns and mortality) and so contains both h_t and h_{t-1}.

$$
w_j\left(c_t, l_t, a_t^b, \mathfrak{s}_t\right) = \begin{cases} u_j(c_t, l_t) & \text{if } h_t \leq 3 \\ b_j(a_t^b) & \text{if } h_t = 4 \text{ and } h_{t-1} < 4 \\ 0 & \text{otherwise} \end{cases} \tag{30}
$$

Let $q_t(\mathfrak{s}_t)$ be the cumulative distribution function of $\mathfrak{s}_t$. Let $c_t(\mathfrak{s}_t)$, $l_t(\mathfrak{s}_t)$ and $beq_t(\mathfrak{s}_t)$ be the policy functions – the mappings from any state of the world at age t to, respectively, optimal consumption, optimal leisure and optimal bequests. I can now give (in equation (11)) the expression for expected utility. $\tilde{V}_0^j$ in equation (31) represents the utility that such a household can expect immediately after its type, j, has been revealed but before the realisation of any other stochastic components.

$$\tilde{V}_0^j(\mathbf{c}, \mathbf{l}, \mathbf{a^b}) = \Sigma_{t=1}^T \beta^t \int w_j \left(c_t(\mathfrak{s}_t), l_t(\mathfrak{s}_t), beq_t(\mathfrak{s}_t), \mathfrak{s}_t \right) q_t(\mathfrak{s}_t) ds_t \tag{31}$$

I.2 The government's optimisation problem

I.2.1 Policy instruments

Let mt be the level of the means-tested old-age income floor in the model's baseline. The government chooses λ_{mt}, where $\lambda_{mt} mt$ is a new level of the income floor. $\lambda_{mt} = 1$ represents the pre-reform level of the income floor, while the two reforms illustrated in Figure 6 would have $\lambda_{mt} = 0.75$ and $\lambda_{mt} = 1.25$. The government will also choose the level of one of four instruments (in four different experiments) to ensure that the budget is balanced. As outlined in Section 5.2, these instruments are i) a proportional tax/subsidy on all income, ii) the contributory public pension, iii) the basic rate of tax[42] and iv) both the UK's main income tax rates. The component of the tax and benefit system that the government is choosing to ensure reforms are revenue neutral is denoted below by tax.

I.2.2 Constraint

Only reforms that maintain the pre-reform government budget balance are allowed. Equation (32) gives R^0, the revenue collected by the government under the base tax and benefit function. $\tau_{it}^0(.)$ represents the net taxes paid by individual i in period t and depends on household behaviour and all the arguments of the tax system (which are not listed). $\bar{\phi}$ is the interest faced by the government - which I set equal to the mean return on the risky assets in which DC pensions are invested.

$$\Sigma_t \Sigma_i \frac{\tau_{it}^0(.)}{(1 + \bar{\phi})^{T-t}} = R^0 \tag{32}$$

When a new income floor is chosen, the level of the budget balance instrument (tax) must be chosen so that household behaviour and new tax function $\tau^1(tax, .)$ where tax is the only argument made explicit, satisfies:

$$\Sigma_t \Sigma_i \frac{\tau_{it}^1(tax, .)}{(1 + \bar{\phi})^{T-t}} = R^0$$

I.2.3 Government's problem

The problem solved by the government is to maximise a weighted-average over j (where the weights - sw_j - determine a social welfare function) of Δ_j, the Consumption/Leisure/Bequest Equivalent Variations. In

---

[42]This is a tax rate levied at 20% on earnings between the 15th and 85th percentile of earnings.

the design problem the weights are set equal to $\frac{1}{4}$ - that is a utilitarian social welfare function in a world where there are equal numbers of each types of household. The constraint on that maximisation problem is that, under any reform, the government balance must be equal to R^0.

$$\Omega \quad = \quad \max_{\lambda_{mt},tax} \Sigma_{j=1}^4 sw_j \Delta_j$$
$$s.t.$$
$$\Sigma_t \Sigma_i \frac{\tau_i^1(tax)}{(1+\phi)^{T-t}} \quad = \quad R^0$$

J Additional tables on counterfactual experiments

To help understand where the welfare effects shown in Table 4 are coming from, it is possible to decompose the CLBEV into an effect coming through each of consumption, leisure and bequests. The decomposition is in the spirit of that of Conesa et al. (2009), extended slightly to allow for bequests.[43] Table 15 gives these effects. Falls in the level of consumption reduce utility for each type, with offsetting increases in utility coming from changes in the distribution of consumption for all but the low educated households who have a DB pension.[44] These effects on the level and distribution of consumption reflect, respectively, the distortions induced by a more generous income floor and the additional insurance that it provides. The effect on utility of the labour-supply distortions, however, are mitigated by increases in the quantity

---

[43]Let Δ^c, Δ^l and Δ^b be defined as:

$$\tilde{V}_0(c^{post}, l, beq) = \tilde{V}\big((1+\Delta^c)c, (1+\Delta^c)l, (1+\Delta^c)beq\big)$$
$$\tilde{V}_0(c^{post}, l^{post}, beq) = \tilde{V}\big((1+\Delta^l)c, (1+\Delta^l)l, (1+\Delta^l)beq\big)$$
$$\tilde{V}_0(c^{post}, l^{post}, beq^{post}) = \tilde{V}\big((1+\Delta^b)c, (1+\Delta^b)l^1, (1+\Delta^b)beq\big)$$

Consumption will be further decomposed into an effect coming from changes in the level and changes in the distribution: Let $\Delta^{c,lev}$ and $\Delta^{c,dist}$ be defined as:

$$\tilde{V}_0(\hat{c}^{01}, l, beq) = \tilde{V}_0\big((1+\Delta^{c,lev})c, (1+\Delta^{c,lev})l, (1+\Delta^{c,lev})beq\big)$$
$$\tilde{V}_0(c^{post}, l, beq) = \tilde{V}_0\big((1+\Delta^{c,dist})\hat{c}^{01}, (1+\Delta^{c,dist})l, (1+\Delta^{c,dist})beq\big)$$

where $\hat{c}^{01} = (1 + g_c)c$ is the consumption allocation in the base scenario scaled by the change in aggregate consumption between the base and the reform scenarios $(1 + g_c)$. If the bequest function had the same curvature as the utility function (i.e. if $K = 0$), it would be the case that $\Delta = (1+\Delta^c)(1+\Delta^l)(1+\Delta^b)$ or $\Delta \approx \Delta^c + \Delta^l + \Delta^b$. The fact that the bequest function has a different curvature from the utility function means that the equality will not be exact, but in all reforms analysed here it very nearly holds.

[44]For this group of households, the (slightly) negative contribution to the change in utility coming from the distribution of consumption is due to the fact that those households with the lowest realised resources choose to respond to more generous income floor by increasing their leisure to a greater extent than their consumption.

of leisure enjoyed. The effect is largest for the low educated/no DB type, for whom the distortions are greatest. All types suffer utility losses from leaving lower bequests (as a more generous income floor reduces the extent to which households hold bequeathable assets).

Table 15: Decomposition of welfare effects in Table 4

	Type				
	Low Ed		High Ed		Average
	No DB	DB	No DB	DB	
Total Change	0.86	-0.04	0.36	-0.54	0.16
Consumption Level	-0.99	-1.17	-2.86	-1.80	-1.70
Consumption Distribution	0.19	-0.17	3.20	0.85	1.01
Leisure	2.00	1.54	0.56	0.62	1.18
Bequest	-0.32	-0.21	-0.45	-0.19	-0.29

Table 16 shows a decomposition of the welfare effects reported in Table 6 which are gained by moving to the optimal old age income floor when private pension subsidies are reformed.

Table 16: Decomposition of welfare effects in Table 6

	Type				
	Low Ed		High Ed		Average
	No DB	DB	No DB	DB	
Total Change	0.96	0.59	0.58	0.33	0.61
Consumption Level	-0.99	0.15	-1.63	-0.48	-0.74
Consumption Distribution	0.87	-0.20	2.09	0.59	0.84
Leisure	1.30	0.74	0.45	0.30	0.70
Bequest	-0.21	-0.09	-0.29	-0.07	-0.17

References

Alan, S. and M. Browning (2010). Estimating intertemporal allocation parameters using synthetic residual estimation. *The Review of Economic Studies 77*(4), pp. 1231–1261.

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

Auerbach, A. J. and L. J. Kotlikoff (1987). *Dynamic fiscal policy.* Cambridge University Press.

Bank of England (2013). Interest rate data. Accessed 19Nov2013.

Banks, J., C. Emmerson, Z. Oldfield, and G. Tetlow (2005). *Prepared for retirement? The adequacy and distribution of retirement resources in England.* Number R67. IFS Reports, Institute for Fiscal Studies.

Barclays Capital (2012, March). The equity gilt study 2012. Technical report, Barclays Capital, London.

Blau, D. M. (2016). Pensions, household saving, and welfare: A dynamic analysis of crowd out. *Quantitative Economics 7*(1), 193–224.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital, and welfare reform. *Econometrica 84*(5), 1705–1753.

Bozio, A., R. Crawford, C. Emmerson, and G. Tetlow (2010). Retirement outcomes and lifetime earnings: descriptive evidence from linked elsa-ni data. DWP Working Papers 81, Department for Work and Pensions.

Braun, R. A., K. A. Kopecky, and T. Koreshkova (2016). Old, sick, alone, and poor: a welfare analysis of old-age social insurance programs. *Review of Economic Studies.*

Conesa, J. C., S. Kitao, and D. Krueger (2009). Taxing capital? not a bad idea after all! *American Economic Review 99*(1), 25–48.

Conesa, J. C. and D. Krueger (1999). Social security reform with heterogeneous agents. *Review of Economic dynamics 2*(4), 757–795.

Crawford, R. and C. O'Dea (2014, Sep). Cash and pensions: Have the elderly in england saved optimally for retirement.

De Nardi, M., E. B. French, and J. B. Jones (2010). Why do the elderly save? the role of medical expenses. *Journal of Political Economy 118*(1), 39–75.

De Nardi, M., S. Imrohoroğlu, and T. J. Sargent (1999). Projected us demographics and social security. *Review of Economic dynamics 2*(3), 575–615.

Deaton, A. S. and C. Paxson (1994). Saving, growth, and aging in taiwan. In *Studies in the Economics of Aging*, pp. 331–362. University of Chicago Press.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review 100*(3), pp. 1238–1260.

French, E. (2005, 04). The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies 72*(2), 395–427.

Golosov, M., A. Shourideh, M. Troshkin, and A. Tsyvinski (2013). Optimal pension systems with simple instruments. *The American Economic Review 103*(3), 502–507.

Golosov, M. and A. Tsyvinski (2006). Designing optimal disability insurance: A case for asset testing. *Journal of Political Economy 114*(2), 257–279.

Gourinchas, P.-O. and J. A. Parker (2002, January). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Greene, W. (2004, 06). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal 7*(1), 98–119.

Guvenen, F. (2009, January). An empirical investigation of labor income processes. *Review of Economic Dynamics 12*(1), 58–79.

HMRC (2015). Registered pension schemes: cost of tax relief. mimeo., Her Majesty's Revenue and Customs.

Hood, A. and L. Oakley (2014, Nov). A survey of the gb benefit system. Technical report.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995, April). Precautionary saving and social insurance. *Journal of Political Economy 103*(2), 360–99.

Huggett, M. and J. C. Parra (2010). How well does the us social insurance system provide social insurance? *Journal of Political Economy 118*(1), 76–112.

Huggett, M. and G. Ventura (1999). On the distributional effects of social security reform. *Review of Economic Dynamics 2*(3), 498–531.

Kitao, S. (2014). Sustainable social security: Four options. *Review of Economic Dynamics 17*(4), 756–779.

Kotlikoff, L. J., K. Smetters, and J. Walliser (1999). Privatizing social security in the united states - comparing the options. *Review of Economic Dynamics 2*(3), 532–574.

Lee, S. Y. T., N. Roys, and A. Seshadri (2015). The causal effect of parents' education on children's earnings. Technical report, Mimeo.

Low, H., C. Meghir, and L. Pistaferri (2010, September). Wage Risk and Employment Risk over the Life Cycle. *American Economic Review 100*(4), 1432–67.

Low, H. and L. Pistaferri (2015, October). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review 105*(10), 2986–3029.

Nardi, M. D. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies 71*(3), pp. 743–768.

Nationwide Building Society (2014). House price index data. Accessed 1Feb2014.

Nelder, J. A. and R. Mead (1965). A simplex method for function minimization. *The computer journal 7*(4), 308–313.

Neumark, D. and E. Powers (1998, May). The effect of means-tested income support for the elderly on pre-retirement saving: evidence from the ssi program in the u.s. *Journal of Public Economics 68*(2), 181–206.

Neumark, D. and E. Powers (2000, October). Welfare for the elderly: the effects of ssi on pre-retirement labor supply. *Journal of Public Economics 78*(1-2), 51–80.

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*(1), pp. 1–32.

Nishiyama, S. (2011). The budgetary and welfare effects of tax-deferred retirement saving accounts. *Journal of Public Economics 95*(11), 1561–1578.

Nishiyama, S. and K. Smetters (2007). Does social security privatization produce efficiency gains? *The Quarterly Journal of Economics*, 1677–1719.

Pope, T. and B. Roantree (2014). A survey of the uk tax system. Institute for Fiscal Studies Briefing Note 9, Institute for Fiscal Studies.

Rendahl, P. (2012). Asset-based unemployment insurance. *International Economic Review 53*(3), 743–770.

Sefton, J. and J. Van De Ven (2009). Optimal design of means tested retirement benefits*. *The Economic Journal 119*(541), F461–F481.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters 20*(2), 177–181.

US Department of the Treasury (2013). Fiscal year 2014 tax expenditures. mimeo., US Department of the Treasury.